



*Объединенный институт ядерных исследований
Лаборатория ядерных проблем, НХП, «Отдел фазотрона»
Группа молекулярной и радиационной генетики*

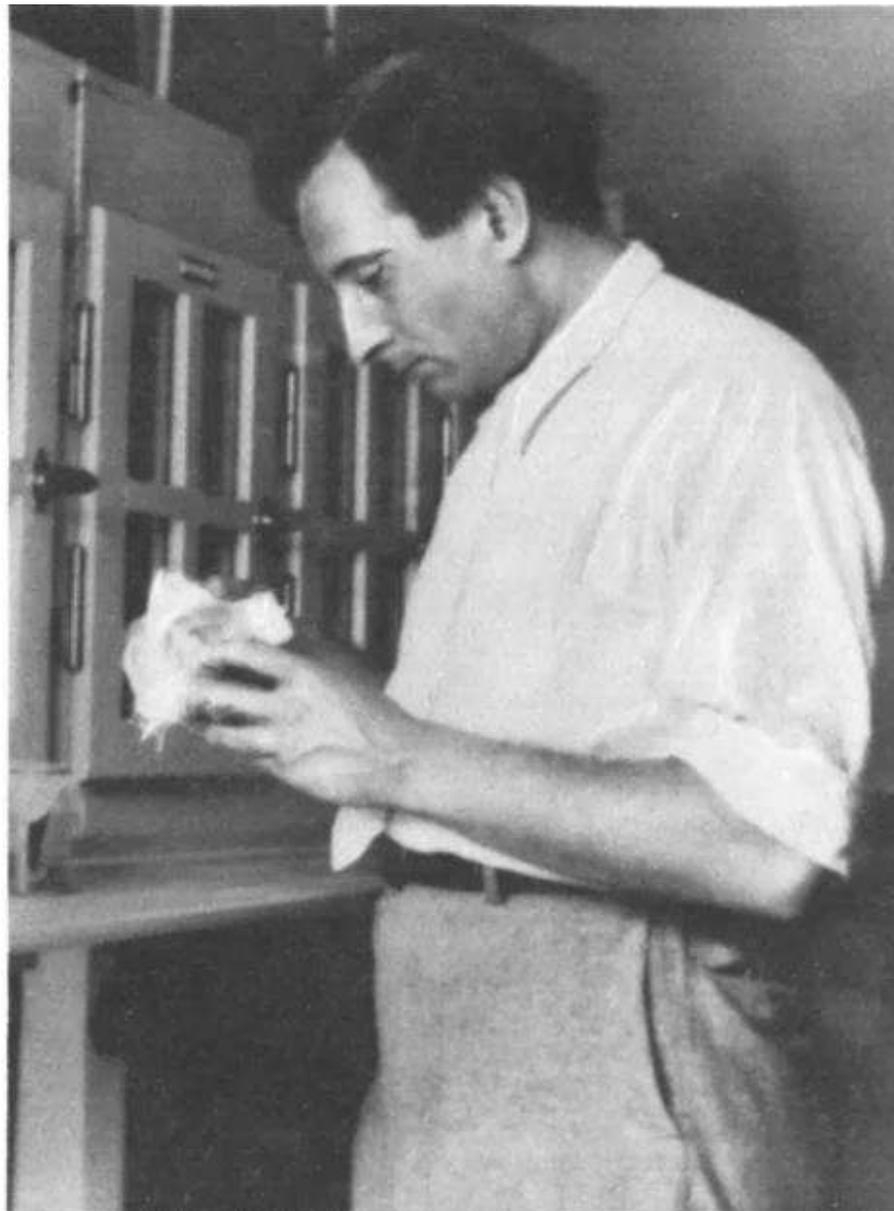
От генетических исследований к полногеномному секвенированию нового поколения в радиационной генетике.

И.Д. Александров А.Н. Русакович

2022 г.



