



Новое секвенирование ДНК

В. Зубов

*Институт теоретической и
экспериментальной биофизики РАН*

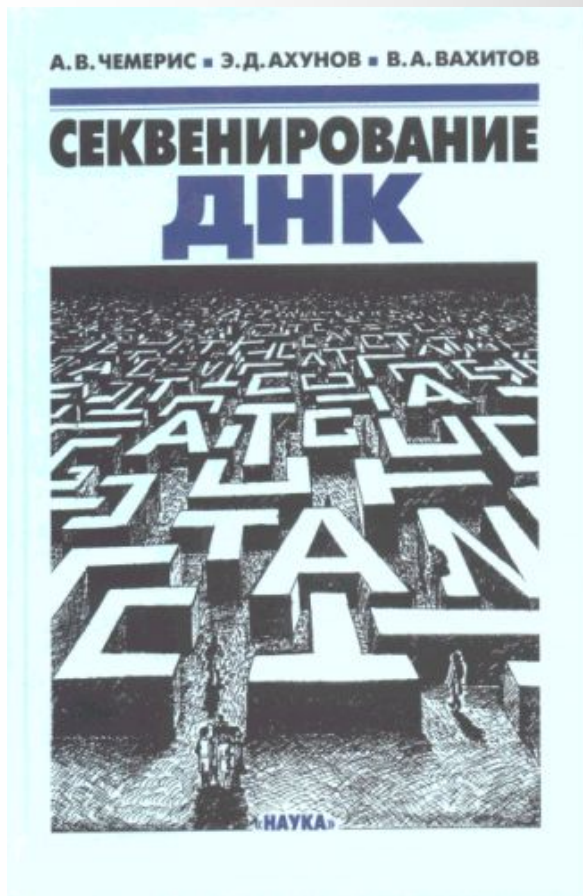
Пушино



Уфа, 10.09.2007



Секвенирование ДНК



Чемерис А.В., Ахунов Э.Д., Вахитов В.А.
М., Наука, 1999. - 429 с.

<http://www.anrb.ru/molgen/chemeris.html>

"... чтение последовательности ДНК
может стать задачей XXI века ..."
(Chargaff, 1968)





Новая старая ДНК

***А.В. Чемерис, С.М. Бикбулатова,
Д.А. Чемерис, В.А. Вахитов***

*К столетнему юбилею
Нобелевских премий
и к 50летию открытия
двойной спирали ДНК*

Издание второе, дополненное и расширенное

Уфа – 2005

<http://www.anrb.ru/molgen/newolddna2005.pdf>





Проект Джим

“Project Jim”



James Watson



**The Nobel Prize
in Physiology or Medicine 1962**

"for their discoveries concerning the molecular structure of nucleic acids and its significance for information transfer in living material"

Первое секвенирование индивидуального генома человека стоимостью менее 1 млн.\$





Стоимость секвенирования генома человека

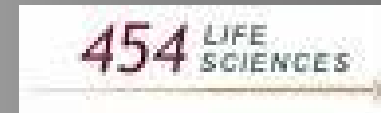
- 1990...2005 г. – 3 млрд.\$
- 2002 г. – 200 млн.\$
- 2006 г. – 20 млн.\$
- 2007 г. – 1 млн.\$
- 2008 – 100 000\$
- 2012 – 1 000\$





Технологии геномного секвенирования

1. Pyrosequencing



2. Cluster-SBS



3. SOLiD





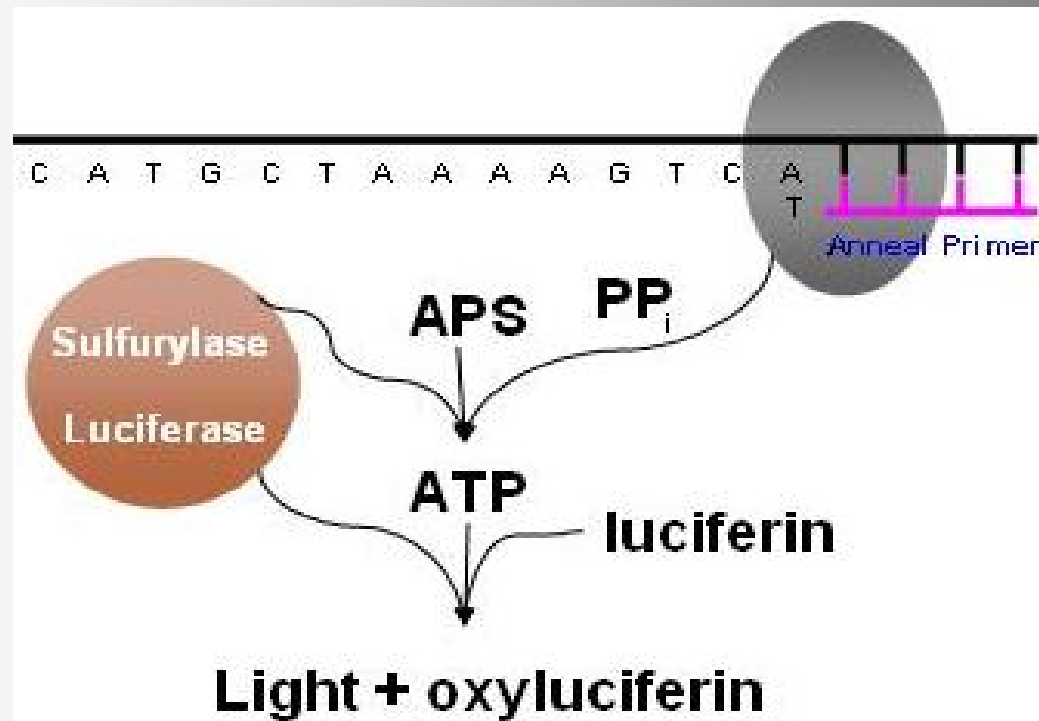
Пиросеквенирование

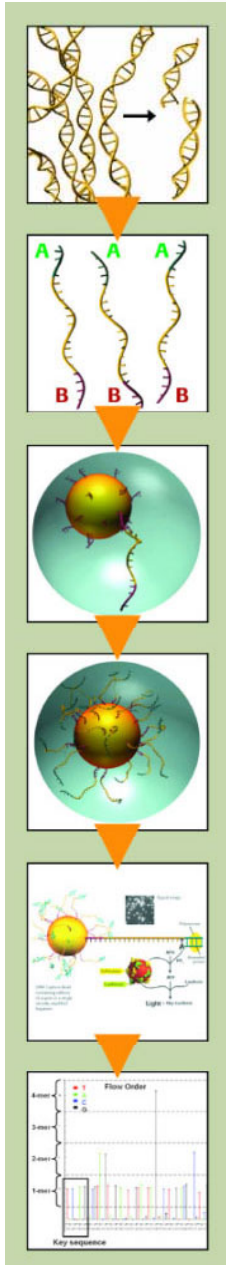
- **1985** год - шведские учёные Pål Nyren и Arne Lundin разработали чувствительный люминометрический метод определения пирофосфата;
- **1987** год – P. Nyren разработал метод анализа ДНК-полимеразной активности, основанный на определении пирофосфата;
- **1988** год – Edward Human предложил использовать этот метод для секвенирования ДНК;
- **1997** год – создана компания «Pyrosequencing» (сейчас – “Biotage AB”);
- **2000** год – в США учреждена компания «454 Life Sciences»;
- **2005** год – появление в продаже пиросеквенатора GS20;
- **2007** год – завершено секвенирование первого индивидуального генома (“Project Jim”).





