



Геномные проекты

Васильев Геннадий Владимирович

Институт цитологии и генетики СО РАН



Третье поколение – одномолекулярные секвенаторы



Имя им - легион





Работающие одномолекулярные секвенаторы

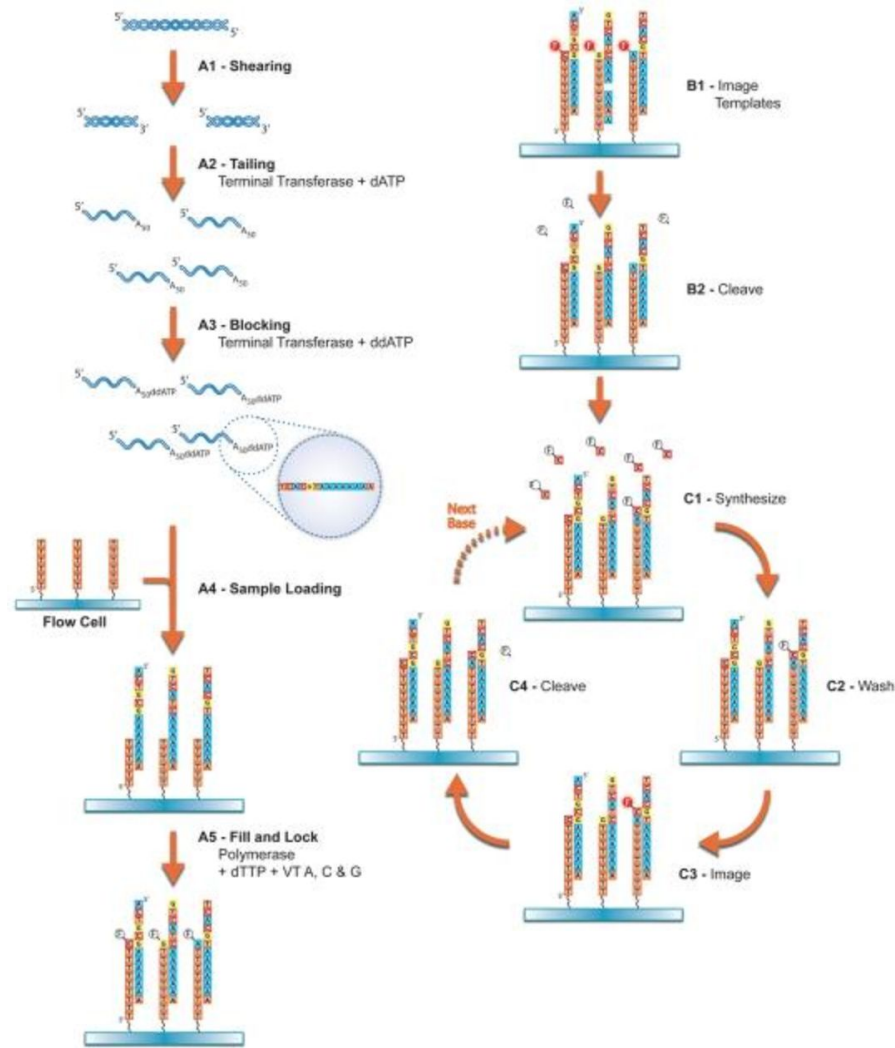


Helicos

Короткие чтения

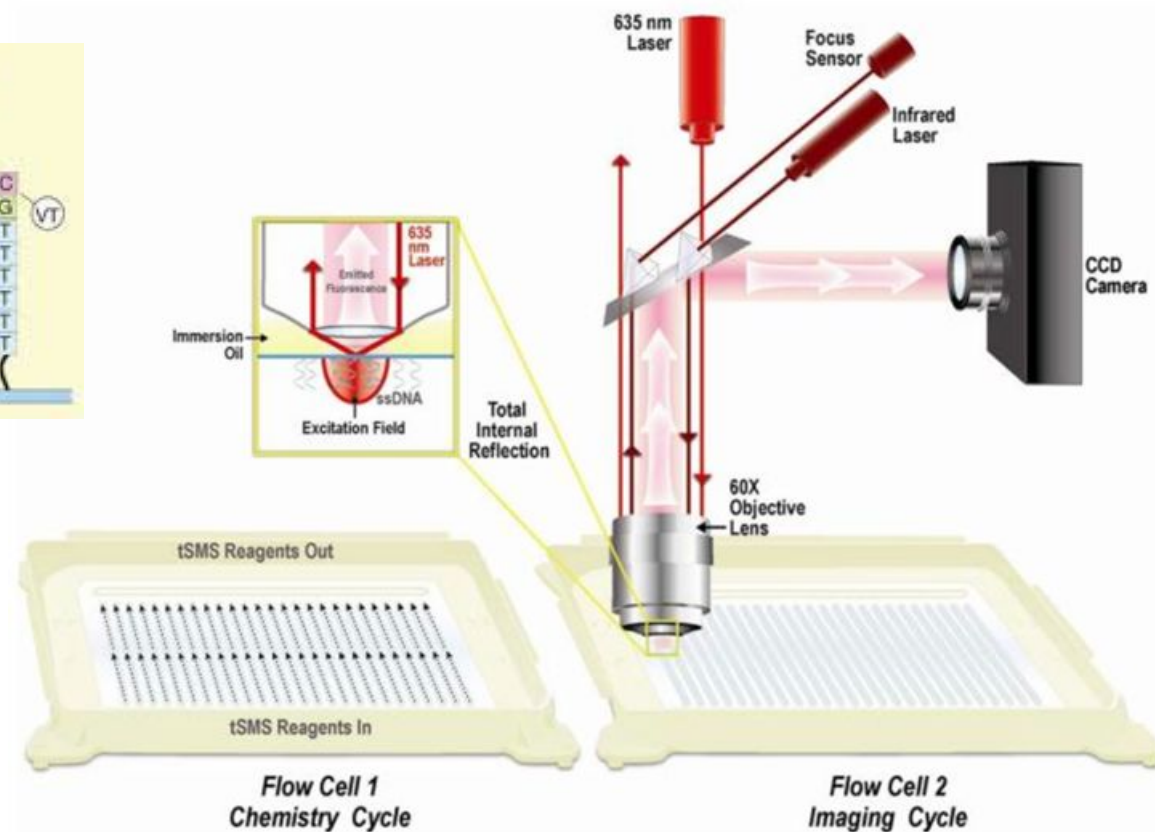
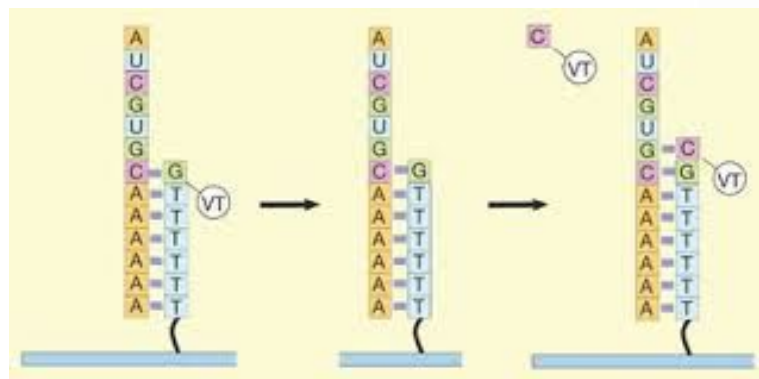
Прямое секвенирование РНК,