Н.В. Тимофеев-Ресовский, 1968г. СССР



Н.В. Тимофеев-Ресовский, 1935г. Германия

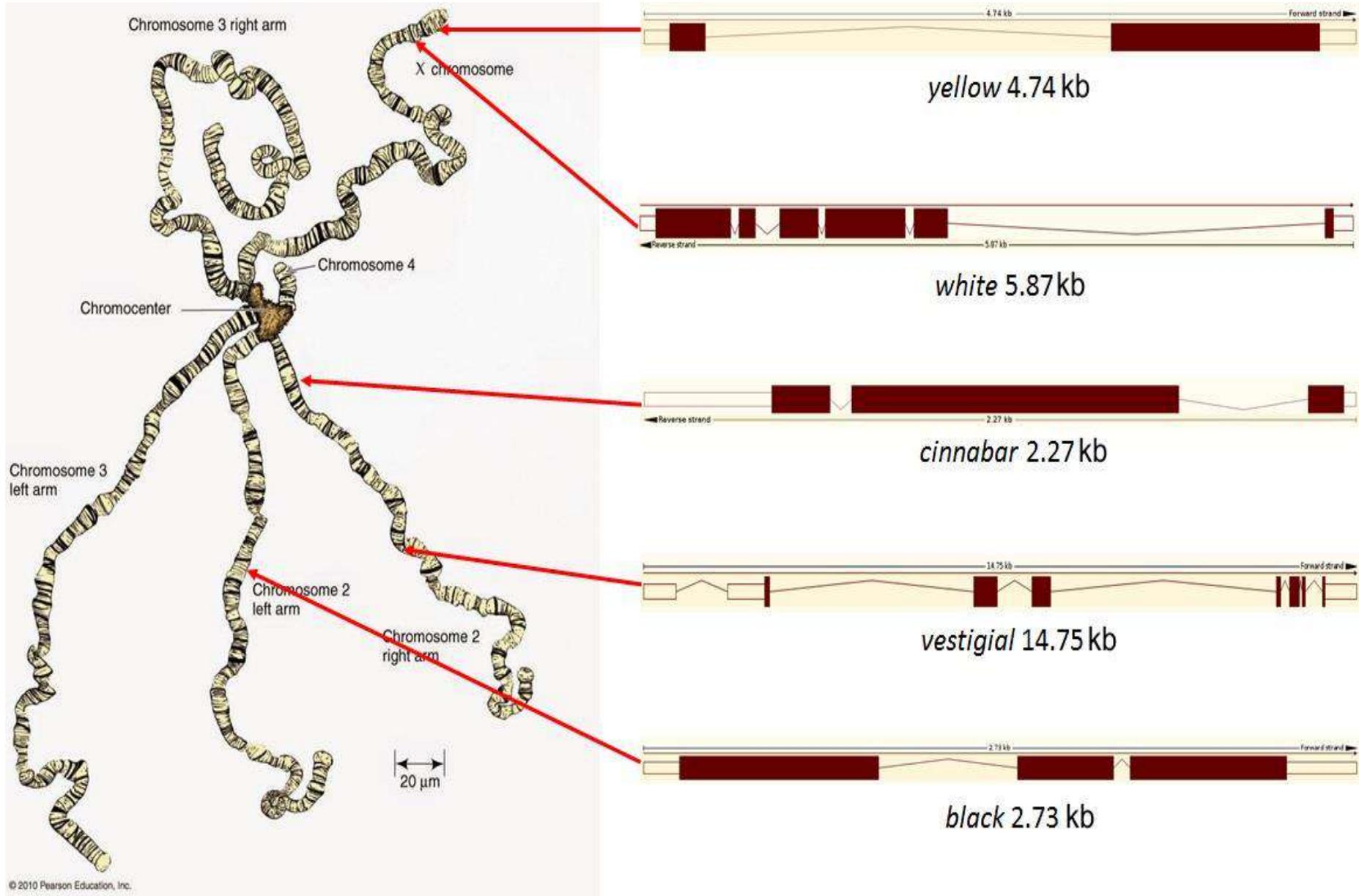
