

СОВРЕМЕННЫЕ МЕТОДЫ СЕКВЕНИРОВАНИЯ ДНК



**Выполнил: магистрант второго года обучения очной
формы**

**по направлению
06.04.01.68 Биология**

программа Общая биология

Группа МБИО21

Маркина Виктория Олеговна

Проверил:

**кандидат биологических наук,
доцент**

Курамшина Зия Мухтаровна



Методы расшифровки нуклеотидной последовательности нуклеиновых кислот в отечественной литературе принято называть методами **секвенирования**.

Еще в 50-е годы прошлого века были разработаны методы, позволяющие определять последовательность аминокислот в полипептидной цепи. Теоретически это несложно, поскольку все аминокислоты, встречающиеся в природных белках, имеют разные свойства. Поэтому, когда был расшифрован генетический код, появилась возможность восстанавливать нуклеотидную последовательность транскрибируемой ДНК по аминокислотной последовательности соответствующего белка. Однако, генетический код является вырожденным. Следовательно, первичная структура ДНК, полученная на основе анализа последовательности аминокислот, не является однозначной. Кроме того, для эукариот таким способом можно восстановить лишь нуклеотидный состав экзонов, тогда как информация о составе интронов теряется в результате сплайсинга.

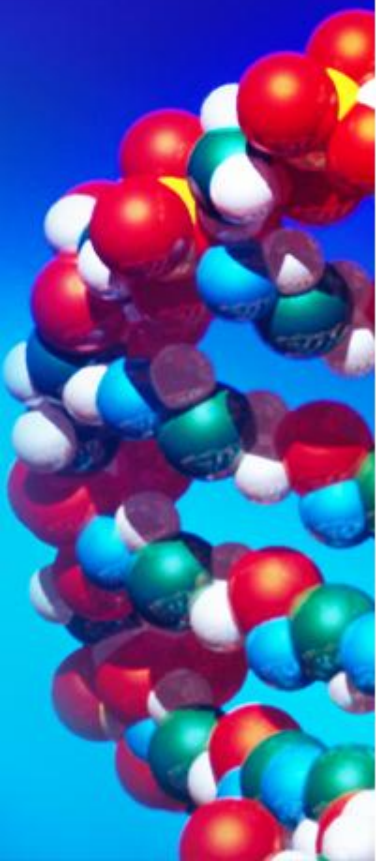


КЛАССИЧЕСКИЕ МЕТОДЫ СЕКВЕНИРОВАНИЯ ДНК

Классические методы секвенирования ДНК были разработаны в конце прошлого века и в настоящее время имеют самое широкое распространение в научно-исследовательских и медико-диагностических учреждениях.

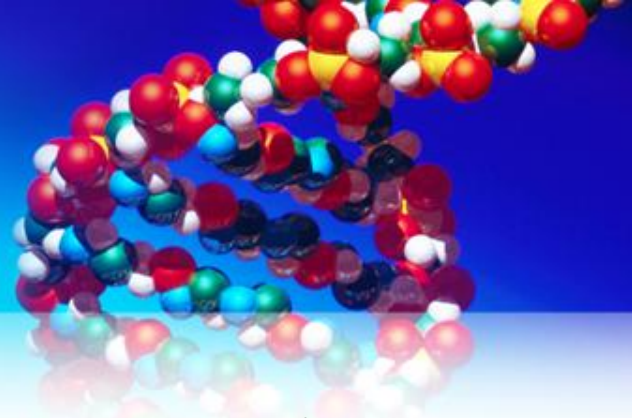
Секвенирование методом терминации цепи с помощью капиллярного электрофореза.

Для проведения секвенирования с помощью капиллярного электрофореза используют самый распространенный на сегодняшний день способ подготовки ДНК – «метод терминации цепи» или «дидезокси метод», разработанный в 70-х годах прошлого века F.Sanger. В 1985 г. этот метод был существенно улучшен, после того как радиоактивную метку смогли заменить светящейся флуоресцентной.





Особенность метода «терминации цепи» заключается в использовании химически модифицированных дезоксирибонуклеотидов, составляющих нуклеотидную последовательность ДНК. Каждая разновидность нуклеотида «помечена» своей флуоресцентной молекулой-маркером. Когда один из таких модифицированных нуклеотидов при синтезе комплементарной цепи ДНК встраивается в нее, дальнейшая достройка этой цепи прекращается. Видоизмененные разновидности нуклеотидов присутствуют в реакционной смеси в значительно меньших количествах, чем обычные нуклеотиды. В результате получается смесь, содержащая полный набор синтезированных фрагментов ДНК, каждый из которых начинается в одном и том же месте, но заканчивается во всех возможных положениях вдоль цепи ДНК-матрицы.



Автоматические секвенаторы разделяют эти фрагменты по величине, как ионы, движущиеся в электрическом поле от «минуса» к «плюсу», пропуская всю смесь через тончайшие капилляры, наполненные гелем. По мере того, как появляется очередной фрагмент, он освещается лазером, что заставляет светиться меченый нуклеотид на его конце. Таким образом, друг за другом определяются разновидности меченых нуклеотидов на концах всех фрагментов и составляется полная нуклеотидная последовательность исследуемого образца ДНК.

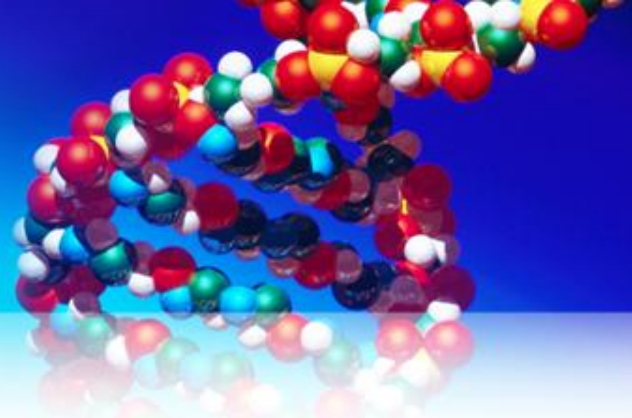


Относительно невысокая стоимость, точность, а также простота автоматизации делают этот метод широко используемым в лабораторной практике. В современной интерпретации данный метод позволяет за один рабочий цикл (ран) в одном капилляре с точностью до 98 % определить последовательность ДНК длиной до 1000 пар нуклеотидов (п.н.)