Секвенирование без ПЦР





Helicos



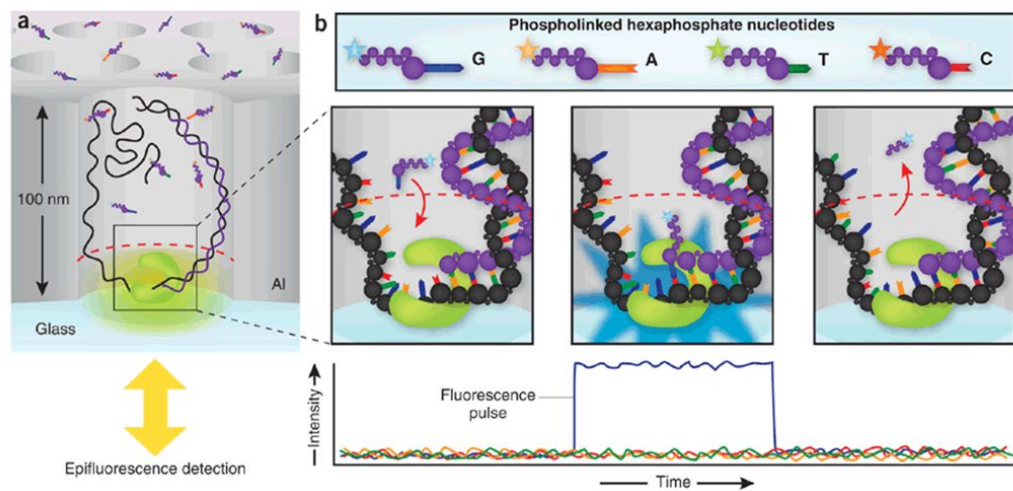


Pacific biosciences

Очень длинные чтения

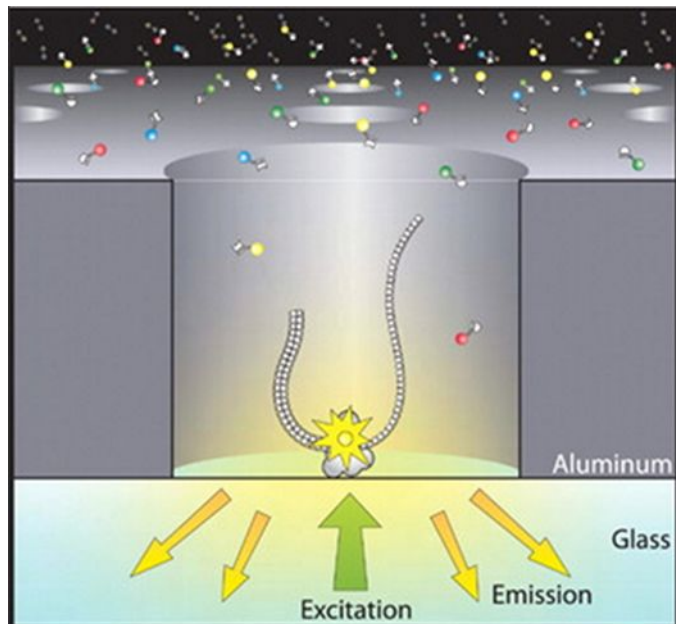
Секвенирование без ПЦР

Высокое число ошибок – до 10%





Pacific biosciences – повышение точности

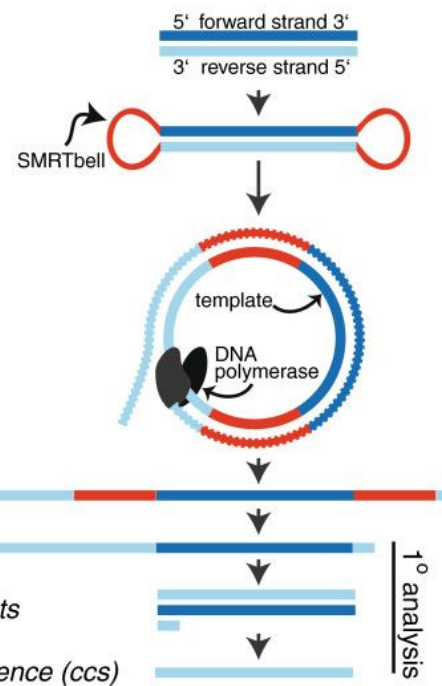


1. generate amplicon

2. ligate adaptors

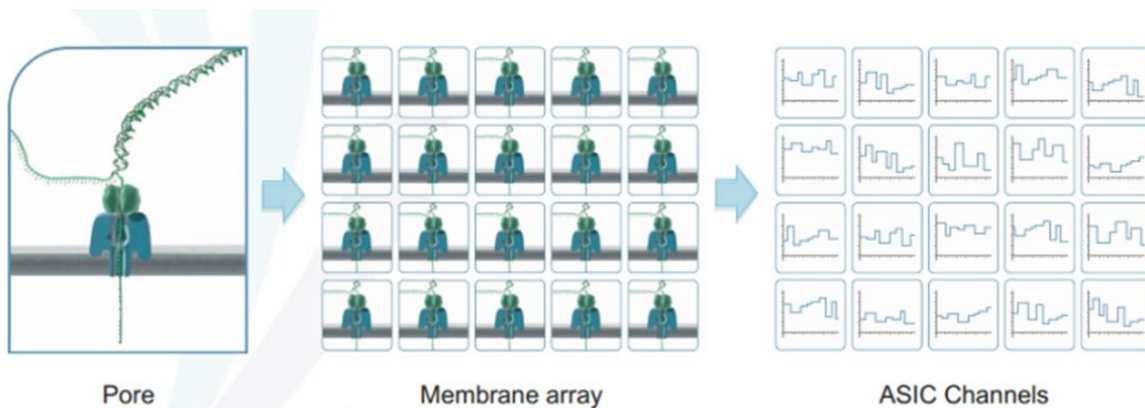
3. sequence

4. data analysis





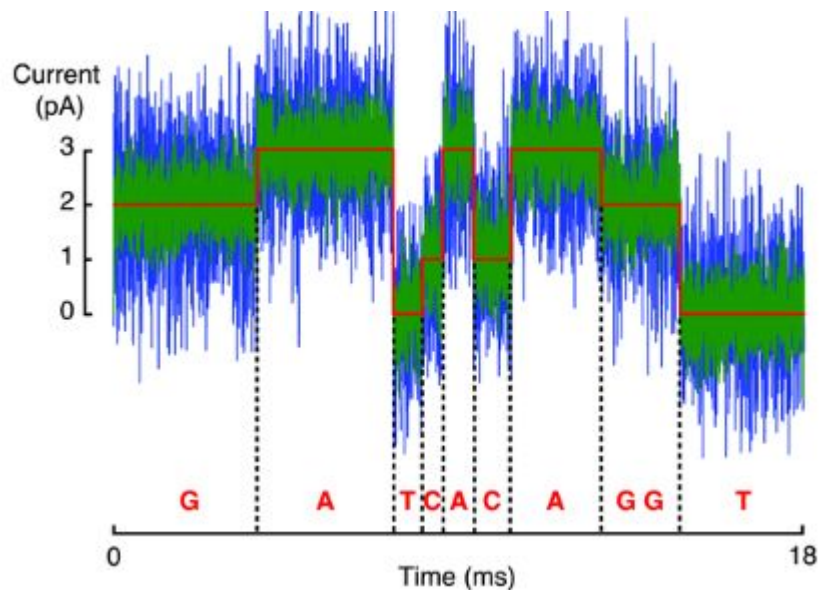
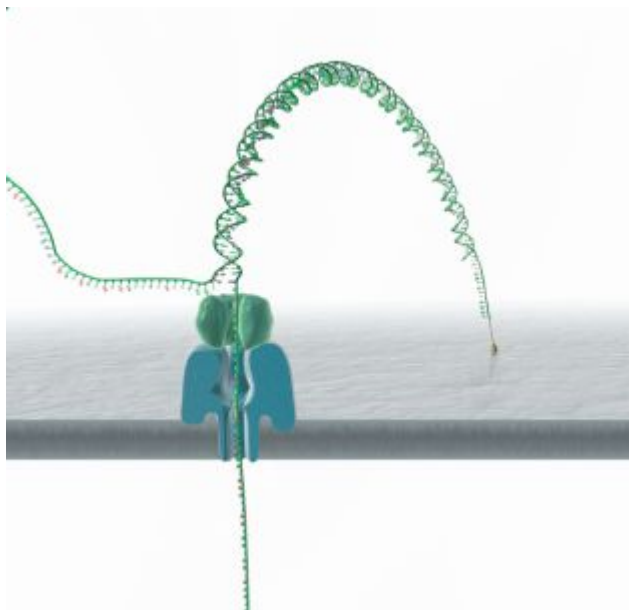
Oxford nanopore minION