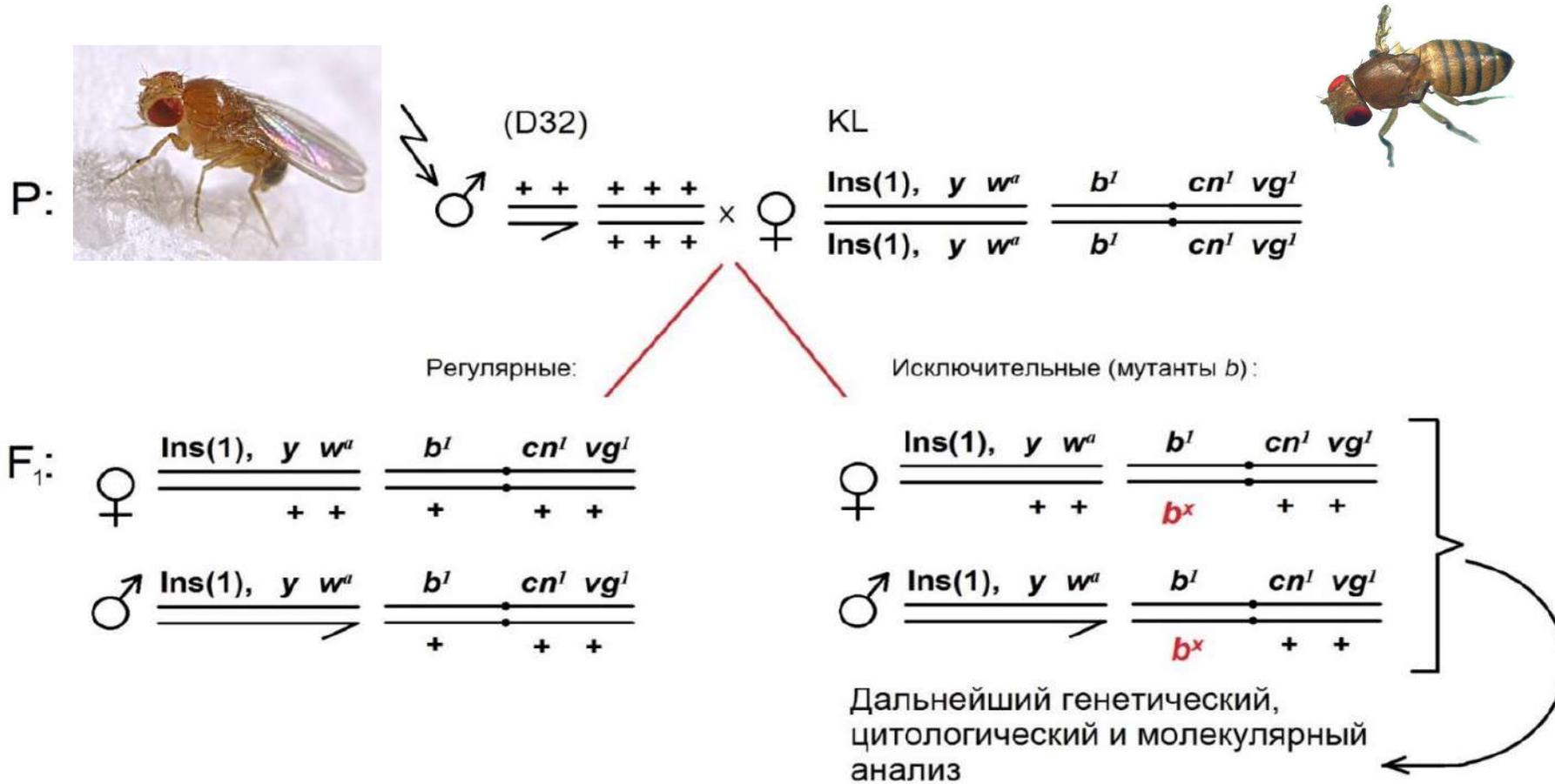
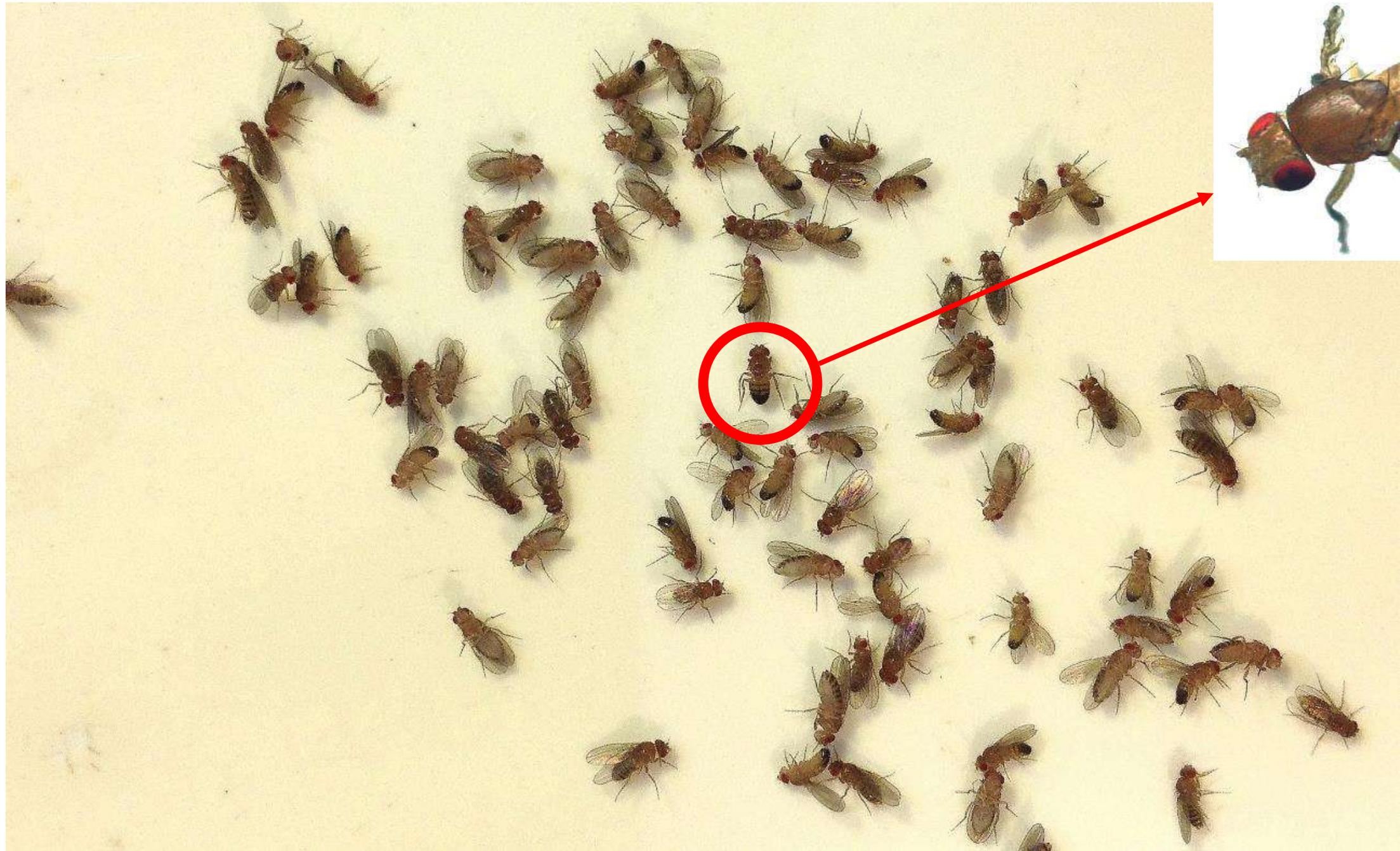