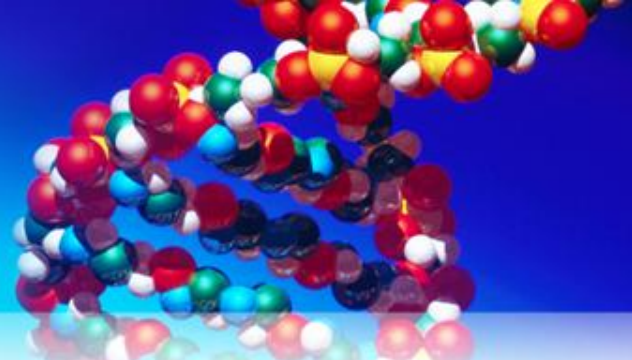


В начале XXI века достигнут предел производительности секвенаторов на основе капиллярного электрофореза, в одном приборе до 384 капилляров. Самые высокопроизводительные модели капиллярных секвенаторов – «MegaBACE 4500» (384 капилляра) и «3730x1 DNA Analyzer» (96 капилляров) – способны секвенировать за одни сутки тысячи фрагментов ДНК длиной до 900 п.н. с точностью 95–98 %. Метод терминации цепи был использован при секвенировании первого генома человека, но при таких масштабах оказался очень дорогостоящим и медленным.

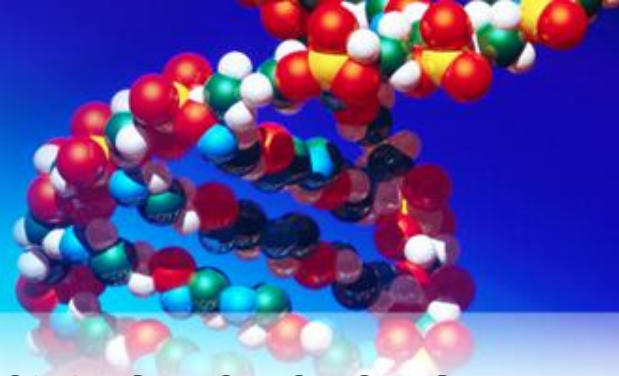
Пиросеквенирование ДНК.



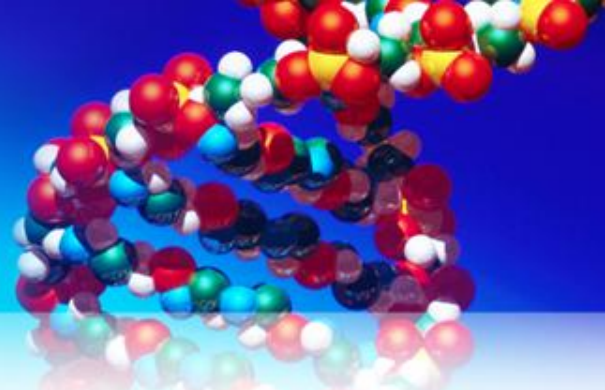
Метод пиросеквенирования основан на определении пирофосфата, образующегося при синтезе комплементарной цепи ДНК. В 1988 г. Е.Нутан впервые предложил использовать этот метод для секвенирования. В 1996 г. Р.Nyrén и М.Ronaghi завершили детальную разработку данного метода.



Согласно этой технологии растворы четырех типов нуклеотидов добавляются последовательно в процессе секвенирования фрагментов ДНК и отмываются после каждой реакции. При включении очередного нуклеотида в синтезируемую комплементарную цепь на подготовленной одноцепочечной ДНК-матрице происходит детекция высвобождающегося пирофосфата путем связывания его с белком люциферазой и в итоге получении светового сигнала. В течение реакции комплементарная цепь ДНК достраивается, а последовательность нуклеотидов определяется по световым сигналам в виде пиков на пирограмме, интенсивность которых пропорциональна количеству встроенных в цепь ДНК нуклеотидов одного вида из очередной реакционной смеси.



Простые химические реакции и надежная система детекции исключают необходимость использования в данном методе гелей, электрофореза и специфических флуоресцентных меток. Лидирующей в этом направлении корпорацией «Biotage AB» разработан ряд систем, в частности «PyroMark Q96 ID», которая имеет высокую скорость определения последовательности ДНК (96 образцов за 1 ч) с точностью до 99 %. В приборе при пиросеквенировании реагенты вводятся головкой струйного принтера в лунки планшета с образцами, после чего не удаляются, а расщепляются специальным ферментом – апиразой.



Однако метод не позволяет качественно определить длину секвенированных фрагментов ДНК более 200 п. н., а также точно определить в последовательности ДНК число одиночных одинаковых нуклеотидов более 6 подряд. Таким образом, метод эффективен для быстрого определения большого числа коротких последовательностей (маркерных зон) ДНК и РНК, выявления единичных мутаций нуклеотидов, повторного секвенирования и т.д. При этом значительно снижается время и стоимость исследования в сравнении с капиллярным электрофорезом.



Методы секвенирования ДНК нового поколения

Чтобы снизить стоимость процедуры секвенирования и увеличить ее производительность, в новых методах определения нуклеотидной последовательности ДНК проводится секвенирование миллионов ее фрагментов одновременно. Отличительной особенностью методов секвенирования ДНК нового поколения является их направленность на получение нуклеотидных последовательностей целых геномов различных организмов, включая человека.