Измерение силы тока, проходящего через каждую нанопору – пикоА.



Oxford nanopore minION



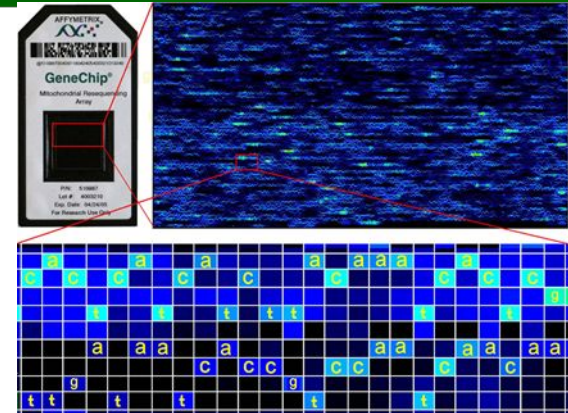
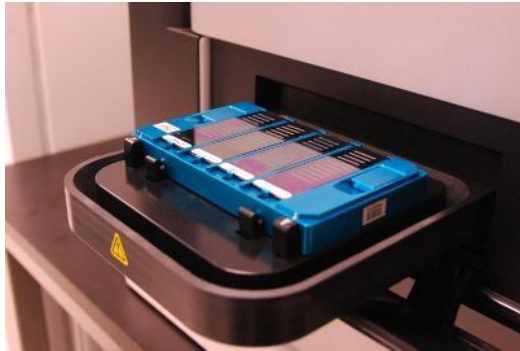
Длинные чтения

Высокий уровень ошибок – до 15%

Систематические ошибки – до 3-5%



Микрочиповые технологии анализа генома и транскриптома



Микрочипы – стеклянные пластинки с иммобилизованными короткими фрагментами ДНК

А) на основе длинных фрагментов НК

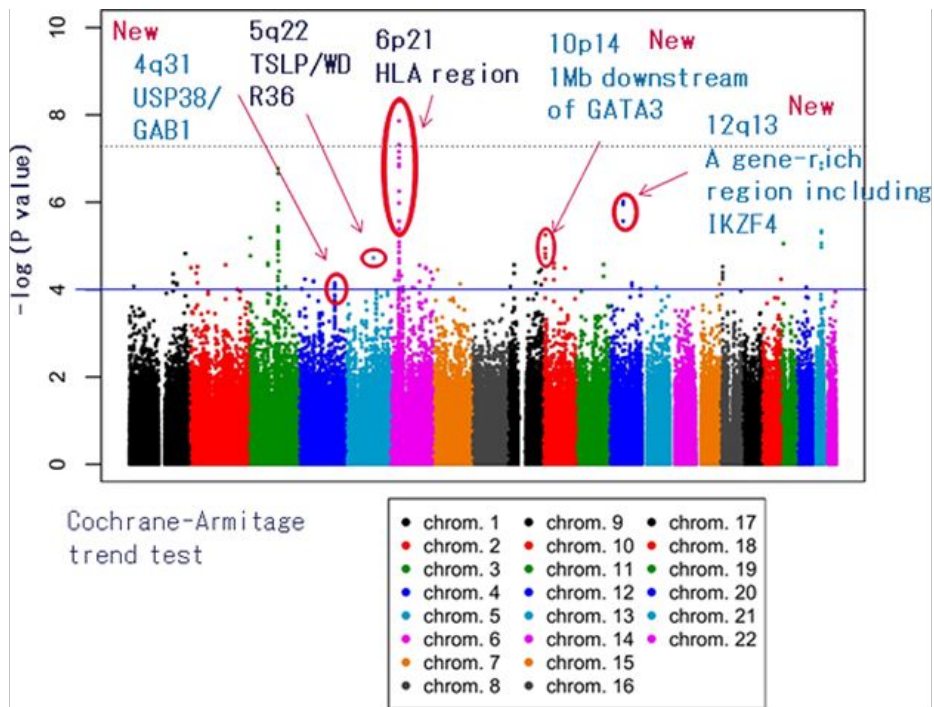
Б) олигонуклеотидные

- короткие
- средней длины
- синтез in situ

С) Чипы Infinium с достройкой 1 п.н.