Схема реакций пиросеквенирования





1. Подготовка ДНК

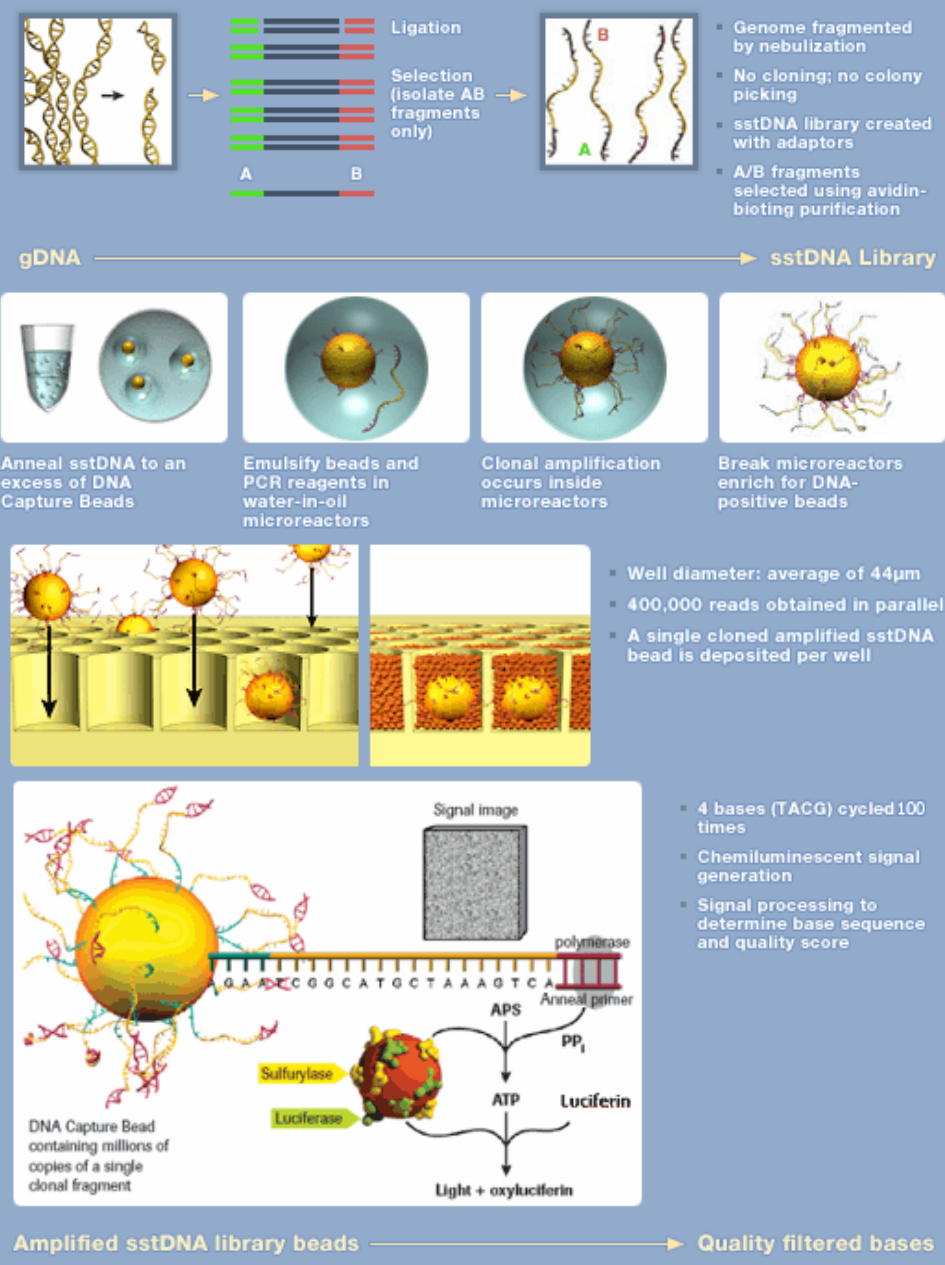
2. Пришивка адапторов

3. Получение эмульсии

4. emPCR

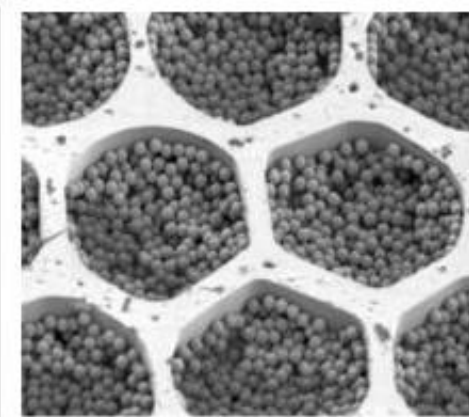
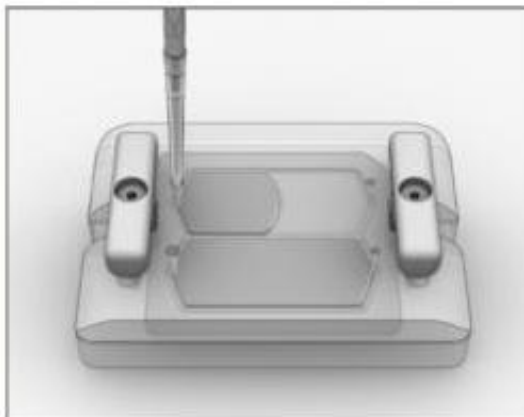
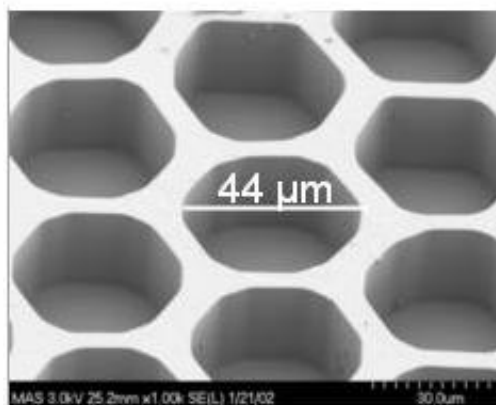
5. Секвенирование

6. Анализ данных



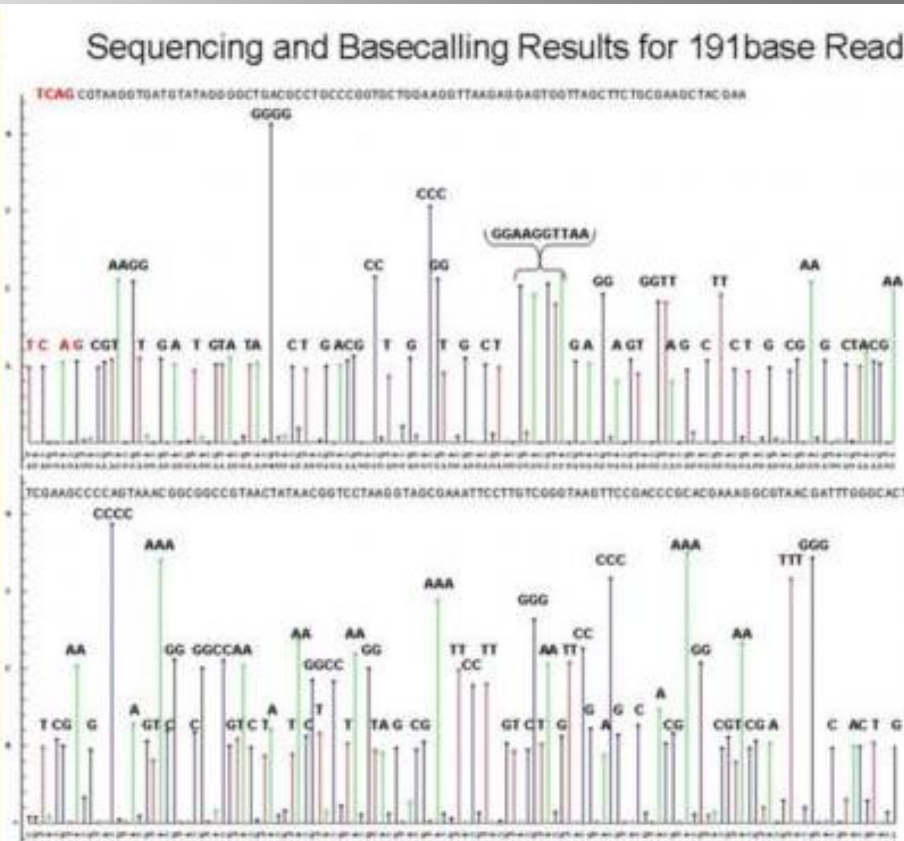
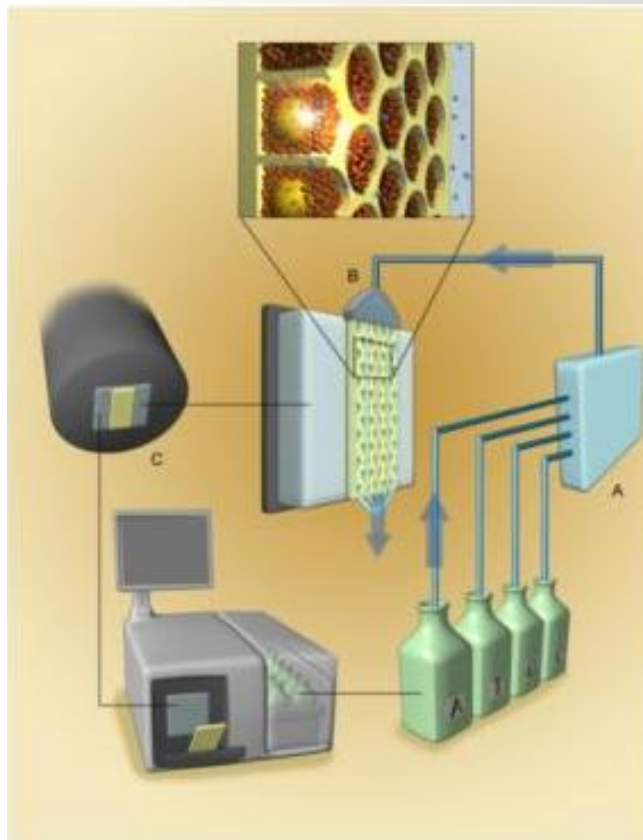


Проточная камера PicoTiterPlate™



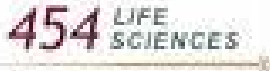





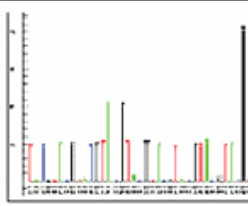

Чтение пирограммы

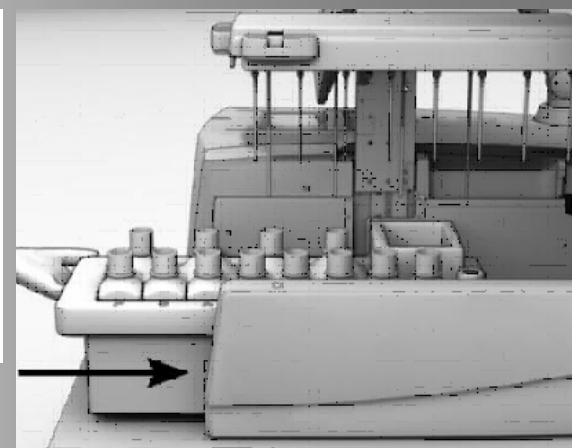




Технические характеристики

 	GS 20	GS FLX
Длина чтения	100 п.о.	200...300 п.о.
Количество последовательностей	100...200 тыс.	400 тыс.
Продолжительность	4 часа	7,5 часов
Суммарная длина	10...20 млн. п.о.	100 млн. п.о.

Genome Sequencer FLX System			
Instrument	Reagent Kits	Software	Accessories
			





Проблемы

- Не различаются гомогенные повторы $\geq 6N$
- Необходимы дНТФ высшей очистки
- Вместо дАТФ используется дАТФ α S
- Спонтанный гидролиз дНТФ и APS
- Термолабильность люциферазы
- Утечка PPi и АТФ из лунок проточной ячейки
- Низкая эффективность регистрации фотонов
- Большие размеры сефарозных микросфер
- Высокий расход дорогих реагентов
- Дороговизна приборов





Решения

1. Применение ячеек с закрывающимися лунками, контактное или аэрозольное введение реагентов
2. Разработка более чувствительных способов регистрации РРi
3. Замена биolumинесценции флуориметрией или денситометрией





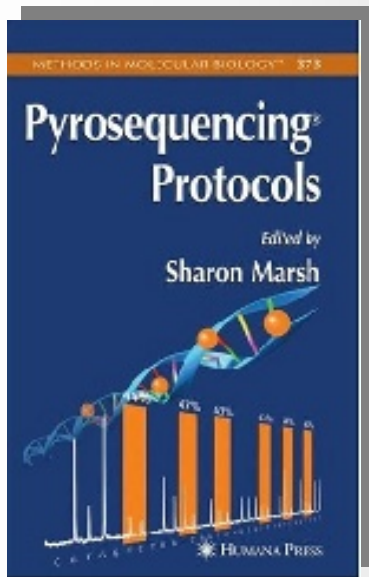
Альтернативные реакции

- γ P-модифицированные субстраты ДНК-полимеразы вместо dNTP (AprrppdN, B₁pppdN и т.п.)
- Киназы вместо люциферазы (NAD + ATP → NADP + ADP и т.п.)
- Амплификация ATP (ATP + AMP → 2ADP; ADP + AcP → ATP)





Pyrosequencing[®] Protocols



Под ред.: Marsh Sharon

Издательство: Humana Press

Дата издания: 2007

Серия: Methods in Molecular Biology

Страницы: 230

Цена: 2989.00 р.