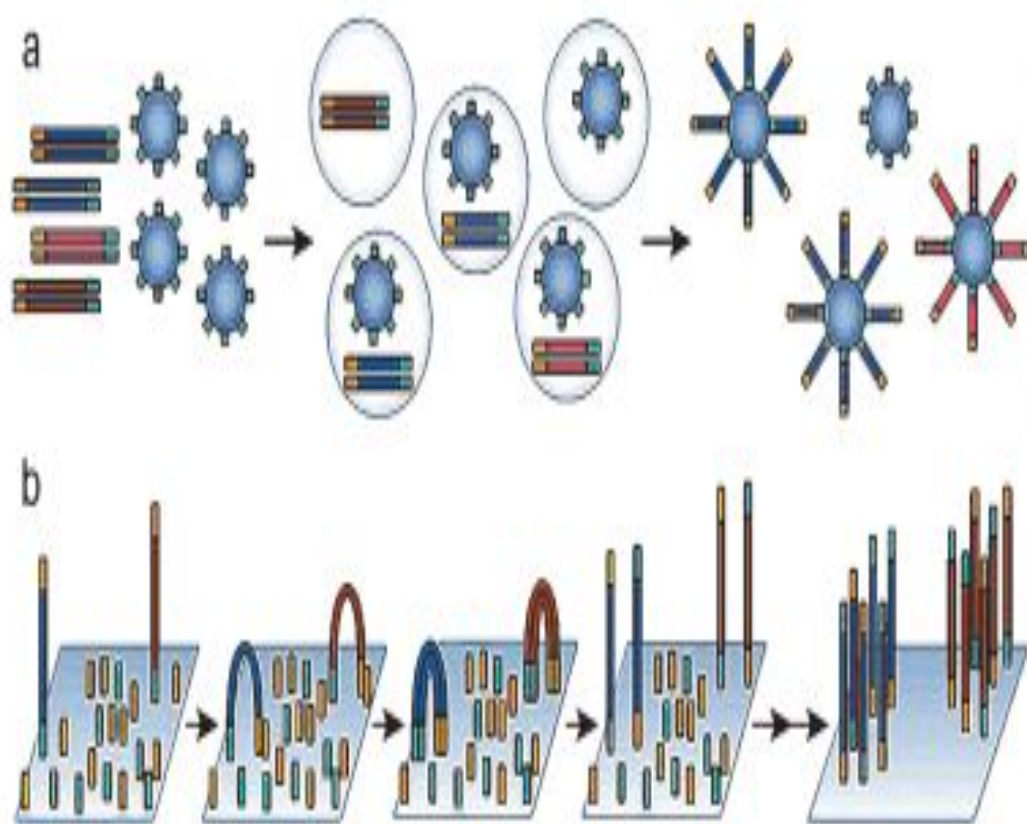


Рис. 1. Клональная амплификация секвенируемых фрагментов в методах NGS:

А – для платформ «GS FLX», «SOLiD» и «PGM» используется эмульсионная ПЦР на поверхности микрошариков. Каждый фрагмент ДНК с обоих концов имеет адаптеры. На поверхности каждой микросферы прикреплен один вид праймеров, которые комплементарны одному из адаптеров на концах фрагментов ДНК; Б – технология «Solexa» использует кластерную ПЦР на подложке, поверхность которой покрывают праймеры двух типов, соединенные с ней через линкеры. Продукты амплификации, происходящие от любого одиночного фрагмента ДНК из библиотеки, остаются локально расположенными около места синтеза и снова амплифицируются по соседству



Сейчас высококачественного секвенирования геномов (точностью 99,999 % и более) добиваются, как минимум, после 30-кратного «прочтения» их последовательностей, секвенирование полного генома человека в этом случае требует прочтения не менее 90 млрд. пар оснований ДНК. Последние модели секвенаторов нового поколения могут выполнить секвенирование до шестнадцати полных геномов человека всего за один рабочий цикл.

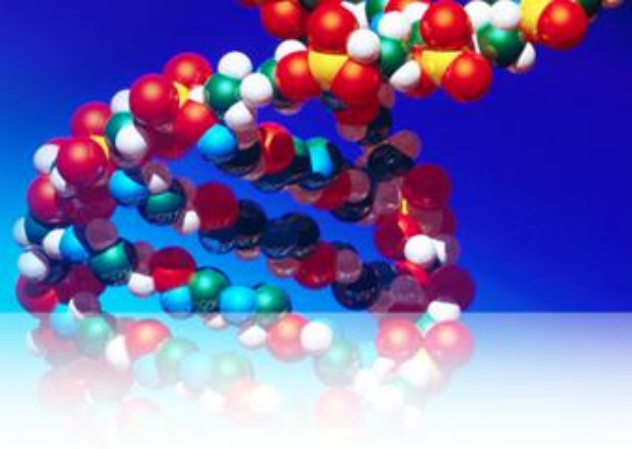


Высокопроизводительное пиросеквенирование ДНК.

В 2004 г. компания «454 Life Sciences» дополнила метод пиросеквенирования возможностями современных технологий. Основой модификации явилось использование эмульсионной ПЦР с одновременной параллельной подготовкой многих сотен тысяч фрагментов ДНК для их секвенирования.



После проведения всех предварительных этапов пробоподготовки каждый из четырех видов нуклеотидов, смешанный с другими реактивами для пиросеквенирования, подается последовательно в проточную камеру, куда помещается подложка, имеющая несколько миллиардов лунок, заполненных микросферами (одна лунка – одна микросфера). Каждая микросфера содержит на своей поверхности после эмульсионной ПЦР многие сотни тысяч копий одного из исходных ДНК-фрагментов. Если в лунку попадает тип нуклеотидов, который комплементарен очередному нуклеотиду в достраиваемой одноцепочечной цепи матрицы ДНК, то полимераза встраивает эти нуклеотиды, что приводит через каскад реакций к высвобождению пирофосфата и генерации общего светового сигнала из лунки. Интенсивность сигнала из каждой лунки пропорциональна количеству нуклеотидов одного вида, встроенных в цепи ДНК за один проход реакционной смеси. Дно подложки находится в оптическом контакте с оптико-волоконным световодом, что позволяет регистрировать излучаемые фотоны со дна многих сотен тысяч лунок одновременно, в которых произошло встраивание очередного известного нуклеотида в достраиваемые комплементарные цепи ДНК в процессе пиросеквенирования.