Микрочиповые технологии анализа генома и транскриптома



Genome-wide association study (asthma)



Структурная геномика изучает последовательность нуклеотидов в геноме, определяет границы и строение генов, межгенных участков, промоторов, энхансеров и др.

Функциональная геномика идентифицирует функцию каждого гена и участка генома, их взаимодействие в клеточной системе.

Эволюционная геномика объясняет пути эволюции геномов, происхождение генетического полиморфизма и биоразнообразия



Секвенирование *de novo*

Исследование ранее не секвенированного генома

Метагеномный проект

Секвенирование суммарной ДНК, выделенной из определённого местообитания, с целью определения видового разнообразия и представленности различных микроорганизмов.

Ресеквенирование генома

Для выполнения обязательно необходима референсная геномная последовательность. Получаемый результат в основном не может выйти за пределы референсной последовательности

Ресеквенирование экзона

Секвенирование всей совокупности или целевой части кодирующих последовательностей генома



Целевое секвенирование (Targeted resequencing) — секвенирование небольшого количества выбранных генов

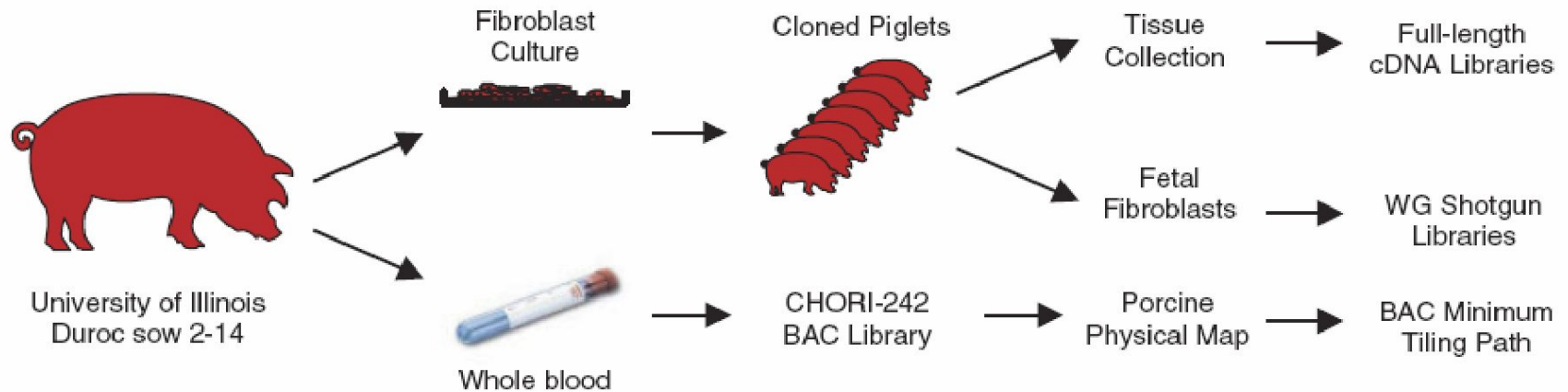
Полногеномное изучение ДНК-белковых взаимодействий
ChIP-Seq, ChIA-PET, 4-C, др. Изучение регуляторных районов генов

Транскриптомный проект
Исследование транскриптома - всего пула мРНК.

Single-cell genome\transcriptome
Стартовым материалом является единичная клетка или хромосома



Геномный проект *de novo* как совокупность подпроектов

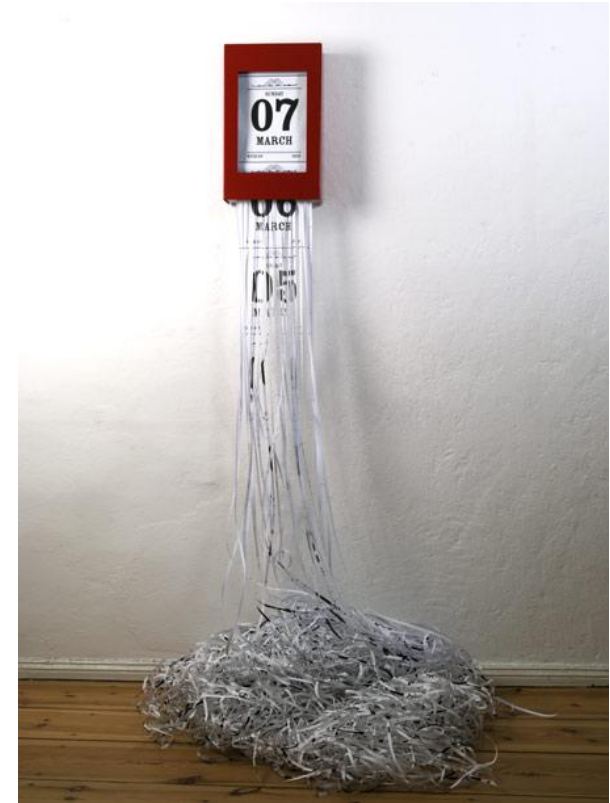


- А) Секвенирование транскриптома
- Б) Секвенирование Shotgun-библиотек с расчётным 3х покрытием, использование библиотек с встройками 3 kb, 10 kb и 50 kb
- В) Создание и секвенирование ВАС-библиотек с расчётным 3х покрытием
- С) Комбинирование данных при биоинформатическом анализе

Schook *et al.*,
2005



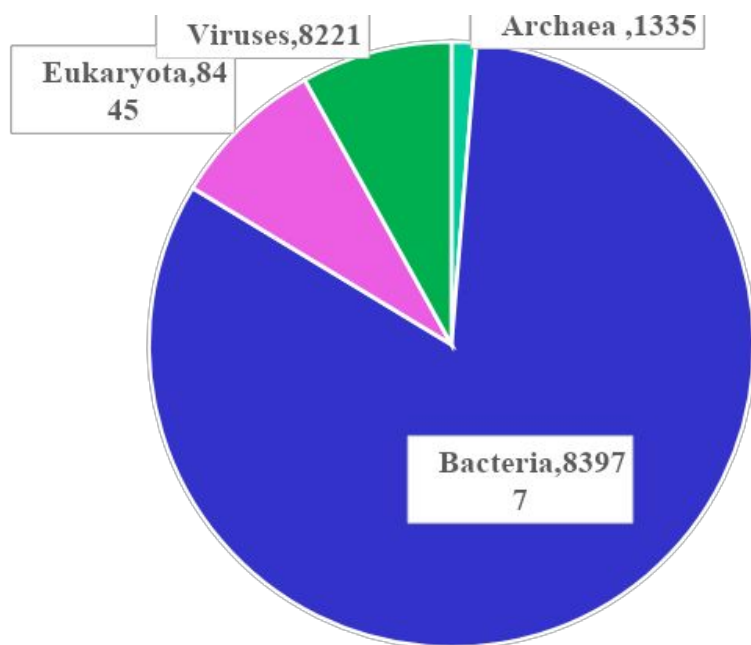
Анализ данных в геномном проекте *de novo*