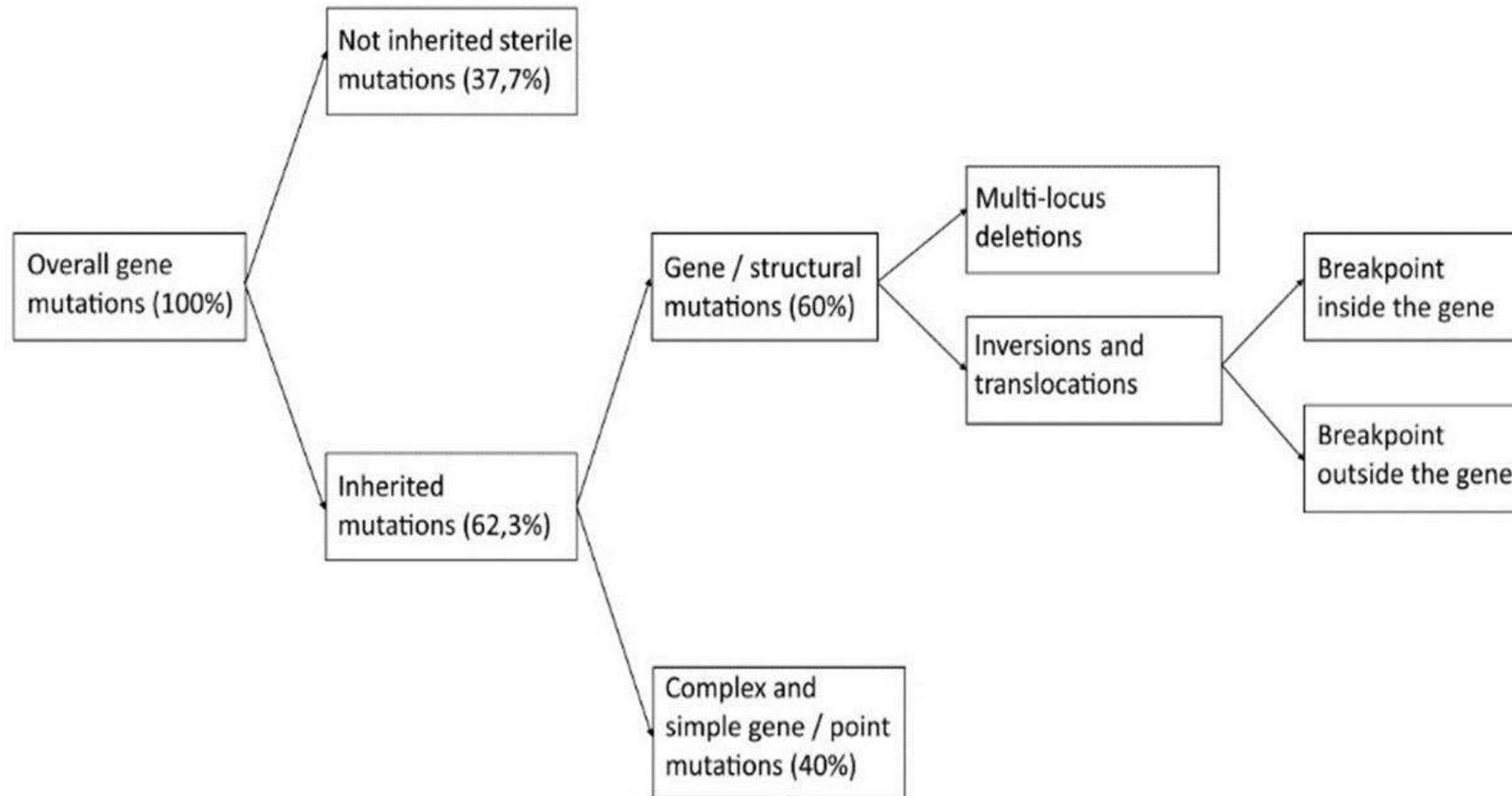