Длина единичного фрагмента ДНК, секвенируемого методом высокопроизводительного пиро-секвенирования, достигает 1000 п.н. с точностью до 99 %. Производительность секвенирования последней из моделей – «GS FLX Titanium XL+» близка к 1 млрд п.н. за один рабочий цикл с консенсусной точностью до 99,997 % (при 15-кратном прочтении).

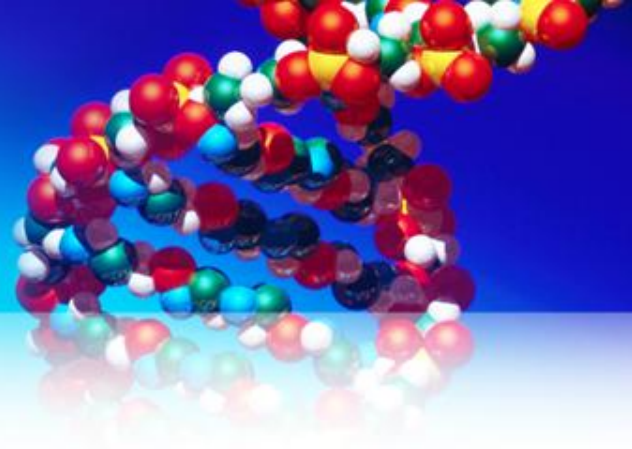


Секвенирование на молекулярных кластерах с использованием флуоресцентно меченных предшественников

В 2005 г. компанией «Illumina» предложен метод секвенирования ДНК на молекулярных кластерах с использованием флуоресцентно меченных предшественников, реализованный в виде платформы «Solexa». В этом методе фрагменты ДНК-матрицы фиксируют на подложке вместе с праймерами и проводят их амплификацию, в результате которой вокруг каждого исходного фиксированного фрагмента ДНК образуется около 1000 его копий. В целом на подложке образуются до миллиарда таких скоплений — кластеров.



Далее ДНК денатурируют и проводят секвенирование ее одноцепочечных фрагментов. Для этого добавляют свободно плавающие праймеры и меченные флуоресцентными красителями разных цветов терминирующие нуклеотиды всех четырех типов, чтобы к каждому одонитевому фрагменту ДНК присоединился праймер и только 1 нуклеотид, цвет которого можно определить с помощью лазера. Затем присоединенный нуклеотид химически модифицируют, чтобы к нему мог присоединиться следующий терминирующий нуклеотид, таким образом, продолжают сиквенс (достраивание) комплементарной цепи, детектируя прибором флуоресценцию из многих сотен миллионов точек-кластеров одновременно.

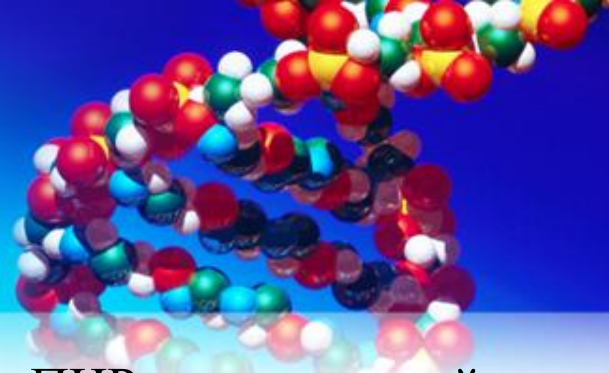


Длина единичного фрагмента ДНК, секвенируемого данным методом, достигает 300 п.н. с точностью 99 %. Последняя выпущенная компанией «Illumina» система секвенирования – «HiSeq X» за один цикл способна секвенировать до 16 полных геномов человека с 30-кратным перекрытием (1600 млрд. п.н.) и консенсусной точностью до 99,999 %.



Циклическое лигазное секвенирование.

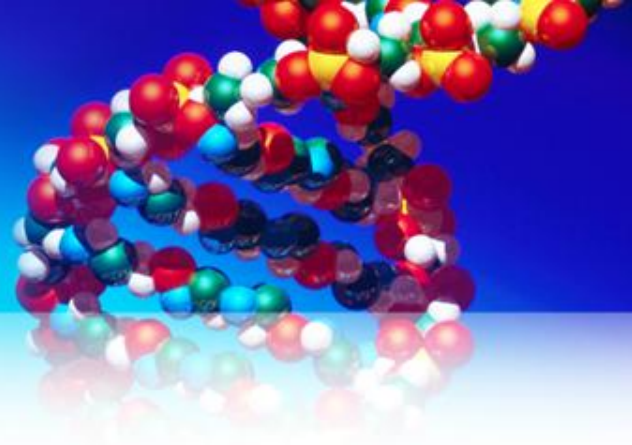
В 2007 г. компанией «Applied Biosystems» (теперь входит в состав корпорации «Life Technologies») была представлена платформа «SOLiD» на основе технологии циклического лигазного секвенирования.