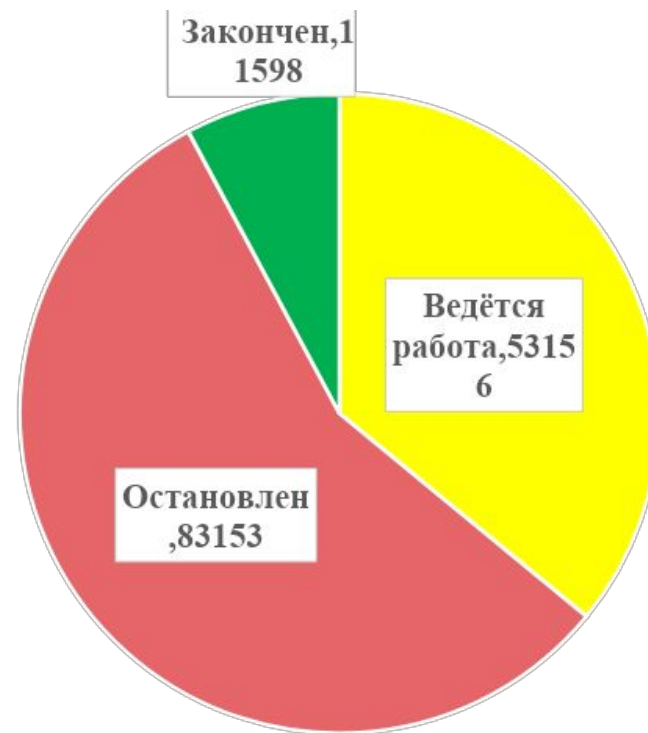
Вот так и выглядят гигабазы коротких последовательностей



Статистика современных геномных проектов



- Archaea
- Bacteria
- Eukaryota
- Viruses



- Ведётся работа
- Остановлен
- Закончен

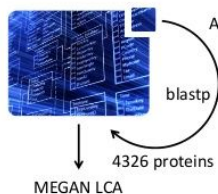
По публикациям



Метагеномика – структурная или функциональная



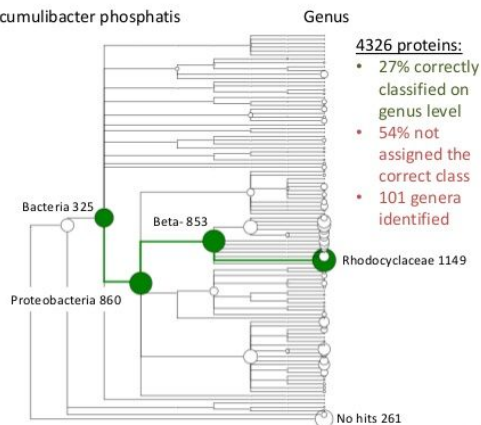
Effect of missing genomes



Lowest common ancestor (LCA) approach:
HR 1: Beta-proteobacteria 80% ID
HR 2: Gamma-proteobacteria 79% ID
HR 3: Actinobacteria 59% ID
Assigned to Proteobacteria

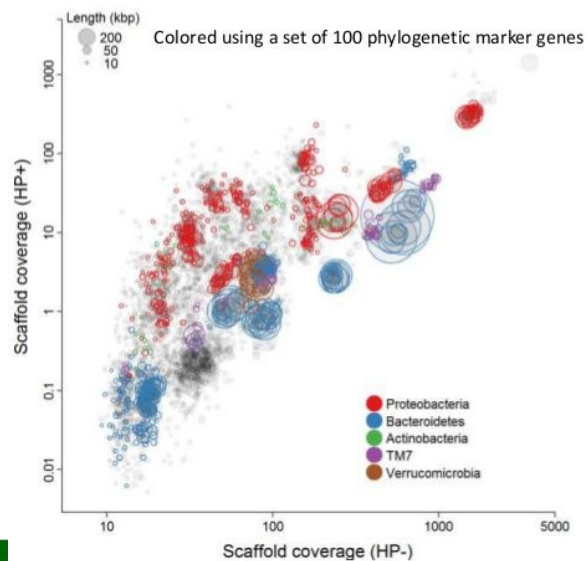
Related genomes

Bacteria	1268
Proteobacteria	564
Betaproteobacteria	84
Rhodocyclales	5
Rhodocyclaceae	5



4326 proteins:
• 27% correctly classified on genus level
• 54% not assigned the correct class
• 101 genera identified

CENTER FOR MICROBIAL COMMUNITIES | AALBORG UNIVERSITY



Albertsen et al., 2013 Nat. Biotech.

CENTER FOR MICROBIAL COMMUNITIES | AALBORG UNIVERSITY

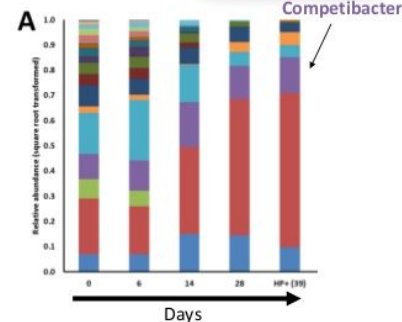


Full-scale EBPR plant

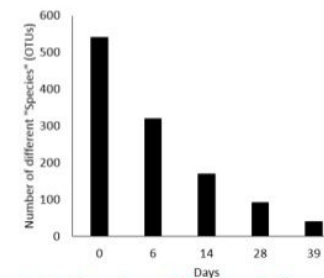


Short term enrichment

SBR reactor



Albertsen et al., 2013 Nat. Biotech.