Схема получения мутаций





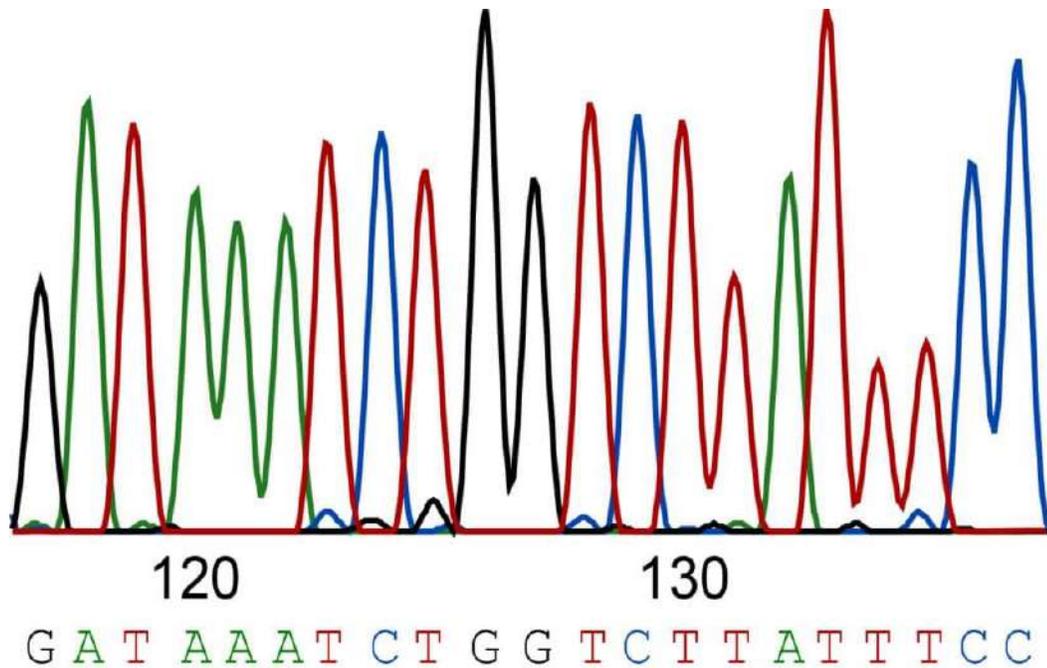
Природа и относительная частота наследуемых мутаций пяти изучаемых генов, индуцированных γ -излучением и нейтронами



Варианты секвенирования

«Классический»

