

Настоящая заявка является частичным продолжением одновременно рассматриваемой заявки на патент США 08/303.058, поданной 8 сентября 1994г.; которая является частичным продолжением заявки на патент США 08/127.420, поданной 27 сентября 1993г.: весь текст и чертежи их специально включены сюда как ссылка без дискламации. Правительство США имеет право на настоящее изобретение согласно гранту Министерства энергетики LDRD 03235 и договору номер W-31-109-ENG-38 между Министерством энергетики США и Чикагским университетом, представляющим Национальную лабораторию Аргонн.

Настоящее изобретение в общем относится к области молекулярной биологии. В частности, изобретение предусматривает создание новых способов и соединений, позволяющих в высшей степени эффективное образование последовательностей молекул нуклеиновой кислоты. Способы согласно этому изобретению пригодны для организации последовательности длинных молекул нуклеиновых кислот, включая хромосомы и РНК, с процессами клонирования и субклонирования.

В настоящее время образование последовательности нуклеиновых кислот составляет неотъемлемую часть научного прогресса. Определение последовательности, т.е. первичной структуры, молекул и сегментов нуклеиновых кислот важно в отношении отдельных проектов, исследующих целый диапазон конкретных прикладных областей. Информация об организации последовательности, влияет на науку, медицину, сельское хозяйство и на все области биотехнологии. Создание последовательности нуклеиновых кислот, конечно, жизненно важно для исследований по геному человека и для других крупномасштабных начинаний, цель которых - продвинуть наше понимание эволюции и функции организмов и помочь в понимании причин различных заболеваний.

Полезность организации последовательности нуклеиновых кислот очевидна. Например, Проект генома человека (ПГЧ), многонациональное начинание, посвященное созданию последовательности всего генома человека, осуществляется в различных центрах. Однако прогресс в этой области в основном и медленен и дорогостоящ. Организация последовательности нуклеиновых кислот обычно определяется на гелях полиакриламида, которые разделяют фрагменты ДНК в диапазоне 1-500 базовых пар с различием по длине на один нуклеотид. Реальное определение последовательности, т.е. порядок отдельных нуклеотидов А, G, C и T, можно достигнуть двумя способами. Во-первых, используя метод Максама и Гилберта химического расщепления фрагмента ДНК у конкретных нуклеотидов (Максам и Гилберт, 1977), [31], или, во-вторых, используя метод создания последовательности окончанием цепочки дидеокси, описанный Сангером и коллегами (Сангер и др. 1977), [41]. Оба метода трудоемки и занимают много времени.

Недавно были предложены другие способы создания последовательности нуклеиновых кислот, в которых не используется этап электрофореза, и эти способы можно в общем назвать созданием последовательности гибридизацией или СПГ (Дрманач и др., 1991, [13]; Кантор и др., 1992, [4]; Дрманач и Черквеняков, патент США 5.202.231), [11].

Развитие некоторых из этих методов вызвало появление новых инструментов создания последовательности типа твердой основы, известных как чипы создания последовательности. О полезности СПГ в основном свидетельствует тот факт, что по этой технологии были выданы патенты США. Однако - хотя СПГ потенциально может повысить скорость, с которой можно образовывать последовательности нуклеиновых кислот, все существующие методы СПГ имеют несколько недостатков.

СПГ можно проводить двумя основными способами, которые часто называют "формат 1" и "формат 2" (Кантор и др., 1992), [4]. В формате 1 олигонуклеотиды неизвестной последовательности обычно около 100-1000 нуклеотидов в длину, выстраиваются на твердой основе или фильтре, так что сами неизвестные образцы связываются (нейтрализуются) (Стречоска и др., 1991, [51]; Дрманач и Черквеняков, патент США 5.202.231), [1]. Реплики ряда затем "опрашиваются" гибридизацией наборами меченых проб длиной приблизительно 6-8 радикалов. В Формате 2 чип образования последовательности образуется из набора олигонуклеотидов с известными последовательностями с длиной приблизительно 6-8 радикалов (Сазерн, WO 89/10977), [48]; Храпко и др., 1991, [28]; Сазерн и др., 1992), [49].

Нуклеиновые кислоты неизвестной последовательности затем мечутся и им позволяют гибридизоваться к связанным олиго.

К сожалению, оба эти формата СПГ имеют несколько ограничений, в частности, требование предшествующих этапов клонирования ДНК. В формате 1 к числу других значительных проблем относится прикрепление разных частей нуклеиновой кислоты, которую нужно организовать в последовательность, к твердой поверхности опоры или приготовление большого набора более длинных проб. В формате 2 к числу основных проблем относятся метки нуклеиновых кислот неизвестной последовательности, обычно высокое получающееся отношение сигнал-шум и то, что можно определить только короткий последовательности.

Еще одна проблема в формате 2 это образование вторичной структуры, которая предотвращает доступ к некоторым целям, и разные условия, необходимые для проб с различным содержанием GC. Поэтому эта технология очевидным образом выигрывает от создания новой процедуры для организации последовательности нуклеиновых кислот и, в частности, такой процедуры, которая позволяет избежать утомительные процессы клонирования или/и субклонирования.

Настоящее изобретение создано для преодоления этих и других недостатков, присущих предыдущим разработкам, путем создания новых способов и соединений для образования последовательности нуклеиновых кислот. Новейшие методы, описанные здесь, обычно назывались изобретателями "форматом 3" и они представляют собой заметное улучшение по сравнению с существующими методами СПГ формат 1 и формат 2. В Формате 3 образования последовательности, обеспечиваемые этим изобретением, последовательности нуклеиновых кислот определяются посредством гибридизации с двумя наборами небольших проб олигонуклеотидов с известными последовательностями. Способ согласно изобретению позволяет получать высоко избирательную последовательность чрезвычайно больших молекул нуклеиновой кислоты, включая хромосомный материал или РНК, без предварительного клонирования, субклонирования или усиления. Кроме того, настоящие способы не требуют наличия большого количества проб, комплексного синтеза более длинных проб или мечения сложной смеси сегментов нуклеиновой кислоты.

Для того, чтобы определить последовательность нуклеиновой кислоты согласно способам этого изобретения, в основном находятся последовательности от нуклеиновой кислоты путем гибридизации с дополняющими последовательностями от двух наборов небольших проб олигонуклеотидов (олиго) с определенной длиной и известной последовательностью, которые включают в себя большинство комбинаций последовательностей для этой длины пробы. Затем анализируются найденные последовательности для определения перекрывающихся протяженностей найденных последовательностей и реконструируется или собирается полная последовательность нуклеиновой кислоты из таких перекрывающихся последовательностей.

Способы образования последовательности можно выполнять с использованием последовательной гибридизации с дополнительными последовательностями из двух наборов небольших олиго. В качестве альтернативы, можно применять способ, описываемый как "проведение цикла (чередование)", когда два набора небольших олиго гибридизуются одновременно с неизвестными последовательностями. Термин "проведение цикла" употребляется, поскольку избирательная часть метода возникает из последующего повышения температуры для "расплавления" тех гибридов, которые некомплементарны. Такие методы проведения цикла обычно применяются в других областях молекулярной биологии, таких, как PCR, и их легко поймут специалисты при чтении этого описания.

Это изобретение применимо для образования последовательности молекул нуклеиновой кислоты очень большой длины. На практике, молекула нуклеиновой кислоты, последовательность которой нужно образовать, в основном разделяется на фрагменты для получения фрагментов нуклеиновой кислоты малой или промежуточной длины, которыми легко манипулировать.

Понятие фрагмент нуклеиновой кислоты, как он здесь используется, обычнее всего означает молекулу нуклеиновой кислоты длиной приблизительно от 10 базовых пар (бп) до 100бп. Считается, что самые предпочтительные способы это те, в которых молекула нуклеиновой кислоты, последовательность которой нужно организовать, обрабатывается для получения фрагментов нуклеиновой кислоты промежуточной длины, т.е. приблизительно от 10бп до 40бп. Однако следует подчеркнуть, что настоящее изобретение это не способ полного образования последовательности небольших фрагментов нуклеиновой кислоты, скорее это способ образования последовательности молекул нуклеиновой кислоты как таковой, что предполагает определение частей последовательности изнутри молекулы - будь то осуществление с использованием целой молекулы, или, для простоты, достигается начальным фрагментированием молекулы на участки меньшего размера приблизительно от 4 до 1000 баз.

Последовательности из молекул нуклеиновой кислоты определяются путем гибридизации к небольшим пробам олигонуклеотидов с известной последовательностью. Когда говорят о "небольших пробах олигонуклеотидов", термин "небольшие" означает пробы с длиной менее 10бп и предпочтительно пробы длиной около 4бп и около 9бп. В одном примере образования последовательности особенно пригодными считаются пробы длиной около 6бп. Для наборов олиго, включающих в себя все комбинации последовательностей для выбранной длины пробыв их число будет представлено как 4^F , где F это длина пробы. Например, для 4-мера набор будет содержать 256 проб; для 5-мера набор будет содержать 1024 проб; для 6-мера - 4096 проб; для 7- мера - 16384 проб, и т.д. Синтез олиго такой длины является вполне рутинным в технике и его можно осуществлять автоматизированным синтезом.

В способах согласно этому изобретению один набор олигонуклеотидов с известной последовательностью, который можно назвать первым набором, будет прикрепляться к твердой основе, т.е. связываться (нейтрализоваться) на этой основе таким образом, что он может участвовать в реакциях гибридизации. Другой набор небольших проб олигонуклеотидов с известной последовательностью, который можно назвать вторым набором, это пробы, находящиеся в растворе и меченные обнаруживаемой меткой. Наборы олиго могут содержать пробы одинаковой длины или разных длин.

Процесс последовательной гибридизации означает, что молекулы нуклеиновой кислоты или фрагменты с неизвестной последовательностью можно гибридизировать к определенным наборам проб олигонуклеотидов с известными последовательностями в отдельные периоды времени (фиг.1). Молекулы или фрагменты нуклеиновой кислоты в основном будут денатурированы, позволяя гибридизацию, и будут добавляться к первому, связанному набору проб при условиях избирательной гибридизации с обеспечением того, что гибридизуются только фрагменты с дополняющими последовательностями. Фрагменты с некомплементарными последовательностями удаляются и затем проводится следующий цикл избирательной гибридизации путем добавления второго, меченного, набора проб в раствор к уже образованной комбинации фрагментов и проб. Меченные пробы, которые гибридизуются рядом с фиксированной пробой, останутся прикрепленными к основе и могут быть обнаружены, чего не происходит, когда имеется промежуток между фиксированными и мечеными пробами (фиг.1).

Процесс одновременной гибридизации означает, что молекулы нуклеиновой кислоты неизвестной последовательности могут контактировать с определенными наборами проб олигонуклеотидов с известными последовательностями одновременно.

Гибридизация происходит при условиях избирательной гибридизации. Затем фрагменты с некомплементарными (не дополняющими) последовательностями "расплавляются", т.е. удаляются при повышении температуры, и затем проводится следующий этап избирательной гибридизации, позволяющий гибридизоваться любым вторым дополнительным пробам. Затем меченные пробы, которые гибридизовались рядом с фиксированной пробой, обнаруживаются таким же образом.

Последовательности нуклеиновой кислоты, являющиеся "дополняющими (комплементарными)", это те, которые способны к спариванию без согласно стандартным правилам комплементарности Уотсона-Крика и вариантам этих правил, насколько они применимы к модифицированным базам. Это значит, что большие пурины или модифицированные пурины всегда будут спариваться по базе с меньшими пиримидинами с образованием только известных комбинаций. К ним относятся стандартный парис гуанина, спаренный с цитозином (G:C), и аденин, спаренный либо с тиминном (A:T), в случае ДНК, либо с урацилом (A:U), в случае

РНК. Также рассматривается использование модифицированных баз или так называемой "универсальной базы (основы)" (М.Николе и др, 1994), [35].

Используемый здесь термин "дополняющие (комплементарные)" последовательности означает последовательности нуклеиновой кислоты, которые в принципе являются дополняющими по всей своей длине и имеют очень мало несовпадений баз. Например, последовательности нуклеиновой кислоты с 6-ю базами по длине можно назвать дополняющими, если они гибридизируются в пяти их шести положении только с одним несовпадением. Естественно, последовательности нуклеиновой кислоты, которые являются "полностью дополняющими" это те последовательности нуклеиновой кислоты, которые являются полностью дополняющими по всей своей длине и не имеют несовпадений баз.

После идентификации путем гибридизации к олиго с известными последовательностями разных отдельных последовательностей, составляющие часть фрагментов нуклеиновой кислоты, затем эти отдельные последовательности анализируются для нахождения перекрывающихся протяженностей последовательностей. Например, участки последовательностей, в которых конец 5' тот же самый, что и коней 3' другом последовательности или наоборот, идентифицируются. Полная последовательность молекулы или фрагмента нуклеиновой кислоты затем может оконтуриваться, т.е. она может реконструироваться из определенных таким образом перекрывающихся последовательностей.

Процессы нахождения перекрывающихся последовательностей и реконструирования полной последовательности обычно определяются вычислительным анализом. Например, если меченная проба 5'-TTTTTT-3' гибридизируется к месту, содержащему фиксированную пробу 5'-AAAAAA-3', определяется 12-мерная последовательность изнутри молекулы нуклеиновой кислоты, а именно 5'-AAAAAATTTTTT-3' (поел. ид. N.1), т.е. последовательность двух гибридизированных проб комбинируется, чтобы обнаружить ранее неизвестную последовательность. Следующий вопрос, на который нужно ответить, это какой нуклеотид идет следующим после только что определенной последовательности 5'-AAAAAATTTTTT-3' (поел. ид. N.1). Имеется четыре возможности, представленные фиксированной пробой 5'-AAAAAT-3' и меченными пробами 5'-TTTTTA-3' для А; 5'-TTTTTT-3' для Т; 5'-TTTTTC-3' для С; и 5'-TTTTTG-3' для Г. Если, например, проба 5'-TTTTTC-3' положительна, а другие три отрицательны, то собранный последовательность расширяется до 5'-AAAAAATTTTTTC-3' (поел. ид. N.2). На следующем этапе алгоритм определяет, какие из меченных проб TTTTCA, TTTTCT, TTTTCC или TTTTCG положительны в месте, содержащем фиксированную пробу AAAATT.

Этот процесс повторяется пока все положительные (F+P) последовательности олигонуклеотидов не используются или не будут определены как ложно положительные.

Таким образом, это изобретение создает очень эффективный способ организовывать последовательности фрагментов и молекул нуклеиновой кислоты большой длины. Как здесь определено, большие молекулы нуклеиновой кислоты это такие молекулы, которые нужно разделять на фрагменты до создания последовательности. Их длина обычно составляет, как минимум, приблизительно 45 или 50 базовых пар (БП), а чаще всего и больше. Фактически, способы согласно изобретению можно использовать, чтобы организовывать последовательность молекул нуклеиновой кислоты практически без верхнего предела длины, так что можно образовывать последовательности около 100бп, 1 килобазы (кб), 100кб, 1 Мегабазы (Мб) и 50 Мб и более, до полных хромосом, включительно, таких, как хромосомы человека, имеющие в длину около 100Мб. Такое большое количество вполне находится в рамках настоящего изобретения, и образование последовательности такого количества баз потребует 2 набора 8-меров или 9-меров (так что F+P приблизительно равно 16-18). Подлежащие образованию последовательности нуклеиновые кислоты могут быть ДНК, например, сДНК, геномная ДНК микрорабоченные хромосомные полосы, космидные ДНК или вставки YAC, или могут быть РНК, включая mРНК, rРНК, tРНК или snРНК.