454 sequencing[®]

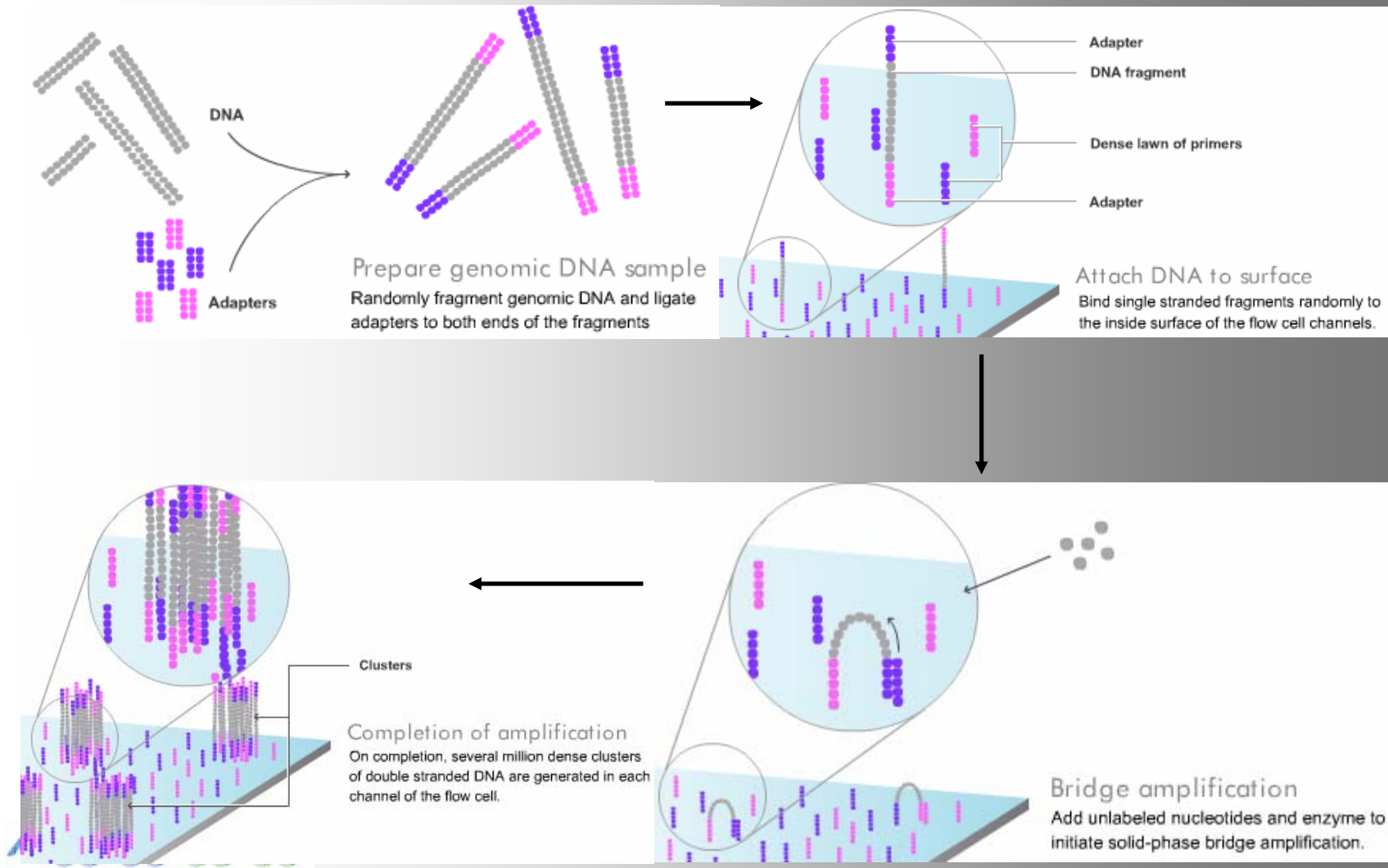
www.454.com





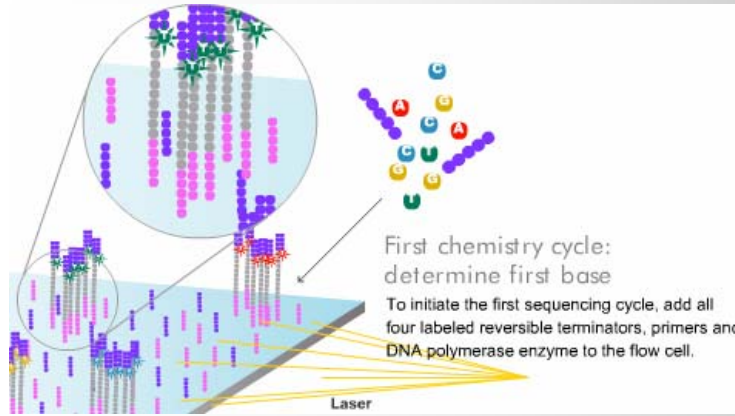
Cluster-SBS Technology

(получение кластеров ДНК)





Cluster-SBS Technology (секвенирование)



Wash off all unincorporated reagents

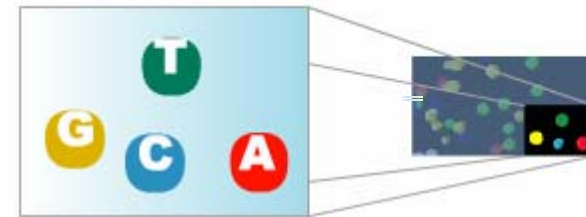
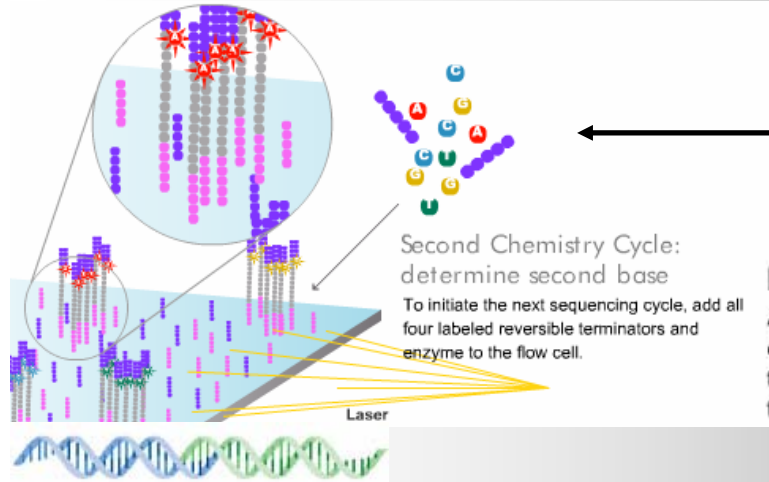


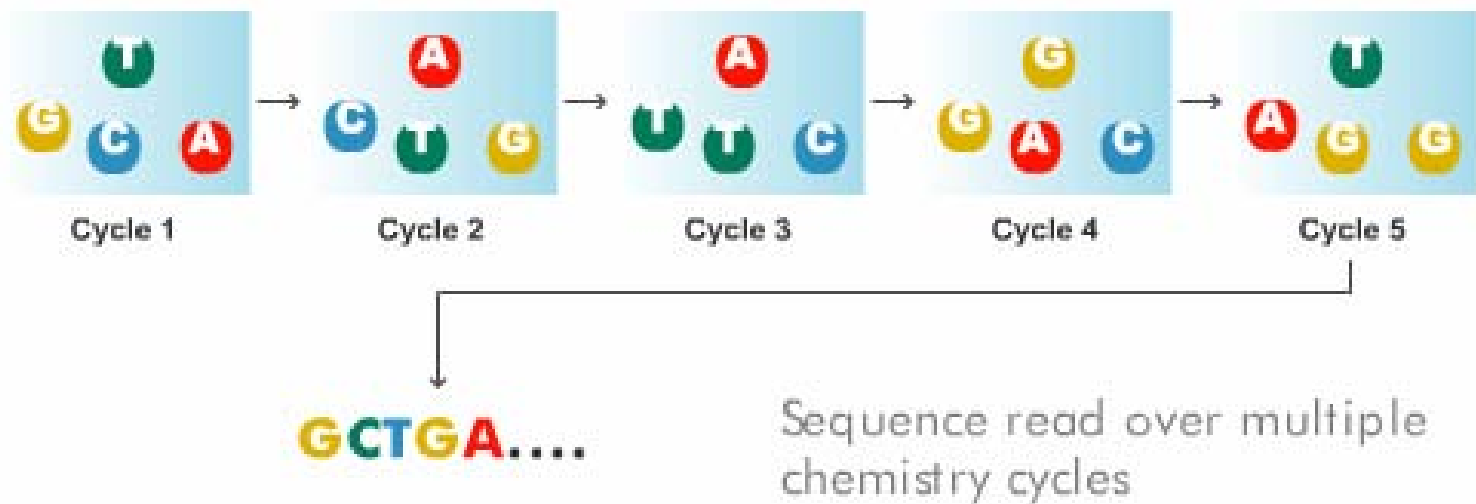
Image of first chemistry cycle
After laser excitation, capture the image of emitted fluorescence from each cluster on the flow cell. Record the identity of the first base for each cluster.

Before initiating the next chemistry cycle
The blocked 3' terminus and the fluorophore from each incorporated base are removed.