- Секвенирование по Сенгеру



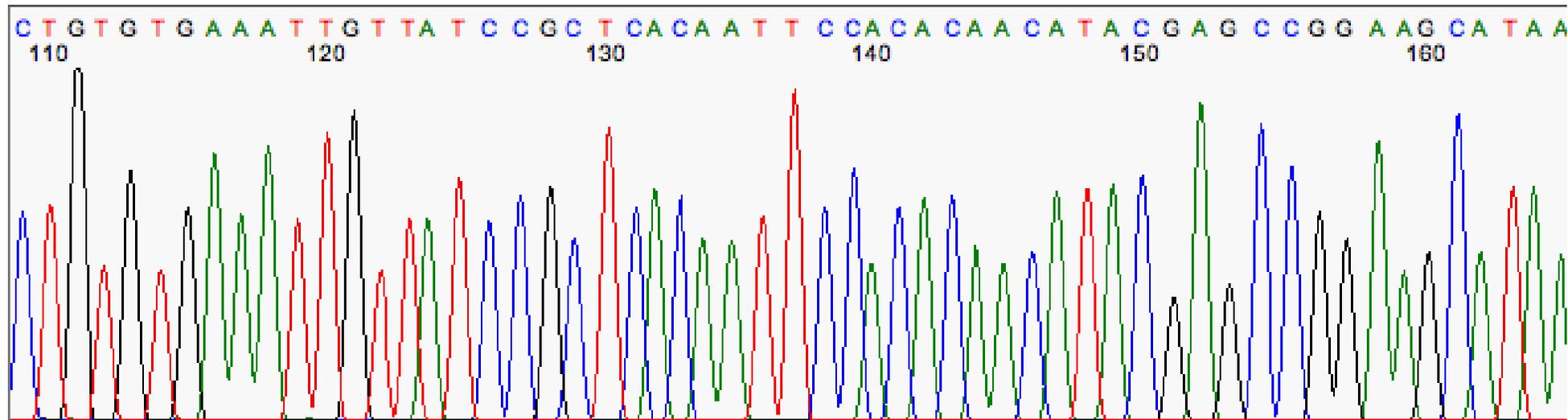
Next-generation sequencing (NGS)