Процесс определение последовательности длинной молекулы нуклеиновой кислоты предполагает просто нахождение последовательностей с длиной F+P из молекулы и комбинирование последовательностей с использованием подходящего алгоритма. На практике, сначала скорее всего разделяют на фрагменты молекулу нуклеиновой кислоты, последовательность которой нужно сформировать, для получения меньших фрагментов, таких, как фрагменты нуклеиновой кислоты с промежуточной длиной. Затем находятся последовательности с длиной F+P гибридизацией, например, последовательной гибридизацией фрагментов к дополняющим последовательностям из 2-х наборов небольших проб олигонуклеотидов с известной последовательностью, как описано выше. Таким образом, можно реконструировать полную последовательность нуклеиновой кислоты чрезвычайно больших молекул из перекрывающихся последовательностей с длиной F+P.

Независимо от того, является ли нуклеиновая кислота, последовательность которой нужно организовать, сама фрагментом с промежуточной длиной или ее вначале обработали, чтобы получить фрагменты с такой длиной, процесс нахождения последовательностей из таких фрагментов нуклеиновых кислот путем гибридизации двух наборов небольших проб олигонуклеотидов с известной последовательностью является основным для описываемых здесь методов образования последовательности. Этот процесс обычно включает в себя следующие этапы:

а) приведение набора или ряда прикрепленных или связанных (нейтрализованных) проб олигонуклеотидов в контакт с фрагментами нуклеиновой кислоты в условиях гибридизации; эффективных для того, чтобы позволить фрагментам с дополняющей последовательностью достаточно гибридизоваться к пробе, что образует первичные комплексы, где фрагмент имеет как гибридизованные, так и негибридизованные, или "свободные", последовательности;

б) приведение первичных комплексов в контакт с набором меченных проб олигонуклеотидов в растворе в условиях гибридизации, эффективных для того, чтобы позволить пробам с дополняющими последовательностями гибридизоваться к негибридизованной или свободной последовательности фрагмента, что образует вторичные комплексы, где фрагмент гибридизован и у прикрепленной (связанной) пробе и к меченной пробе;

в) удаление из вторичных комплексов любых меченных проб, которые не были гибридизованы рядом с

прикрепленной пробе, что оставляет только соседние вторичные комплексы;

г) обнаружение соседних вторичных комплексов путем обнаружения наличия метки в меченной пробе; и

д) нахождение последовательностей олигонуклеотидов из фрагментов нуклеиновой кислоты в соседних вторичных комплексах путем комбинирования или соединения известных последовательностей гибридизированных и меченных проб.

Гибридизация или "условия промывки", выбранные для проведения одного или двух этапов гибридизации, может осуществляться согласно конкретному выбранному примеру образования последовательности. Например, оба условия гибридизации могут быть организованы так, чтобы позволить пробам олигонуклеотида гибридизироваться к данному фрагменту нуклеиновой кислоты, когда они содержат дополняющие последовательности, например, в принципе совпадающие последовательности, например, такие последовательности, которые гибридизируются у 5-ти или 6-ти положений. Предпочтительно этапы гибридизации проводят, используя простое робототехническое устройство, которое обычно применяется в существующих процедурах образования последовательностей.

В качестве альтернативы, условия гибридизации могут быть организованы так, чтобы позволить гибридизацию только тем пробам и фрагментам олигонуклеотидов, которые имеют полностью дополняющие последовательности. Эти более избирательные или "строгие" условия можно использовать для обоих отдельных этапов процесса последовательной гибридизации или только для каждого одного этапа. В таких случаях пробы олигонуклеотидов, будь то связанные или меченные пробы, смогут гибридизироваться только к данному фрагменту нуклеиновой кислоты, если они имеют общие полностью дополняющие последовательности с фрагментом.

Выбранные условия гибридизации в основном диктуют степень сложности, требуемую для анализа полученных данных. В равной степени, компьютерные программы, имеющиеся для анализа любых сформированных данных, могут диктовать условия гибридизации, которые нужно применять в данной лаборатории. Например, в самом избирательном процессе оба этапа гибридизации будут проводиться в условиях, позволяющих гибридизировать только олиго и фрагменты с полностью дополняющими последовательностями. Поскольку несовпадающих баз не будет, этот способ подразумевает наименее сложный вычислительный анализ и по этой причине это в настоящее время предпочтительным способом применения изобретения на практике. Однако применение менее избирательных условий для одного или обоих этапов- гибридизации также входит в рамки этого изобретения.

Подходящие условия гибридизации для применения на одном из этапов или на обоих могут обычно определяться процедурами оптимизации или первыми и-г серии исследований. Разные виды первых исследований обычно проводятся специалистами по организации последовательностей нуклеиновых кислот при установлении рабочих процедур и адаптации процедуры для данной лаборатории. Например, такие условия, как температура; концентрации каждого из компонентов; продолжительность времени этапов; применяемые буферы и их pH и прочность на ионы могут меняться и тем самым оптимизироваться.

В предпочтительных примерах осуществления способ образования последовательности нуклеиновой кислоты согласно этому изобретению включает в себя выборочным этап для отбора вторичных комплексов гибридизации, которые содержат непосредственно соседствующие связанные и меченные пробы, в отличие от тех, которые не являются непосредственно соседними и разделены одной, двумя или несколькими базами.

Имеется много разных процессов для удаления меченных проб, которые не гибридизировались непосредственно рядом с прикрепленной пробой, т.е. не гибридизировались задними сторонами, каждым из которых оставляет только непосредственно соседние вторичные комплексы.

Такие избирательные процессы могут основываться исключительно на этапах промывки с контролируемой строгостью, где применяемые условия гибридизации организованы так, что непосредственно соседние пробы остаются гибридизованными благодаря повышенной стабильности, создаваемой взаимодействиями между наборами соседних нуклеотидов. Опять же, такие условия промывки, как температура, концентрация, время, буферы, pH прочность на ионы и т.п. можно менять, чтобы оптимизировать удаление меченных проб, которые не являются непосредственно соседними.

В предпочтительных примерах осуществления изобретения непосредственно соседние связанные и меченные пробы будут связаны, т.е. объединены ковалентной связью до выполнения этапов промывки для удаления любых не образовавших связи проб. Образование связей может достигаться путем обработки раствором, содержащим химическое связующее вещество, например, растворимый в воде карбодимид или цианоген бромид. Предпочтительно можно использовать фермент лигазы, такой, как лигаза T4 ДНК из бактерии фага T4, который промышленно поставляется многими фирмами (например, Biolabs). В любом случае, затем можно будет удалить не непосредственно соседние меченные пробы посредством более строгих условий промывки, которые не могут воздействовать на ковалентно соединенные меченные и фиксированные пробы.

Остающиеся соседние вторичные комплексы будут обнаруживаться наблюдением за нахождением метки из меченных проб, присутствующих внутри комплексов.

Пробы олигонуклеотидов могут метиться химически обнаруживаемой меткой, такой как флуоресцентным красителем, или адекватно модифицированными для их обнаружения процедурами хемилюминесцентного проявления. или радиоактивными метками, такими, как ³⁵S, ³²P, ³³P или ³³P, хотя сейчас предпочтителен ³³P. Пробы могут также метиться нерадиоактивными изотопами и обнаруживаться масс-спектрометрией.

В настоящее время наиболее предпочтительный рассматриваемый способ осуществления на практике этого изобретения предполагает осуществление этапов гибридизации в условиях, предназначенных для того, чтобы позволить гибридизировать только тем пробам и фрагментам олигонуклеотидов, которые имеют полностью дополняющие последовательности и позволить оставаться гибридизованными только непосредственно соседним пробам. Этот способ вследствие этого требует наименее сложного вычислительного анализа.

Если молекула нуклеиновой кислоты с неизвестной последовательностью длинее, чем около 45 или 50бп, один эффективный способ определения ее последовательности обычно предполагает обработку молекулы для формирования фрагментов нуклеиновой кислоты промежуточной длины и определение последовательностей из фрагментов. Молекула нуклеиновой кислоты, будь то ДНК или РНК, может быть разделена на фрагменты любым из многочисленных способов, включая, например, разрезанием ограничительным ферментным вывариванием, сдвигом (скалыванием) физическими средствами, такими, как сверхзвуковая обработка, обработкой NaOH или сдвигом при низком давлении.

В некоторых примерах осуществления изобретения, например, при наличии небольших проб олигонуклеотидов длиной приблизительно от 4бп до 9бп можно иметь целью получить фрагменты нуклеиновой кислоты длиной приблизительно от 10бп до 40бп.

Естественно, более длинные пробы в основном будут использоваться в сочетании с образованием последовательности более длинных фрагментов нуклеиновой кислоты и наоборот. В некоторых предпочтительных примерах осуществления используемые небольшие пробы олигонуклеотидов будут иметь длину около 6бп, а фрагменты нуклеиновой кислоты, последовательность которых нужно организовать, будут иметь длину в основном около 20бп. При желании фрагменты можно разделять по размерам, чтобы получить фрагмента соответствующей длины, например, фрагмента могут обрабатываться на геле, таком, как гель агарозы, и фрагменты приблизительно желаемой длины могут отрезаться.

Способ определения последовательности молекулы нуклеиновой кислот можно, также пояснить на примере используя следующие условия. Вначале произвольно фрагментируется некоторое количество нуклеиновой кислоты, последовательность которой нужно образовать для получения смеси фрагментов нуклеиновой кислоты длиной Т. Приготавливается ряд связанных проб олигонуклеотидов с известными последовательностями и длиной F и набор меченных проб олигонуклеотидов в растворе с известными последовательностями и длиной Р, где $F + P = T$, предпочтительно когда Т приблизительно равно 3F.

Затем ряд связанных проб олигонуклеотидов приводится в контакт со смесью фрагментов нуклеиновой кислоты в условиях гибридизации, эффективных для того, чтобы позволить формирование первичных комплексов с гибридизованными, дополняющими последовательностями длиной F и негибридизованными последовательностями фрагментов длиной Т-F. Предпочтительно гибридизованные последовательности длиной F будут содержать только полностью дополняющие последовательности.

Затем первичные комплексы приводятся в контакт с набором меченных проб олигонуклеотидов в условиях гибридизации, эффективных для того, чтобы позволить образование вторичных комплексов с гибридизованными, дополняющими последовательностями длиной F и соседними гибридизованными, дополняющими последовательностями длиной Р. В предпочтительных примерах осуществления изобретения только меченым пробам с полностью дополняющими последовательностями будет позволено гибридизировать. И только тем пробам, которые гибридизуют непосредственно рядом со связанной пробой, будет позволено оставаться гибридизованными. В самых предпочтительных примерах на этом этапе соседние связанные и меченные пробы олигонуклеотидов будут также образовывать связи.

Затем обнаруживаются вторичные комплексы путем обнаружения наличия метки и идентифицируются последовательности длиной F+P из фрагментов нуклеиновой кислоты ко вторичных комплексах путем комбинирования известных последовательностей гибридизованных связанных и меченных проб. Затем перекрывающиеся протяженности последовательностей длины F+P будут идентифицироваться, что позволяет реконструировать или собирать полную последовательность нуклеиновой кислоты из определенных перекрывающихся последовательностей.

В способах согласно этому изобретению олигонуклеотиды первого набора могут прикрепляться к твердой основе, т.е. связываться (нейтрализовываться) любым из методов, известных специалистам. Например - прикрепление может производиться посредством адресуемой приводимой в действие лазером фотодепротекции (Фодор и др. 1991,[22]; Пиз и др., 1994,[36]). Один в общем случае предпочтительный способ эти прикрепление олиго через фосфатную группу с использованием таких реагентов, как нуклеозид фосфорамидит или нуклеозид водород фосфат, как описано Сазерном и Маскосом (заявки РСТ W090/03382),[48] с использованием основ из стекла, нейлона или тефлона.

Другой предпочтительный способ это метод генерированного светом синтеза, описанный Пизом и др. (1994),[36]. Можно также купить связанные с основой ряды олигонуклеотидов, например, такие, которые продаются фирмами Affymetrix и Beckman.

Связанные олигонуклеотиды могут образовываться в ряд, содержащий все пробы или подмножества проб с данной длиной (предпочтительна приблизительно 4-10 баз), и более предпочтительно во множество рядов) связанные олигонуклеотидов, расположенные с образованием так называемого "чипа создания последовательности". Один пример чипа это когда гидрофобные сегменты используются для создания отдельных пространственных участков. Чипы создания последовательности можно сконструировать для разных областей применения, таких, как планирование, частичное создание последовательности, создание последовательности намеченных участков для целей диагностики, создание последовательности mРНК и создание широкомасштабной последовательности генома. Для каждой области применения можно сконструировать конкретный чип с пробами разных размеров или с неполным набором проб.

В одном примере осуществления изобретения обоими наборами проб олигонуклеотидов будут пробы длиной в 6 баз, т.е. 6-меры. В этом случае каждый набор олиго содержит 4096 отдельных проб. Пробы первого набора предпочтительно фиксированы в ряду на микрочипе, наиболее удобно при размещении в 64 ряда и 64 столбца. Второй набор из 4096 олиго будет метиться обнаруживаемой меткой и распределяться на набор отдельных трубочек. В этом примере 4096 чипов будут комбинироваться в большой ряд или в несколько рядов. После гибридизации фрагментов нуклеиновой кислоты небольшое количество меченных олигонуклеотидов будет добавляться к каждому микрочипу для второго этапа гибридизации и только один из каждых 4096 нуклеотидов будет добавляться к каждому микрочипу.

К числу других объектов этого изобретения относятся наборы материалов для создания последовательностей нуклеиновых кислот. В основном такие наборы включают в себя твердую основу, и которой прикреплен ряд проб олигонуклеотидов с известными последовательностями, как показано на фиг.2А, фиг.2В и фиг.2С, где олигонуклеотиды способны участвовать в реакциях гибридизации, и набор емкостей, содержащих растворы меченных проб олигонуклеотидов с известными последовательностями. Рассматриваются также такие схемы, как те, которые представлены на фиг.4. Она описывает использование "универсальной базы", либо как способа прикрепления или у конечного положения, чтобы придать новое измерение гибридизации фрагментов.

В этих наборах материалов прикрепленные пробы олигонуклеотидов и пробы в растворе могут иметь длину приблизительно 4-9бп, причем предпочтительны пробы длиной около 6бп. Олиго могут быть мечены химически обнаруживаемыми или радиоактивными метками, причем в общем случае предпочтительны пробы, меченные ³²P, и еще более предпочтительна пробью меченные ³³P. Наборы материалов могут также включать в себя химическое или другое связующее вещество, такое, как фермент лигазы ДНК. В состав наборов материалов могут включаться разные другие соединения и материалы, такие, как устройства с 96 кончиками или 96 штырьками, буферы, реагенты для разрезания длинных молекул нуклеиновой кислоты и инструменты для выбора размеров фрагментов ДНК. Наборы материалов могут также включать меченные пробы РНК, так что пробы можно удалять обработкой РНКазы и повторно используемыми чипами создания последовательности.

На чертежах проиллюстрировано настоящее изобретение.

На Фиг.1 показаны основные этапы процесса гибридизации. Этап 1: Немеченная целевая ДНК-последовательность, в которой нужно образовать (Т), гибридизуется в избирательных условиях к ряду прикрепленных проб олигонуклеотидов. Описаны места с пробой Fx и Fy. Дополняющие последовательности для Fx и Fy находятся на разных положениях Т. Этап 2: Меченные пробы, Pi (одна проба на чип) гибридизуются к ряду. Описана проба, имеющая дополняющую цель на Т, которая соседствует с Fx, но не с Fy. Этап 3: Путем применения избирательных условий или реагентов селективно расплавляются комплексы без соседних проб. Конкретный пример это образование связи меченной пробы α фиксированной пробой, когда меченная проба гибридизуется "задней стороной к задней стороне" с прикрепленной пробой. Положительные сигналы обнаруживаются только в случае соседних проб, таких, как Fx и Pi, и в конкретном примере только в случае образовавших связи проб.

На Фиг.2А, Фиг.2В и Фиг.2С представлены компоненты примерного набора материалов для образования последовательностей.