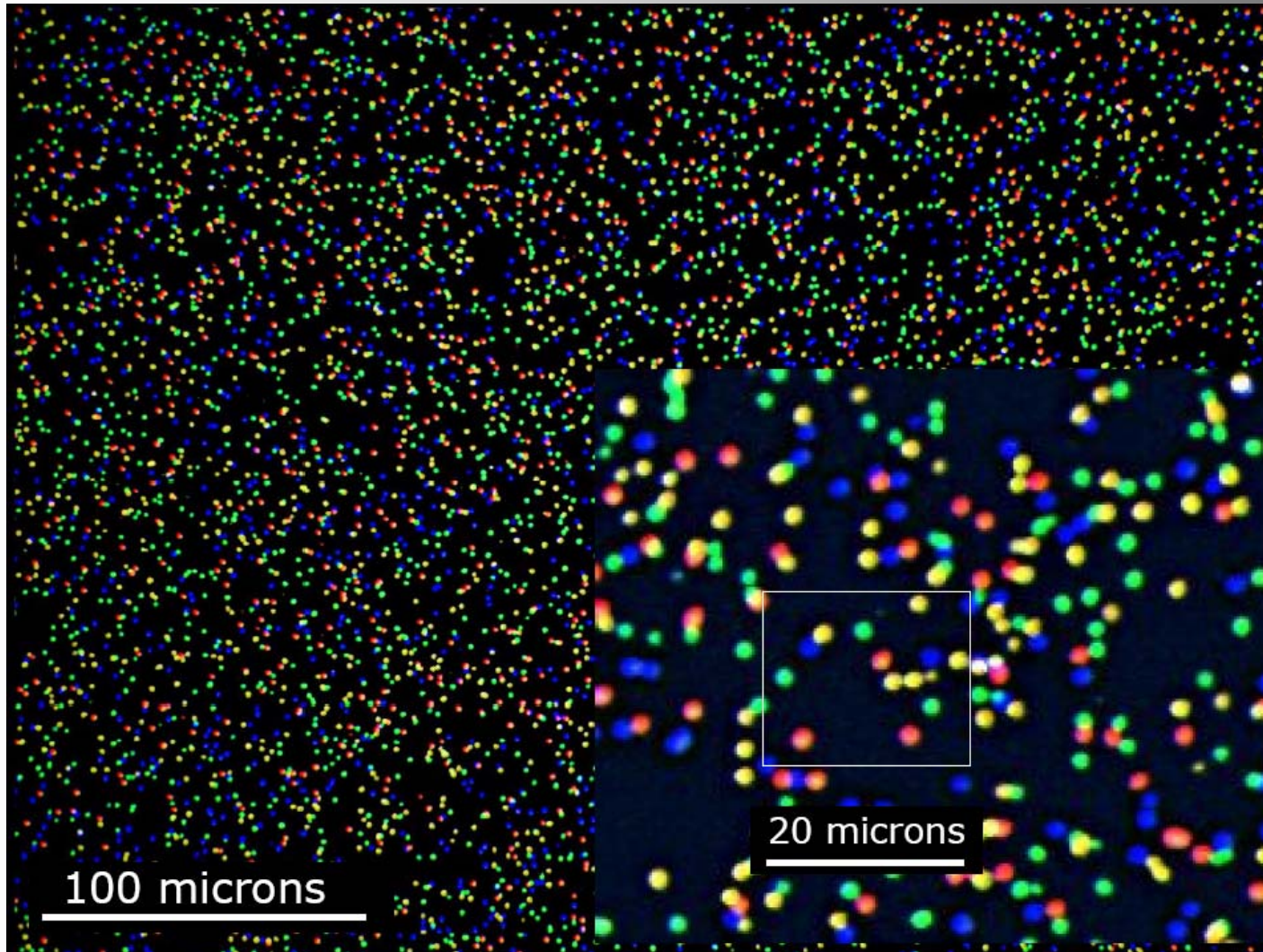
Cluster-SBS Technology

(считывание последовательностей)



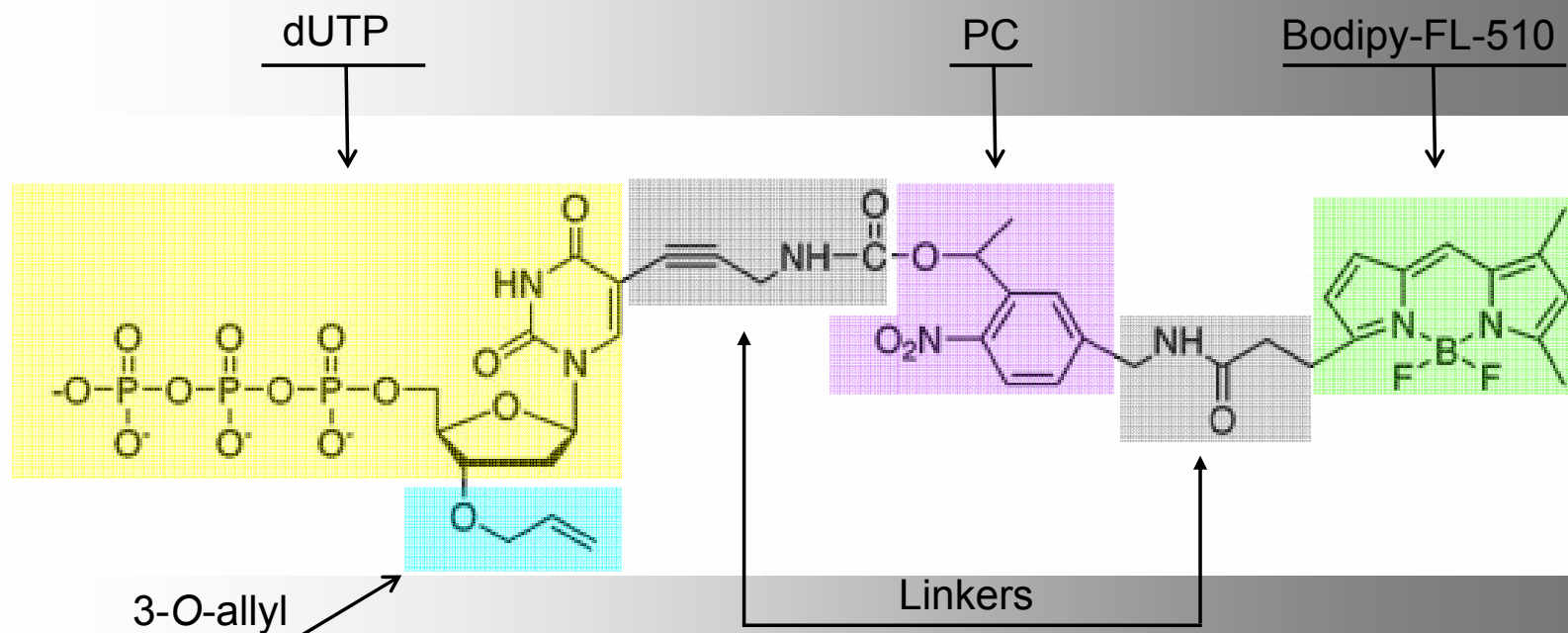


Кластеры ДНК





3-O-allyl-dUTP-PC-Bodipy-FL-510

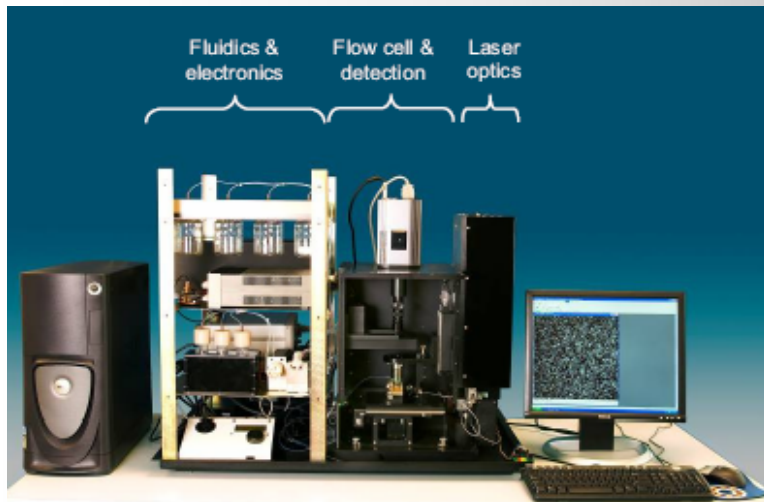


Design and synthesis of a 3-O-allyl photocleavable fluorescent nucleotide as a reversible terminator for DNA sequencing by synthesis / H. Ruparel et. al. // PNAS. – 2005. - Vol.102, №17. – P.5932–5937.

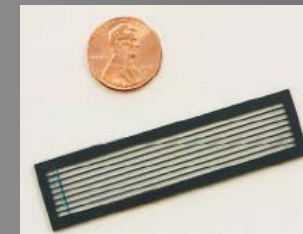




1G Genome Analyser



- Производительность - ~1 млрд.п.о. за рабочий цикл
- Диаметр кластеров ДНК – 1...2 мкм
- Расстояние между кластерами - ~4 мкм
- Проточная ячейка на 8 образцов по 5 млн. кластеров ДНК
- Всего считывается до 40 млн. кластеров ДНК
- Длина секвенирования – 24 п.о. (до 32 п.о.)





Проблемы

- Сложность синтеза ДНК-полимеразных субстратов.
- Необходимость использования специальной ДНК-полимеразы (9°Nт и т.п.).
- Дороговизна реагентов.
- Непрочность связи праймеров со стеклом.
- Неупорядоченность расположения кластеров ДНК.
- Небольшая длина получаемых сиквенсов.
- Дороговизна секвенатора.

