



«Анализ транскриптомных данных»

# Лекция #6. Транскриптомика одиночных клеток

**Серёжа Исаев**

аспирант **ФБМФ МФТИ**  
аспирант **MedUni Vienna**

# Содержание курса

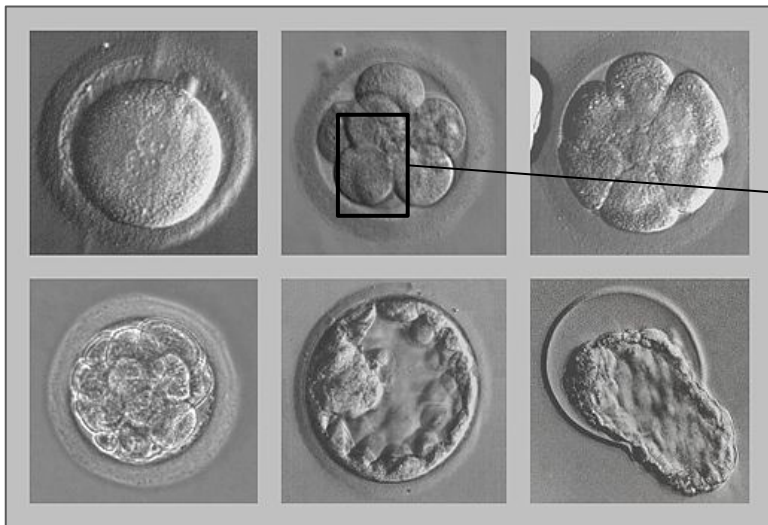
## 1. Bulk RNA-Seq:

- a. экспериментальные подходы,
- b. выравнивания и псевдовыравнивания,
- c. анализ дифференциальной экспрессии,
- d. функциональный анализ;

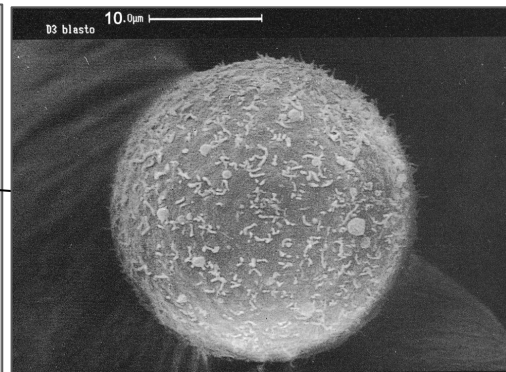
## 2. Single-cell RNA-Seq:

- a. экспериментальные подходы,**
- b. отличия от процессинга bulk RNA-Seq,
- c. методы снижения размерности,
- d. кластера и траектории,
- e. мультимодальные омики одиночных клеток.

# Tang et al., 2009 — первая работа

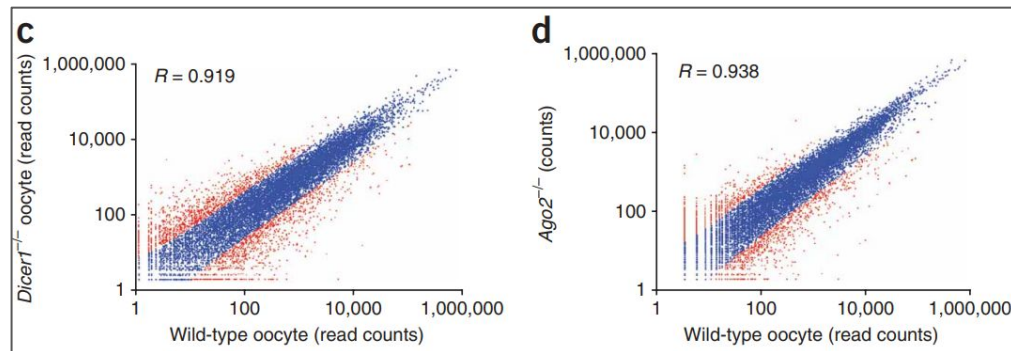


Оплодотворённая яйцеклетка, восьмиклеточная стадия, стадия адгезии, морула, бластоцист, зона вылулления.  
Источник: <http://nobelprize.org/>



Бластомер. Источник:  
<https://www.ehd.org/>

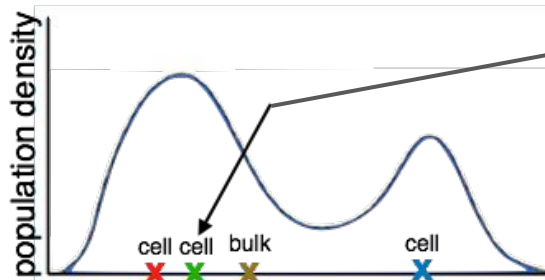
Сравнение профилей  
экспрессии нормального  
бластомера и бластомера с  
нокаутами из Tang et al. 2009



	<b>Bulk RNA-Seq</b>	<b>scRNA-Seq</b>
<b>Начало</b>	2008	2009
<b>Экспрессия</b>	средний уровень экспрессии	распределение уровней экспрессии
<b>Количество транскриптов</b>	~15-20 000 на образец	~200-10 000 на клетку
<b>% Транскриптома</b>	80-95%	10-50%

## Bulk RNA-Seq

	<b>Exp</b>
<b>Gene1</b>	100
<b>Gene2</b>	5.5
...	
<b>Gene N</b>	0.5

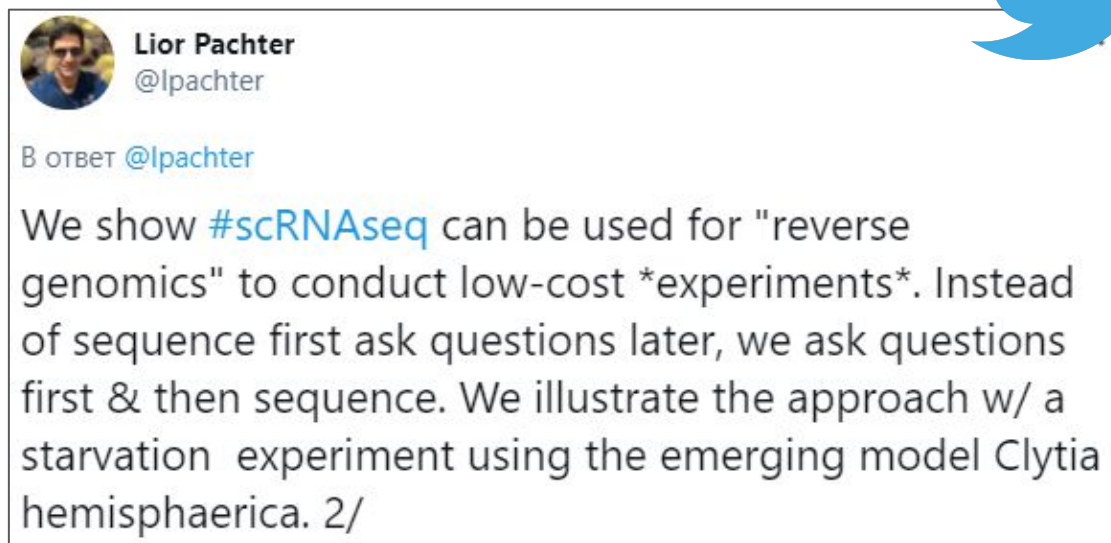


## scRNA-Seq

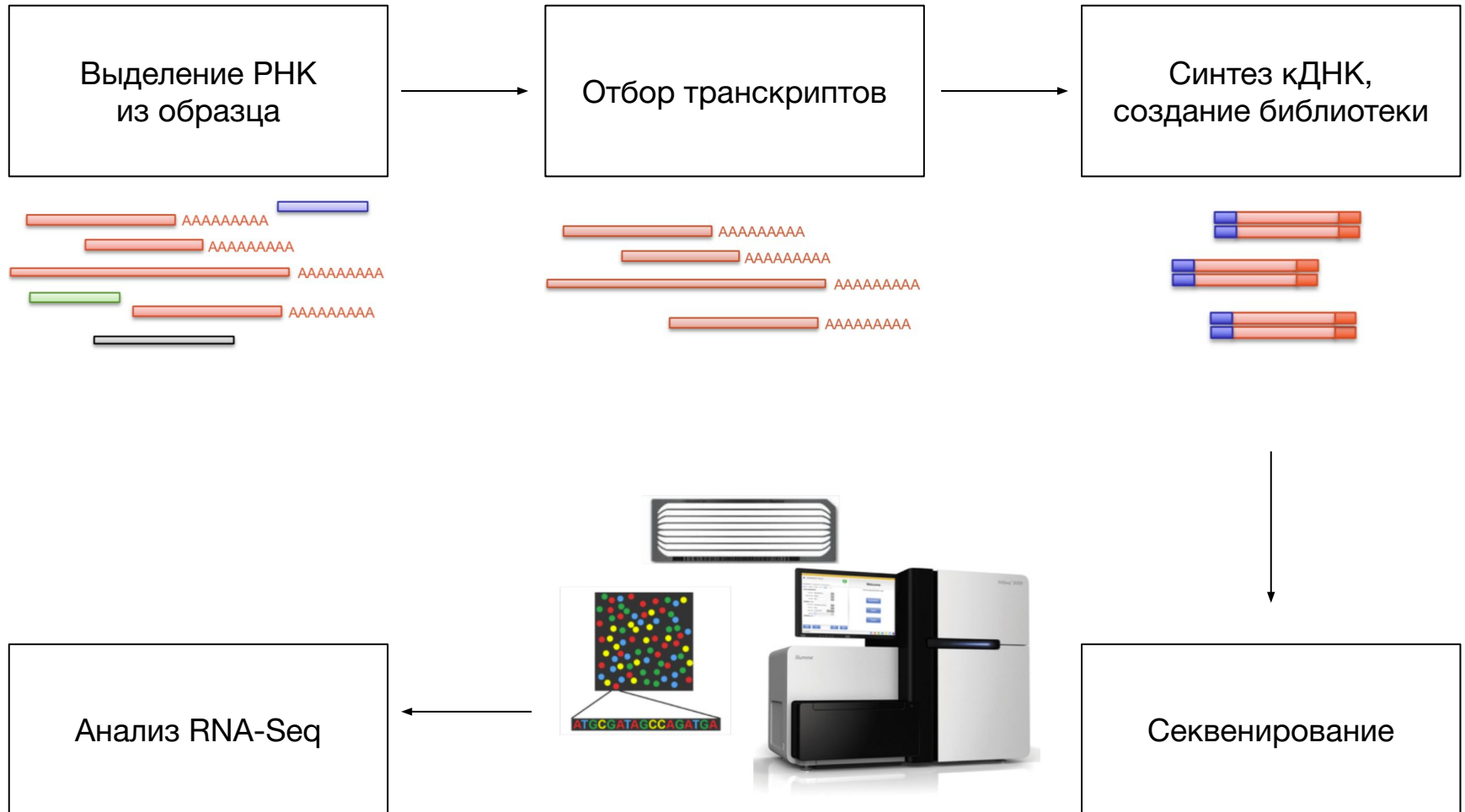
	<b>Cell1</b>	<b>Cell2</b>	...	<b>Cell K</b>
<b>Gene1</b>	3	0		2
<b>Gene2</b>	0	2		1
...				
<b>Gene N</b>	0	13		2