Фиг.2А. Чипы создания последовательности, представляющие собой ряд 4^P тождественных секций, каждая из которых содержит тождественные (или разные) ряды олигонуклеотидов. Секции могут разделяться физическими барьерами или гидрофобными полосками. Считается, что в ряду имеется 4000-16000 олигонуклеотидов.

Фиг.2В - это вид в увеличении секции чипа, содержащей 4^F участков, каждый с конкретной пробой олигонуклеотида (4000-16000), синтезированной или установленной в этой месте. Участки могут быть размером всего в несколько микрон, а размер секции приблизительно 1-10мм.

Фиг.2С- представляет набор трубочек или одну или несколько пластин с множеством карманов с соответствующим числом карманов (в этом случае 4^P карманов). Каждым карманом содержит некоторое количество конкретно меченного олигонуклеотида.

Дополнительное количество проб можно хранить немеченными, если мечение не производится во время синтеза; в этом случае набор материалов для создания последовательности будет содержать необходимые компоненты для мечения проб. Линии, соединяющие трубочки/карманы и секциями чипа, описывают этап в процедуре организации последовательности, когда некоторое количество меченной пробы переносится в секцию чипа. Перенос может осуществляться пипеткой (простой или многоканальной) или рядом штырьков (шпилек), переносящих жидкость посредством поверхностного натяжения. Инструменты для переноса также могут включаться в набор материалов.

Фиг.3А, Фиг.3В и Фиг.3С. Гибридизация фрагментов ДНК, полученных произвольным разрезанием некоторого количества молекулы ДНК. На Фиг.3А фрагмент Т1 ДНК такой, что содержит полные цели для фиксированных и нефиксированных меченных проб. Фиг. 3В представляет случаи, когда фрагмент Т ДНК неправильно разрезан. На Фиг.3С имеется достаточно места для гибридизации пробы Р но соседняя последовательность не комплементарна ей. И в случае В, и в случае С сигнал будет уменьшен из-за насыщения молекул прикрепленной пробы F. Одновременная гибридизация фрагментами ДНК и меченными пробами и циклическое проведение процесса гибридизации это некоторые возможные способы увеличения выхода правильных соседних гибридизаций.

Фиг.4. Использование "универсальной базы" как средства связи или в конечном положении для гибридизации. Универсальные базы (М-база, Николо и др. 1994, [34] или все 4 базы можно добавлять в синтезе пробы. Это один из способов увеличения длины проб и, таким образом, стабильности дуплексов без увеличения количества проб. Кроме того, использование универсальных баз на свободном конце проб создает прокладку, которая позволяет прочитывать последовательность в другом кадре (фрейме).

Определение последовательностей молекул нуклеиновых кислот находит важное использование во всех фундаментальных и прикладных биологических исследованиях (Дрманач и Черквеняков, 1990, [10]). Настоящее изобретение создает новые и эффективные способы для создания последовательностей и анализа молекул нуклеиновой кислоты. Одно из намечаемых использований этой методологии, в сочетании с другими способами образования последовательностей, это использование в работе над Проектом генома человека (ПГЧ).

В настоящее время известны 2 метода создания последовательности гибридизацией (СПГ). В первом, формат 1, неизвестные геномные ДНК или олигонуклеотиды длиной до 100-200 выстраиваются на твердой подложке. Затем эти ДНК "опрашиваются" гибридизацией набором меченных проб, которые обычно 6-8-меры.

В обратном методе, Формат 2, олигомеры в 6-й нуклеотидов связываются (нейтрализуются) на твердой основе, и им позволяют отжигаться до частей клонированной и меченой ДНК.

В любом из 2-х видов анализа СПГ нужно включать много этапов, чтобы прийти к определенной последовательности. Конкретные проблемы, связанные с существующими методами СПГ это те, которые ассоциируются с синтезом большого числа проб и с трудностями эффективной избирательной гибридизации. Полное распознавание совпадения-несовпадения затруднено по двум основным причинам. Во-первых, конечное несовпадение проб длиннее чем 10 баз очень неизбирательно, и во-вторых сложная смесь печенных сегментов ДНК, которая получается при анализе длинных фрагментов ДНК, генерирует высокий фон.

Настоящее изобретение обеспечивает эффективную избирательную гибридизацию без большого числа проб или проб с увеличенной длиной и также исключает многие из этапов нанесения меток и клонирования, которые являются конкретными недостатками каждого из известных методов СПГ. Описываемые в высшей степени эффективные способы создания последовательности нуклеиновой кислоты, называемые созданием последовательности в формате 3, основаны на гибридизации с двумя наборами небольших проб олигонуклеотидов с известными последовательностями, и таким образом можно определять, как минимум, в два раза большую длину последовательности. Эти способы позволяют организовывать последовательность чрезвычайно больших молекул нуклеиновой кислоты, включая хромосомы, и решать разные другие проблемы СПГ, такие, как, например, прикрепление или мечение многих фрагментов нуклеиновой кислоты. Изобретение имеет чрезвычайно большую потенцию. т.к. его можно также использовать для создания последовательности РНК и даже неувеличенных образцов РНК.

После настоящего изобретения, как описано в заявке США 08/127.402 и у Дрманача (1994), [17], был представлен другой вариант СПГ, названный позиционным СПГ (ПСПГ) (Бруд и др., 1994), [2]. ПСПГ по сути является вариантом СПГ формата 2 (в котором олигонуклеотиды известных последовательностей связываются (нейтрализуются) и используются для гибридизации к нуклеиновым кислотам неизвестной последовательности, которые предварительно метились). В ПСПГ связанные пробы, не будучи простыми. односкрученными пробами, это дуплексы - содержащие односкрученные 3' свесы. Биотинилированные дуплексные пробы связываются на покрытых стрептавидином магнитных шариках с образованием некоторого вида связанной пробы и затем перемешиваются с 32p-мечеными целевыми нуклеиновыми кислотами, последовательность которых нужно создать.

Затем добавляется лигаза Т: ДНК для связывания любой гибридизированной целевой ДНК к более короткому концу дуплексной пробы.

Однако, хотя это представляет собой интересный подход, ПСПГ (как оно излагается Брудом и др., 1994), [2], не отражает значительного улучшения по сравнению с существующей технологией СПГ. Например, в отличие от методологии формата 3 этого изобретения, ПСПГ не удлиняет длины последовательности, которую можно определить за один цикл способа. В ПСПГ также поддерживается обременительное требование нанесении меток на неизвестную целевую ДНК, чего не требуется в формате 3. В общем, ПСПГ предлагается для использования в сравнительных исследованиях или в планировании, а не в новом создании последовательности генома. Таким образом, это значительно отличается от формата 3, который, хотя и широко применим во всех областях создания последовательностей, является очень мощным инструментом для образования последовательностей даже самых крупных геномов.

Нуклеиновые кислоты, последовательность которых нужно организовать, можно сначала разделить на фрагменты. Этого можно достигнуть любыми способами, включая, например, разрезание ограничительным ферментным вывариванием, особенно с Cvi I, как описано Фицджералдом и др. (1992), [21]; сдвигом физическими средствами, такими, как ультразвуковая обработка; обработкой NaOH и т.п. При желании фрагменты соответствующей длины, например, приблизительно 10-40bp, могут вырезаться из геля. Полная последовательность нуклеиновой кислоты первоначальной молекулы, такой, как хромосома человека, будет определяться путем определения последовательностей F+P, присутствующих в начальной молекуле, и сборки частей перекрывающихся последовательностей F+P.

Поэтому это не требует промежуточного этапа определения последовательностей фрагмента, скорее последовательность всей молекулы сконструируется из оконтуренных последовательностей F+P.

Для цели следующего обсуждения в общем случае будет приниматься, что 4 базы составляют последовательности нуклеиновых кислот, которые нужно организовать. Это А, G, C и T для ДНК и А, G, C и U для РНК. Однако может быть полезным в некоторых примерах осуществления использовать модифицированные базы в небольших пробах олигонуклеотидов. Чтобы выполнить это изобретение, сначала обычно готовится ряд небольших проб олигонуклеотидов определенной длины, который включают в себя все комбинации последовательностей для этой длины пробы. Это число представляется 4^N (4 в степени N), где длина пробы обозначена N. Например, имеются 4096 возможных последовательностей для 6-мерной пробы $4^6 = 4096$).

Один набор таких проб длиной F (4^F) будет фиксироваться в квадратном расположении на микрочипе, который может быть в диапазоне от 1мм^2 до 1см^2 . В настоящем примере они будут расположены в 64 ряда и 64 столбца. Естественно, обеспечивается, чтобы пробы олиго были прикреплены или иначе связаны с поверхностью микрочипа, так чтобы могли участвовать в реакциях гибридизации. Другой набор олиго с длиной P, числом 4^P будет тоже синтезирован. Олиго в этом наборе "P" будут метиться обнаруживаемой меткой и будут распределяться по набору трубочек (фиг. 2A, фиг.2B и фиг.2C).

4^P из чипов будут комбинироваться в большой ряд (или несколько рядов приблизительно $10\text{-}100\text{см}^2$ для удобных размеров); где P соответствует длине олигонуклеотидов во втором наборе олигомера (фиг.2B и фиг.2C). Снова, в качестве удобного примера, P выбирается как 6 ($P=6$).

Нуклеиновые кислоты, последовательности которых нужно организовать, будут разделяться на фрагменты с получением меньших фрагментов нуклеиновых кислот с неизвестной последовательностью. Средняя длина этих фрагментов, нарываемая T обычно должна быть больше, чем объединенная длина F и P и может приблизительно в три раза больше длины F (т.е. $F+P \leq T$ и T приблизительно равно $3F$). В

настоящем примере целью будет получение фрагментов нуклеиновой кислоты длиной приблизительно 30 базовых пар. Эти фрагменты будут денатурированы и добавляться к большим рядам в условиях, которые облегчают гибридизацию дополняющих последовательностей.

В самой простой и предпочтительной в настоящее время форме изобретения будут выбираться условия гибридизации, которые позволят произойти значительной гибридизации только если 6 последовательных нуклеотидов в фрагменте нуклеиновой кислоты являются дополняющими для всех 6 нуклеотидов F пробы олигонуклеотида. Такие условия гибридизации будут определяться обычными начальными исследованиями оптимизации, в которые определяются такие условия как температура концентраций разных компонентов, длительность этапов и используемые буферы, включая pH буфера.

На этой стадии каждый микрочип содержит определенные гибридизированные комплексы. Они будут в виде пробы: комплексы фрагмента, в которых вся последовательность пробы гибридизована к фрагменту, но в которых фрагмент, будучи более длинным, имеет некоторые негибридизованные последовательности, образующие "хвост" или "хвосты" комплекса. В этом примере дополняющие гибридизированные последовательности будут длиной F, а негибридизованные последовательности будут иметь общую длину T-F.

Дополняющая часть фрагмента может располагаться у соответствующего конца или по направлению к нему, так что можно образовать один более длинный негибридизованный хвост. В качестве альтернативы, дополняющая часть фрагмента может располагаться по направлению к противоположному концу, так что образуются 2 негибридизованных хвоста (фиг.3А, фиг.3В и фиг.3С).

После промывки с целью удаления недополняющих фрагментов нуклеиновой кислоты, которые не гибридизовали, небольшое количество меченных олигонуклеотидов в наборе Р добавляется к каждому микрочипу для гибридизации к хвостам фрагмента нуклеиновой кислоты неизвестной последовательности, которые выдаются из пробы: комплекса фрагмента. Только один из каждого из 4 нуклеотидов будет добавляться к каждому микрочипу. Сейчас предпочтительно использовать условия гибридизации, которые позволяют иметь место значительному образованию связей только, если все 6 нуклеотидов меченной пробы являются дополняющими для 6 последовательных нуклеотидов хвоста фрагмента нуклеиновой кислоты. Условия гибридизации определяются описанными выше начальными исследованиями, в которых оптимизируются такие элементы, как температура, концентрация, время, буферы и т.п.

На этом этапе каждый микрочип будет содержать определенные "вторичные гибридизированные комплексы". Они будут в форме проба: фрагмент: комплексы пробы, в которой вся последовательность каждой пробы гибридизована к фрагменту и в которой фрагмент, вероятно, имеет некоторые негибридизованные последовательности. В этих вторичных гибридизированных комплексах связанная проба и меченая проба могут быть гибридизованы к фрагменту, так что две пробы являются непосредственно соседними или расположены друг к другу задними сторонами.

Однако, учитывая то, что фрагменты будут в общем случае более длинными, чем сумма длин проб, связанная проба и меченная проба могут гибридизоваться к фрагменту в не соседних положениях, отделенных одной или несколькими базами.

Затем большие ряды обрабатываются посредством, процесса для удаления негибридизованных меченных проб. В предпочтительных примерах осуществления используемый процесс удалит из ряда не только негибридизованные меченные пробы, но и не соседние гибридизованные меченные пробы. В процессе используются избирательные условия, чтобы позволить этим вторичным комплексам гибридизации, которые содержат соседние связанные и меченные пробы, дискриминироваться из тех вторичных комплексов гибридизации, в которых фрагмент нуклеиновой кислоты гибридизован к 2-м пробам, но эти пробы не соседние. Это является важным аспектом изобретения в том смысле, что это позволит конечное оконтуривание секции последовательности фрагмента, соответствующей комбинированной последовательности связанной пробы и меченной пробы.

Процесс избирательности, используемый для удаления негибридизованных и не соседних гибридизованных проб из ряда при оставлении соседних гибридизованных проб прикрепленными может быть снова контролируемым процессом промывки. Соседние гибридизованные пробы не будут подвергаться влиянию выбранных условий из-за своей повышенной стабильности благодаря реакциям между наборами (слоями) соседних нуклеотидов. Однако в предпочтительных примерах осуществления будут обрабатываться крупные ряды, так что любые соседние пробы будут ковалентно связаны например, путем наработки раствором, содержащим химическое связующее вещество или, более предпочтительно, фрагмент лигазы, такой как лигаза T4 ДНК (Ландегрен и др. 1988, [30]: Ву и Уоллас, 1989, [52]).

В любом случае, полный ряд будет подвергаться строгому промыванию, так что единственная оставшаяся метка, связанная с рядом, будет в виде двойных скрученных комплексов проба-фрагмент-проба соседними гибридизованными частями с длиной F+P (т.е. 12 нуклеотидов и данным примере). С использованием этого двухэтапной реакции гибридизации возможна очень высокая избирательность, потому что в расчет принимаются 3 или 4 независимых избирательных процесса: избирательная гибридизация фрагмента Т к пробе длиной F баз; избирательная гибридизация пробы длиной Р баз к фрагменту Т; избирательная стабильность полного совпадения (F+T+P) гибрида по сравнению с Р гибридами или даже несовпадающими гибридами, содержащими не соседствующие пробы F+P; и избирательное образование связей 2-х конечных баз F и Р.

Затем обнаруживаются так называемые соседние вторичные комплексы путем наблюдения местоположения остающейся метки на ряде. С положения метки можно определить последовательность длиной F+P (например 12) нуклеотидов путем комбинирования известных последовательностей связанных (нейтрализованных) и меченных проб. Затем можно реконструировать или собрать полную последовательность нуклеиновой кислоты первоначальной молекулы, такой как хромосома человека, из определенных таким образом перекрывающихся последовательностей F+P.

Если образование связей используется в процессе создания последовательности, как сейчас

предпочтительно, то обычный чип олигонуклеотидов нельзя использовать повторно. Изобретатель считает, что это не будет ограничивать, т.к. имеются разные методы утилизации. Например, можно сформировать специфически легко расщепляющуюся связь между пробами и затем расщепить связь после обнаружения.

В качестве альтернативы, можно использовать рибонуклеотиды для второй пробы, пробы Р, или использовать рибонуклеотид для соединения базы в пробе Р, так что эту пробу можно затем удалить обработкой РНКазой или урацил-ДНК гликозилатом (Крейг и др., 1989), [17]. Другие рассматриваемые способы можно установить связи путем химического образования связей, которые можно, селективно обрывать (Долинный и др., 1988), [9].