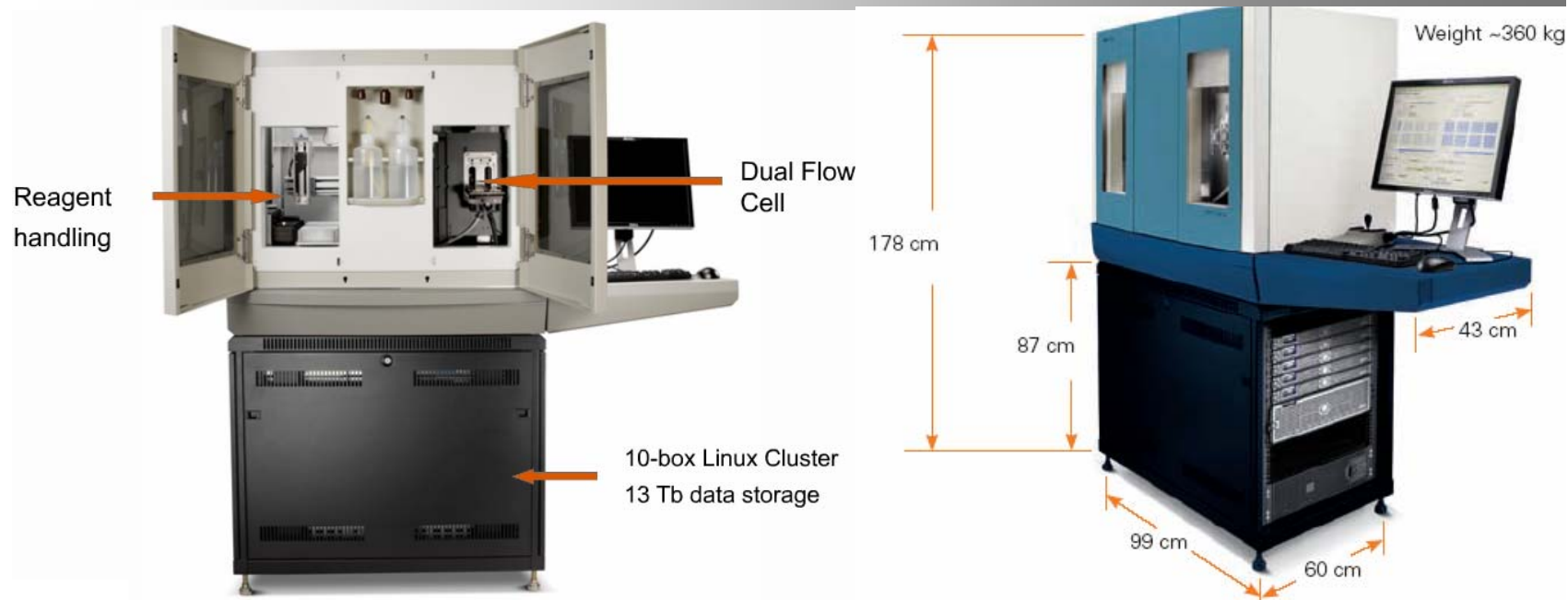
www.illumina.com





SOLiD™ System

(Sequencing by **O**ligonucleotide **L**igation and **D**etection)

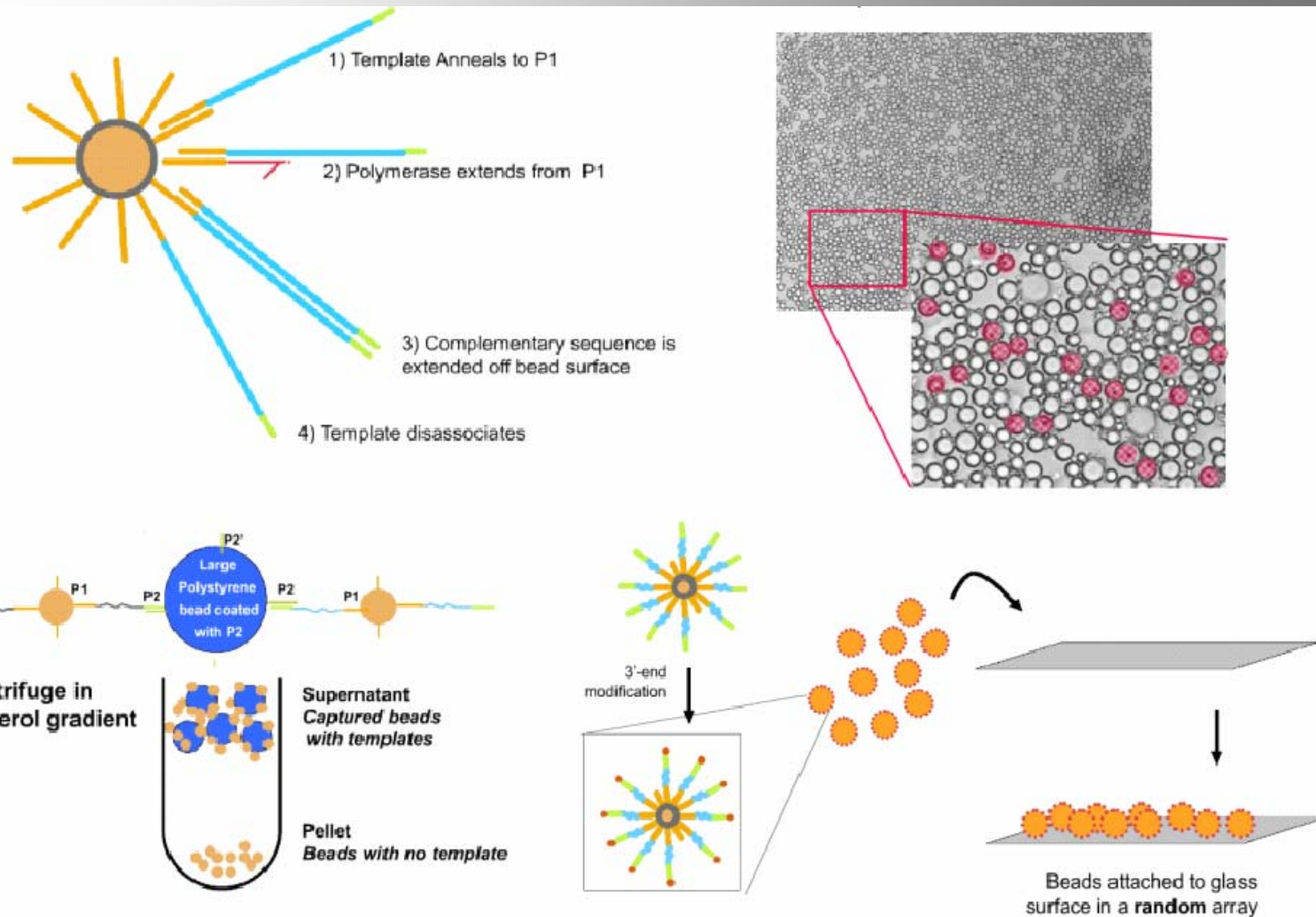


- Производительность - до 500 млн.п.о. в день (до **3 млрд.п.о.** за двойной цикл).
- Продолжительность цикла – 3 дня (**6 дней** при двойном секвенировании).
- Потребляемая мощность – до **5 Квт.**
- Длина чтения – до **35 п.о.**



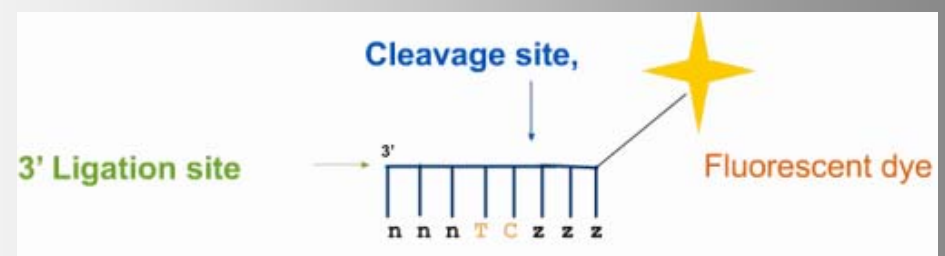
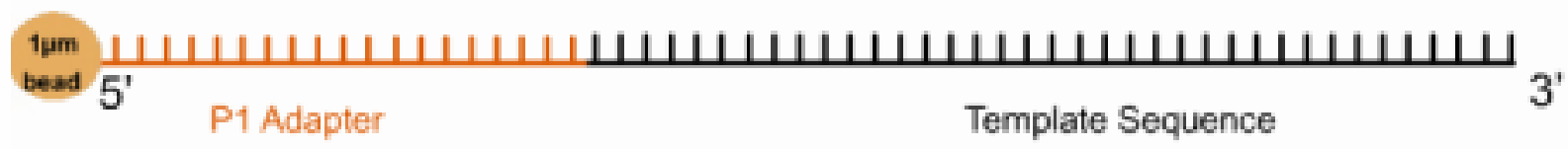
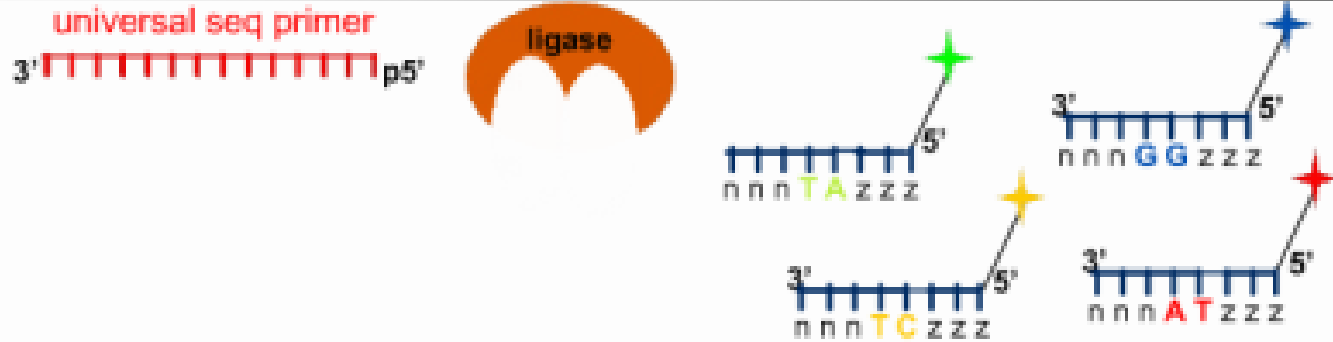