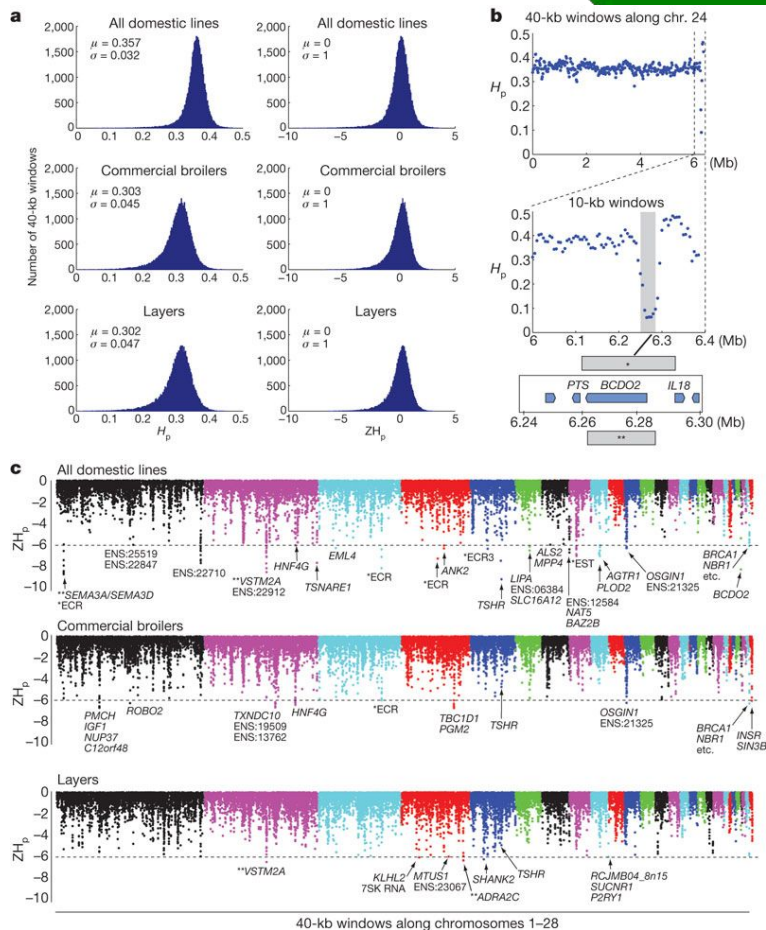


1. Reduction of (micro)-diversity

CENTER FOR MICROBIAL COMMUNITIES | AALBORG UNIVERSITY



Ресеквенирование генома



2010 г. – завершён проект «тысяча геномов», для каждого персонального генома человека секвенировано около 85% последовательности.

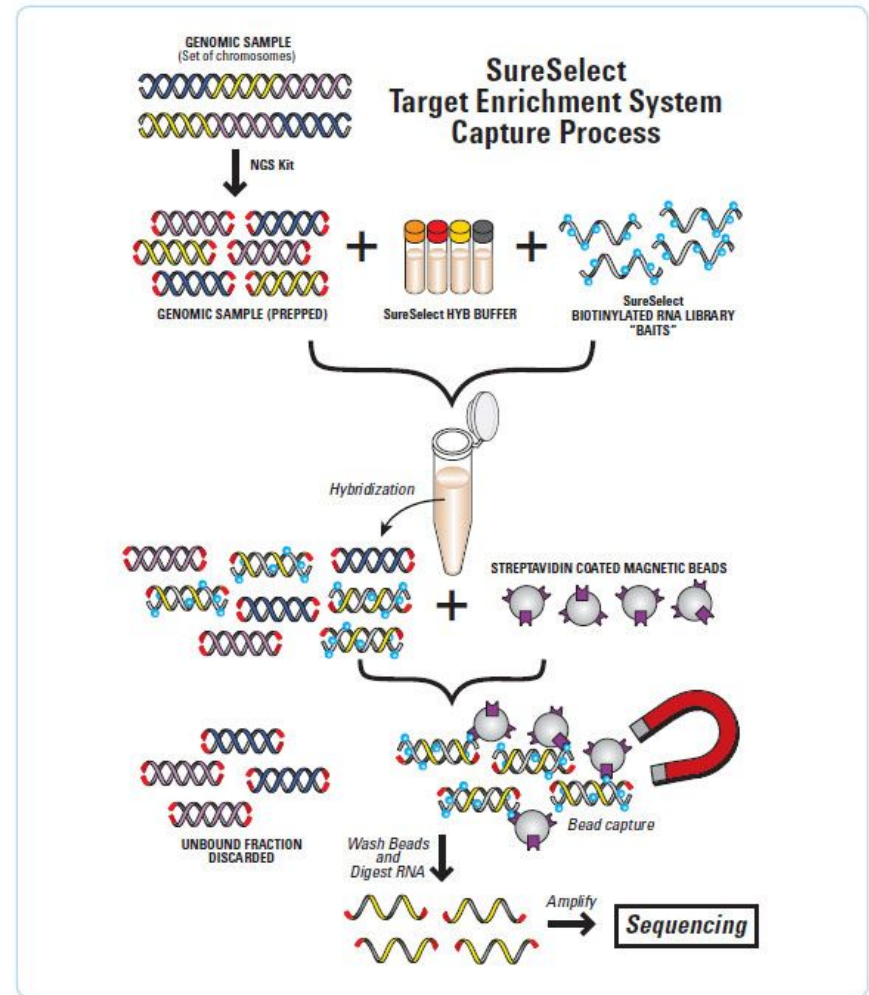
Ресеквенирование генома различных пород кур позволило выделить гены, изменённые в ходе селекции



Секвенирование экзона

Экзом – совокупность всех кодирующих частей генов. Около 1% от всего генома. Часто включают функционально важные фланкирующие и регуляторные последовательности – тогда занимает до 1.5-2%

Экзом может быть полным либо частичным





Целевое ресеквенирование - панели генов для анализа мутаций

The Ion AmpliSeq™ Cancer Panel targets 50 genes

<i>ABL1</i>	<i>EZH2</i>	<i>JAK3</i>	<i>PTEN</i>
<i>AKT1</i>	<i>FBXW7</i>	<i>IDH2</i>	<i>PTPN11</i>
<i>ALK</i>	<i>FGFR1</i>	<i>KDR</i>	<i>RB1</i>
<i>APC</i>	<i>FGFR2</i>	<i>KIT</i>	<i>RET</i>
<i>ATM</i>	<i>FGFR3</i>	<i>KRAS</i>	<i>SMAD4</i>
<i>BRAF</i>	<i>FLT3</i>	<i>MET</i>	<i>SMARCB1</i>
<i>CDH1</i>	<i>GNA11</i>	<i>MLH1</i>	<i>SMO</i>
<i>CDKN2A</i>	<i>GNAS</i>	<i>MPL</i>	<i>SRC</i>
<i>CSF1R</i>	<i>GNAQ</i>	<i>NOTCH1</i>	<i>STK11</i>
<i>CTNNB1</i>	<i>HNF1A</i>	<i>NPM1</i>	<i>TP53</i>
<i>EGFR</i>	<i>HRAS</i>	<i>NRAS</i>	<i>VHL</i>
<i>ERBB2</i>	<i>IDH1</i>	<i>PDGFRA</i>	
<i>ERBB4</i>	<i>JAK2</i>	<i>PIK3CA</i>	