Другие варианты и улучшения этой методологии создания последовательности также рассматриваются и входят в рамки этого изобретения. Сюда относится использование модифицированных олигонуклеотидов для повышения специфичности или эффективности способов, аналогичное описанному Хоханзелем и Лерахом (1990), [23]. Можно применять и гибридизации с проведением циклов для увеличения сигнала гибридизации, как это используется в технологии PRC. В этих случаях будут использоваться циклы с разными температурами для повторения гибридизации некоторых проб. Изобретение предусматривает также определение сдвигов в кадрах считывания с использованием равномолярных количеств проб, которые имеют разную базу у коцевого положения. Например, используя равномолярные 7-меры, в которых первые 6 баз это та же самая определенная последовательность, а последние положения могут быть А, Т, С или G альтернативе.

Включаются следующие примеры для демонстрации предпочтительных примеров осуществления изобретения. Специалисты должны понимать, что описываемые в следующих примерах методы представляют собой методы, открытые изобретателем для нормального функционирования при практическом использовании изобретения, и таким образом их можно считать составляющими предпочтительные режимы такого практического использования.

Однако специалисты должны понять в свете этого описания, что можно сделать много изменений в конкретных примерах осуществления, которые здесь описаны, и все же получить похожий или аналогичный результат, не отходя от духа и объема изобретения.

ПРИМЕР 1

ПРИГОТОВЛЕНИЕ СВЯЗАННЫХ С ОСНОВОЙ ОЛИГОНУКЛЕОТИДОВ

Олигонуклеотиды, т.е. небольшие сегменты нуклеиновой кислоты, можно легко приготовить, например, путем прямого синтеза олигонуклеотида химическими средствами, как обычно происходит на практике с использованием автоматизированного синтезатора олигонуклеотидов.

Связанные с основой нуклеотиды могут готовиться любым из способов, известным специалистам, с использованием любой пригодной основы, такой, как стекло, полистирен или тефлон. Одна стратегия состоит в точном нахождении места олигонуклеотидов, синтезированных стандартными синтезаторами. Связывание (нейтрализацию) можно достичь пассивной адсорбцией (Инуи и Хондо, 1990), [24]; использованием УФ света (Нагата и др., 1985, [24]; Дален и др., 1987, [8]; Морри и Киллинс, 1989), [32], или образованием ковалентной связи базовой модифицированной ДНК (Келлер и др., 1988, [26]; 1989, [27]).

Можно применять другую стратегию – использовать в качестве образователя связи сильное взаимодействие биотин-стрептавидин.

Например, Бруд и др., 1991, [2], описывают использование биотинилированных[проб, хотя это дуплексные пробы, которые связаны на покрытых стрептавидином магнитных шариках. Покрытые стрептавидином шарики можно купить в фирме Dynal, Осло. Конечно, эта же химия образования связей применима к покрытию любой поверхности стрептавидином Биотинилированные пробы можно купить в газных фирмах, например, в Oregon Technologies (Аламеда, Калифорния).

Фирма Nunc Laboratories (Нейпервилл, Иллинойс) также продает подходящий к использованию материал. Nunc Laboratories разработала метод, посредством которого ДНК можно ковалентно связать с поверхностью микро-карманами, называемой CovaLink NH. CovaLink NH это поверхность из полистирена, к которой привить вторичные амино-группы (>NH), которые служат предмостными позициями для образования дальнейших ковалентных связей. Модули CovaLink можно закупить в Nunc Laboratories. Молекулы ДНК могут быть связаны с CovaLink исключительно у 5'-конца посредством связи фосфорамидата, что позволяет связывание (нейтрализацию) более чем 1 пмоль ДНК (Расмуссен и др., 1991), [40].

Использование полосок CovaLink NH для ковалентного связывания молекул ДНК у 5'-конца было описано (Расмуссен и др., 1991), [40]. В этой технологии используется связь фосфорамидата (Чу и др., 1983), [6]. Это полезно, так как предпочтительна связывание (нейтрализация) с использованием только одной ковалентной связи. Связь фосфорамидата связывает ДНК со вторичными амино-группами CovaLink NH, которые расположены у конца плеч прокладки, ковалентно привитых на поверхность полистирена через плечо прокладки длиной 2 нм.

Чтобы связать олигонуклеотид с CovaLink NH через связь фосфорамидата, окончание олигонуклеотида должно иметь 5'-конец фосфатную группу. Вероятно, даже можно ковалентно CovaLink и затем использовать стрептавидин для связывания проб.

Более конкретно, способ образования связи включает в себя растворение ДНК в воде (7,5 нг/мкл) и денатурирование в течение 10 мин при 95°C и охлаждение на льду в течении 10 мин.

0,1 М 1-метилимидазол с температурой льда, pH 7,0 (1-Melm7) затем доавляется к конечной концентрации 10 mM 1-Melm7. Затем раствор ss ДНК наносится на полоски CovaLink (75 мкл/карман), стоящие на льду.

Карбодиимид 0,2 М 1-этил-3-(3- диметиламинопропил) карбодиимид (EDC), растворенным в 10 mM 1-Melm7, освежается и 25 мкл добавляется на карман. Полоски выдерживаются в термостате 5 часов при 50°C. После выдержки полоски промываются с использованием, например Nunc-Immuno Wash; сначала полоски промываются 3 раза, затем они промываются промывающим раствором в течении 5 мин и наконец они промываются 3 раза (причем промывающий раствор это 0,4 N NaOH, 0,25% SDS, нагретого до 50°C).

Считается, что еще один пригодный метод для использования с этим изобретением описан в заявке WC 90/03382 (Сазерн и Маскос), [48]. Этот способ приготовления олигонуклеотида, связанного с основой, предполагает прикрепление нуклеозида 3'-реактанта через фосфотную группу ковалентной связью сложного двойного эфира фосфора к алифатическим гидроксильным группам, несомым основой. Затем олигонуклеотид синтезируется на поддерживаемом нуклеозиде и защищающие группы удаляются из цепочки синтетического олигонуклеотида в стандартных условиях, которые не отщепляют олигонуклеотида от основы. К числу пригодных реагентов относятся нуклеозид фосфорамидит и нуклеозид водород фосфат.

Более конкретно, чтобы использовать этот метод, основа, например стеклянная пластина, дериватизируется при контакте со смесью ксилена, глицилоксипропилтриметоксилана и малым количеством диизопропилэтиламина при 90°C в течение полусуток. Затем она тщательно промывается метанолом, эфиром и высушивается на воздухе.

Затем дериватизированная основа, нагревается перемешиванием в гексаэтиленгликоле, содержащем каталитическое количество концентрированной серной кислоты, в течение полусуток в атмосфере аргона при 80°C с получением алкил гидроксил дериватизированной основы. После промывания метанолом и эфиром основа высушивается в вакууме и хранится в аргоне при -20°C.

Затем синтез олигонуклеотида производится вручную в стандартных условиях с использованием дериватизированной стеклянной пластинки в качестве твердой основы. Первый нуклеотид будет 3'-водород фосфат, используемый в виде соли триэтиламмония. Этот метод приводит к получению связанных с основой олигонуклеотидов высокой чистоты.

Можно применять стратегию "на чипе" для приготовления рядов проб ДНК. Например, адресуемую приводимую в действие лазером фотолептацию можно использовать в химическом синтезе олигонуклеотидов прямо на стеклянной поверхности, как описано Фодором и др., (1991), [22]. Пробы также можно связать (нейтрализовать) на нейлоновых основах, как описано Ван Нессом и др. (1991), [51]; или связать с тефлоном с использованием метода Данкена и Кавальера (1988), [19].

Фодор и др. (1991), [22] описывают направляемый светом синтез динуклеотидов, который применим к пространственно направленному синтезу сложных соединений для применения в микро-производстве устройств. Он основан на методе. Используя свет для направления одновременного синтеза химических соединений на твердой основе. Структура воздействия света или других видов энергии через маску или другими пространственно адресуемыми средствами определяет то, какие области основы активизируются для создания химической связи.

Активизация светом происходит из удаления фотолабильных защитных групп из выбранных участков. После депротекции (снятия защиты) первое соединение, содержащее фотолабильную защищающую группу, подвергается воздействию по всей поверхности, однако реакция происходит только с участками, на которые попал свет на предыдущем этапе. Затем подложка освещается через вторую маску, которая активизирует другую область для реакции со вторым защищенным стандартным блоком. Структура масок, используемых в таком освещении, и последовательность реагентов определяют конечные продукты и их размещение. При методе Фодора возможна высокая степень миниатюризации, потому что плотность мест синтеза связана только физическими ограничениями на пространственную адресуемость, т.е. дифракцией света. Каждое соединение доступно и его положение точно известно. Поэтому олиго-чип, получаемый таким образом, можно легко использовать в СПГ.

Фодор и др. (1991), [22] описывают активизированное светом образование динуклеотида следующим образом. 5'-Нитровератрил тимидин синтезировался из 3'-O-тимидин ацетата. После депротекции с базой 5'-нитровератрил тимидин прикреплялся к аминированной подложке через связь с 3'-гидроксильной группой. Защищающие группы нитровератрила удалялись освещением через 500мкм маску типа шахматной доски. Затем подложка обрабатывалась активизированным фосфорамидитом 2'-диоксицидином. Чтобы отслеживать реакцию флуорометрически, деоксицитидин модифицировался образователем связи, защищенным FMOC аминокислотой, прикрепленным к экзоциклическому амину. После удаления FMOC-защищающей группы с базой области, которые содержали динуклеотид были флуоресцентно мечены обработкой подложки FITC. Поэтому, следуя этому методу, можно синтезировать связанные с основой олигонуклеотиды.

Образование связи олигонуклеотида с нейлоновой основой, как описано Ван Нессом и др. (1991), [51] требует активации нейлоновой поверхности через алкиляцию и селективную активацию 5'-амин олигонуклеотидов цианур-хлоридом следующим образом. Нейлоновая поверхность этилируется с использованием триэтилоксоний тетрафторобората с образованием амин-реактивных имидат-сложных эфиров на поверхности нейлона и 1-метил-2-пирролиллон используется как растворитель. Поверхность нейлона не полируется, чтобы была максимально возможная площадь поверхности.

Затем активизированная поверхность реагирует с поли(этиленмин) (Mn приблизительно 10K-70K) с образованием полимерного покрытия, которое создает расширенную поверхность амина для прикрепления олиго. Олигонуклеотид(ы) с аминным хвостом селективно реагирует с избытком цианур-хлорида, исключительно в аминном хвосте, давая 4,6-дихлоро-1,3,5-триазилил-олигонуклеотид(ы) в количественном выходе. Смещение одной хлорной среды цианур-хлорида амино-группой значительно уменьшает реактивность остающихся групп хлора. Это приводит к повышенной гидролитической стабильности 4,6-дихлоро-1,3,5-триазилил-олигонуклеотида (олигонуклеотидов), который стабилен значительные периоды времени в содержащих буфер водных растворах (pH 8,3,4 градуса C, 1 неделя) и легко изолируется и считается хроматографией размерного элюирования или ультрафильтрацией.

Эта реакция специфична для аминного хвоста без очевидной реакции на нуклеотидных средах. Затем покрытая PEI поверхность нейлона реагирует с активизированным цианур-хлоридом олигонуклеотидом. Высокие концентрации "захватывающей" последовательности легко связываются на поверхности и непрореагировавшие амины покрываются сукциновым ангидридом на конечном этапе процесса дериватизации.

Один конкретный способ приготовления олигонуклеотидов, связанных с основой, заключается в

использовании генерируемого светом синтеза, описанного Пизом и др. (1994), [36]. Эти авторы использовали существующие методы фотолитографии для формирования рядов связанных проб олигонуклеотидов (чипов ДНК). Эти методы, в которых свет используется, чтобы направлять синтез проб олигонуклеотидов в миниатюризованных рядах высокой плотности, используют фотолабильные 5'-защищенные N-ацил-деокси нуклеозид фосфорамидиты, химию поверхностного создателя связей и универсальные стратегии комбинаторного синтеза. Матрица их 256 пространственно определенных проб олигонуклеотидов может формироваться таким образом и затем использоваться в удобной организации последовательности в формате 3, как здесь описано.

Пиз и др. (1994), [36] представили стратегию, пригодную для использования в направляемом светом синтезе олигонуклеотидов. В этом способе поверхности твердой опоры, модифицированная фотолабильными защищающими группами, освещается через фотолитографическую маску, давая реактивные гидроксильные группы в освещенных участках. Затем 3'-фосфорамидит-активизированный деоксинуклеозид (защищенный как 5"-гидроксил фотолабильной группой) подается на поверхность и связи образуются на местах, на которые воздействовал свет. Вслед за покрытием окислением подложка промывается и поверхность освещается через 2-ю маску для воздействия на дополнительные гидроксильные группы для образования связей. Второй 5'-защищенный, 3'-фосфорамидит-активизированный деоксинуклеозид наносится на поверхность. Селективная фотодепротекция и циклы образования связей повторяются пока не получается желаемый набор продуктов. Поскольку применяется фотолитография, процесс можно миниатюризовать для формирования рядов с высокой плотностью проб олигонуклеотидов-последовательность которых известна в каждом месте.

Путь синтеза для приготовления необходимых 5'-O-(альфа-метил-6-нитропиперонилоксикарбонил)-N-fwbk-2'-деоксинуклеозид фосфорамидитов (MeNPoc-N-ацил-2'-деоксинуклеозид фосфорамидитов) предполагает на первом этапе наличие N-ацил-2'-деоксинуклеозида, который реагирует с 1-(2-нитро-4,5-метилendioксифенил)-этил-1-хлороформатом с получением 5'-MeNPoc-N-ацил-2'-деоксинуклеозида. На втором этапе 3'-гидроксил реагирует с 2-цианозтил N,N'-диизопропилхлорофосфорамидитом с использованием стандартных процедур с получением 5'-MeNPoc-N-ацил-2'-деоксинуклеозид -3'-O-(2-цианозтил-N,N'-диизопропил)фосфорамидитов. Светозащищающая группа стабильна в обычных условиях синтеза фосфорамидита и может удаляться водной базой. Эти реагенты можно длительно хранить в аргоне при 4S00C.

Сообщалось о периодах полупроведения фотолиза 28 с, 31 с, 27 с и 18 с для MeNPoc-dT, MeNPoc-dc^{ibu}, MeNPoc-dG^{PAC}, MeNPoc-dA^{PAC} соответственно (Пиз и др., 1994), [36]. Поэтому в литографическом синтезе рекомендуется время освещения 4,5мин (9 x t1/2MeNPoc-dC) для обеспечения >99% удаления MeMPoc защищающих групп.

Пригодная синтетическая основа это состоит из 5,1 X 7,6см стеклянной подложки, приготовленной путем чистки в концентрированной NaOH, за чем следует исчерпывающее полоскание в воде. Затем поверхности дериватизируются 2 часа раствором 10% (объем/объем) бис(2-гидроксизтил) аминопропилтриэтоксилана (фирма Petrarch Chemicals, Бристол, Пенсильвания) в 95% этанола, промываются тщательно этанолом и эфиром, высушиваются в вакууме при 40°C и нагреваются при 100°C в течение 15мин. В таких исследованиях образователь связи синтеза прикрепляется реакцией дериватизированных подложек с 4,4'-диметокситритил (DMT)-гексаэтилокси-О-цианозтил фосфорамидитом.

Резюмируя, чтобы инициировать синтез пробы олигонуклеотида, подходящая производная деоксинуклеозид фосфорамидита прикрепляется к синтетической основе через образователь связи. Затем области основы активизируются для синтеза освещением через, например, отверстия 800 x 12800мкм фотолитографической маски. Могут проводиться дополнительные циклы синтеза фосфорамидита (с DMT-защищенными деоксинуклеозидами) для формирования любой требуемой последовательности, такой как любая 4-,5-,6-,7-,8-,9- или даже 10-мерная последовательность. После удаления фосфата и экзоциклических амин-защищающих групп концентрированной NH₄OH в течение 4 часов подложку можно установить в термостатически контролируемую камеру гибридизации с водяной рубашкой, готовую к работе.