В данном методе также используют эмульсионную ПЦР, после которой каждая из микросфер содержит на своей поверхности около миллиона копий одного из исходных ДНК-фрагментов. Затем фрагменты ДНК денатурируют и модифицируют по 3' концу, что позволяет ковалентно связать миллиарды микросфер, покрытых фрагментами ДНК, с подложкой. Затем следует процесс циклического лигазного секвенирования. К одноцепочечной матрице ДНК присоединяют праймер и добавляют синтезированные флуоресцентно меченные восьмичленные олигонуклеотидные зонды. Каждый зонд в позиции 1 и 2 содержит два известных нуклеотида, которые в сочетании со следующими тремя основаниями должны иметь комбинацию последовательности, комплементарную соответствующим коротким фрагментам матрицы. Последние 3 основания, одинаковые (универсальные) для всех зондов, имеют на конце флуоресцентную метку. После того, как один из зондов комплементарно связывается с первыми пятью основаниями от праймера на одноцепочечной матрице ДНК, лигаза сшивает его с праймером. При этом метка детектируется и удаляется вместе с последними тремя универсальными нуклеотидами зонда.



Затем происходит сшивание (лигирование) другого зонда, комплементарно соединившегося со следующими пятью основаниями, от зонда уже присоединенного ранее и так далее до конца матрицы ДНК. После чего построенная комплементарная цепь удаляется, а к матрице добавляется праймер, который смещен по месту посадки на один нуклеотид относительно посадки предыдущего праймера, и весь процесс достройки второй цепи методом лигирования повторяется. Для определения полной последовательности секвенируемого фрагмента нужно провести эту процедуру с пятью праймерами, каждый из которых смещен относительно предыдущего праймера по месту посадки на один нуклеотид.



Длина единичного фрагмента ДНК, секвенируемого методом лигирования, достигает 75 п.н. с точностью до 99,99 %. Выпущенный компанией «Applied Biosystems» прибор «5500xl SOLiD System» способен за один цикл секвенировать 3 полных генома человека с 30-кратным покрытием и точностью 99,9999 % .



Полупроводниковое секвенирование.

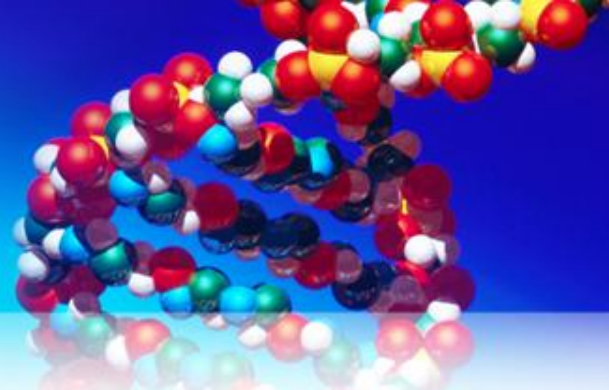
В 2010 г. компания «Ion Torrent» (США) (теперь входит в состав корпорации «Life Technologies») представила метод секвенирования ДНК на полупроводниковом чипе (Semiconductor sequencing).



Данный метод не требует модификации нуклеотидных оснований, хемилюминесцентных или флуоресцентных меток. Когда нуклеотид включается в комплементарную достраиваемую цепь ДНК полимеразой, происходит выделение водородного иона как побочного продукта, что изменяет рН реакционной смеси, определение этого изменения рН положено в основу детекции последовательности нуклеотидов ДНК при полупроводниковом секвенировании. Пробоподготовка ДНК в данном методе практически не отличается от уже описанной для лигазного секвенирования. Проведение предварительной эмульсионной ПЦР здесь необходимо для усиления сигнала об изменении в значении рН. Это изменение рН в каждой отдельной лунке на чипе будет давать около миллиона копий одного из исходных ДНК-фрагментов, прикрепленных к поверхности микросферы. Для этого подготовленную таким образом библиотеку ДНК денатурируют и распределяют на микроэлектронный чип, имеющий миллионы микролунок (одна лунка – одна микросфера) с иончувствительной поверхностью, связанной с датчиками.



Затем последовательно добавляют реакционные ПЦР смеси, отличающиеся по составу наличием в них только одного из 4 видов нуклеотидов, и детектируют изменение рН в лунках, где произошло встраивание этих нуклеотидов в комплементарные цепи ДНК-матриц. Основанный на данной технологии секвенатор «Personal Genome Machine» (PGM) в настоящее время за один цикл определяет до 2 млрд п.н., при этом длина единичного прочтения превышает 400 п.н. с точность 99,5 %. Метод полупроводникового секвенирования быстро совершенствуется и за последние 2 года увеличил свою производительность в 100 раз. Выпущенный новый прибор «Ion Proton» на основе данной технологии нацелен на секвенирование генома человека. В 2014 г. данная платформа за один цикл будет определять до 60 млрд п. н.



Описанные системы секвенирования второго поколения значительно отличаются по производительности, стоимости и продолжительности рабочего цикла. По сочетанию всех этих параметров наиболее оптимальной для персонализации медицины выглядит система «HiSeq X», «NextSeq 500» и «Ion Proton», а для секвенирования бактериальных геномов – «MiSeq», или «PGM» (таблица).

Сравнительная характеристика NGS и NNGS секвенаторов (январь 2014 г.)

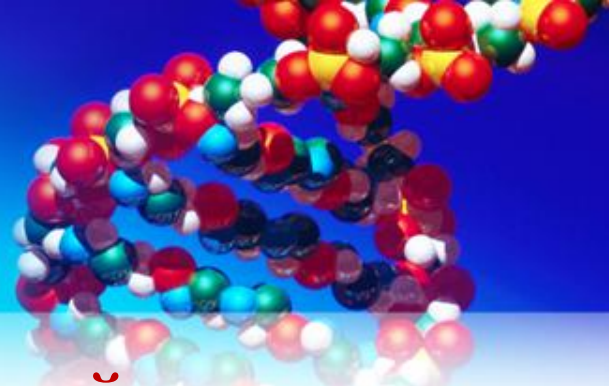
Характеризуемый показатель	Модели секвенаторов							
	NGS							NNGS
	GS Junior	FLX Ti+	MiSeq	HiSeq X	5500 xl	Ion Proton	PGM	PacBio RS
	Roche		Illumina		Life Technologies			Pacific Biosciences
Длина единичного «прочтения», п.н.	450	800	300	150	60	200	400	8500
Продолжительность этапа секвенирования, сут	0,42	0,42	0,3	3	7	0,17	0,3	0,1
Число секвенируемых нуклеотидов за цикл, млрд. п.н.	0,035	1	15	1600	180	10	2	6
Стоимость одного цикла (без приготовления библиотек), тыс. \$*	0,8	8	0,7	12	4	1,3	0,5	0,1

*Приблизительная стоимость одного цикла работы прибора в США (без приготовления библиотек).