Ion 314™ Chip: 2 samples, ~1,400x average coverage
Ion 316™ Chip: 8 samples, ~1,400x average coverage
Ion 318™ Chip: 16 samples, ~1,400x average coverage

Ion AmpliSeq Comprehensive Cancer Panel

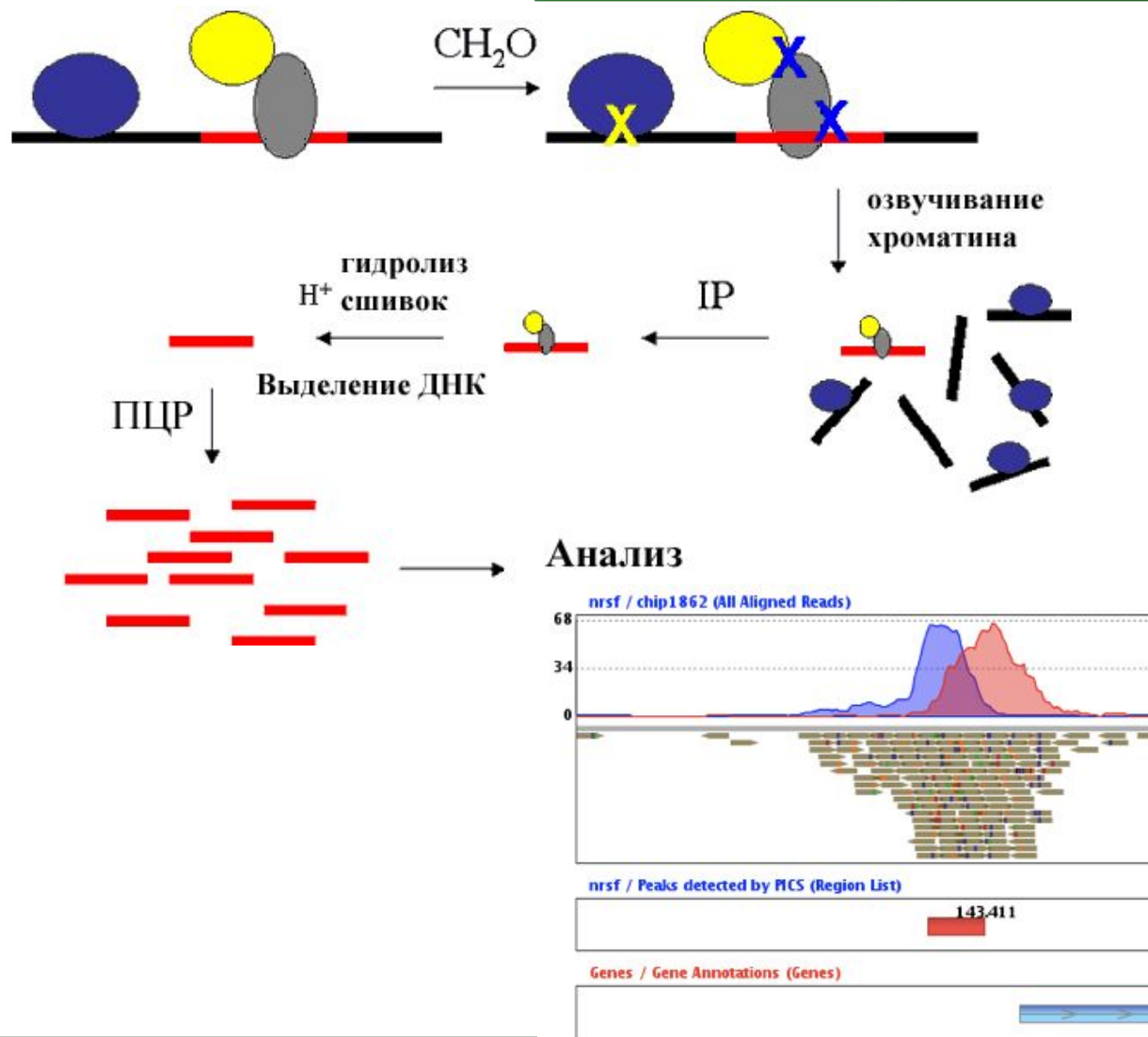
Экзоны 409 генов-супрессоров, онкогенов и регуляторных белков: регуляторы апоптоза, репарации ДНК, ростовые и транскрипционные факторы

~16,000 в 4-х смесях, мультиплексирование на чип IonProton 4 образца при 1000x покрытия

Ключевая особенность — возможность существенной автоматизации анализа для конечного пользователя