Особенности технологии 1





Особенности технологии 2

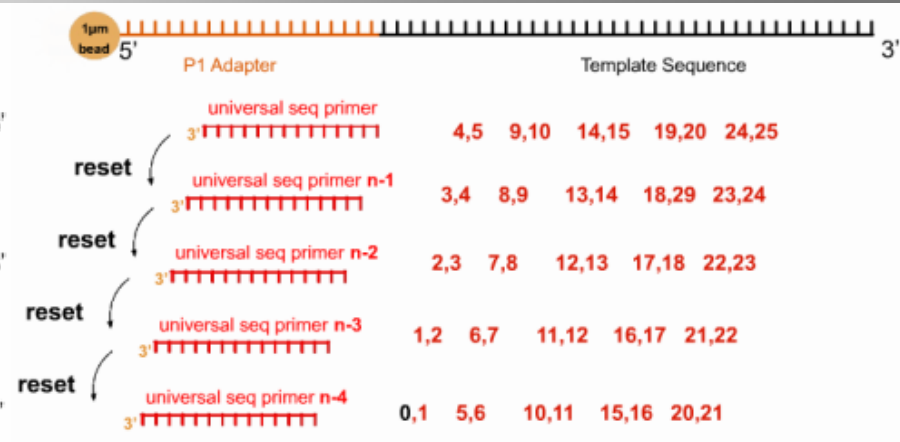
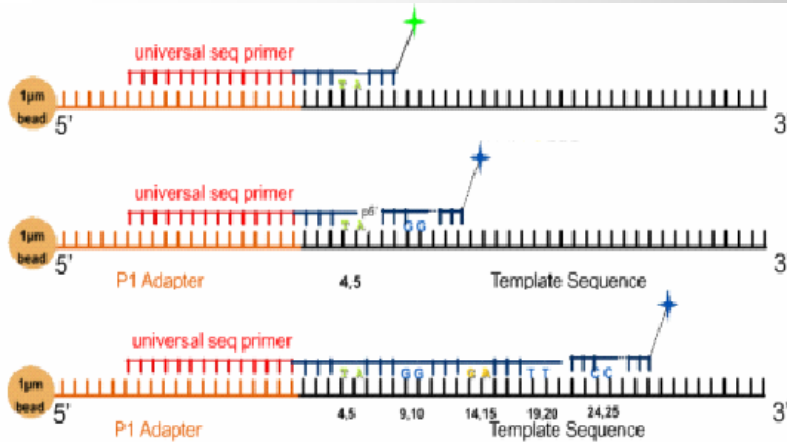


1,024 Octamer Probes (4⁵)
 4 Dyes, 4 dinucleotides, 256 probes per dye
 N= degenerate bases Z= Universal bases





Особенности технологии 3



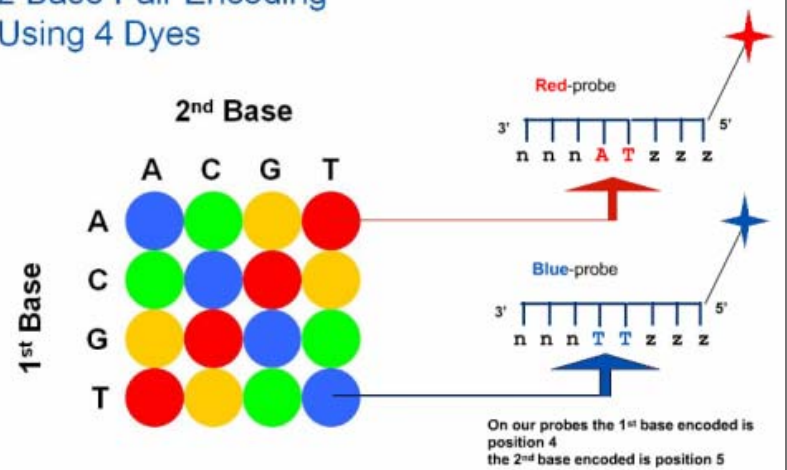
Advantages of 2 base pair encoding

A C G G T C G T C G T G T G C G T

A C G G T C G T C G T G T G C G T



2 Base Pair Encoding Using 4 Dyes

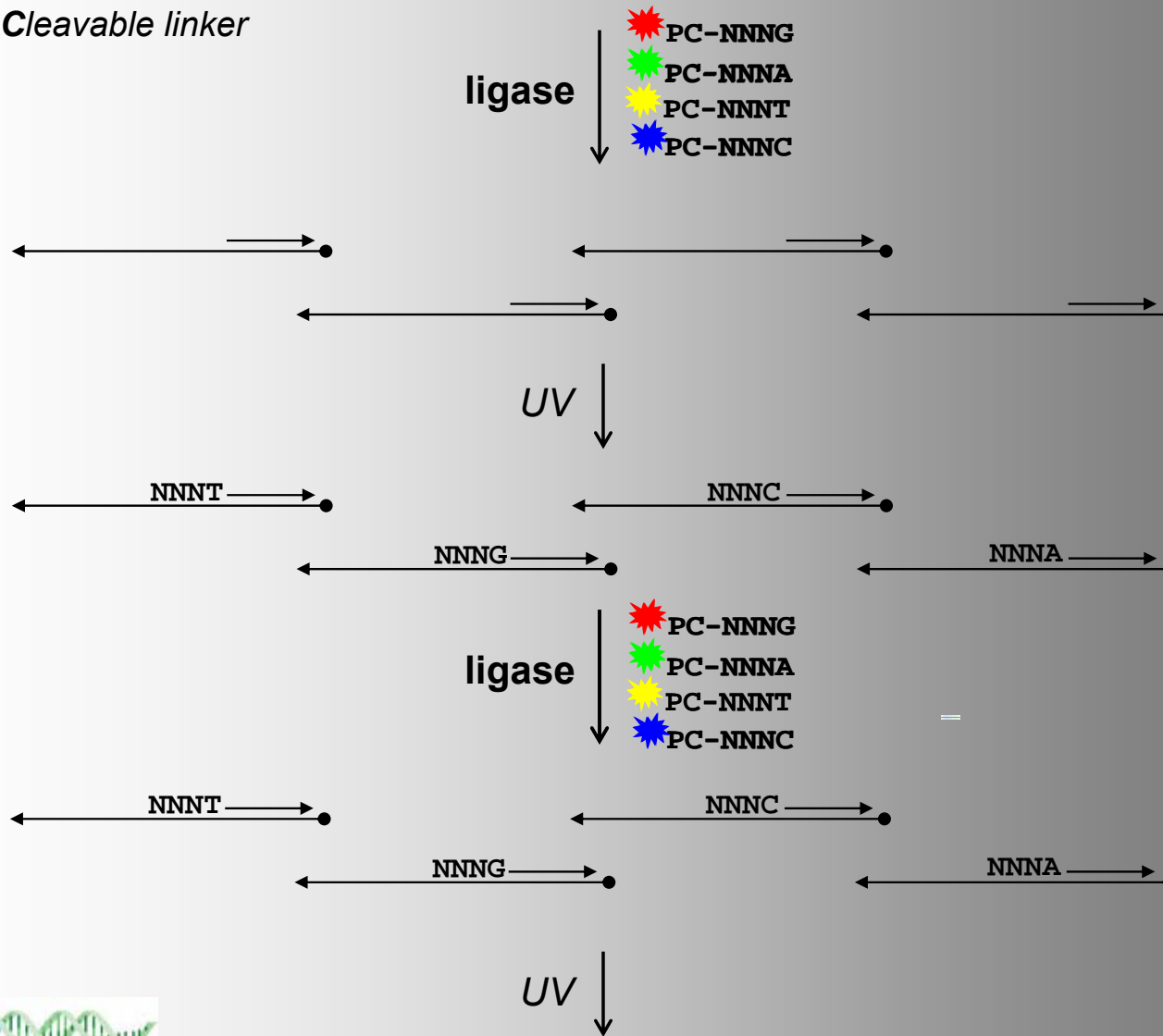




Фотוליгазное секвенирование



PC – *PhotoCleaveable linker*





Проблемы

- Высокий уровень ошибок (~1/60)
- Необходимость многократного чтения (~15 раз) для получения достоверных (99,99%) результатов
- Медлительность (3...6 суток на рабочий цикл)
- Избыточный вес (360 кг)
- Чрезмерная мощность (5 Квт)
- Недолговечность газовых лазеров
- Для установки требуется вентилируемое помещение
- Избыточная производительность компьютера
- Чрезмерная ёмкость накопителя информации (13 Tb)
- Высокая стоимость



solid.appliedbiosystems.com

+ DVD



True Single Molecule Sequencing (tSMS)



HeliScope™

- Two glass flow cells
- Each flow cell is divided into 25 channels
- 50 individual samples per run
- Proprietary surface chemistry
- ~25 bases in length
- Multi-pass sequencing

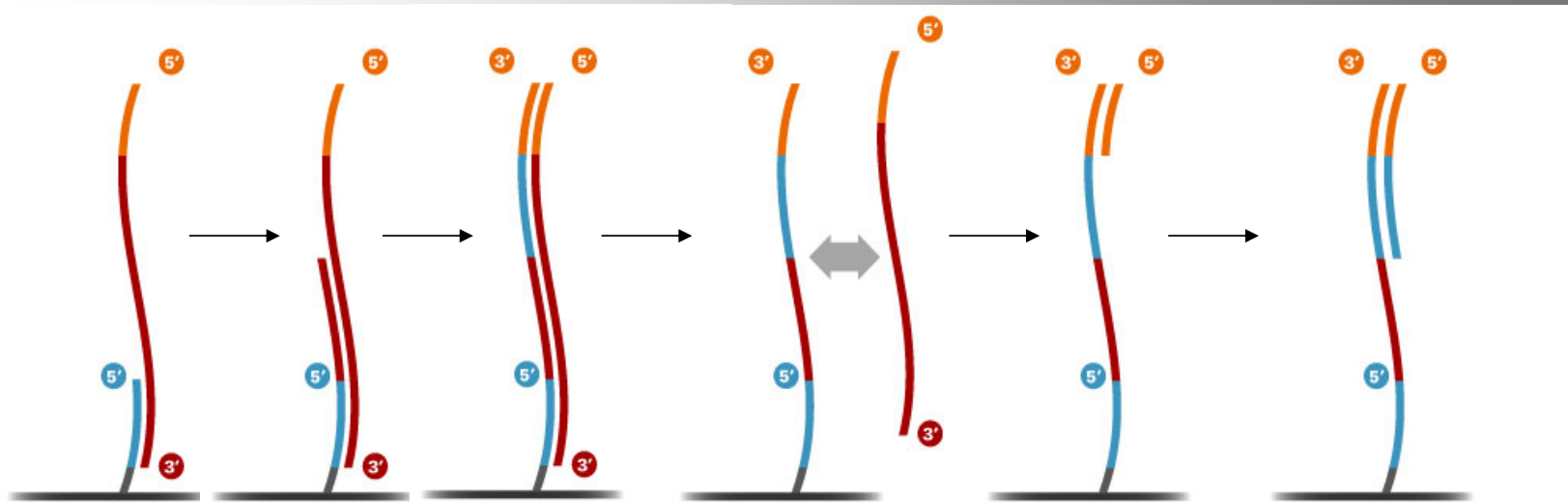


One human genome at 20x coverage in 2.5 days



Prep-Free Paired-End Reads

- The genomic sample is digested/sheared.
- Fragments within a desired size range are selected.
- An adaptor sequence, containing a universal priming site, is ligated to the 5' ends of the fragments.
- Poly(A) tails are generated on the 3' ends of the fragments.