Методы секвенирования новейшего поколения

Методы секвенирования третьего поколения (NNGS) призваны исправить основные недостатки методов второго поколения, а именно: сложную пробоподготовку, небольшую длину единичных прочтений, потребность в усилении сигнала от каждого из анализируемых фрагментов ДНК путем их амплификации, длительное время цикла, необходимость многочисленного повторного секвенирования.



Технология секвенирования одной молекулы.

В 2008 г. компания «Helicos BioSciences» впервые продемонстрировала на своем приборе метод секвенирования нуклеотидной последовательности единичных фрагментов ДНК без их предварительной амплификации. Эта технология – «tSMS» (true Single Molecule Sequencing) обладает очень высокой чувствительностью, благодаря разработанному компанией сверхнизкому фоновому свечению поверхности подложки, реагентам для секвенирования и способу визуализации. Согласно новому методу, ДНК разделяется на небольшие фрагменты, и к одному из концов каждого фрагмента с помощью специального фермента прикрепляют длинную последовательность из молекул аденина, имеющую в конце светящуюся метку. На специальную подложку прикрепляют последовательности из молекул тимина на расстоянии, позволяющем анализировать каждый фрагмент по отдельности. Когда образцы распределяют на подложку, то происходит комплементарное связывание последовательностей из тимина, прикрепленных на ее поверхности, с адениновыми последовательностями, связанными с фрагментами ДНК. Далее подложка сканируется, и по свечению меток находят положение каждого присоединенного отрезка молекулы ДНК.



Затем метку удаляют и добавляют по очереди каждый из четырех типов меченых нуклеотидов, детектируя свечение в местах связывания, после чего метки от присоединенных нуклеотидов удаляют и вводят следующий тип меченых нуклеотидов. Так, нуклеотид за нуклеотидом последовательность достраивается по комплементарной цепи ДНК, что позволяет секвенировать фрагменты длиной до 55 п.н. Компьютер записывает положение миллионов вспышек после каждой реакции. Такая система обеспечивает прочтение миллиардов нуклеотидов в сутки с точностью до 99,999 % при 20-кратном перекрывании. Однако прибор, основанный на данной технологии («Heli Score»), получился слишком дорогим, расход реагентов высоким при малой длине чтения нуклеотидных последовательностей. В ноябре 2012 г. компания «Helicos Bioscience» была признана банкротом и прекратила свое существование.

Секвенирование единичных молекул в реальном времени.

- В 2009 г. был представлен метод секвенирования «Single Molecule RealTime» (SMRT), разработанный компанией «Pacific Biosciences». Секвенатор этой компании позволяет «читать» каждый фрагмент ДНК десятки раз, при этом консенсусная точность определения его нуклеотидной последовательности достигает 99,9 %.

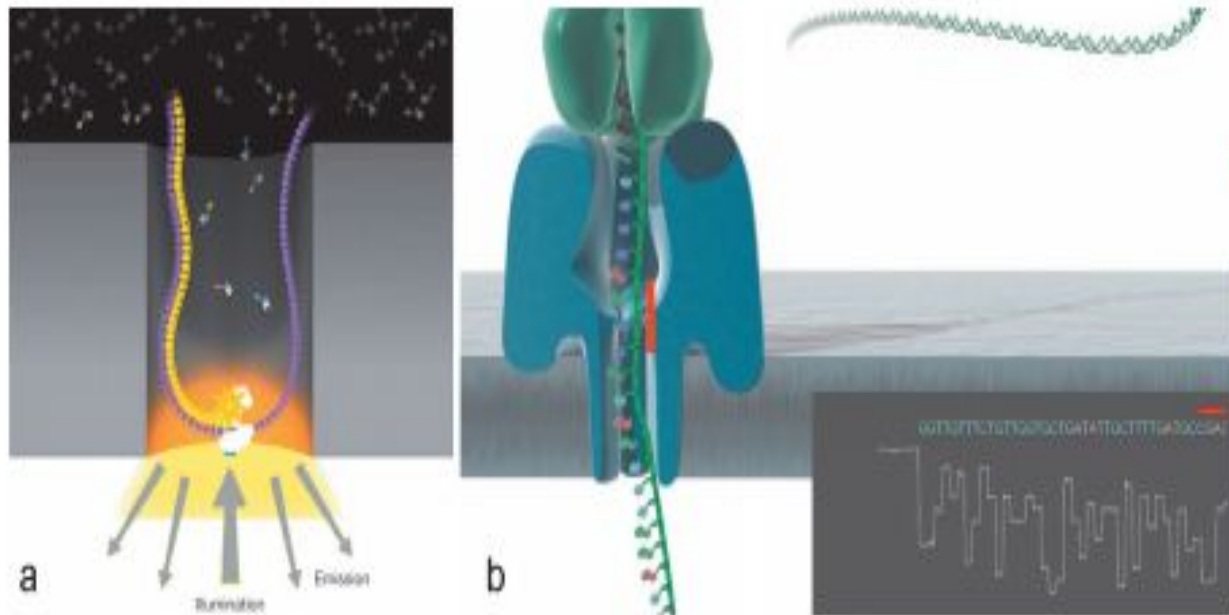
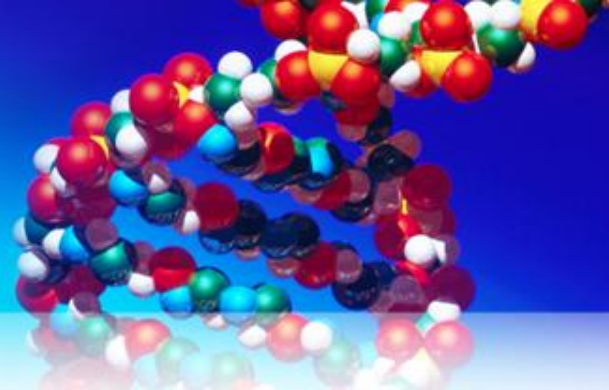