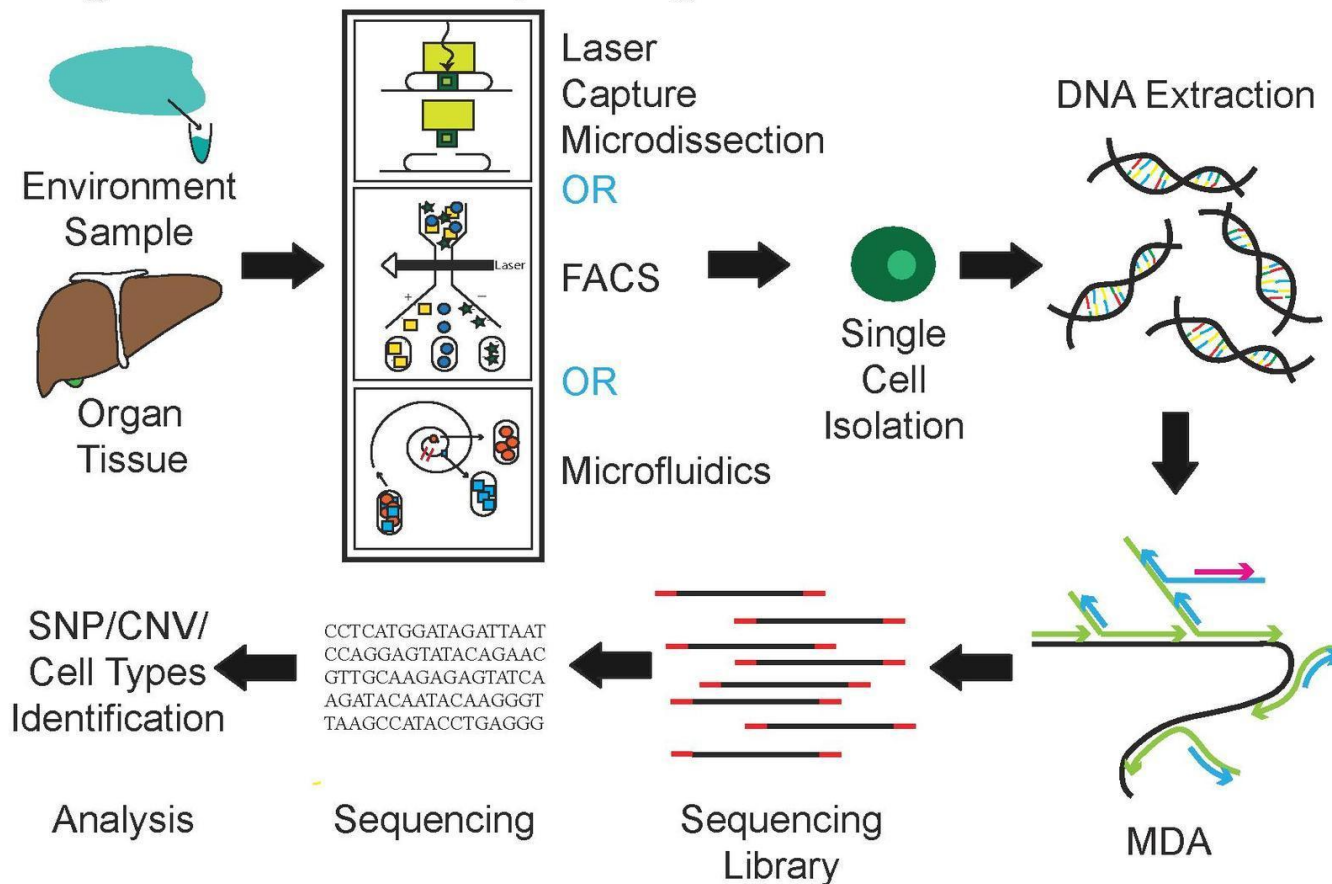


ChIP-Seq — полногеномный анализ взаимодействий



Анализ генома одной клетки (Single-cell sequencing)

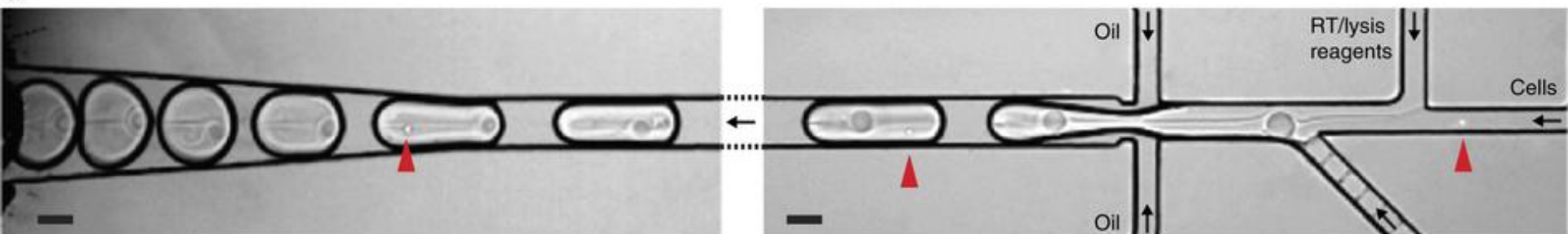
Single Cell Genome Sequencing Workflow





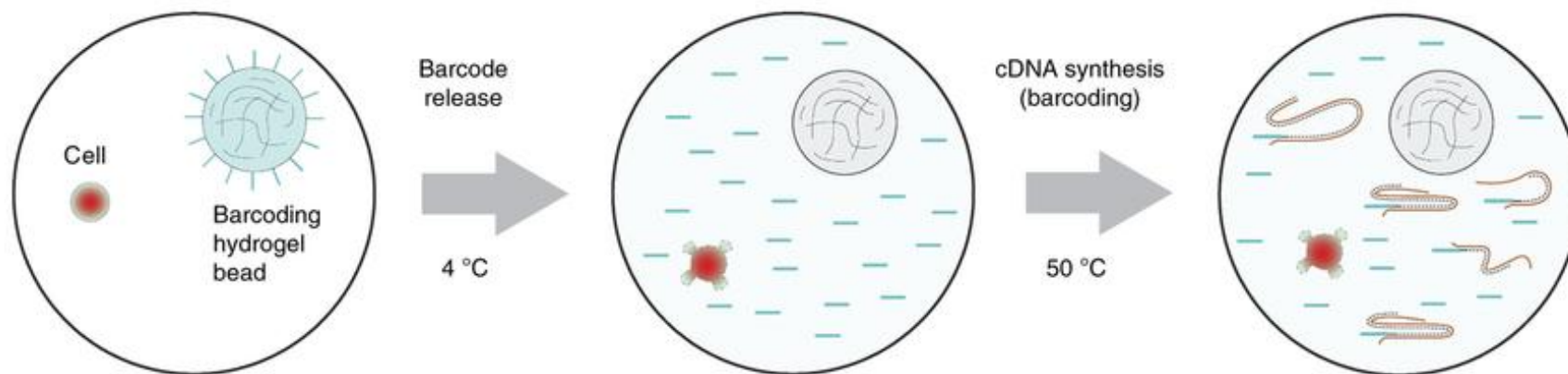
Single cell transcriptome

a

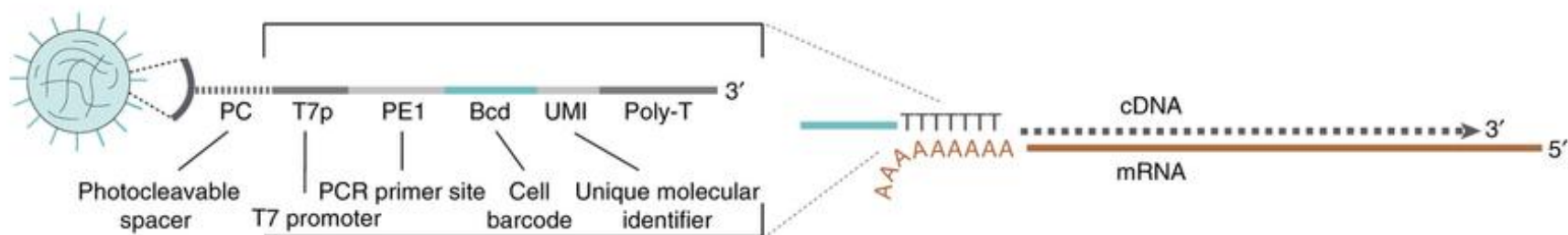


Barcoding hydrogel beads

b



c





Спасибо за внимание