Once the first 25-base read is obtained, the template is copied to the end, in a single step, using a polymerase and all four natural nucleotides.

Another round of ~25 tSMS quads is carried out to sequence the 5' end of the original template, i.e. the 3' end of the reverse complement





Геномные секвенаторы

Cluster-SBS



1G



HeliScope™


tSMS



454 sequencing®



GS FLX



SOLiD™

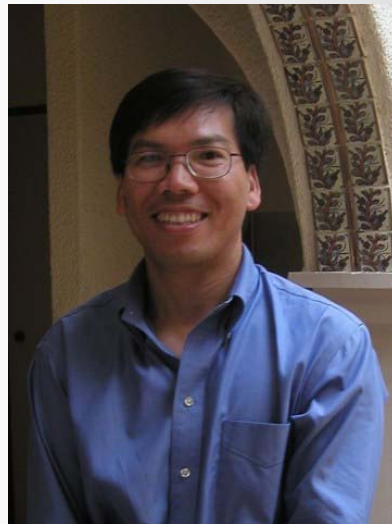




Sequencing by denaturation (SBD)



- Parallel clonal amplification of individual DNA molecules
- Isothermal rolling circle amplification (RCA)
- Sequencing by denaturation technology
- **Project Period:** 08/01/2005 - 05/31/2008



PI: Xiaohua Huang
Ph.D, Assistant Professor
Department of Bioengineering
Jacobs School of Engineering
University of California, San Diego

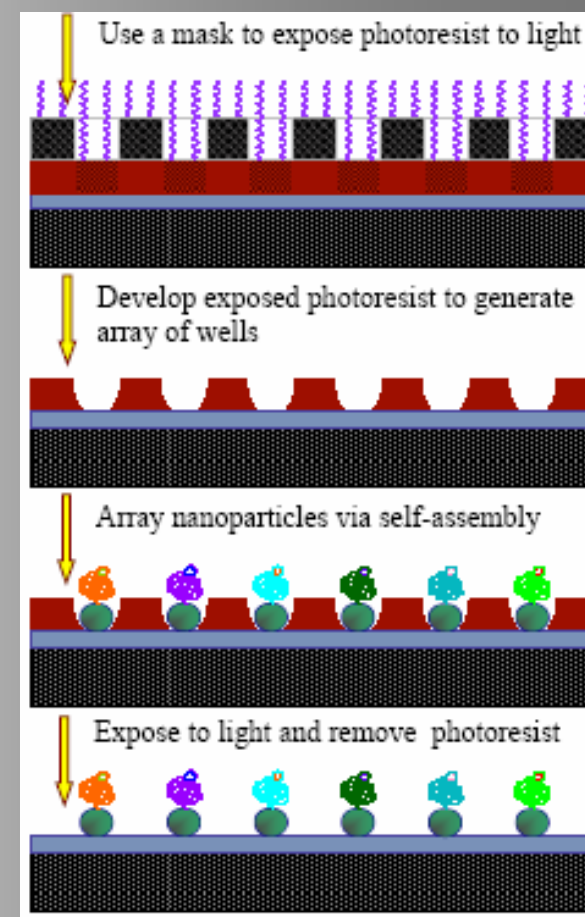
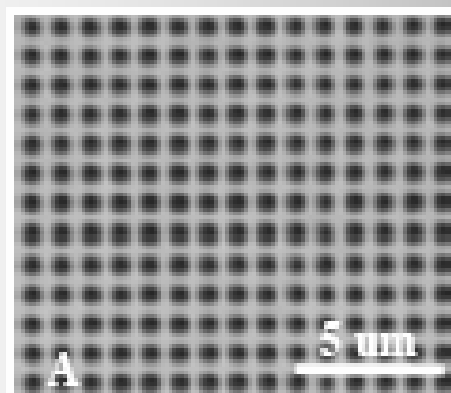
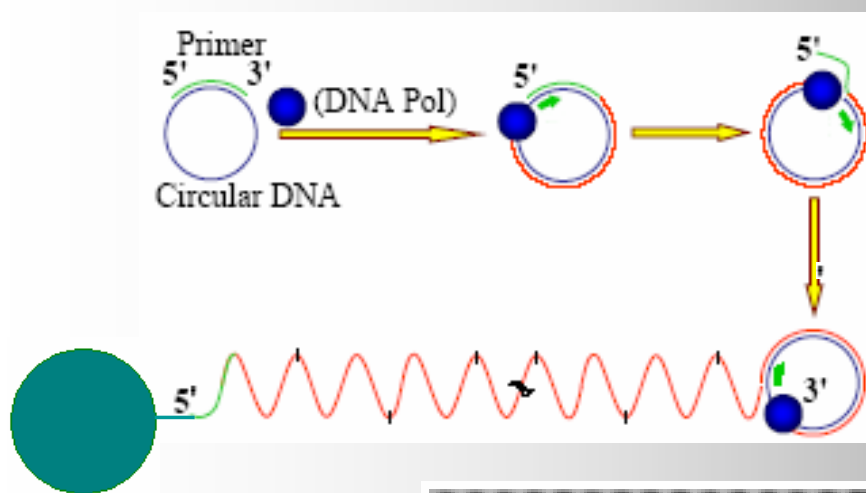




Fabrication Biomolecular Nano-Array

Kristopher Barbee and Xiaohua Huang

20th ANNUAL BIOENGINEERING GRADUATE RESEARCH SYMPOSIUM March 11, 2006



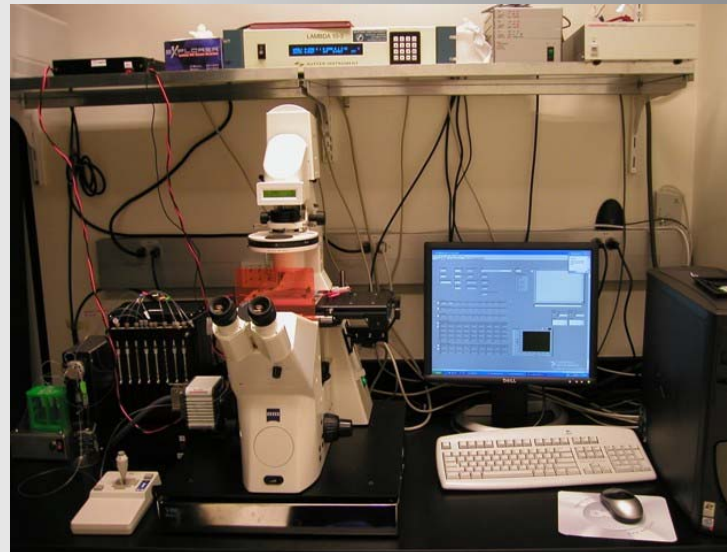


Single molecule sequencing by ligation (SM-SBL)



Pavel Pevzner
Professor of
Computer Science

- Cyclic sequencing by ligation
- Reading 3 bases per sequencing step
- Using photocleavable oligonucleotide probes
- Sequencing from paired-ends of single DNA molecules
- >1 billion individual DNA molecules can be sequenced
- 64 oligonucleotide probes with fluorescence nanoparticles
- **Project Period: 09/29/2006 - 08/31/2007**





Cyclic Ligation and Cleavage (CLiC)

21 мая 2007 г.

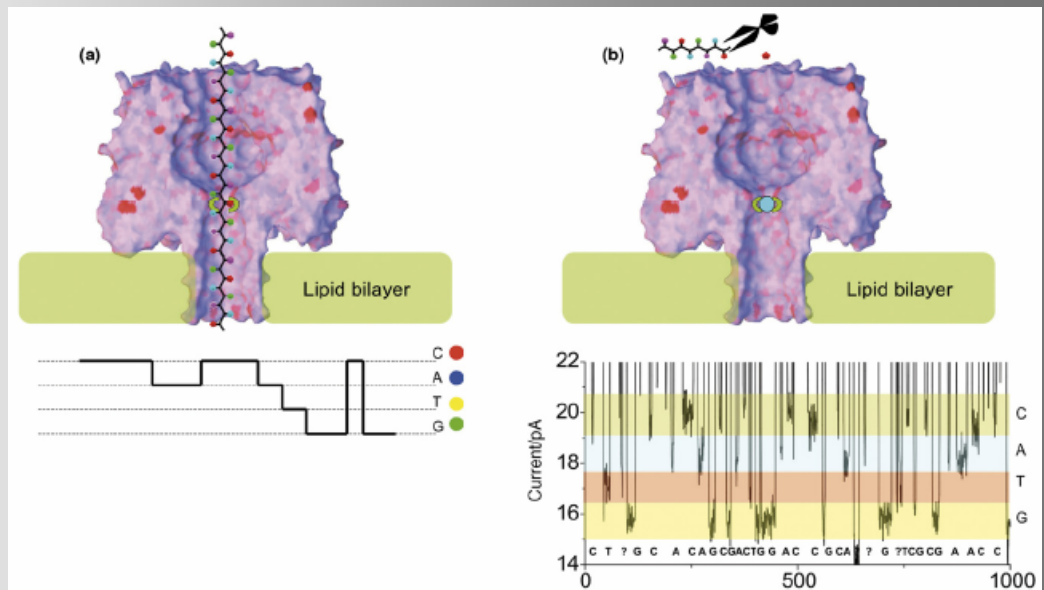
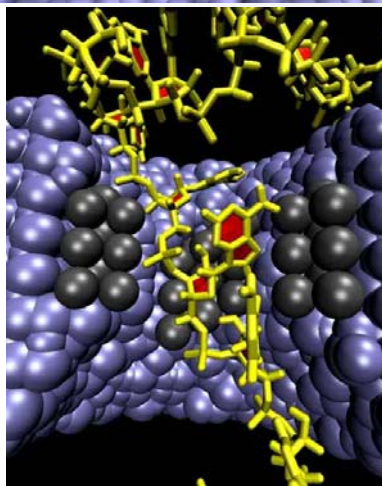
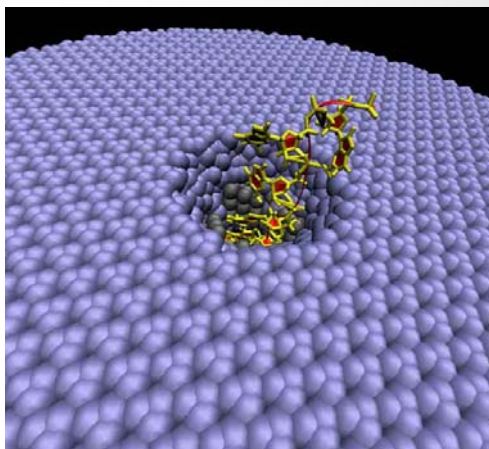
Сан-Диего