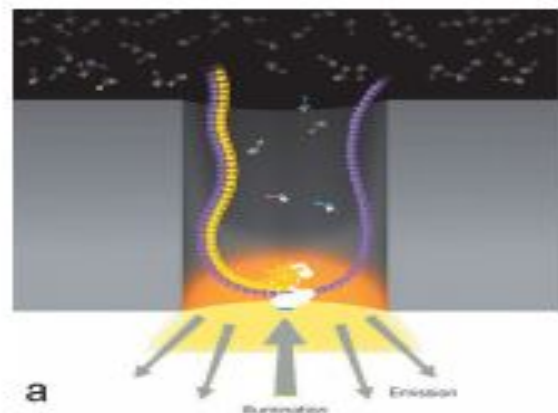
Рис. 2. Детекция сигнала при секвенировании:

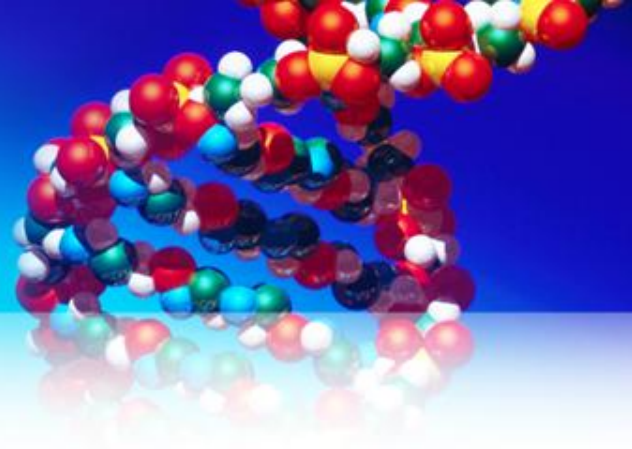
А – на платформе «SMRT»; Б – по технологии «Oxford Nanopore»



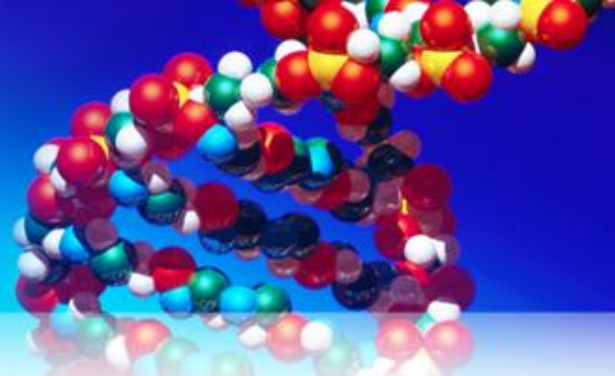
Технология данного метода основана на секвенировании в реальном времени одно-нитевых фрагментов ДНК длиной до 10000 п.н. и более с помощью ДНК-полимеразы. На дно специальных ячеек, расположенных на прозрачном стекле, прикрепляются одиночные молекулы ДНК-полимеразы. Область вокруг закрепленного фермента просвечивается с помощью специального лазера. В каждую ячейку добавляют нуклеотиды всех четырех типов, помеченные разными светящимися маркерами. Область, которую анализирует лазер, столь мала, что меченые нуклеотиды не задерживаются в ней достаточно долго, чтобы их свечение было зафиксировано прибором. Если же ДНК-фрагмент удерживается полимеразой, то во время достройки его комплементарной цепи сигнал от каждого присоединившегося нуклеотида фиксируется в сканируемой области. После присоединения очередного нуклеотида его светящаяся метка удаляется (рис. 2, А).

Рис.2. Детекция сигнала при секвенировании
А – на платформе «SMRT»





Прибор «РасВіо RS», основанный на этой технологии, имеет высокую собственную стоимость, но в то же время экспрессную пробоподготовку, высокую скорость и низкую стоимость секвенирования ДНК.



Секвенирование через нанопоры.

Возможность использовать матрицы с нанопорами для быстрого секвенирования ДНК исследуется уже более 15 лет в научных центрах Европы и США.



Нанопоры представляют собой nanoотверстия, которые могут быть биологическими, например порообразующий белок в мембране как липидный бислой, или твердотельными (из таких синтетических материалов, как нитрид кремния или графен). При секвенировании через нанопоры используется физическое различие между нуклеотидными основаниями для их идентификации в молекуле ДНК. Отрицательно заряженный одноцепочечный фрагмент ДНК длиной в несколько тысяч нуклеотидов протягивают через пору в мембране диаметром 2–5 нм, регистрируя изменение электропроводности нанопоры при помощи электродов по мере поочередного прохождения через нее нуклеотидов. Каждому типу основания соответствует свое изменение электропроводности из-за различия между ними по размерам, поэтому они закрывают пору в большей или меньшей степени и на разную продолжительность. Соответственно этому изменяется и электропроводность (рис. 2, b).