Конечно, можно легко приобрести чип ДНК, такой как один из описанных выше активизируемых светом чипов, из промышленного источника. В этом отношении можно связаться с фирмами Affymetrix из Санты-Клары, Калифорния 95051, и Beckman.

ПРИМЕР II

МОДИФИЦИРОВАННЫЕ ОЛИГОНУКЛЕОТИДЫ ДЛЯ ИСПОЛЬЗОВАНИЯ В ПРОБАХ

Модифицированные олигонуклеотиды могут использоваться во всех процедурах этого изобретения для повышения специфичности или эффективности гибридизации. Способ достижения этого - замена природных нуклеотидов модификацией базы. Например, можно использовать пиримидины с галогеном у положения C5. Считается, что это улучшает стабильность дуплекса путем влияния на установку базы слоями. Можно также применять 2,6-диаминопурин для придания -3-ей связи галогена в его базовом спаривании с тиминном, что термически стабилизирует ДНК-дуплексы. Сообщалось, что использование 2,6-диаминопурина приводит к значительному улучшению стабильности дуплекса коротких олигомеров.

Его включение, как предлагается, предоставляет более строгие условия отжига иницирующего вещества, тем самым улучшая специфичность образования дуплекса и подавляя фоновые проблемы или исключая использование более коротких олигомеров.

Синтез вариантов трифосфата этих модифицированных нуклеотидов описан Хохайзелем и Лерахом (1990), [23]. 5-хлоро-2'-деоксиуридин и 2,6-диаминопурин 2'-деоксинуклеозид приобретаются, например, в фирме Sigma. Фосфорилирование выполняется следующим образом: 50мг сухого 2-NH₂-dAdo принимается в 500мкл сухого триэтил фосфата при перемешивании в аргоне. Добавляется 25мкл POCl₃ и смесь выдерживается у термостате при -20°C. Тем временем 1ммоль пирофосфорной кислоты растворяется в 0,95мл три-н-бутиламина и 2мл метанола и высушивается во вращающемся испарителе. Затем она высушивается испарением дважды из 5мл пиридина, причем перед вторым разом также добавляется 70мкл

три-п-бутиламина. Наконец, это растворяется в 2мл сухого диметил формамида.

Спустя 90мин при -20°C смесь фосфорилирования испаряется для удаления избытка POCl_3 : добавляется три-п-бутиламмоний пирфосфат в диметил формамиде. Выдерживание проводится в течение 1,5мин при комнатной температуре. Реакция останавливается путем добавления 5мл 0,2М триэтиламмоний бикарбоната (pH 7,6) и смесь держится на льду 4 часа. Для 5-Cl-dUrd условия будут тождественны, но добавляется 50мкл POCl_3 и фосфорилирование проводится при комнатной температуре в течение 4 часов.

После гидролиза смесь испаряется, pH устанавливается на 7,5, и экстрагируется 1 объемом диэтил эфира. Разделение продуктов производится например, на (2,5x20см) Q-Sepharose колонне с использованием линейного градиента 0,15М к 0,8М триэтиламмоний бикарбонату. При хранении замороженными нуклеотиды стабильны длительное время. Можно также использовать неизбирательный аналог базы или универсальную базу, как она сконструирована Николсом и др. (1994), [35]. Этот новый аналог, 1-(2'-деокси-бета-D-рибофураносил)-3-нитропиррол (обозначаемый М) был сформирован для использования в пробах олигонуклеотидов и иницирующих веществах для решения проблем проектирования, которые возникают в результате вырождения генетического кода, или когда имеются только данные по фрагментарной последовательности пептидов. Этот аналог повышает укладку слоями и одновременно минимизирует взаимодействия с образованием водородных связей, не разрушая пространственно дуплекс ДНК. Аналог М нуклеозида был сконструирован чтобы максимально увеличить взаимодействия при укладке слоями с использованием апротических полярных заместителей, связанных с гетероароматическими кольцами, повышая внутри- и вне-рядевого взаимодействия при укладке, чтобы уменьшить роль водородных связей в специфичности базового спаривания. Николс и др. (1994), [35] предпочитают 3-нитропиррол 2'-деоксирибонуклеозид из-за его структурной и электронной схожести с р-нитроанилином, чьи производные относятся к самым малым известным пропласткам двойной свитой ДНК. Диметокситритил-защищенный фосфорамидит нуклеозида М также может включаться в нуклеозиды, используемые как иницирующие вещества для образования последовательности и цепной реакции полимеразы (PCR). Николс и др. (1994), [35] показали, что значительное количество нуклеотидов могут заменяться М без потери специфичности иницирующего вещества. Уникальным свойством М –это его способность замещать длинные звенья смежных нуклеозидов и все же давать функциональные иницирующие вещества для образования последовательностей. Как сообщалось, все последовательности с 3, 6 и 9 замещениями М дают читаемые "лестницы" образования последовательностей, и все PRC с 3-мя разными содержащими М иницирующими веществами привели к расширению правильного продукта (Николс и др., 1994), [35].

Способность содержащих 3-нитропиррол олигонуклеотидов функционировать в качестве иницирующих веществ уверенно показывает, что дуплексная структура должна образовываться с дополняющими прядями. Оптические тепловые профили, полученные для пар олигонуклеотидов $d(5'-C_2-T_5XT_5G_2-3')$ и $d(5'-C_2A_5YA_5G_2-3')$ (где X и Y могут быть A, C, G, T или M), как сообщалось, подходят под обычный сигмоидальный рисунок, наблюдаемый для перехода от двойной к одной пряди (нити) ДНК. Значения T_m олигонуклеотидов, содержащих базовые пары X^*M (где X было A, C, G или T и Y было M), как сообщалось, все попадают в диапазон 3°C (Николс и др. 1994), [35].

ПРИМЕР III

ПРИГОТОВЛЕНИЕ ОБРАЗУЮЩИХ ПОСЛЕДОВАТЕЛЬНОСТИ ЧИПОВ И РЯДОВ

Этот пример описывает физические варианты осуществления образующих последовательности чипов, рассматриваемых изобретателем.

Основной пример это использование 6-меров, прикрепленных к поверхностям к 50 микрон для получения чипа с размерами 3 x 3мм, который можно комбинировать с получением ряды 20 x 20см. Другой пример это использование 9-мерных олигонуклеотидов, прикрепленных к поверхности в 10 x 10 микрон для создания 9-мерного чипа с размерами 5 x 5мм. 4000 единиц таких чипов можно использовать для создания ряда 30 x 30см. На фиг.2A, фиг.2B и фиг.2C показан еще один пример ряда, в котором 4000-16000 олигонуклеотидов расположены по квадрату. Пластина или набор трубочек, как также указано, можно упаковать с этим рядом как часть набора материалов для создания последовательностей.

Ряды можно отделить физически друг от друга или с помощью гидрофобных поверхностей. Один возможный способ использовать разделение гидрофобной полоской - это применить такую технологию, как Iso-Grid Microbiology System, выпускаемую фирмой QA Laboratories, Торонто, Канада.

Гидрофобные сеточные мембранные фильтры (HGMF) использовались в аналитической микробиологии пищевых продуктов уже около десятилетия, где они уникально привлекательны тем, что имеют расширенный числовой диапазон и автоматизированный подсчет колоний. Одна промышленно поставляемая сетка это ISO-GRID(TM) от фирмы QA Laboratories Ltd. (Торонто, Канада), которая состоит из квадрата (60 x 60см) полисульфонного полимера (Gelman Tuffryn HT-450, размер пор 0,45мк), на которое отпечатана краской черная гидрофобная сетка, состоящая из 1600 (40 x 40) квадратных ячеек. HGMF был предварительно модифицирован добавкой бактериальных суспензий вакуумной фильтрацией и выдерживался на дифференциальной или селективной среде по выбору.

Поскольку рост микробов ограничен ячейками сетки с известным положением и размером мембраны, HGMF функционирует более как аппарат MPN, чем обычная пластина или мембранный фильтр. Петеркин и др. (1987), [37] сообщили, что эти HGMF можно использовать для распространения и хранения геномных библиотек при использовании с репликатором HGMF. Один такой инструмент реплицирует рост от каждой из 1600 ячеек ISO-GRID и позволяет сделать много копий образцового HGMF (Петеркин и др., 1987), [37].

Шарп и др. (1989), [46] также применяли ISO-GRID HGMF от QA Laboratories и автоматизированный счетчик HGMF (MI-100 Interpreter) и репликатор RP-100. Они сообщили о методе поддержания и сортировки многих культур микробов.

Петеркин и его коллеги после описали способ сортировки проб ДНК с использованием гидрофобного сеточно-мембранного фильтра (Петеркин и др., 1989), [38].

Эти авторы сообщили о способах эффективной гибридизации колоний прямо на HGMF. Раньше плохие

результаты получались из-за низкой связывающей ДНК способности полисульфонного полимера, на котором печатаются HGMP. Однако Петеркин и др. (1989), [38] сообщили, что связывание ДНК с поверхностью мембраны улучшилось при обработке реплицированных и выдержанных в термостате HGMP полиэтиленгликолем, поликаатионом, до контакта с ДНК. Хотя в этой ранней работе использовалось ячеистое прикрепление ДНК и цель была отличной от цели этого изобретения, описанную методологию можно легко адаптировать для формата 3 СПГ.

Для того, чтобы быстро идентифицировать полезные последовательности, Петеркин и др. (1989), [38] использовали радио-меченную плазмидную ДНК из разных клонов и испытывали ее специфичность относительно ДНК на приготовленных HGMP. Таким образом, ДНК из рекомбинантных плазмидов быстро сортировалась гибридизацией колоний относительно 100 организмов на репликатах HGMP, которые можно легко и воспроизводимо приготовить.

Нужно решить две основные проблемы. Манипулирование с малыми (2-3мм чипами и параллельное выполнение тысяч реакций. Решение, которое предлагает это изобретение, состоит в том, чтобы содержать чипы и пробы в соответствующих рядах в одном примере чипы, содержащие 250000 9-меров синтезируются на кремниевой подложке в форме пластин 8 x 8mm (15мкМ/олигонуклеотид, Пиз и др., 1994), [36] выстроенных в формате 8 x 12 (96 чипов) с желобком между ними в 1 мм. Пробы добавляются либо многоканальной пипеткой, либо рядом штырьков, по одной пробе на 1 чип. Для расчета всех 4000 6-меров нужно применять 42 ряда чипов, или используя разные ряды, или повторно используя один набор рядов чипов несколько раз.

В указанном случае, при использовании ранней номенклатуры этой заявки, $F=9$; $P=8$; $F+P=15$.

Чипы могут иметь пробы с формулой B_xN_n , где x это число специфицированных баз B ; а n это число неспецифицированных баз, так что $x=4-10$ и $n=1-4$. Чтобы достичь более эффективной гибридизации и избежать потенциального влияния любых олигонуклеотидов основы, специфицированные базы можно окружить неспецифицированными базами, что представляется такой формулой как $(N)_nB_x(N)_m$ (фиг.4).

ПРИМЕР IV

ПРИГОТОВЛЕНИЕ ФРАГМЕНТОВ НУКЛЕИНОВОЙ КИСЛОТЫ

Нуклеиновые кислоты, последовательности которых нужно создать, можно получить из любого подходящего источника, такого, как сДНК, геномная ДНК, хромосомная ДНК, микрорассеченные хромосомный полосы, космидные или YAC вставки и РНК, включая мРНК без каких-либо этапов расширения. Например, Сэмбрук и др. (1989), [41] описывают 3 протокола для изолирования ДНК с высоким молекулярным весом от клеток млекопитающих (стр.9.14-9.23).

Затем нуклеиновые кислоты фрагментируются любым из способов, известных специалистам, включая, например, использование ограничительных ферментов, как описано на стр.9.24-9.28 у Сэмбрука и др. (1989), [41], сдвиг ультразвуком и обработку NaOH.

Сдвиг при низком давлении также уместен, как описано Шриффером и др. (1990), [43]. В этом способе образцы ДНК пропускаются через небольшой французский датчик давления с набором давлений от малого до промежуточного. Рычажное устройство позволяет производить контролируемое приложение от малого до промежуточного давления к элементу. Результаты этих исследований указывают, что сдвиг при малом давлении это полезная альтернатива звуковых и ферментных методов фрагментации ДНК.

Один особенно подходящий способ для разделения ДНК на фрагменты, как считается, это тот, где используются две эндонуклеазы с распознаванием базы, CviJI, описанный Фицджералдом и др. (1992), [21]. Эти авторы описали подход для быстрой фрагментации и фракционирования ДНК на конкретные размеры, которые они считали пригодными для "дробового" клонирования и создания последовательностей. Настоящий изобретатель предусматривает, что это будет также особенно пригодно для генерирования произвольных, но относительно небольших фрагментов ДНК для использования в настоящей технологии образования связей.

Ограничительная эндонуклеаза CviJI обычно расщепляет распознавательную последовательность PuGCPu между G и C с оставлением тупых концов. Нетипичные условия реакции, которые меняют специфичность этого фермента (CviJI**), дают квази-произвольное распределение фрагментов ДНК из малых молекул PUC19 (2688 базовых пар). Фицджералд и др. (1992) количественно оценили произвольность этой стратегии фрагментирования с использованием CviJI** продукта вываривания pUC19, который был фракционирован по размерам способом быстрой гелевой фильтрации и непосредственно связан, без ремонта концов, вектором клонирования lacZ minus M13. Анализ последовательности 76 клонов показал, что CviJI** ограничивает PuGCPu и PuGCPu в дополнение к местам PuGCPu и что новые данные и последовательность накапливаются со скоростью, согласующейся с произвольной фрагментацией.

Как сообщается в литературе, преимущества этого подхода по сравнению фракционированием соникацией (звуком) и агарозным гелем включают в себя: требуются меньшее количество ДНК (0,2-0,5мкг вместо 2-5мкг); и имеется меньше этапов (не нужно предварительного связывания, ремонта концов, химической экстракции или электрофореза агарозного геля и элюирования). Предлагается, что эти преимущества также будут полезны при приготовлении ДНК для создания последовательностей по формату 3.

Независимо от того, как получают или готовятся фрагменты нуклеиновой кислоты, важно денатурировать ДНК, чтобы получить для гибридизации одно-завитые кусочки. Это достигается выдерживанием раствора ДНК в течение 2-5 минут при 80-90°C. Затем раствор быстро охлаждают до 2°C для предотвращения повторной натурации фрагментов ДНК до того, как они контактируют с чипом. Фосфатные группы тоже нужно удалить из геномной ДНК, как описано в примере VI.

ПРИМЕР V

ПРИГОТОВЛЕНИЕ МЕЧЕННЫХ ПРОБ

Пробы олигонуклеотидов можно готовить автоматизированным синтезом, что является рутинным для специалистов, например, с использованием системы Applied Biosystems. В качестве альтернативы, пробы

можно приготовить используя методы Biotechnologies Inc. с помощью слоев пористых подложек тефлона.

Пробы олигонуклеотидов можно метить, например, радиоактивными метками (³⁵P, ³²P, ³³P и предпочтительно ³³P) для рядов с участками 100-200мкм или 100-400мкм; нерадиоактивными изотопами (Якобсен и др., 1990), [25]; или флуорофорами (Брамбо и др., 1988), [3]. Все такие способы нанесения меток являются обычными в технике, как показано в соответствующих разделах у Сэмбука и др. (1989), [41] и в другой литературе, например, у Шуберта и др., (1990), [44]; Мураками и др. (1991), [33] и Кейта и др. (1991), [5].

В отношении радио-мечения обычные способы это мечение концов с использованием T4 полинуклеотидной киназы или нанесение меток с высокой специфической активностью с использованием Кленов или даже T7 полимеразы. Это описывается следующим образом.