# Планирование эксперимента

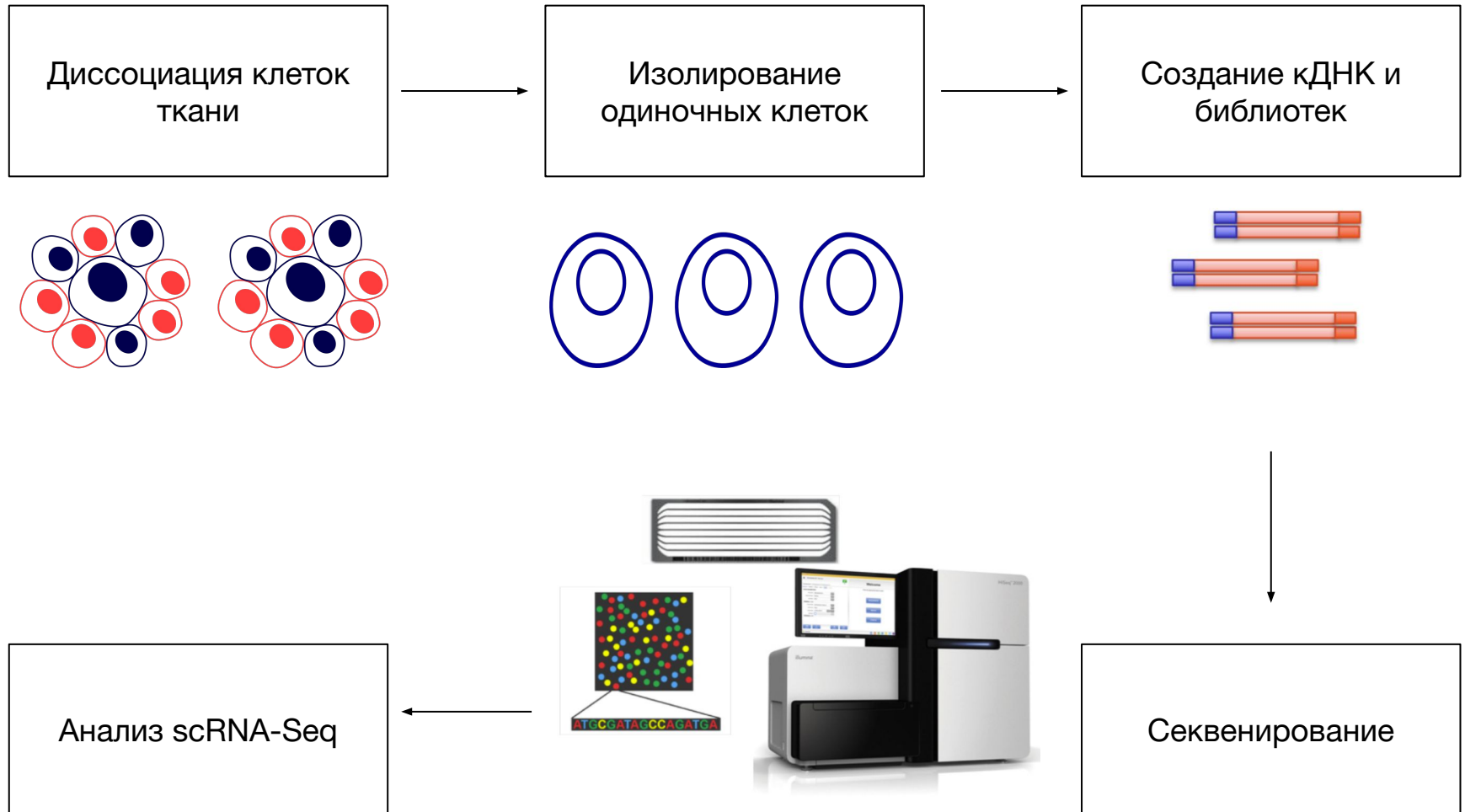
Сначала задайте вопросы — **какое именно явление я хочу изучить?** и **какая у меня гипотеза?** — и только потом ставьте эксперимент. Это касается не только scRNA-Seq, но и вообще любых исследований.



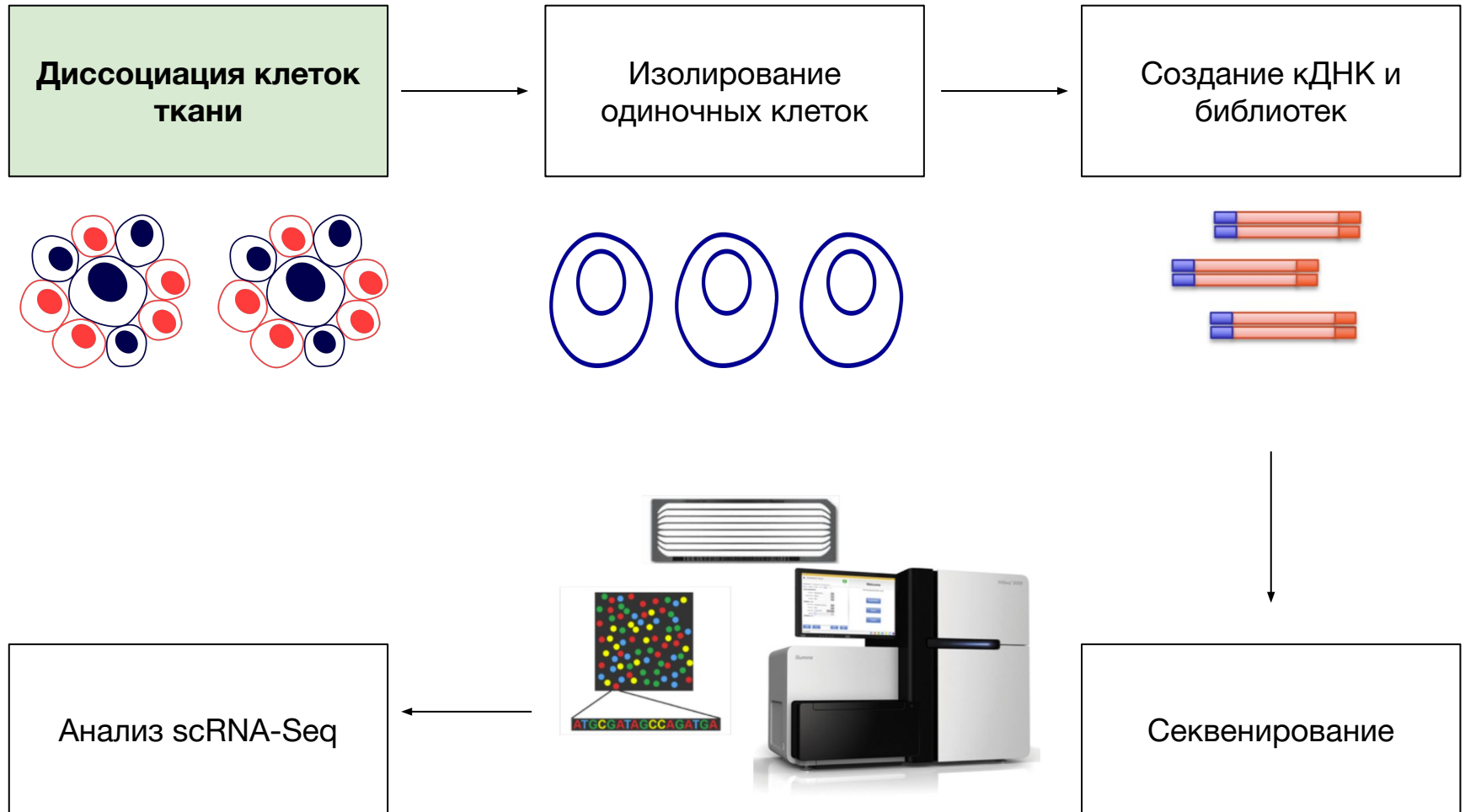
# Общая схема эксперимента RNA-Seq



# Общая схема эксперимента scRNA-Seq



# Общая схема эксперимента scRNA-Seq



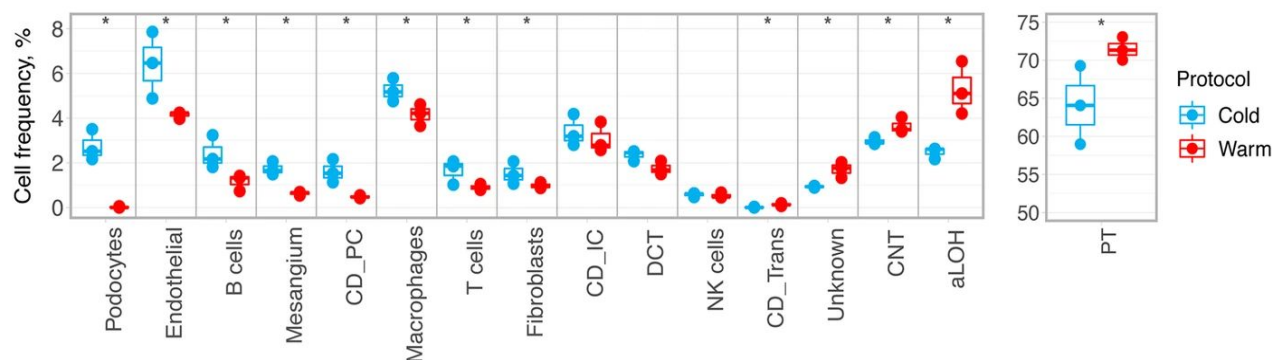


# Диссоциация клеток

В основном клетки в изучаемых тканях находятся в “сцепленном” состоянии (они соединены при помощи молекул адгезии и т. п.). Для того, чтобы их “расцепить”, необходимо провести диссоциацию ткани:

- диссоциация при нагревании (напр., Multi-tissue dissociation kit 2) — может вызвать активацию транскрипции генов теплового шока,
- диссоциация на холоду (с использованием протеазы *Bacillus Licheniformis*).

Сравнение методов диссоциации  
ткани из Denisenko et al., 2020.  
Некоторые клеточные типы не  
детектируются при диссоциации  
“горячим” методом



# Создание клеточного атласа всего организма

Диссоциация ткани — это один из самых важных шагов при пробоподготовке scRNA-Seq. Во время этой процедуры клетки могут

1. умереть (и тогда мы увидим смещённый клеточный состав),
2. изменить свой экспрессионный профиль (и тогда мы увидим тот же клеточный состав, но не в нативном состоянии),
3. диссоциировать неполностью (и тогда мы увидим большое количество дублетов).