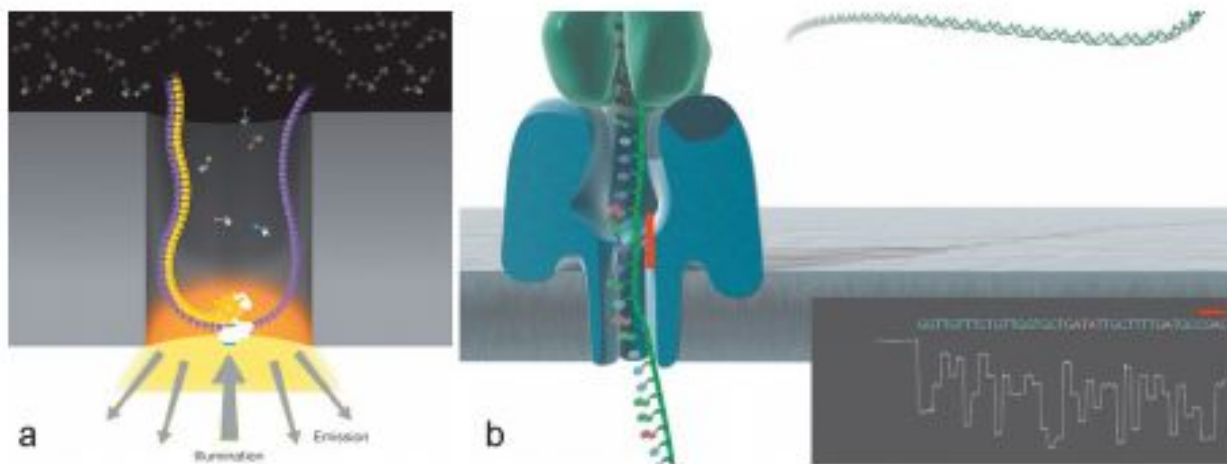


Рис. 2. Детекция сигнала при секвенировании:

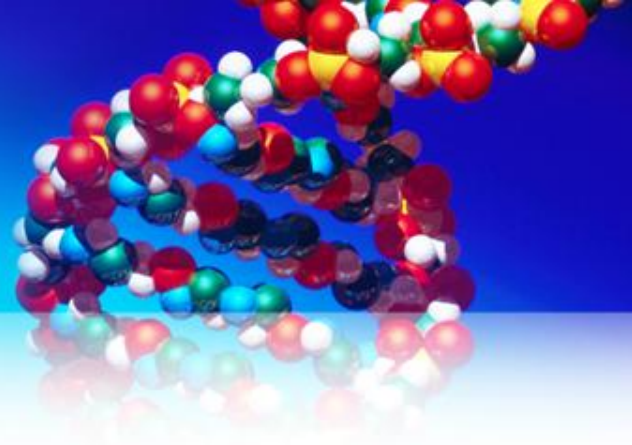
А – на платформе «SMRT»; Б – по технологии «Oxford Nanopore»



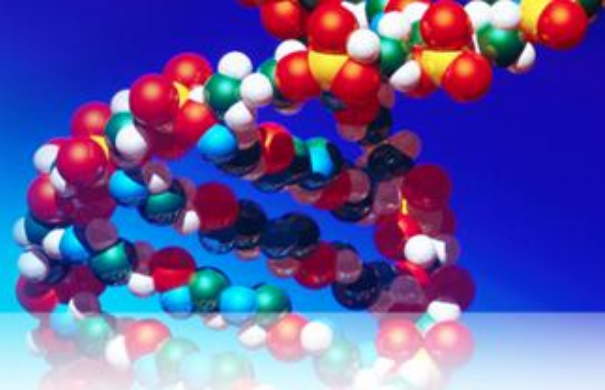
Однако данная технология имеет, по крайней мере, два серьезных технических препятствия для реализации: отсутствие надежного подхода к контролю продвижения ДНК через нанопоры и технические трудности в создании достаточно малых датчиков. В одном из вариантов для замедления прохождения фрагмента ДНК через нанопоры на его конец крепят магнитную микросферу. Чтобы тянуть молекулу за микросферу, используют магнит, включенное электрическое поле в то же время тянет ДНК в противоположную сторону, затягивая ее в отверстие нанопоры. Таким образом, скорость движения ДНК через нанопоры определяется балансом этих сил. При этом чтение последовательности нуклеотидов происходит со скоростью в сотни тысяч раз быстрее, по сравнению со стандартными методами секвенирования. Секвенирование генома человека займет в этом случае всего 20 ч, поскольку многократной амплификации матрицы не понадобится.



Другой предложенный фирмой «IBM» вариант технологии – «ДНК Транзистор», использует многослойную (диэлектрик-металл) мембрану. Изменение напряжения между адресными слоями металла в мембране создает электрические поля внутри нанопор, циклически включая и выключая которые можно двигать ДНК через нанопоры с шагом в один нуклеотид за цикл, улавливая датчиками разницу между возможными четырьмя вариантами оснований. Существуют также другие варианты реализации секвенирования ДНК через нанопоры, но в настоящее время научно-исследовательские разработки нанопорных секвенаторов находятся на опытно- конструкторской стадии.



Недавно компания «Oxford Nanopore Technologies» заявила о том, что практически доработала технологию секвенирования на основе нанопор до ее коммерческого аппаратного воплощения. Выход первых нанопорных секвенаторов данной компании под названием «GridION» и «MinION» может состояться уже в 2015 г. Уровень ошибок данной технологии составит 0,1–1 %, а длина читаемых единичных фрагментов ДНК достигнет многих десятков тысяч нуклеотидов . Это будет очередным прорывом в области секвенирования ДНК, так как время, затрачиваемое на данный процесс, и его стоимость уменьшатся на порядок.



На сегодняшний день из практически реализованных методов секвенирования третьего поколения функционирует только платформа «SMRT» (Pacific Biosciences). Вероятно, через несколько лет метод секвенирования ДНК через нанопоры станет доминирующим и вытеснит практически используемые методы второго поколения.

Таким образом, классические, новые и новейшие методы секвенирования направлены на выполнение разноплановых задач, связанных с определением нуклеотидной последовательности ДНК. Методы секвенирования третьего поколения могут в ближайшее время заменить своих предшественников, в том числе стать практической альтернативой и для классических методов секвенирования ДНК.



Спасибо за внимание!