Синтетические олигонуклеотиды синтезируются без фосфатной группы у своих окончаний 5' и поэтому легко метятся переносом гамма-³²P или гамма-³³P из (гамма-³²P)_{АТР} или (гамма-³³P)_{АТР} с использованием фермент бактериофаг T4 полинуклеотид киназы. Если реакция проводится эффективно, специфическая активность таких проб может быть такой высокой, как специфическая активность самих (гамма-³²P)_{АТР} или (гамма-³³P)_{АТР}. Описанная ниже реакция предназначена метить 10 пмоль олигонуклеотида до высокой специфической активности. Нанесение меток на разные количества олигонуклеотида можно легко достичь путем увеличения или уменьшения размера реакции при поддержании постоянными концентрации всех компонентов.

Реакционная смесь будет создаваться с использованием 1,0мкл олигонуклеотида (10 нмоль/мкл); 2,0мкл 10х бактериофаг T4 полинуклеотид киназного буфера; 5,0мкл (гамма-³²P)_{АТР} или (гамма-³³P)_{АТР} (пр. акт. 5000Ci/ммоль; 10mCi/мл в водном растворе) (10 пмоль); и 11,4мкл воды 8 единиц (около 1мкл) бактериофаг T4 полинуклеотидной киназы добавляется к реакционной смеси, хорошо перемешивается и выдерживается 45 минут при 37°C. Реакционная смесь нагревается в течение 10 минут при 68°C для инактивации бактериофаг T4 полинуклеотидной киназы.

Затем определяется эффективность переноса ³²P или ³³P к олигонуклеотиду и его специфическая активность. Если специфическая активность пробы приемлема, она очищается. Если специфическая активность слишком низка, добавляются еще 8 единиц фермента и выдерживаются еще 30 минут при 37°C до нагревания реакционной смеси в течение 10 минут при 68°C для инактивации фермента.

Очистка радио-меченных олигонуклеотидов может достигаться путем осаждения с этанолом; осаждением с цетилпиридиний бромидом; хроматографией через био-гель P-60; или хроматографией на колонне Sep-Pak C18.

Пробы с более высокими специфическими активностями можно получить, используя фрагмент Кленова E.coli. ДНК полимеразы I для синтеза пряди ДНК дополняюще к синтетическому олигонуклеотиду. Короткое иницирующее вещество гибридизует к шаблону олигонуклеотида, чья последовательность является дополнением желаемой радио-меченной пробы. Затем иницирующее вещество расширяется с использованием фрагмента Кленова E.coli ДНК полимеразы I для включения (альфа-³²P)dNTP или (гамма-³³P)dNTP в направляемой шаблоном манере. После реакции шаблон и продукт разделяются денатурированием, за чем следует электрофорез через полиакриламидный гель в условиях денатурации. С этим способом при желании можно формировать пробы олигонуклеотидов, содержащие несколько радиоактивных атомов на молекулу олигонуклеотида.

Чтобы использовать этот метод, нужно смешать в трубе микрофуги рассчитанные количества (альфа-³²P)dNTP или (альфа-³³P)dNTP, необходимые для достижения желаемой специфической активности и достаточные для получения полного синтеза всех прядей шаблона. Концентрация dNTP не должна быть меньше 1мкМ на любом этапе в ходе реакции. Затем в трубу добавляются соответствующие количества иницирующего вещества и шаблона ДНК, причем иницирующее вещество в молярном отношении в 3-10 раз превышает шаблон.

Затем добавляется 0,1 объема 10 х буфера Кленова и хорошо перемешивается. Затем добавляются 2-4 единицы фрагмента Кленова E.coli ДНК полимеразы I на 3мкл реакционного объема, перемешивается и выдерживается 2-3 часа при 4°C.

При желании, течение реакции можно отслеживать путем удаления небольших (0,1мкл) кратных и измерения доли радиоактивности, которая стала способна осаждаться, 10% трихлоруксусной кислотой (TCA).

Реакционная смесь разжижается равным объемом наполняющего гелем буфера, нагревается до 80°C в течение 3 минут и затем вся проба загружается на денатурирующий полиакриламидный гель. После электрофореза гель автордиографируется, что позволяет найти пробу и удалить ее из геля. Имеются также следующие разные способы флуорофобного нанесения меток. Брамбо и др. (1988), [3], описывают синтез флуоресцентно меченных иницирующих веществ. Аналог деоксиуридина с первичным аминовым "плечом образователя связи" из 12 атомов, прикрепленный у С-5, синтезируется. Синтез аналога состоит из дериватизации 2'-деоксиуридина через органометаллические промежуточные вещества с получением 5'(метил пропеноил)-2'-деоксиуридина. Реакция с диметокситритил-хлоридом дает соответствующий привод 5'-диметокситритил. Сложный эфир метила гидролизруется, активизируется и реагирует с соответствующим образом моноацелированным алкил диамином. После очистки результирующие нуклеозиды плеча образователя связи преобразуются в аналоги нуклеозида, способные для химического синтеза олигонуклеотида.

Затем получают олигонуклеотиды, которые содержат одну или 2 базы с плечом образователя связи с использованием модифицированной химии фосфоридитов. В раствор 50нмоль олигонуклеотида с плечом образователя связи в 25мкМ бикарбоната натрия 500мМ (рН 9,4) добавляются 20мкл FITC 300 мМ в диметил сульфоксиде. Смесь перемешивается при комнатной температуре 6 часов. Олигонуклеотид отделяется от свободного FITC путем элюирования из колонны Sephadex G-25 1 x 30см с 20vV ацетата аммония (рН 6) с комбинированием фракций в первом пике УФ-поглощения.

В общем случае, флуоресцентное нанесение меток на олигонуклеотид у его 5'-конца в начале предполагало 2 этапа. Сначала N-защищенная производная аминоксил фосфорамидита добавляется к 5'-концу олигонуклеотида во время автоматизированного синтеза ДНК. После удаления всех защищающих групп сложный эфир NHS соответствующего флуоресцентного красителя соединяется с 5'-амино-группой за полсуток, за чем следует очистка меченного олигонуклеотида от избытка красителя с использованием обратной фазы HPLC или PAGE.

Шуберт и др. (1990) [44], описали синтез, фосфорамидита, который позволяет получать олигонуклеотиды, меченные флуоресцеином во время автоматизированного синтеза ДНК. Метил-сложный эфир флуоресцеина алкилируется 4-хлоро(4,4'-диметокситритил) бутанолом-1 в присутствии K_2CO_3 и KI в DMF в течение 17 часов. После удаления группы тритила 1% TFA_в в хлороформе продукт фосфитилируется стандартными процедурами посредством бис (диизопропиламино)метоксифосфина. Фосфорилирование указанной полученной производной флуоресцеина приводит к достаточному выходу Н-фосфоната. Получающийся амидит (раствор 0,1М в сухом ацетонитриле) используется для автоматизированного синтеза разных иницирующих веществ с использованием бета-цианоэтил фосфорамидитной химии и синтезатора ДНК. Отщепление от основы и депротекция осуществляются с помощью 25% водного раствора аммиака в течение 36 часов при комнатной температуре. Необработанный продукт очищается посредством PAGE, и меченное иницирующее вещество видимо как бледно-зеленая флуоресцентная полоса на 310нм. Элюирование и обессоливание с использованием патронов RP 18 дает желаемый продукт.

Флуоресцентное мечение 5'-конца пробы в методе Шуберта непосредственно достигается во время синтеза ДНК в последнем цикле образования связей,

Выход связей так же высок, как и с обычными фосфорамидитами. После депротекции и удаления аммиака путем лиофилизирования с использованием скоростного вакуума или осаждения этанолом меченные флуоресцентно олигонуклеотиды могут прямо применяться для образования последовательностей ДНК в формате 3 СПГ.

Мураками и др.[33] также описывали приготовление меченных флуоресцеином олигонуклеотидов. Этот синтез основан на методе поддерживаемого полимером фосфорамидита и Фосфоната водорода. Этилендиамин или гексаметилендиамин используется как привязь. Они вводятся через Фосфорамидатную связь, которая была образована окислением водород-фосфонатного промежуточного вещества в растворе CCl_4 . Модифицированные олигонуклеотиды подвергаются мечению с использованием первичного амин-ориентирующего реагента, FITC, на шариках. Получающийся модифицированный олигонуклеотид отщепляется от шариков и затем очищается RPLC.

Кейт и др. (1991), [5] описывают использование проб олигонуклеотидов, непосредственно сопряженных с алкалин фосфатазой в комбинации с прямой хемилюминесцентной подложкой (AMPPD), чтобы позволить обнаружение проб. Фосфатазу алкалина можно ковалентно соединить с модифицированной базой олигонуклеотида. После гибридизации олиго выдерживается в термостате с AMPDD. Фермент алкалин фосфатазы разбивает AMPDD с получением соединения, которое дает флуоресценцию без возбуждения, т.е. лазер не нужен. Считается, что сильный сигнал может формироваться с использованием такой технологии.

Меченные пробы можно свободно приобрести в разных промышленных источниках, включая GENSET, а не синтезировать.

ПРИМЕР VI

УДАЛЕНИЕ ФОСФАТНЫХ ГРУПП

Как алкалин фосфатаза бактерий (BAP), так и кишечная алкалин фосфатаза телят (CIP) катализируют удаление радикалов 5'-фосфата из ДНУ и РНК. Поэтому они уместны для удаления 5' фосфатов из ДНК или/и РНК для предотвращения образования связей и неправильной гибридизации. Удаление фосфатов, как оно описано Сэмбруком и др. (1989), [41], осуществляется после разрезки или иного сдвига геномной ДНК.

BAP более активен из 2-х алкалин фосфатаз, но он также намного более стоек к нагреву и чистящим веществам. Поэтому трудно замедлить BAP полностью у конца реакций дефосфорилирования. Протеиназа K применяется для вываривания CIP, который должен быть полностью удален, если последующие связи должны эффективно работать. Альтернативный метод этг инактивация CIP нагреванием до 65°C в течение 1 часа (или 75°C в течение 10 минут) в присутствии 5mM EDTA (pH 8,0) с последующей очисткой дефосфорилированной ДНК экстрагированием с фенол: хлороформом.

ПРИМЕР VII

ПРОВЕДЕНИЕ ОБРАЗОВАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ ПУТЕМ ДВУХЭТАПНОЙ ГИБРИДИЗАЦИИ

Ниже следуют некоторые примеры для описания выполнения методологии образования последовательностей, рассматриваемой изобретателем. Сначала, весь чип гибридизируется смесью ДНК сложностью 100 миллионов бп (1 хромосома человека). Руководство по проведению гибридизации можно найти в статьях, таких как Дрманач и др. (1990), [18]; Храпко и др. (1991), [28] и Бруд и др. (1994), [2]. В этих статьях указаны диапазоны температур гибридизации, буферы и этапы промывки, которые пригодны для использования на начальном этапе формата 3 СПГ.

Данный изобретатель в особенности рассматривает проведение гибридизации в течение до нескольких часов при высоких концентрациях соли и при низкой температуре (от -2 до 5°C из-за относительно малой концентрации получаемой целевой ДНК. Для этой цели применяется буфер SSC вместо буфера фосфата натрия (Дрманач и др., 1990), [18] которым осаждается при 10°C.

Промывка не обязательно должна быть экстенсивной (несколько минут) из-за второго этапа и может быть полностью исключена, если проведение циклов гибридизации используется для создания последовательности в высшей степени сложных образцов ДНК. Этот же буфер используется для этапов гибридизации и промывки, чтобы можно было продолжать второй этап гибридизации с мечеными пробами.

После должной промывки с использованием простого робототехнического устройства на каждом ряде, например, ряде 8 x 8мм (пример III), добавляется одна меченная проба, например, 6-мер. Используется 96-кончиковое или 96-штырьковое устройство, выполняя это за 42 операции. Снова используется диапазон

избирательных условий, как ранее описывалось в научной литературе.

Данным изобретателем в особенности рассматривается использование следующих условий. Сначала, после добавления меченных проб и выдержки только в течение нескольких минут (из-за высокой концентрации добавленных олигонуклеотидов) при низкой температуре (0-5°C) температура повышается до 3-10°C в зависимости от длины F+P и добавляется промывающий буфер. В это время используемый промывающий буфер совместим с любой реакцией образования связей (например, в диапазоне концентрации соли 100 mM). После добавления лигазы температура снова повышается до 15-37°C, чтобы позволить быстрому образованию связей (менее 30 мин) и дальнейшему избиранию гибридов с полным совпадением и несовпадениями.

Использование катионных чистящих (моющих) веществ также рассматривается для использования в формате 3 СПГ, как описано Понтиусом и Бергом (1991), [32]. Эти авторы описывают использование двух простых катионных чистящих веществ, дедecil- и цетилтриметиламмоний бромид (DTAB и CTAB) в повторной натурации ДНК.

DTAB и CTAB это варианты четвертичного аминтетраметиламмоний бромид (TMAB), в котором одна из групп метила замещена либо 12-углеродной (DTAB), либо 16-углеродной (CTAB) алкильной группой. TMAB это соль бромид иона тетраметиламмония, реагент, применяемый в экспериментах по повторному натурированию нуклеиновых кислот для уменьшения сдвига содержания G-C температуры расплавления. DTAB и CTAB аналогичны по структуре натрий додецил сульфату (SDS) с замещением отрицательно заряженного сульфата SDS положительно заряженным четвертичным амином. Хотя SDS обычно применяется в буферах гибридизации для уменьшения неспецифического образования связей и для замедления нуклеаз, он не сильно влияет на скорость повторного натурирования.

При проведении процесса образования связей можно добавить фермент в меченные пробы или после этапа должной промывки для снижения фона.

Хотя ее ранее не предлагали для использования с каким-либо методом СПГ, технология лигазы хорошо установлена в области молекулярной биологии. Например. Худ и коллеги описали метод обнаружения гена посредством лигазы (Ландегрен и др., 1988), [30] методологию которого можно легко приспособить для использования в формате 3 СПГ. Ландегрен описывает анализ присутствия данных последовательностей ДНК на основе способности 2-х олигонуклеотидов отжигаться непосредственно рядом друг с другом на дополняющей целевой молекуле ДНК.

Затем 2 олигонуклеотида соединяются ковалентно под действием ДНК лигазы при условии, что нуклеотиды у места соединения правильно спарены базами. Хотя это ранее не рассматривалось, эта ситуация теперь возникает в образовании последовательности в формате 3.

Ву и Уоллас также описывают использование бактериофаг Т4 ДНК лигазы для соединения двух соседних коротких синтетических олигонуклеотидов. Их реакции образования связи олиго проводились в 50mM Tris HCl pH 7,6, 10 mM MgCl₂, 1mM ATP, 1mM DTT и 5% PEG. Реакции образования связи проводились при нагревании до 100°C в течение 5-10 мин, за чем следовало охлаждение до 0°C до добавления Т4 ДНК лигазы (1 единица; Bethesda Research Laboratory). Большинство реакций образования связей проводились при 30°C и заканчивались нагреванием до 100°C в течение 5 мин.

Затем проводится конечная промывка, пригодная для избирательного обнаружения гибридизированных соседних, или образовавших связь, олигонуклеотидов о длиной (F+P). Этот этап промывки проводится в воде в течение нескольких минут при 40-60°C для вымывания всех необразовавших связи меченных проб и всех других соединений, чтобы максимально уменьшить фон. Благодаря ковалентно связанным меченым олигонуклеотидам обнаружение упрощается (с ним не связаны ограничения по времени и низкой температуре).

В зависимости от используемой метки получение изображений чипов производится на разных аппаратах. Для радиоактивных меток можно использовать технологию хранения фосфора с экраном и PhosphorImager в качестве сканера (фирма Molecular Dynamics, Саннивейл, Калифорния). Чипы помещаются в кассету и накрываются фосфористым экраном. После 1-4 часов воздействия экран сканируется и файл изображения хранится на твердом диске компьютера.