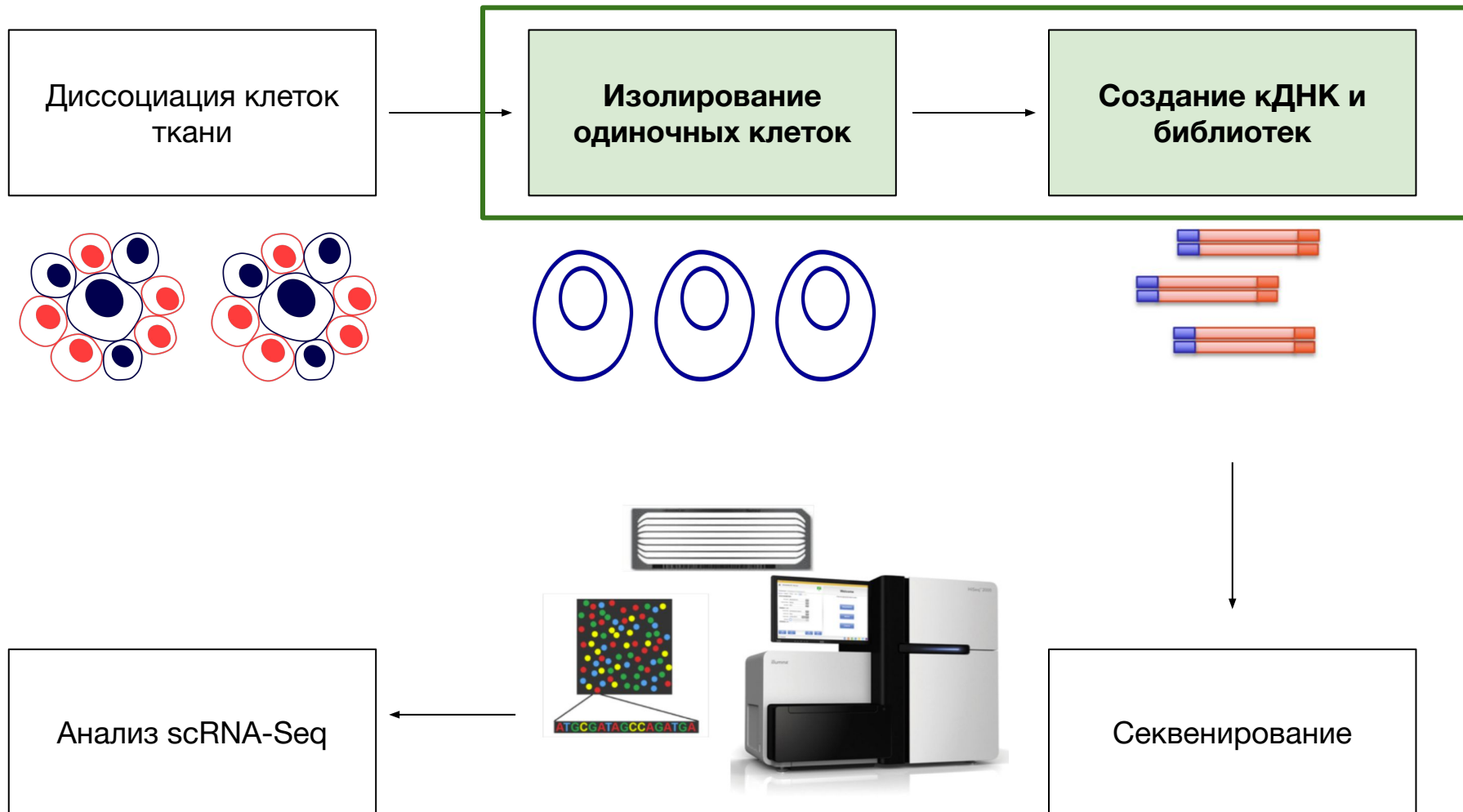
Для каждой ткани используется собственный протокол диссоциации. И это является очень большой проблемой для создания пан-тканевого клеточного атласа.

# scRNA-Seq vs. snRNA-Seq

- Профилирование ядер из крупных клеток ( $> 40$  мкм), которые не проходят через микрофлюидику
- Позволяет профилировать отдельные ядра, выделенные из замороженных тканей, отделяя получение ткани от немедленной обработки образца
- snRNA-Seq может также обрабатывать образцы, которые не могут быть успешно диссоциированы, даже если они свежие, из-за хрупкости клеток

Ядра имеют меньшее количество мРНК по сравнению с клетками, и их сложнее обогатить для конкретных интересующих типов клеток.

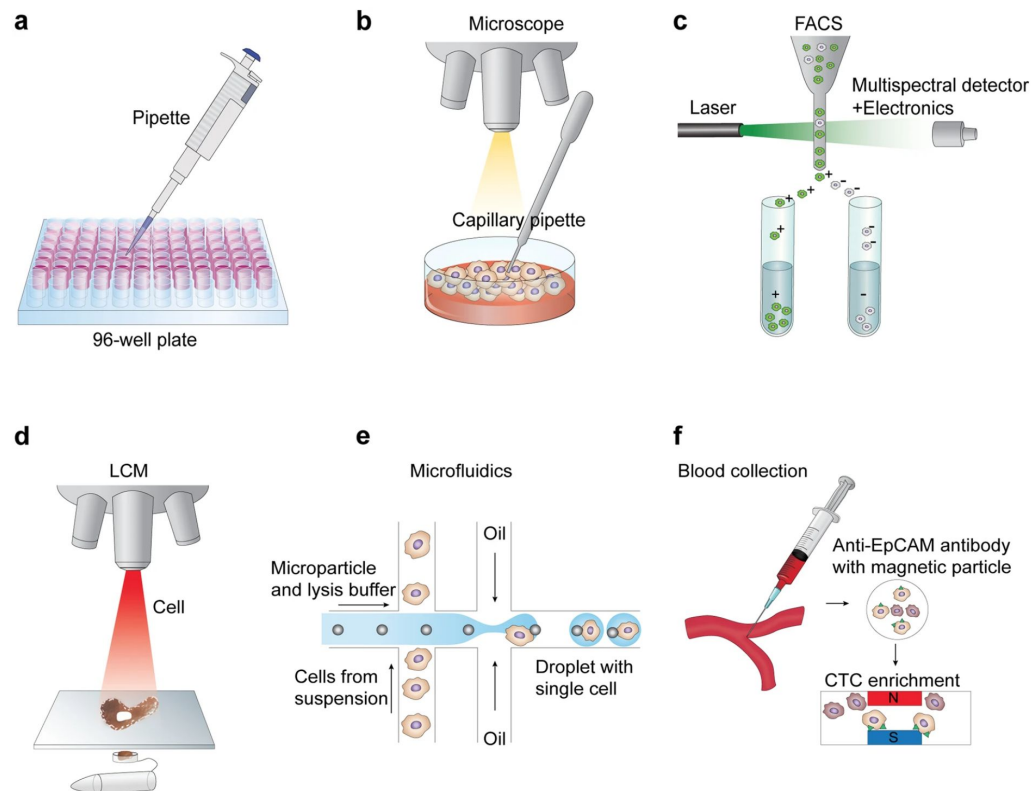
# Общая схема эксперимента scRNA-Seq



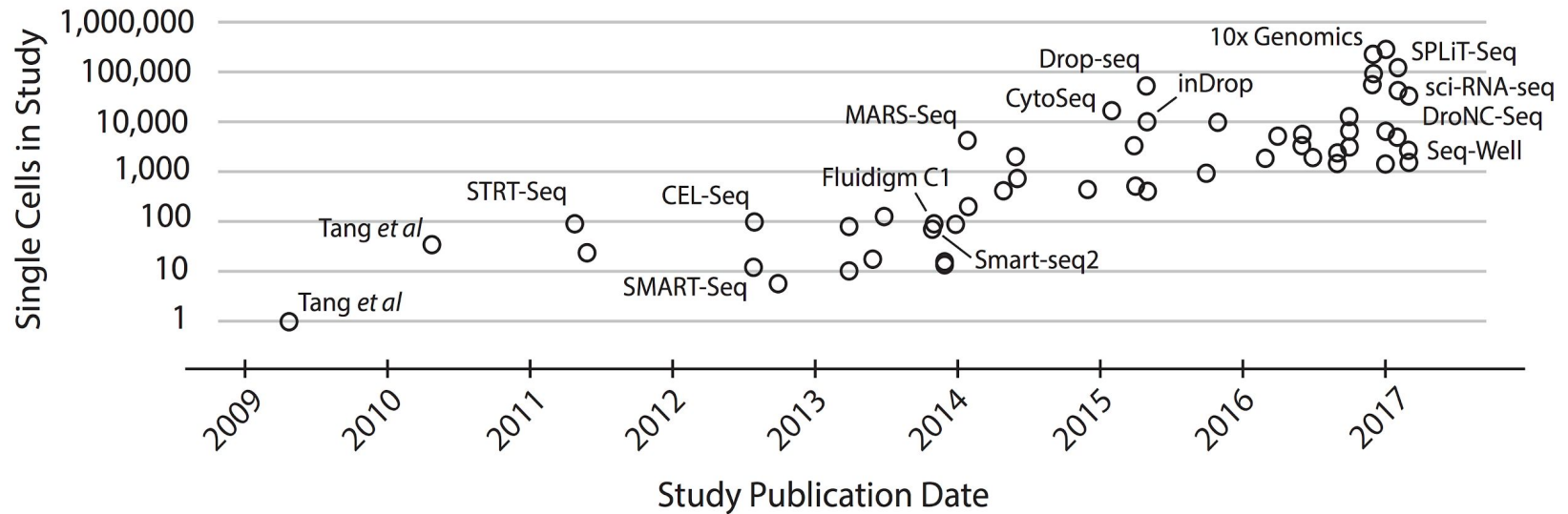
# Изолирование одиночных клеток

Существует множество различных способов изолировать одиночные клетки

Как правило, стадия изолирования одиночных клеток очень тесно связана с дальнейшими стадиями подготовки библиотек, поэтому рассмотрим их вместе