*Kalim Mir, Ph.D.,
The Wellcome Trust Centre for Human Genetics,
University of Oxford*

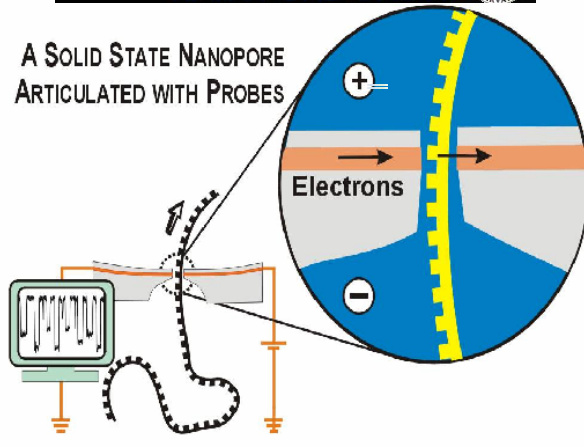




Nanopore Sequencing

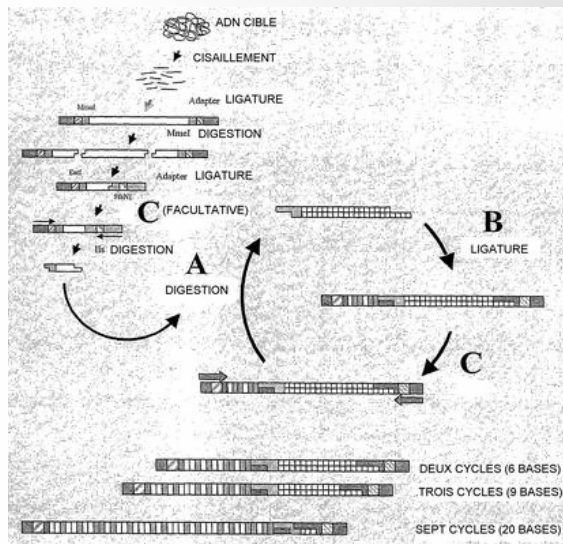


THE HARVARD NANOPORE GROUP





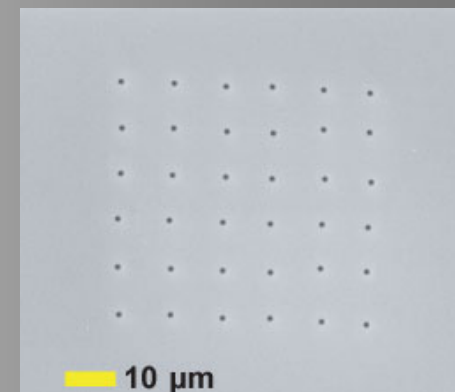
Design Polymer Technology



Advanced "DNA molecules" designed solely for the purpose of single molecule sequencing.

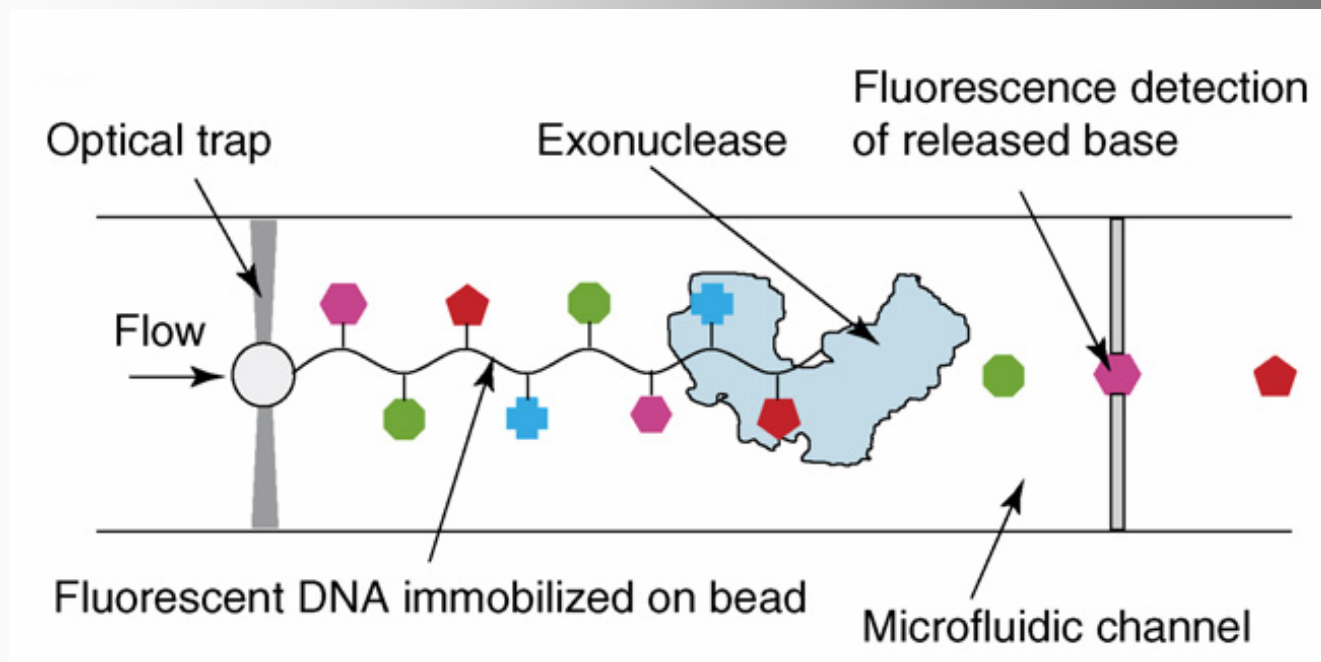


Meller Group
Single Molecule Biophysics
& Nano-biotechnology





Экзонуклеазное секвенирование





Прочие технологии



Ultra-Fast Real Time sequencer



Single Molecule Sequencing



U.S. GENOMICS
pioneers in single molecule biology

GeneEngine™ technology



VisiGen® sequencing system



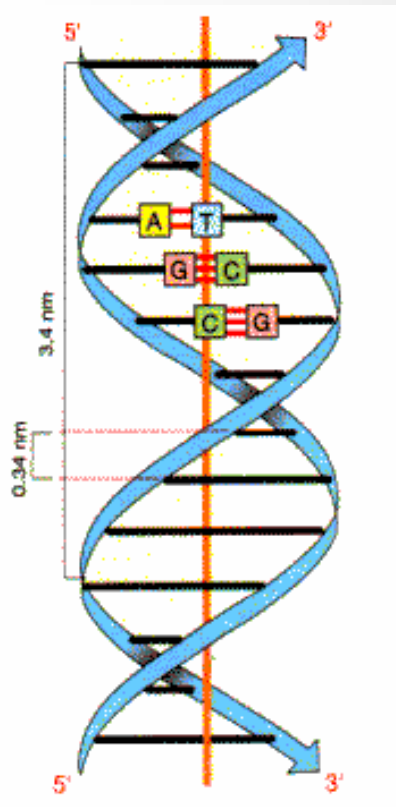
GE Global Research

-
-
-





Недостатки мономолекулярных методов



- Высокая частота ошибок
- Неработоспособность
- Чрезмерная прецизионность приборов
- Сложность параллельного сканирования
- Дороговизна секвенирования

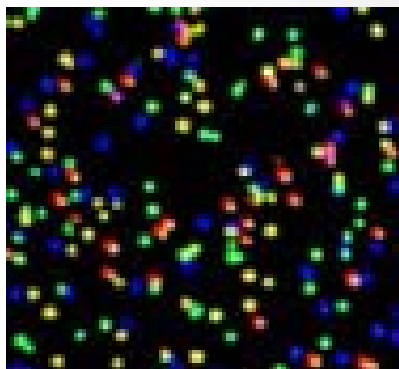
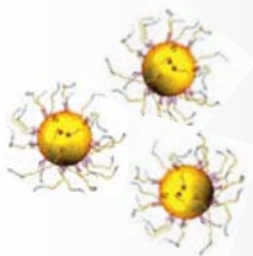
Исключения:

1. tSMS (многократное чтение последовательностей)
2. SBD (олигомерная матрица)





Преимущества мультимолекулярных технологий



- Работоспособность не требует доказательств.
- Острота конкурентной борьбы.
- Параллельность подготовки ДНК и сканирования.
- Низкий расход реагентов.
- Возможность значительного удешевления.
- Сравнительная простота оборудования
- Доступность.



Основные этапы развития ВТ и ГС (возможные аналогии)

БЭСМ-6

первая супер-ЭВМ второго поколения



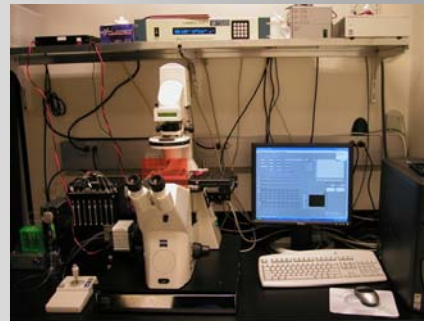
Персональный компьютер

модели IBM PC



Classmate PC

персональный компьютер ученика





Выводы

1. Началом эры геномного секвенирования можно считать 24 мая 2007 года.
2. Существенная особенность технологий геномного секвенирования - параллельность обработки препаратов ДНК.
3. Миниатюризация - ведущий фактор снижения стоимости секвенирования.
4. Чрезмерная миниатюризация контрпродуктивна.
5. Конкурентная борьба на рынке секвенирования ДНК постоянно обостряется.
6. Стоимость геномного секвенирования быстро снижается.
7. Секвенирование ДНК в США стало задачей XXI века.
8. Америку нельзя догнать, но можно перегнать.





Заключение

- 2012 г. - 1 000 руб.