Для обнаружения флуоресцентных меток используются камеры с ПЗС и эпифлуоресцентная или конфокальная микроскопия. Для чипов, сформированных прямо на элементах изображения камеры с ПЗС обнаружение может проводиться, как описано Эггерсом и др., (1994), [20].

Детекторы с приборами с зарядовой связью (ПЗС) служат в качестве активных твердых основ, которые количественно обнаруживают и формируют изображение распределения меченных целевых молекул в анализах на основы проб. В этих устройствах используются присущие микроэлектронике характеристики, которые сочетают в высшей степени параллельные анализы, сверхчувствительное обнаружение, высокую производительность (пропускную способность) интегральный сбор данных и вычисление. Эггерс и др. (1994) описывают применение ПЗС для основанных на пробах анализов, таких, как формат 3 СПГ настоящего изобретения, которые позволяют количественную оценку с точностью до секунд благодаря высокой чувствительности и применяемой прямой связи.

Подход интегрального обнаружения с ПЗС позволяет проводить обнаружение событий образования молекулярных связей на чипах. Детектор быстро формирует двумерную картину, которая уникально характеризует образец. В конкретной работе молекулярного детектора на базе ПЗС отдельные биологические пробы связываются (нейтрализуются) непосредственно на элементах изображения ПЗС и могут прикрепляться к одноразовой полоске-крышке, помещенной на поверхность ПЗС. Молекулы образца можно метить радиоизотопом, хемилюминесцентными или флуоресцентными ярлыками.

После воздействия на образец ряда проб на основе ПЗС Фотоны или продукта затухания радиоизотопа испускаются у местоположений элементов изображения, где образец образовал связи, в случае формата 2, с двумя дополняющими пробами.

В свою очередь, пары электрон-дырка генерируются в кремнии, когда заряженные частицы или излучение

от меченных образцов падают на управляющие электроды ПЗС. Затем электроны собираются под соседними управляющими электродами ПЗС и последовательно считываются на модуле дисплея. Количество фотоэлектронов, генерируемое у каждого элемента изображения, прямо пропорционально количеству молекулярных событий образования связей в такой ближайшей области. Следовательно, образование молекулярных связей можно определять количественно (Эггерс и др., 1994), [20].

Как недавно сообщалось, ПЗС на базе кремния имеют преимущества как полупроводниковые датчики обнаружения и формирования изображений, прежде всего благодаря высокой чувствительности устройств в широком диапазоне длин волн (от 1 до 10000 ангстрем). Кремний очень чувствителен к электромагнитному излучению от видимого спектра до мягких рентгеновских лучей. Для видимого света один фотон, падающий на управляющий электрод ПЗС, приводит к появлению одного заряженного электроном пакета под электродом. Одна бета-частица мягкого рентгеновского излучения (обычно в диапазоне КэВ до МэВ) генерирует от тысяч до десятков тысяч электронов. В добавок к высокой чувствительности, ПЗС, описанные Эггерсом и др. (1994), [20] предоставляют широкий динамический диапазон (4-5 порядков величины), поскольку обнаруживаемый пакет заряда может включать от нескольких до 10^5 электронов. Характеристика обнаружения является линейной в широком динамическом диапазоне.

Путем размещения создающего изображение ряда вблизи от образца эффективность сбора улучшается, как минимум, в 10 раз по сравнению с методами на основе линзы, такими, какие применяются в обычных камерах на ПЗС. То есть образец (эмиттер) находится в близком контакте с детектором (матрицей формирования изображения) и это исключает обычную оптику формирования изображений, такую, как линзы и зеркала.

Когда радиоизотопы прикреплены кап сообщающие группы к целевым молекулам, энергетические частицы обнаруживаются. Несколько сообщающих групп, которые излучают частицы с разной энергией, успешно использовались с микро-изготовленными детекторами, такие как ^{32}P , ^{33}P , ^{35}S , ^{14}C , ^{125}I . Частицы с более высокой энергией, такие как от ^{32}P , обеспечивают самую высокую чувствительность молекулярного обнаружения, тогда как частицы с более низкой энергией, такие как от ^{35}S , обеспечивают лучшее разрешение. Поэтому, выбор сообщающего радиоизотопа может быть сделан по желанию. Когда конкретная радиоизотопная метка выбрана, характеристики обнаружения можно предсказать путем расчета отношения сигнал-шум, как описано Эггерсом и др. (1994), [20].

Альтернативная процедура люминесцентного обнаружения предполагает использование флуоресцентных или хемилюминесцентных сообщающих групп, прикрепленных к целевым молекулам. Флуоресцентные метки можно прикреплять ковалентно или через взаимодействие. Флуоресцентные красители, такие, как бромид этидия с сильными полосами поглощения в диапазоне ближнего УФ (300-350нм) и основными полосами излучения в видимом диапазоне (500-650нм) наиболее пригодны для используемых устройств с ПЗС, поскольку квантовая эффективность на несколько порядков величины ниже на длине волны возбуждения, чем на длине волны флуоресцентного сигнала.

С точки зрения обнаружения люминесценции, многокремниевые управляющие электроды (схемы) ПЗС имеют "встроенную" способность отфильтровывать вклад падающего света в УФ-диапазоне, но они очень чувствительны к видимой люминесценции, генерируемой флуоресцентными сообщающими группами.

Такая присущая большая избирательность относительно УФ-возбуждения позволяет достигать высоких отношений сигнал-шум (более 100) с помощью ПЗС, как указывалось в статье Эггерса и др. (1994), [20].

Для связывания проб на детекторе можно получить матрицы гибридизации на недорогих подложках из SiO_2 , которые затем помещают на поверхность ПЗГ после гибридизации и сушки. Этот формат экономически эффективен, т.п. гибридизация ДНК проводится на недорогих односторонних подложках из SiO_2 , что позволяет повторно использовать более дорогой детектор на ПЗС. В качестве альтернативы, пробы можно связывать прямо на ПЗС с созданием специализированной матрицы проб.

Чтобы связать пробы на покрытии SiO_2 , однослойный оксидный слой связывается с поверхностью пленки с использованием эпоксидно-силанового реагента и стандартной химии модификации SiO_2 . Затем амин-модифицированные пробы олигонуклеотидов связываются с поверхностью SiO_2 посредством вторичного образования аминов с эпоксидным кольцом. Получающаяся связь создает 17 способных вращаться связей разделения между 3' базой олигонуклеотида и поверхностью SiO_2 . Чтобы обеспечить полную депротонацию аминов и минимизировать образование вторичной структуры во время создания связи, реакцию проводят в 0,1М КОН и выдерживают при 37°C 6 часов.

В формате 3 СПГ вообще сигналы считаются на каждую из миллиарда точек. Будет необходимо гибридизировать все ряды, например, 4000 5 x 5мм за 1 раз и возможна последующее использование меньшего числа рядов.

Гибридизации с проведением циклов это один возможный способ увеличения сигнала гибридизации. За 1 цикл большинство фиксированных проб гибридизуют с фрагментами ДНК с хвостовыми последовательностями, некомплементарными для меченных проб. При увеличении температуры эти гибриды будут расплавляться (фиг.3). За следующий цикл некоторые из них (около 0,1%) гибридизуются с соответствующим фрагментом ДНК и дополнительные меченные пробы образуют связи. В этом случае, происходит избирательное плавление гибридов ДНК с несовпадениями одновременно для обоих наборов проб.

В гибридизации с проведением циклов все компоненты добавляются до начала циклов при 37°C для Т4 или при более высокой температуре для термостабильной лигазы. Затем температура снижается до 15-37°C и чип выдерживается до 10 минут, а потом температура повышается до 37°C или выше в течение нескольких минут, и потом снова уменьшается. Циклы можно повторять до 10 раз. В одном варианте оптимальную более высокую температуру (10-50°C) можно использовать без циклов и можно провести более длинную реакции-образования связей (1-3 часа).

Описанная здесь процедура позволяет изготавливать сложные чипы с использованием стандартного синтеза и точного определения мест олигонуклеотидов, потому что нужно относительно небольшое

количество олигонуклеотидов. Например, если все 7-мерные олиго синтезированы (16384 пробы), можно определить списки 256 миллиона 14-меров.

Один важный вариант изобретенного способа заключается в применении более чем 1 по-разному меченую пробу на базовый ряд. Это можно выполнить имея в виду 2 цели; мультиплексирование для снижения числа отдельно гибридизированных рядов; или определение списка еще более длинных олиго-последовательностей, таких, как 3 x 6 или 3 x 7. В этом случае, если используются 2 метки, специфичность 3 последовательных олигонуклеотидов может быть почти абсолютной, потому что положительные места должны иметь достаточное количество сигналов обоих меток.

Еще один дополнительный вариант - использовать чипы, содержащие пробы B_xN_y , где y составляет от 1 до 4. Эти чипы позволяют считывать последовательности в разных кадрах. Этого можно также достигнуть путем использования соответствующих наборов меченных проб или обе пробы F и P могут иметь некоторые неспецифицированные концевые позиции (т.е. некоторый элемент с окончательным вырождением). Можно также использовать универсальные базы как часть образователя связи для соединения проб с определенной последовательностью с твердой основой. Это делает пробу более готовой к гибридизации и делает строение более стабильным. Если проба имеет 5 баз, можно, например, использовать 3 универсальные базы как образователь связи (фиг.4).

ПРИМЕР VIII

АНАЛИЗ ПОЛУЧЕННЫХ ДАННЫХ

Файлы с изображениями анализируются программой анализа изображений, такой, как программа DOTS (Дрманач и др., 1993), [15,16] и шкалируются и оцениваются включенными статистическими функциями, например, в программу SCORES (Дрманач и др., 1994), [17]. Из распределения сигналов определяется оптимальный порог для преобразования сигнала в +/- выход.

Из положения обнаруженной метки определяется F+P последовательности нуклеотидов из фрагментов путем комбинирования известных последовательностей связанных и меченных проб, соответствующих меченым позициям. Полная последовательность нуклеиновой кислоты или подфрагменты последовательности первоначальной молекулы, такой, как хромосома человека, затем собираются из перекрывающихся последовательностей F+P, определенных вычислительным вычитанием.

Один вариант состоит в преобразовании сигналов гибридизации, например меток, в +/- выход во время процесса сборки последовательности. В этом случае сборка начинается с последовательности F+P с очень большим счетом, например, с F+P последовательности AAAAAATTTTTT (посл. ид. N.1). Счета всех 4-х возможных перекрывающихся проб AAAAAATTTTTT (посл. ид. N.3), AAAAAATTTTTT (посл. ид. N.4), AAAAAATTTTTT (посл. ид. N.5) и AAAAAATTTTTT (посл. ид. N.6) и 3 дополнительные пробы, которые отличны в начале (TAAAAATTTTTT, посл. ид. N.7; CAAAAATTTTTT, посл. ид. N.8; GAAAAATTTTTT, посл. ид. N.9) сравниваются и определяются 3 результата: (1) только начальная проба и только 1 из 4-х перекрывающихся проб имеют счета, которые значительно положительны по отношению к другим 6 пробам, в этом случае последовательность AAAAAATTTTTT (посл. ид. N.1) будет продлена на 1 нуклеотид вправо; (2) ни одна проба, за исключением начальной пробы, имеет значительно положительный счет, сборка остановится, например, последовательность AAAAAATTTTTT (посл. ид. N.10) находится на конце молекулы ДНК, последовательность которой создается, (3) обнаружена более чем 1 значительно положительная проба среди перекрывающихся или/и других 3-х проб; сборка останавливается из-за ошибки или разветвления (Дрманач и др., 1989), [12].

Процессы вычислительного вычитания используют компьютерные программы с применением существующих алгоритмов (см., например, Певзнер, 1989; Дрманач и др., 1991, [14]; Лабат и Дрманач, 1993, [29]).

Если, вдобавок к F+P, F(пространство 1)P, F(пространство 2)P, F(пространство 3)P или F(пространство 4)P определяются, алгоритмы будут использоваться, чтобы совместить все наборы данных для корректировки потенциальных ошибок или для разрешения ситуации, когда имеется проблема разветвления (см., например, Дрманач и др., 1989, [12]; Бейнс и др., 1988, [1]).

ПРИМЕР IX

ПОВТОРНОЕ ИСПОЛЬЗОВАНИЕ ЧИПОВ СОЗДАНИЯ ПОСЛЕДОВАТЕЛЬНОСТИ

Если образование связей используется в процессе создания последовательности, то обычный чип олигонуклеотидов нельзя сразу же повторно использовать. Изобретатель рассматривает, что это можно решить разными способами.

Можно использовать рибонуклеотиды для второй пробы, пробы P, так что эту пробу можно затем удалить обработкой РНКазой. В обработке РНКазой может использоваться РНКаза А, эндорибонуклеаза, которая специфически разъедает одно-свитую РНК 3' до радикалов пиримидина и расщепляет связь фосфата до соседнего нуклеотида. Конечными продуктами являются пиримидин 3' фосфаты и олигонуклеотиды с окончательными пиримидин 3' фосфатами. РНКаза А работает в отсутствии адьюнктов и двухвалентных катионов.

Чтобы использовать РНКазу, обычно выдерживают чип в любом пригодном содержащем РНКазу буфере, как описано Сэмбруком и др., (1989), [41]. Уместно использовать 30-50мкл содержащего РНКазу буфера на 8 x 8мм или 9 x 9мм матрицы при 37°C от 10 до 60 минут. Затем промывают буфером гибридизации.

Хотя это не применяется широко, можно также использовать базу урацила, как описали Крейг и др. (1989), [7], в конкретных примерах осуществления. Разрушение комбинации связанной пробы для получения повторно используемого чипа можно достичь вывариванием ремонтным ферментом *E.coli*, ураци-ДНК гликозилазой, которая удаляет урацил из ДНК.

Можно также сформировать специфично расщепляемую связь между пробами и затем расщепить связь после обнаружения. Например, этого можно достичь химическим образованием связи, как описано Шабаровой и др. (1991), [45]; и Долинной и др. (1988), [9].

Шабарова и др. (1991), [45] описывают конденсацию олигодеоксирибо нуклеотидов цианоген бромидом

как конденсирующим веществом. В их одноэтапной химической реакции образования связи олигонуклеотиды нагреваются до 97°C, медленно охлаждаются до 0°C, затем добавляется 1мкл 10 М BrCN в ацетонитриле.

Долинная и др. (1988), [9] показывают как включать межнуклеотидные связи фосфорамидата и пирофосфата в дуплексы ДНК. Они также используют метод образования химической связи для модификации сахара-фосфатного основания ДНК с растворимым в воде карбодиимидом (CDI) как связующим веществом. Селективное расщепление фосфоамидных связей предполагает контакт с 15% CH₃COOH в течение 5 мин при 95°C. Селективное расщепление, связи пирофосфата предполагает контакт со смесью пиридин-вода (9:1) и с только что дистиллированным (CF₃CO)₂O.

Хотя соединения и способы согласно этому, изобретению были описаны в терминах предпочтительных примеров его осуществления, для специалистов будет очевидно, что можно изменить соединение, способы и этапы или последовательность этапов описанного здесь способа, не отходя от замысла, духа и объема изобретения. Более конкретно очевидно, что некоторые вещества, которые связаны химически и физиологически, могут заменяться описанными здесь веществами, и будут получены те же самые или похожие результаты. Все такие аналогичные заменители и модификации, очевидные для специалистов, считаются находящимися в духе, рамках и замысле изобретения, определяемого прилагаемой формулой изобретения. Все заявленное, вещества и способы, могут получаться и выполняться без ненужного экспериментирования.