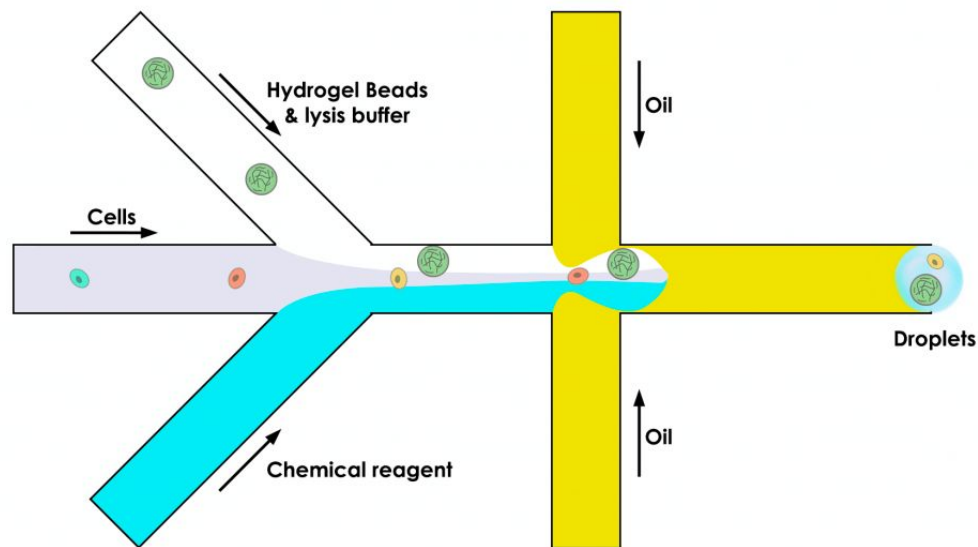


# История развития методов scRNA-Seq



# Droplet-based (капельные) методы

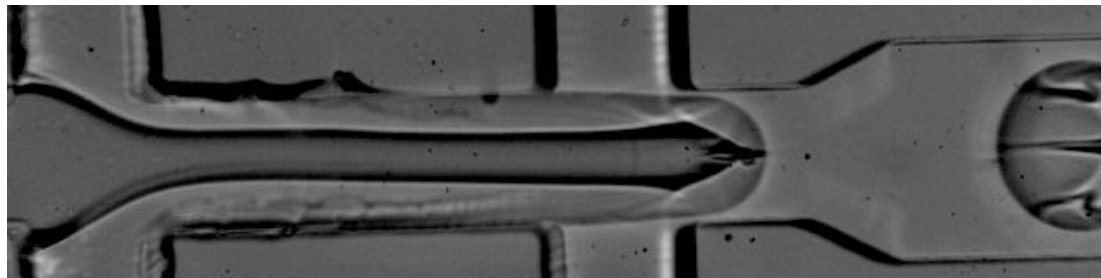
Капельные методы основаны на том, что клетки изолируются друг от друга, поступая по капиллярам в масляную фракцию и образуя там отдельные компартменты, содержащие необходимые реагенты и одну клетку



Источник: <http://mccarrolllab.org/dropseq/>

# Droplet-based (капельные) методы

Капельные методы основаны на том, что клетки изолируются друг от друга, поступая по капиллярам в масляную фракцию и образуя там отдельные компартменты, содержащие необходимые реагенты и одну клетку



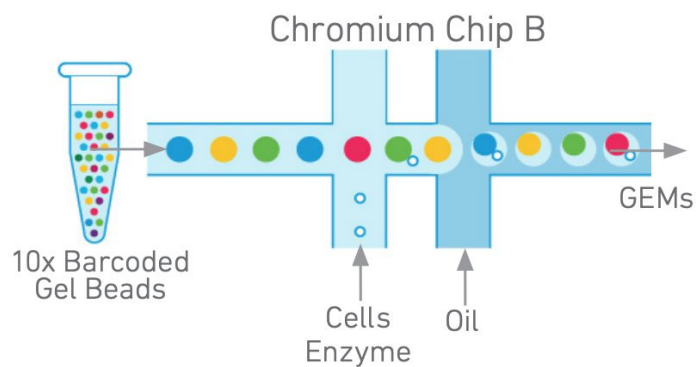
Источник: [https://www.elveflow.com/microfluidic-reviews/droplet-digital-microfluidics/drop-seq/#\\_ftn4](https://www.elveflow.com/microfluidic-reviews/droplet-digital-microfluidics/drop-seq/#_ftn4)



# 10x Chromium

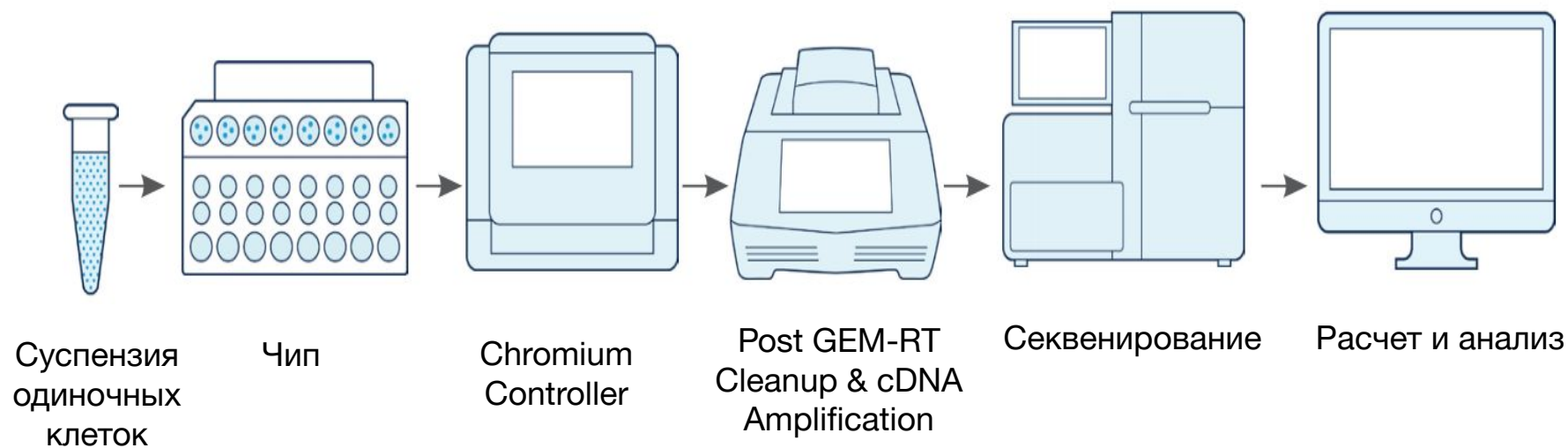
Источник: [10xgenomics.com](https://10xgenomics.com)

Контроллер 10x Chromium является сейчас одной из самых популярных платформ для создания библиотек scRNA-Seq

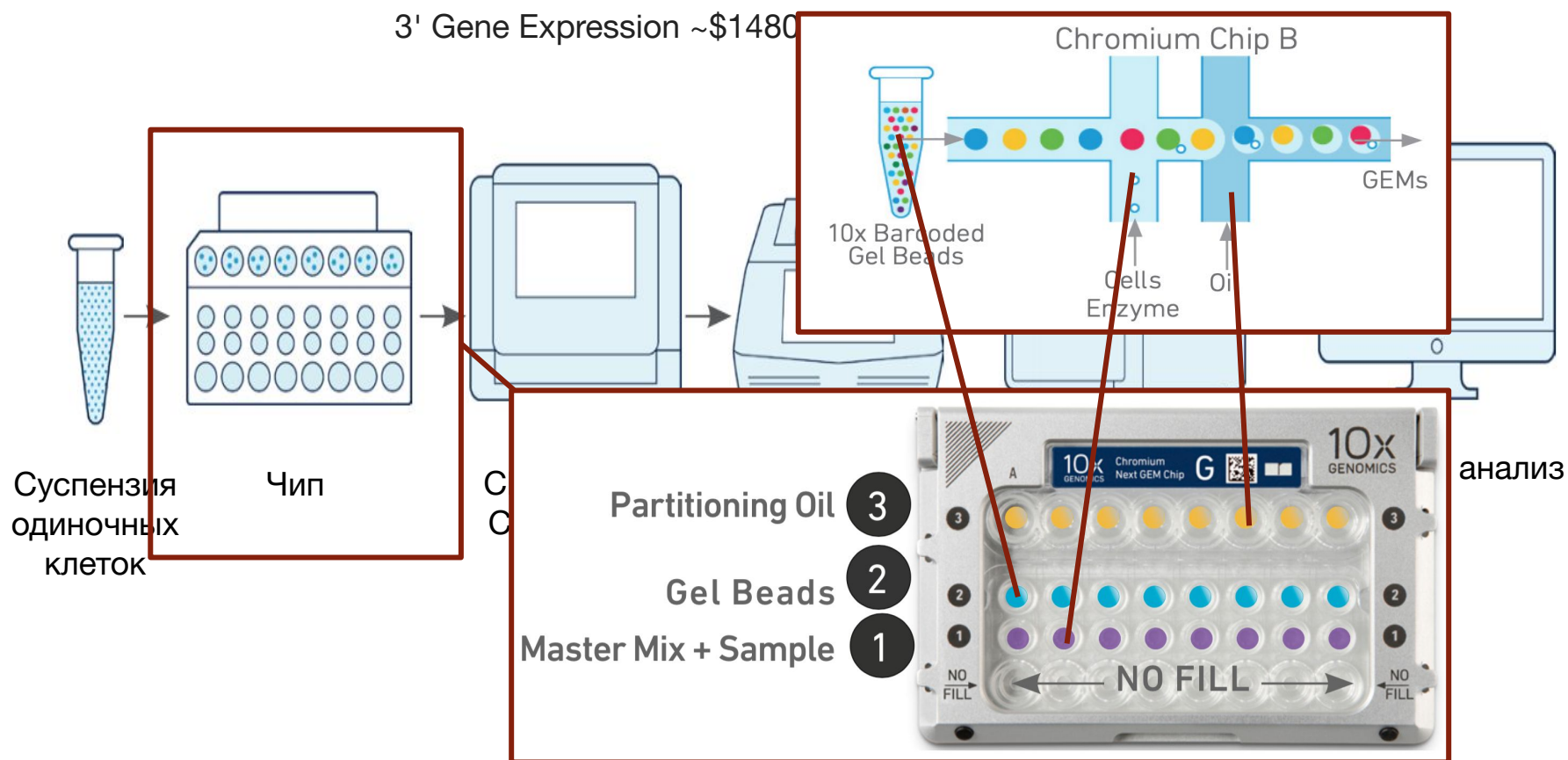


# Процесс пробоподготовки

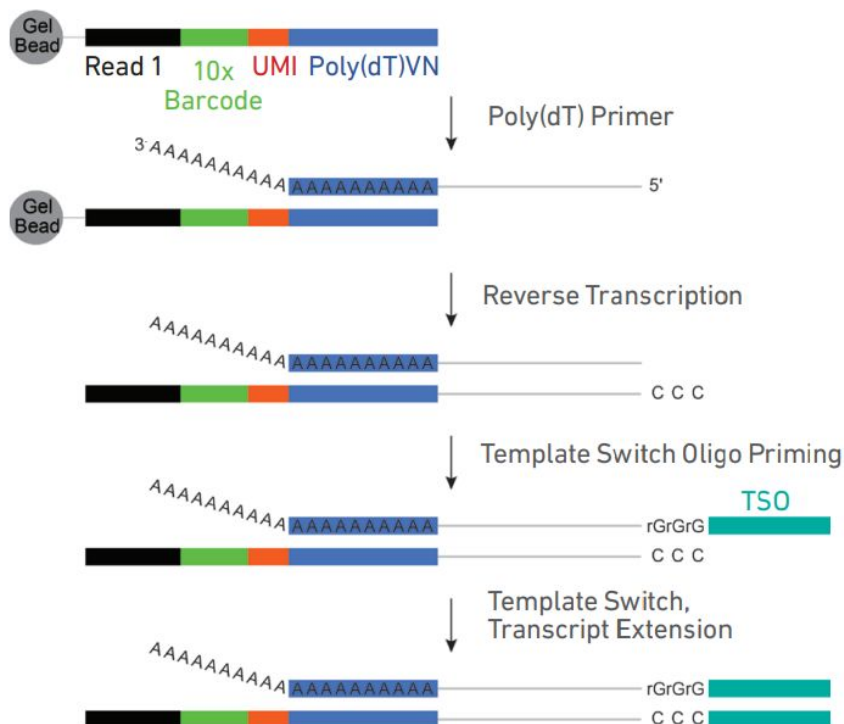
3' Gene Expression ~\$1480/образец (WES ~150\$/образец)



# Процесс пробоподготовки



# 10x v3 3' Gel Beads

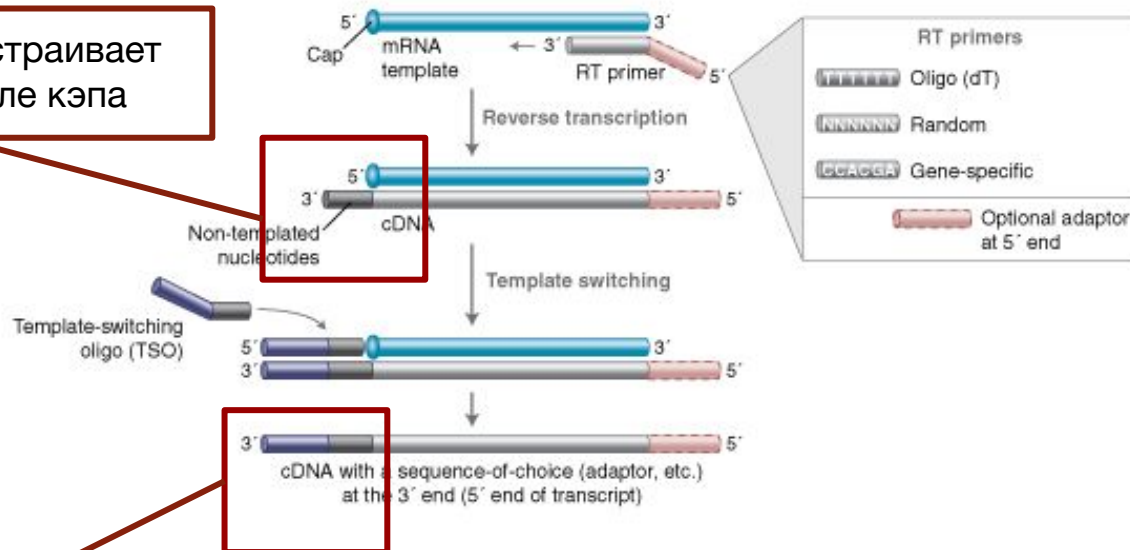


К каждому шарiku прикреплён уникальный праймер, который состоит из

1. Праймера Illumina TruSeq Read 1,
2. **Баркода** (последовательности, которая одинакова у всех праймеров данного шарика, однако различается между всеми шариками),
3. **UMI** (последовательности, которая уникальна для всех праймеров данного шарика, но может повторяться между шариками),
4. Poly(dT)-последовательности.

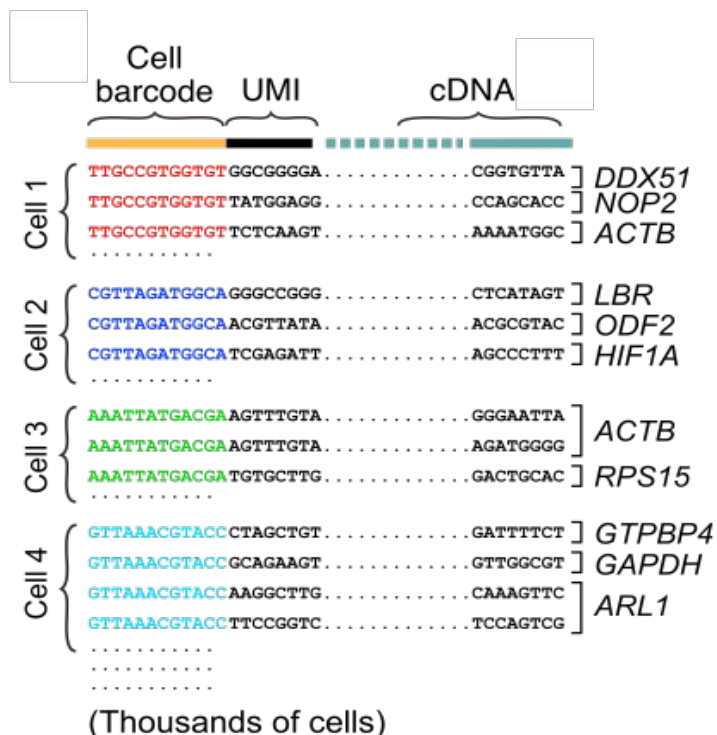
# Template Switch Oligo (TSO)

Полимераза достраивает  
CCC строго после кэпа



После гибридизации TSO и  
CCC происходит смена  
матрицы синтеза

# Баркоды и UMI



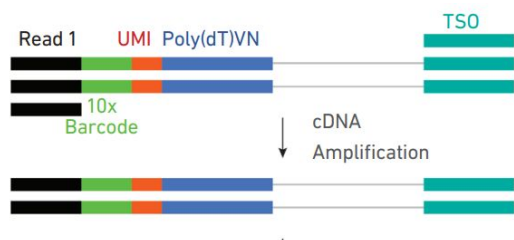
Последовательность баркода, как и UMI, будет в итоге отсеквенирована

Последовательность баркода определяет клетку, к которой мы отнесём данный транскрипт

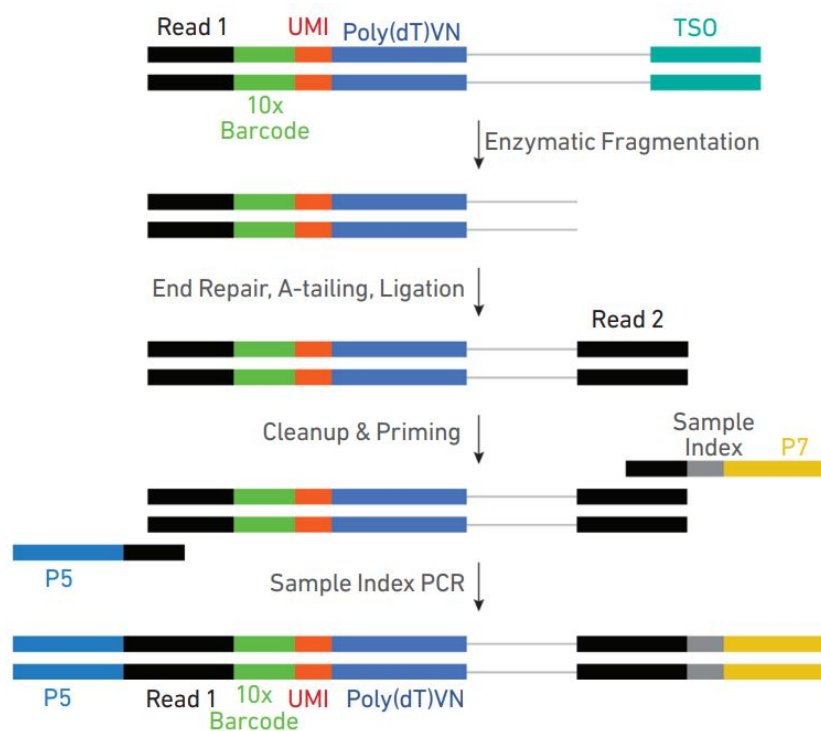
Последовательность UMI позволяет определить ПЦР-дубликаты (amplification bias — это достаточно большая проблема в случае, когда у нас мало РНК)

# Подготовка библиотек для секвенирования

## ПЦР



## Подготовка библиотеки

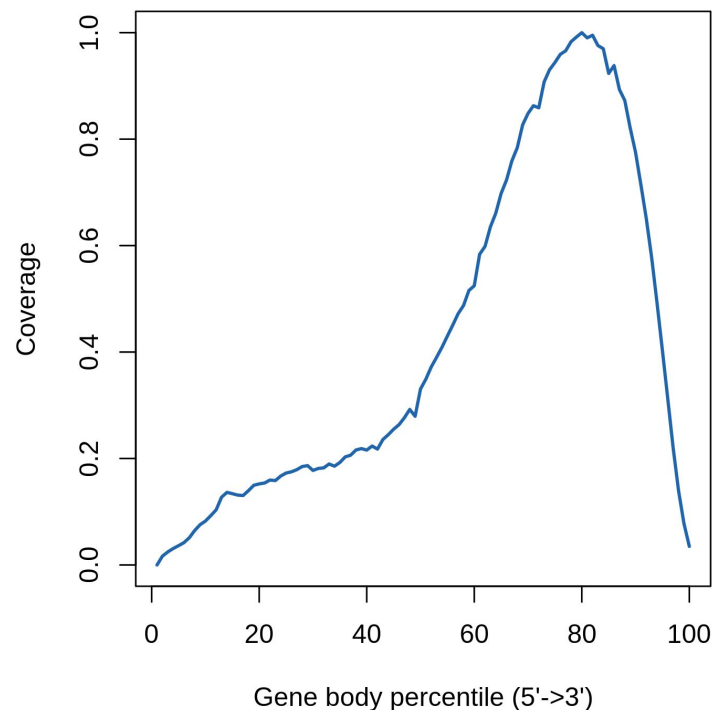


# Покрытие транскрипта ридами (3'-протокол)

Несмотря на TSO, у нас присутствует стадия фрагментации кДНК. Более того, после этой стадии пришивается только один из концов, праймер для секвенирования другого конца всегда зафиксирован на 3'-конце последовательности. Как итог у нас возникает неравномерность в покрытии транскрипта ридами.