ССЫЛКИ

1. Bains et al., 1988, J. Theor. Biol., 135:303-307.
2. Broude et al., 1994, Proc. Natl. Acad. Sci. USA, 91:3072-3076.
3. Brumbaugh et al., 1988, Proc. Natl. Acad. Sci. U.S.A., 85:5610-5614.
4. Cantor et al., 1992, Genomics, 13, 1378.
5. Gate et al., 1991, GATA, 8(3): 102-106.
6. Chu et al., 1983, Nucleic Acids Res., 11:6513-6529.
7. Craig et al., 1989, Nucleic Acids Research, 17(12): 4605.
8. Dahlen et al., 1987, Mol. Cell. Probes 1:159-168.
9. Dolinnaya et al., 1988, Nucleic Acids Research, 16(9):3721-3738.
10. Drmanac & Crkvenjakov, 1990, Scientia Yugoslavica, 16, 97.
11. Drmanac & Crkvenjakov, U.S. Patent 5,202,231.
12. Drmanac et al., 1989, Genomics, 4:114-128.
13. Drmanac et al., 1991, J. Biomol. Struct. & Dyn., 8:1085.
14. Drmanac et al., 1991, In "Electrophoreses, Supercomputers and the Human Genome", pp 47-59, World Scientific Publishing Co., Singapore.
15. Drmanac et al., 1993a, Proceedings of 2nd International Conference on Bioinformatics, Supercomputing, and Complex Genome Analysis, World Scientific Publishing Co./ pp. 121-134.
16. Drmanac, et al., 1993b, DNA Sequence Determination by Hybridization: a Strategy for Efficient Large-Scale Sequencing, Science, 260:1649-1652.
17. Drmanac, 1994, Abstract Book for Genome Mapping and Sequencing; arranged by Richard Myers, David Porteous and Robert Waterstone, Cold Spring Harbor Laboratories, p.60.
18. Drmanac et al., 1994, Proceedings of the 3rd International Workshop of Transcribed Sequences, In Press.
19. Duncan & Cavalier, 1988, Analytical Biochemistry, 169:104-108.
20. Eggers et al., 1994, BioTechniques, 17(3):516-524.
21. Fitzgerald et al., 1992, Nucleic Acids Research, 20(14):3753-62.
22. Fodor et al., 1991, Science, 251:767-768.
23. Hoheisel & Lehrach, 1990, FEBS Lett., 274(1,2):103-106.
24. Inouye & Hondo, 1990, J. Clin. Microb., 28:1469-1472.
25. Jacobsen et al., 1990, Genomics, 8:001-007.
26. Keller et al., 1988, Anal. Biochem., 170:441-450.
27. Keller et al., 1989, Anal. Biochem., 177:27-32.
28. Khrapko et al., 1991, J. DNA Sequencing Mapping, 1, 375.
29. Labat and Drmanac, 1993, Proceedings of 2nd International Conference on Bioinformatics, Supercomputing, and , Complex Genome Analysis, World Scientific Publishing Co., pp. 555-565.
30. Landegren et al. 1988, Science, 241:1077-1080.
31. Maxam & Gilbert, 1977, Proc. Natl. Acad. Sci., 74, 560.
32. Morriey & Collins, 1989, Mol. Cell. Probes 3:189-207.
33. Murakami et al., 1991, Nucleic Acids Research, 19(15):4097-4102.
34. Nagata et al., 1985, FEBS Letters, 183:379-382.
35. Nichols et al., 1994, Nature, 369:492.
36. Pease et al., 1994 Proc. Natl. Acad. Sci., 91:5022-5026.
37. Peterkin et al., 1987, BioTechniques 5(2):132-134.
38. Peterkin et al., 1989, Food Microbiology 5(2):281-284.
39. Pontius & Berg, 1991, Proc. Natl. Acad. Sci. U.S.A., 88:8237-8241.
40. Rasmussen et al., 1991, Analytical Biochemistry, 198:138-142.
41. Sambrook et al., 1989, Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory. Cold Spring Harbor, NY.
42. Sanger, et al., 1977, Proc. Natl. Acad. Sci., 74, 5463.
43. Schrieffer et al., 1990, Nucleic Acids Research, 18(24):7455.
44. Schubert et al., 1990, Nucleic Acids Research, 18(11):3427.
45. Shabarova et al., 1991, Nucleic Acids Research, 19(15):4247-4251.
46. Sharp et al., 1989 Food Microbiology, 6:261-265.

47. Southern, PCT Patent Application WO 89/10977.
48. Southern & Maskos, PCT Patent Application WO 90/03382.
49. Southern et al., 1992, Genomics, 13, 1008.
50. Strezoska et al., 1991, Proc. Natl. Acad. Sci., 88, 10089.
51. Van Ness et al., 1991, Nucleic Acids Research, 19(12):3345.
52. Wu & Wallace, 1989 Gene, 76:245-254.

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:1:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 12 пар оснований
- (B) ТИП: нуклеиновая кислота
- (C) ЦЕННОСТЬ: единичная
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:1:

AAAAAATTTT TT

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:2:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 13 пар оснований
- (B) ТИП: нуклеиновая кислота
- (C) ЦЕННОСТЬ: единичная
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:2:

AAAAAATTTT TTС

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:3:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 12 пар оснований
- (B) ТИП: нуклеиновая кислота
- (C) ЦЕННОСТЬ: единичная
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:3:

AAAAATTTTT TA

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:4:

- (A) ДЛИНА: 12 пар оснований
- (B) ТИП: нуклеиновая кислота
- (C) ЦЕННОСТЬ: единичная
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:4:

AAAAATTTTT TT

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:5:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 12 пар оснований
- (B) ТИП: нуклеиновая кислота
- (C) ЦЕННОСТЬ: единичная
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:5:

AAAAAATTTT TC

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:6:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 12 пар оснований
- (B) ТИП: нуклеиновая кислота
- (C) ЦЕННОСТЬ: единичная
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:6:

AAAAAATTTT TG

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:7:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 12 пар оснований
- (B) ТИП: нуклеиновая кислота
- (C) ЦЕННОСТЬ: единичная
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:7:

TAAAAATTTT TT

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:8:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(A) ДЛИНА: 12 пар оснований

(B) ТИП: нуклеиновая кислота

(C) ЦЕННОСТЬ: единичная

(D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:8:

САААААТТТТ ТТ

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO:9:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(A) ДЛИНА: 12 пар оснований

(B) ТИП: нуклеиновая кислота

(C) ЦЕННОСТЬ: единичная

(D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:9:

ГАААААТТТТ ТТ

ИНФОРМАЦИЯ ДЛЯ SEQ ID NO: 10:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

(A) ДЛИНА: 11 пар оснований

(B) ТИП: нуклеиновая кислота

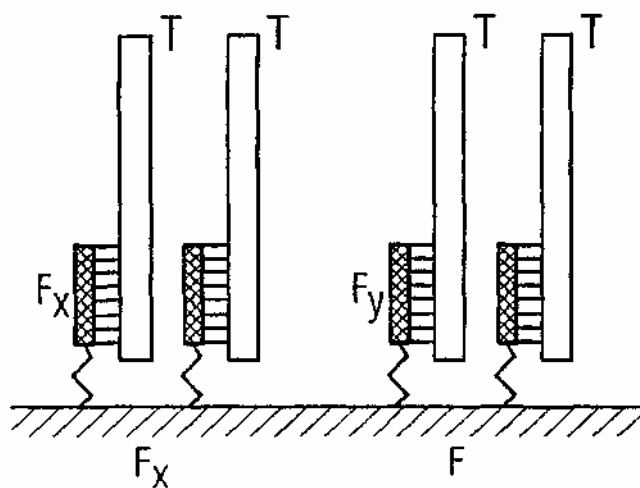
(C) ЦЕННОСТЬ: единичная

(D) ТОПОЛОГИЯ: линейная

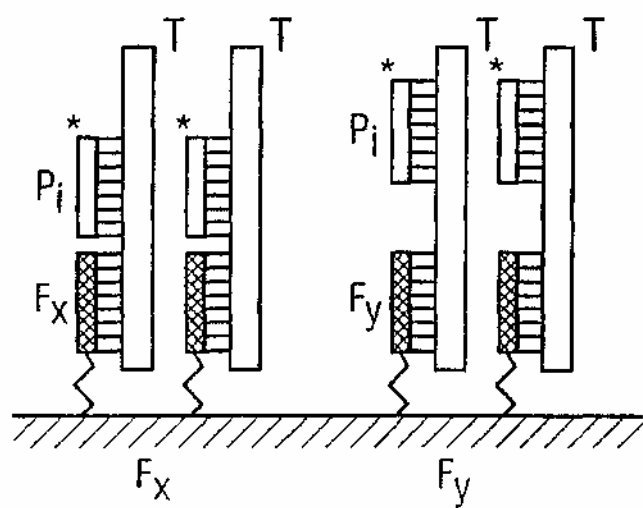
(ii) ТИП МОЛЕКУЛЫ: DNA

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO:10:

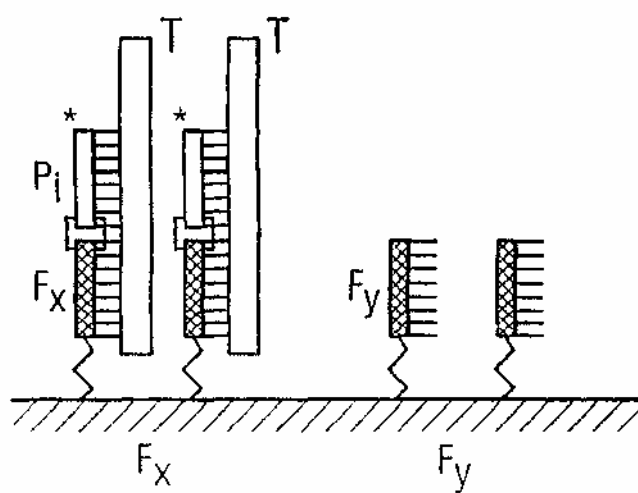
ААААААТТТТ Т



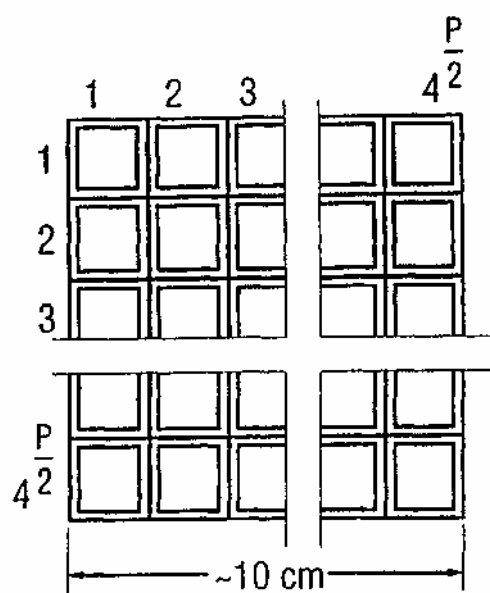
Фиг. 1А



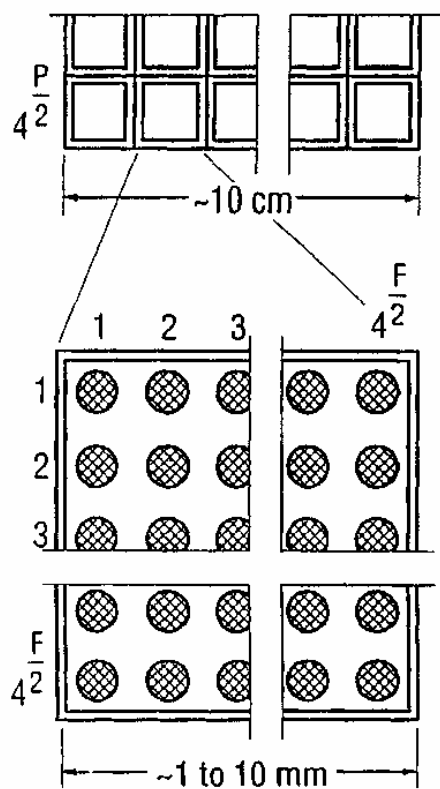
Фиг. 1В



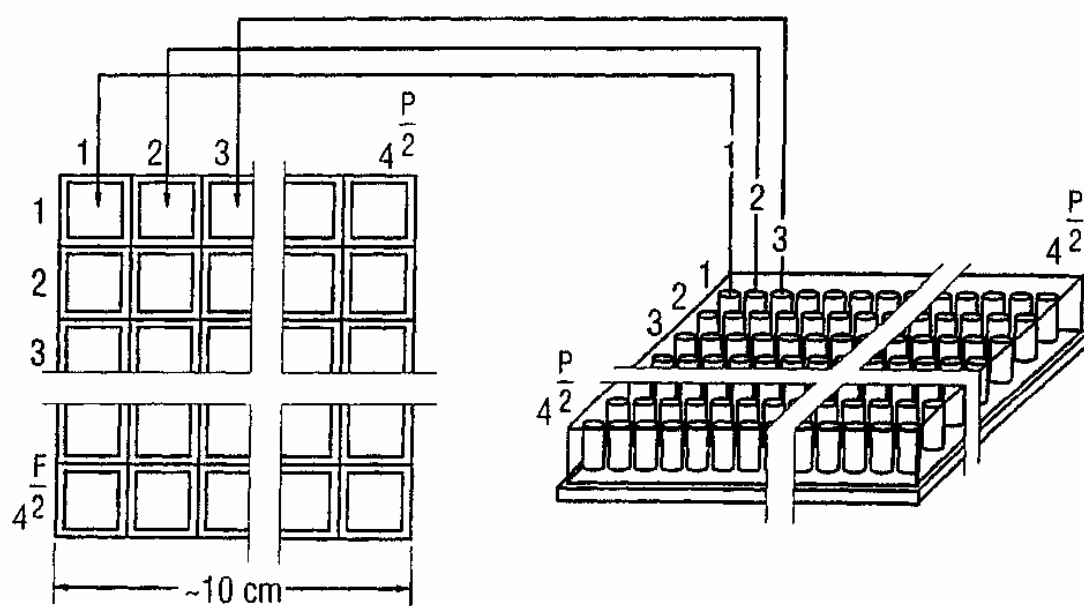
Фиг. 1С



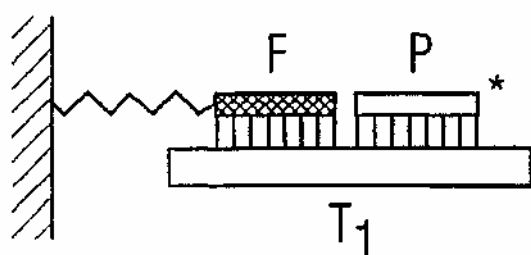
Фиг. 2А



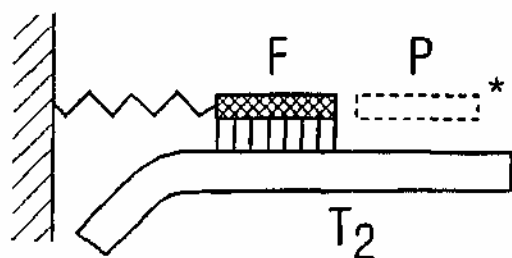
Фиг. 2В



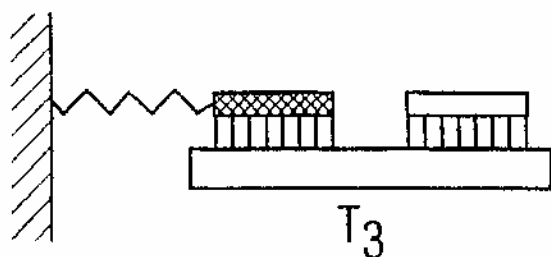
Фиг. 2С



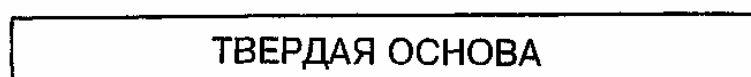
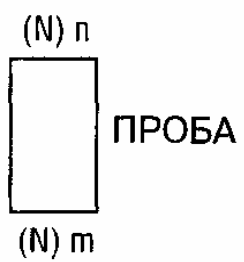
Фиг. 3А



Фиг. 3В



Фиг. 3С



Фиг. 4