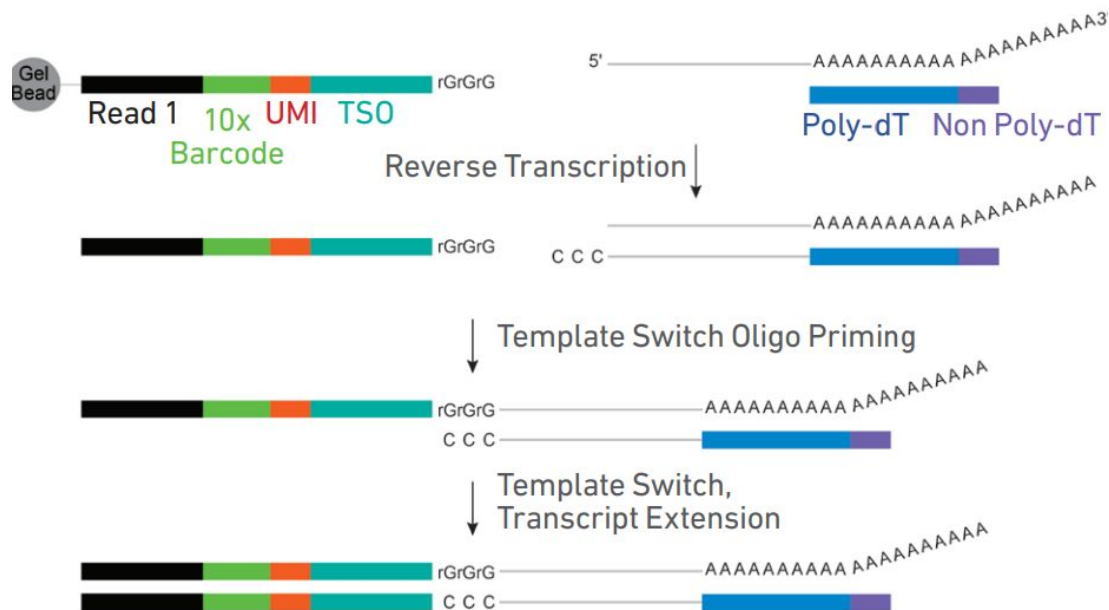
Источник:

[https://liulab-dfci.github.io/MAESTRO/example/RNA\\_infrastructure\\_10x/RNA\\_infrastructure\\_10x.html](https://liulab-dfci.github.io/MAESTRO/example/RNA_infrastructure_10x/RNA_infrastructure_10x.html)





# 10x v3 5' Gel Beads



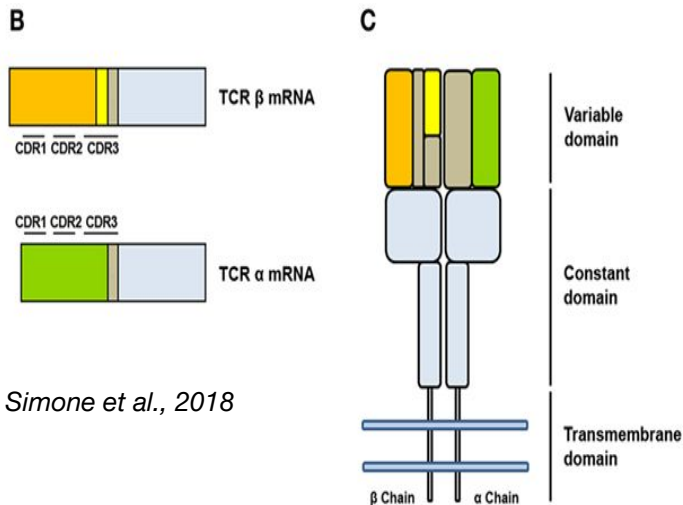
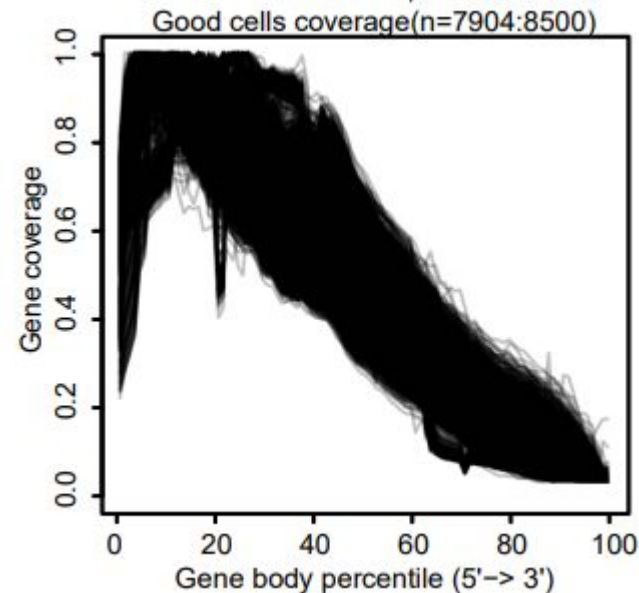
Бывает так (когда? зачем?), что нам катастрофически необходимо хорошее покрытие на 5'-конце транскриптов

Эlegantное решение данной проблемы — это праймер не из баркода-UMI-oligo(dT), а из баркода-UMI-TSO

# Покрытие транскрипта ридами (5'-протокол)

В случае с 5'-протоколом также наблюдается неравномерность покрытия транскрипта прочтениями, однако уже в сторону 5'-конца транскрипта

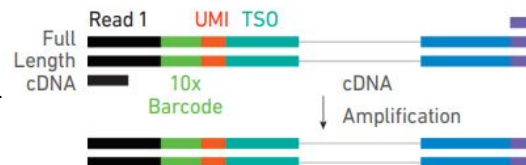
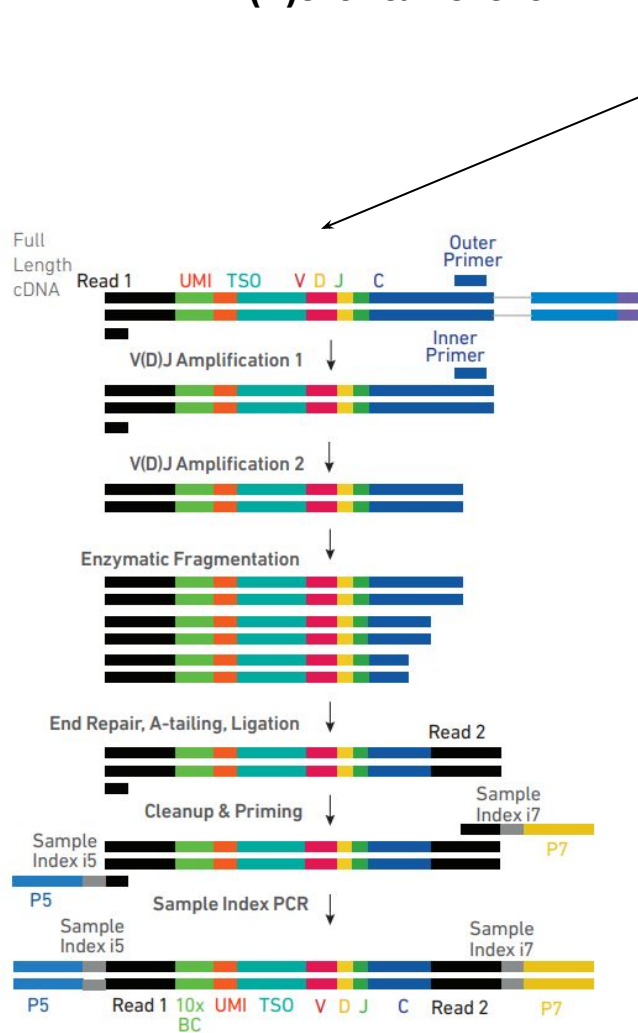
Обогащение 5'-концов транскрипта необходимо тогда, когда нам нужно дополнительно обогатить библиотеку V(D)J-фрагментами для определения Т- или В-клеточных рецепторов иммунных клеток



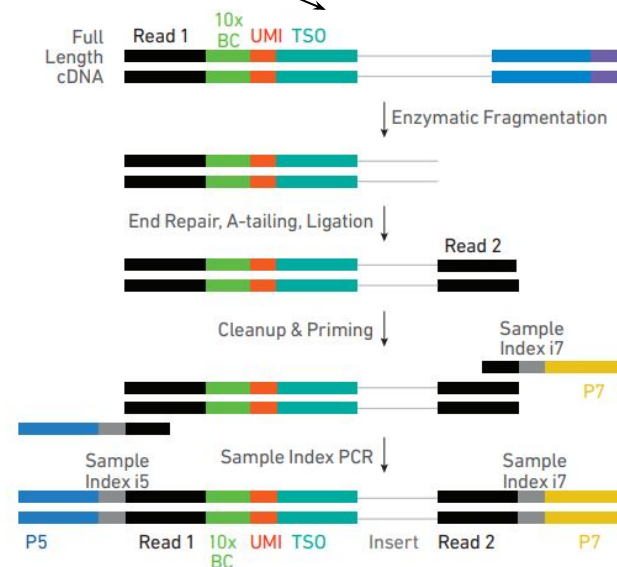
*Simone et al., 2018*

*Abugessaisa et al., 2019*

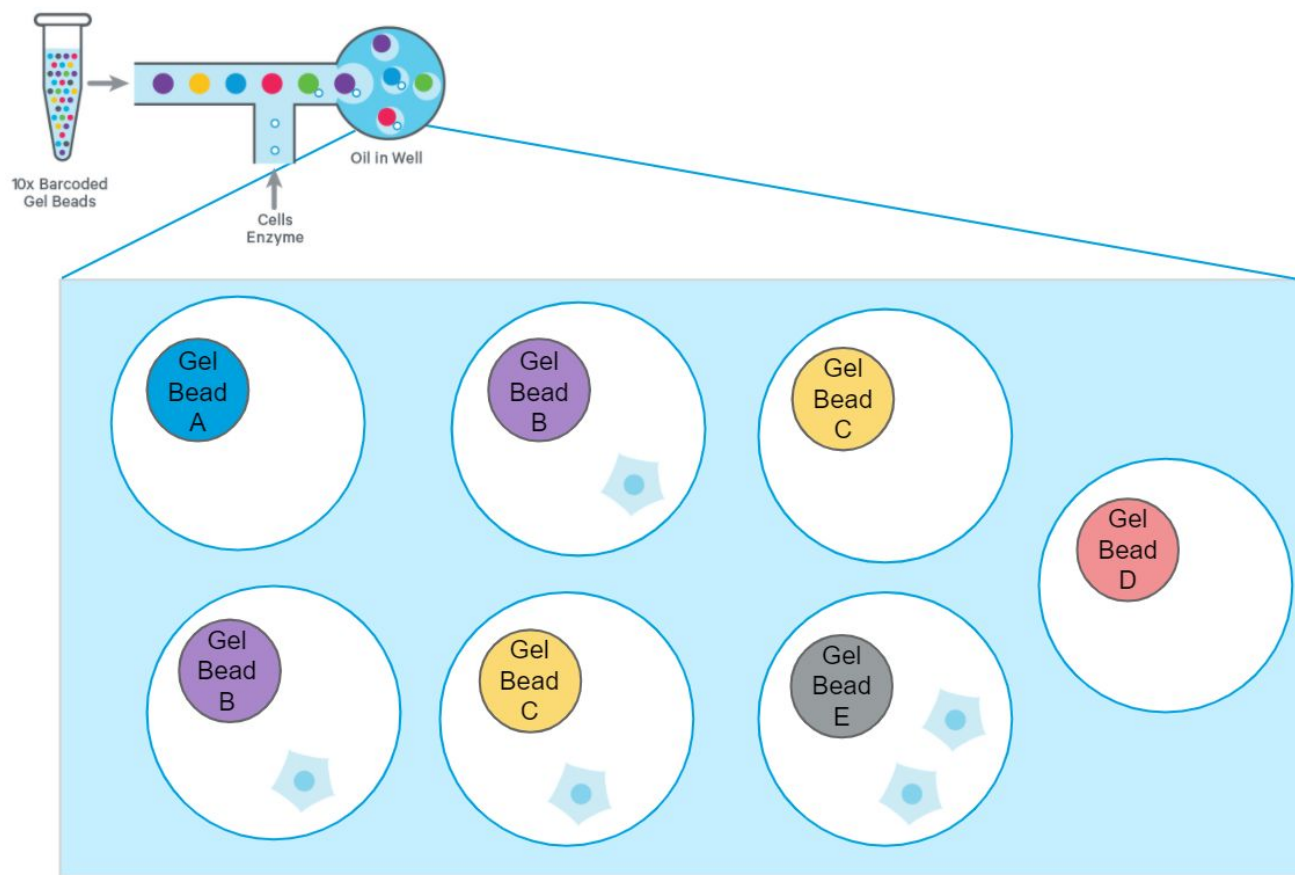
## V(D)J-библиотека



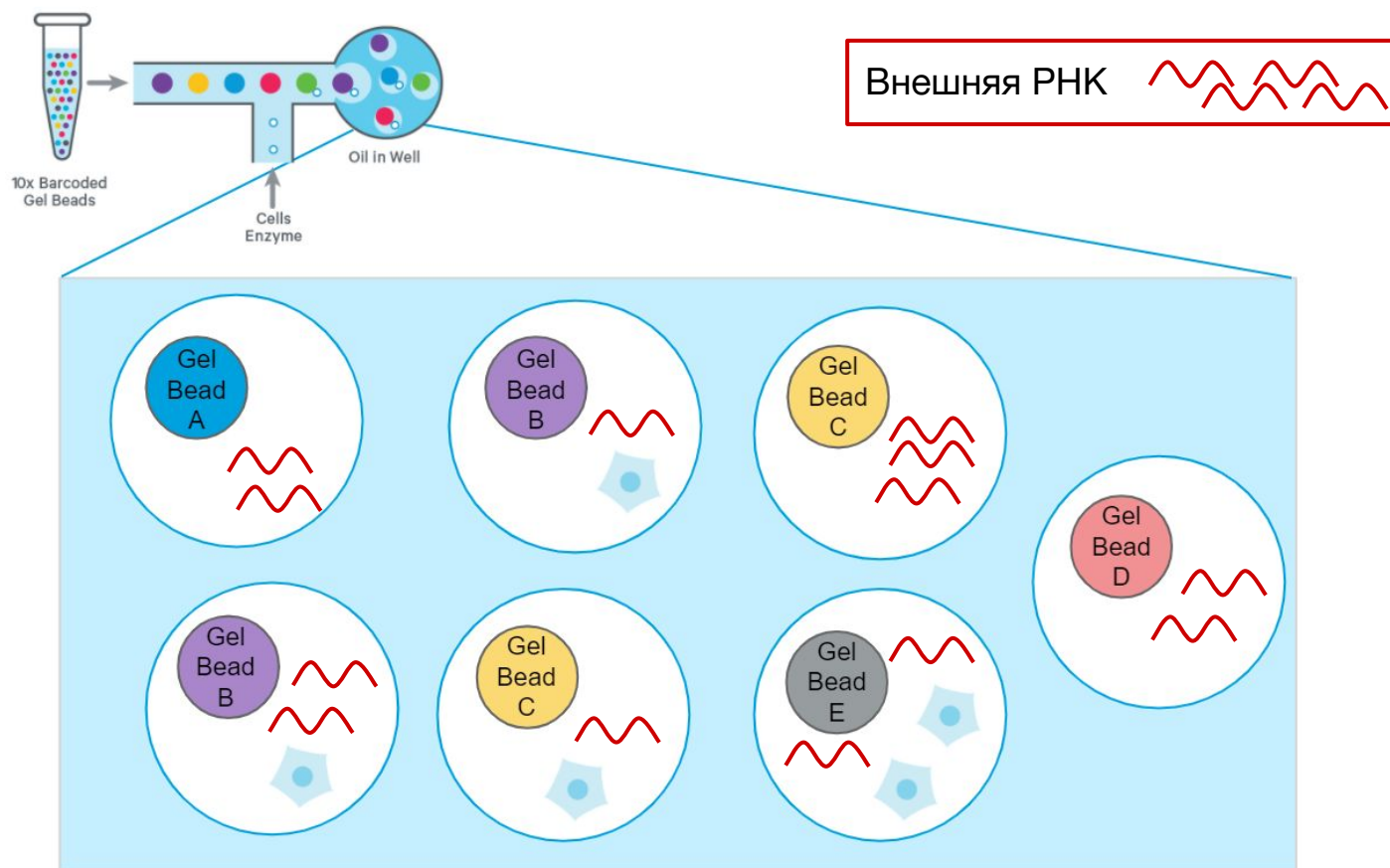
## 5'-GEX библиотека



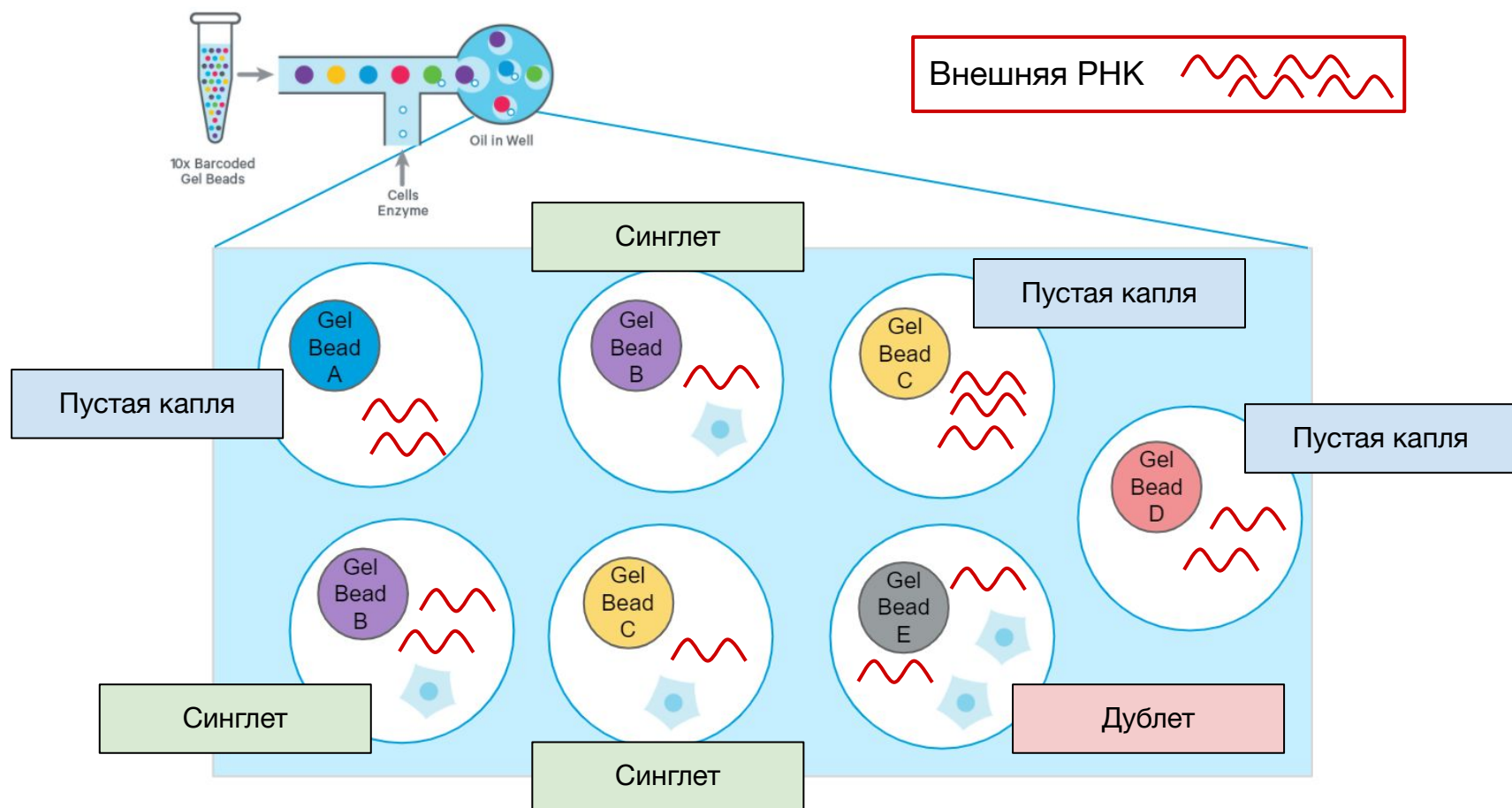
# Пустые капли и дубликаты



# Пустые капли и дублиеты



# Пустые капли и дублиеты

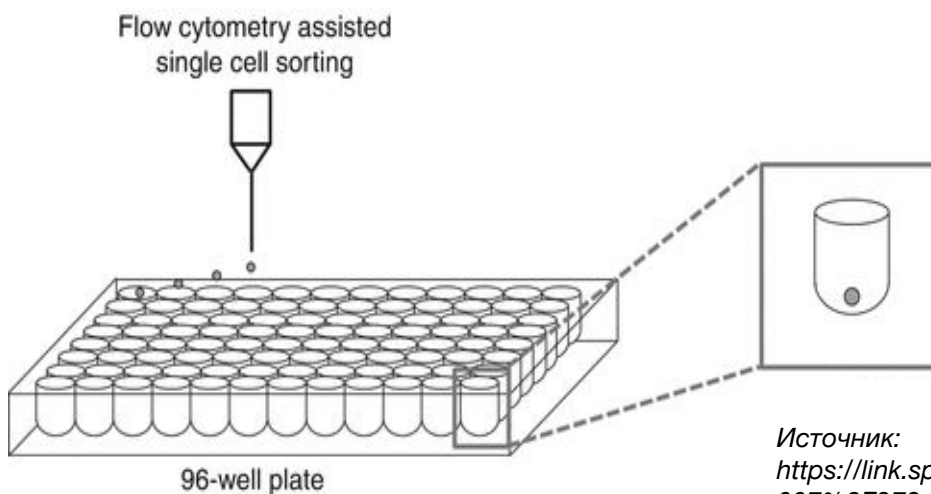


**Smart-seq**

# Разделение клеток на лунки при помощи сортера

Существует ряд методов, которые требуют изолирования одиночных клеток при помощи клеточного сортера (например, методы класса Smart-seq)

Главная проблема таких методов в ограниченном числе клеток, которые можно отсеквенировать за раз

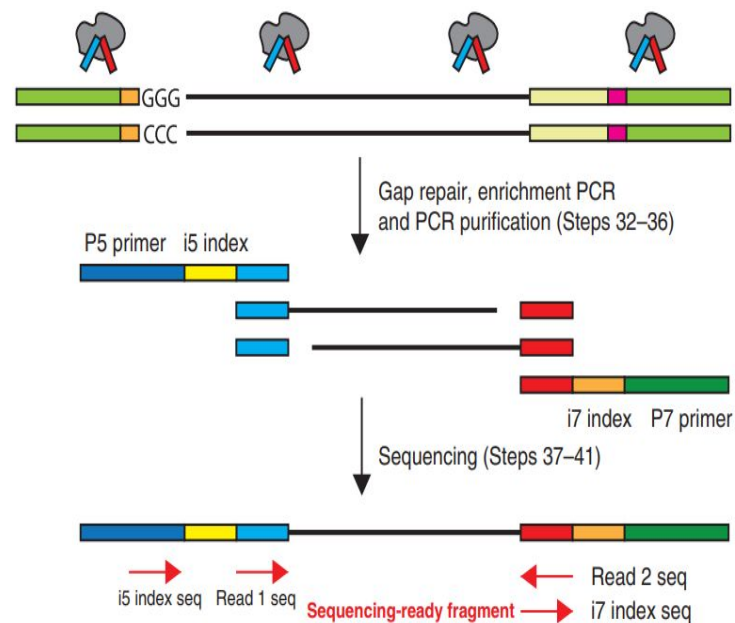
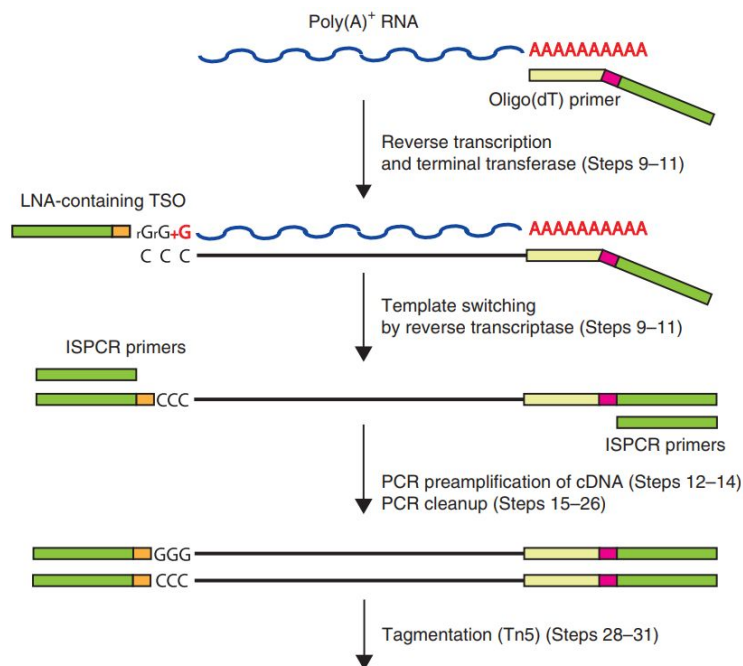


Источник:  
[https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-9863-7\\_625](https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-9863-7_625)

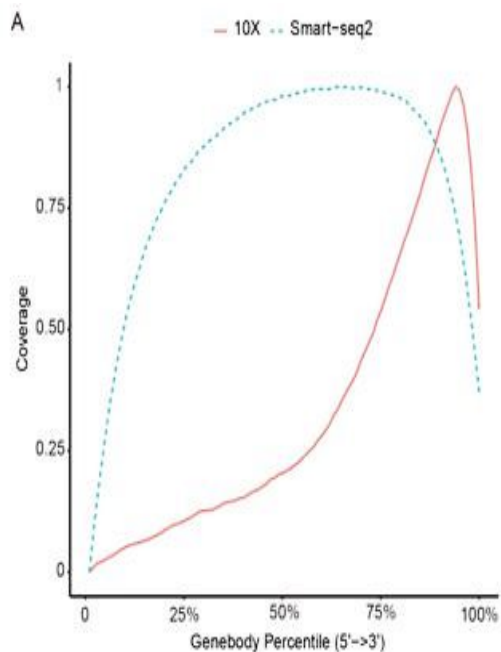


# Smart-seq2

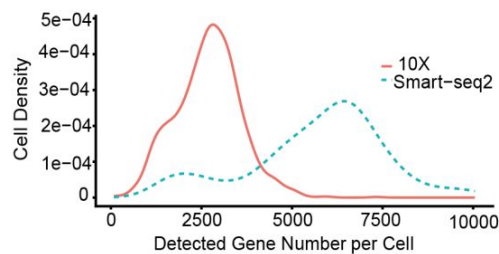
Контроль за ПЦР-дубликатами не ведётся (**отсутствуют UMI**), а идентичность клеток определяется образец-специфичными адаптерами для секвенирования (отдельный на каждую лунку)



# Smart-seq2 vs. 10x Chromium GEX

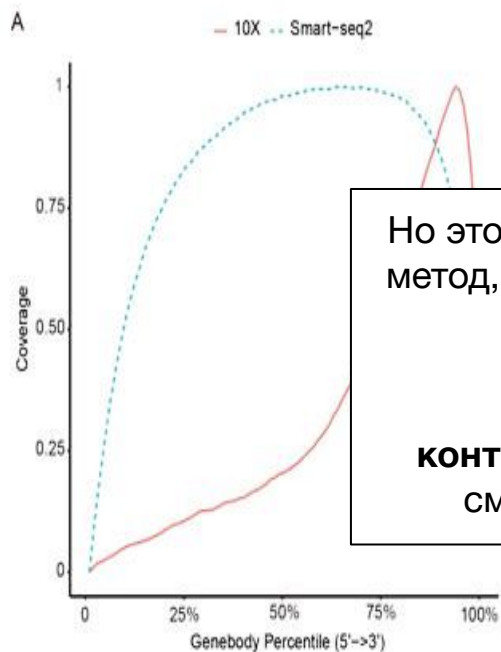


Smart-seq2 позволяет достичь более равномерного покрытия, чем 10x Chromium



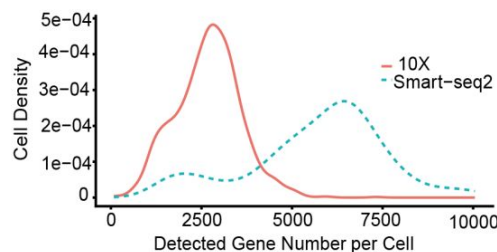
Smart-seq2 охватывает экспрессию большего числа генов, чем 10x Chromium

# Smart-seq2 vs. 10x Chromium GEX



Но это гораздо более **дорогой** и **трудоёмкий** метод, поэтому чаще вместо него используют 10x Chromium

Также Smart-seq2 **не позволяет контролировать ПЦР-дубликаты** и часто смещает анализируемые экспрессии

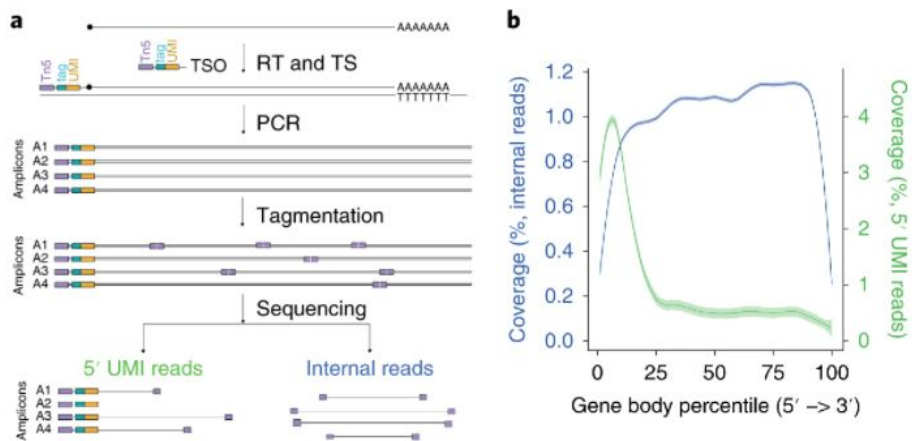


Smart-seq2 охватывает экспрессию большего числа генов, чем 10x Chromium

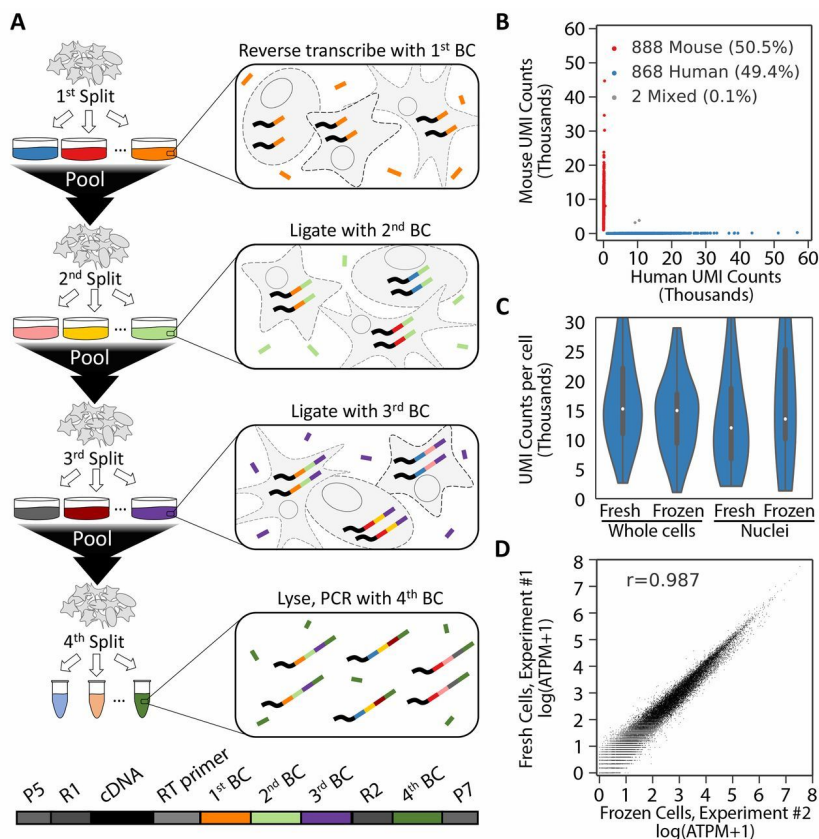
# Smart-seq3

Добавлены UMI, перед которыми содержится специальный тег, который позволяет распознать последовательность UMI

Как итог мы имеем как контроль за ПЦР-дубликатами, так и полное покрытие “внутренними” ридами, что важно, например, в случае, когда нам хочется изучить мутации



# Parse Biosciences Evercode (ex SPLiT-Seq)



PRODUCTS ▼

TECHNOLOGY

## Introducing The Parse Single Cell Whole Transcriptome Solution



The Parse Single Cell Whole Transcriptome Kit is the most scalable single cell RNA-seq solution on the market, allowing you to profile 100,000 cells and 48 samples together in one run.

LEARN MORE

REQUEST A QUOTE