- Illumina
- Nanoball seq.
- Ion-Torrent
- Nanopore

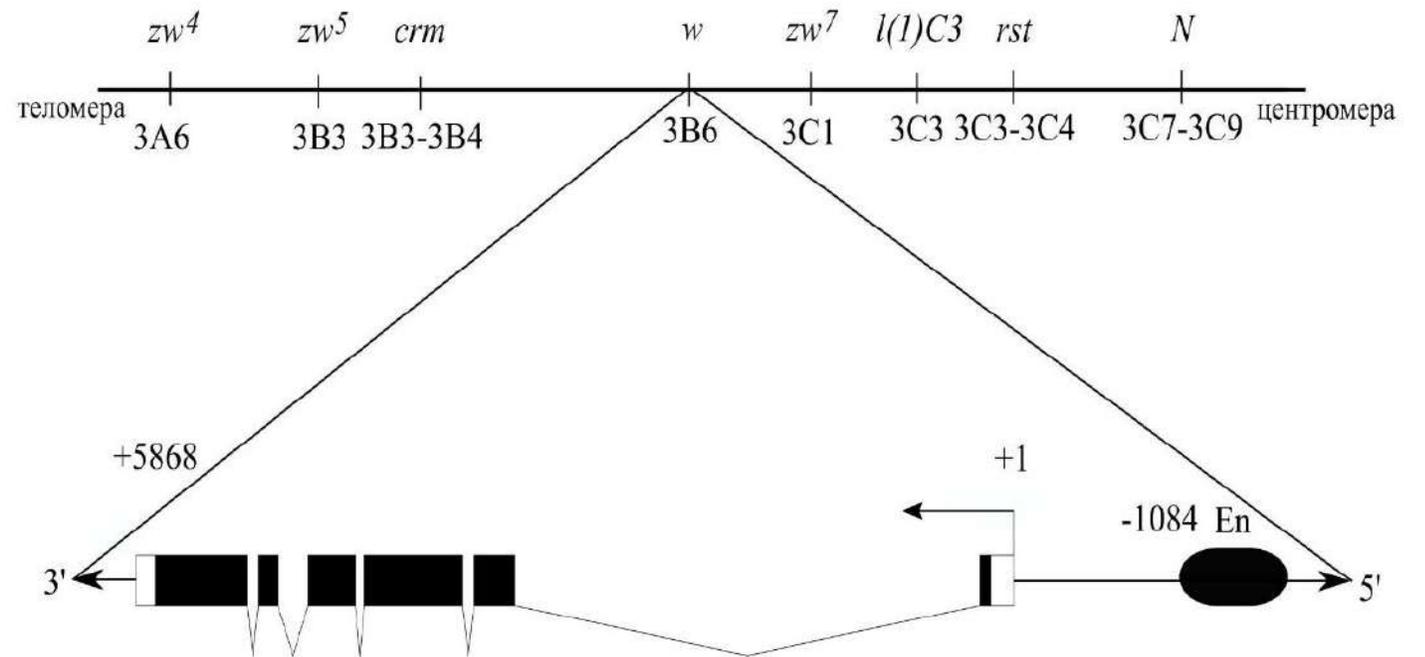


Секвенирование по Сенгеру

- Высокая надежность
- Дешевизна в отношении небольших участков
- Длина прочтений
- Сложность при покрытии неизвестных участков

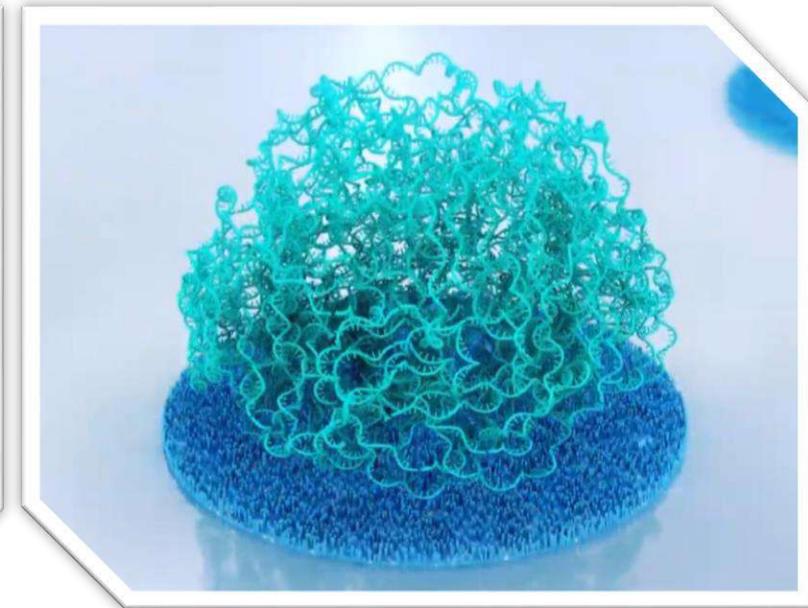
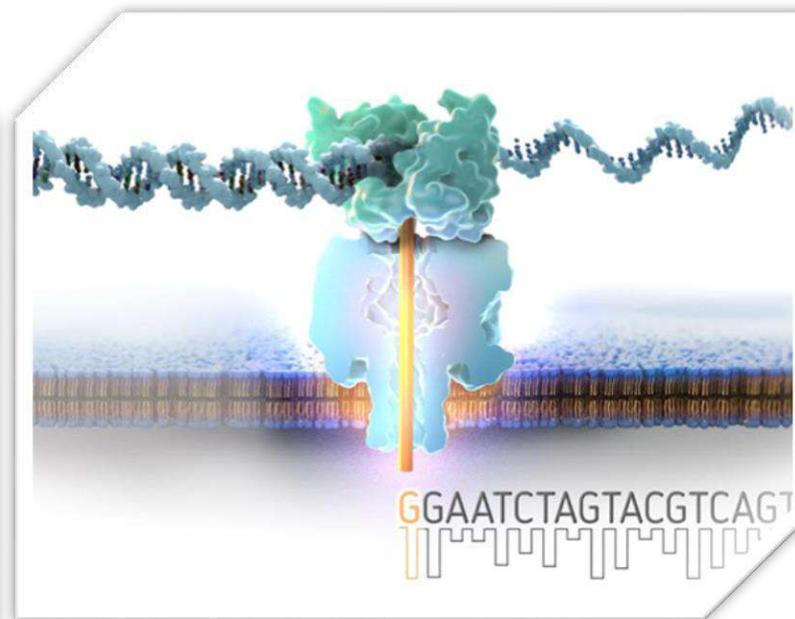
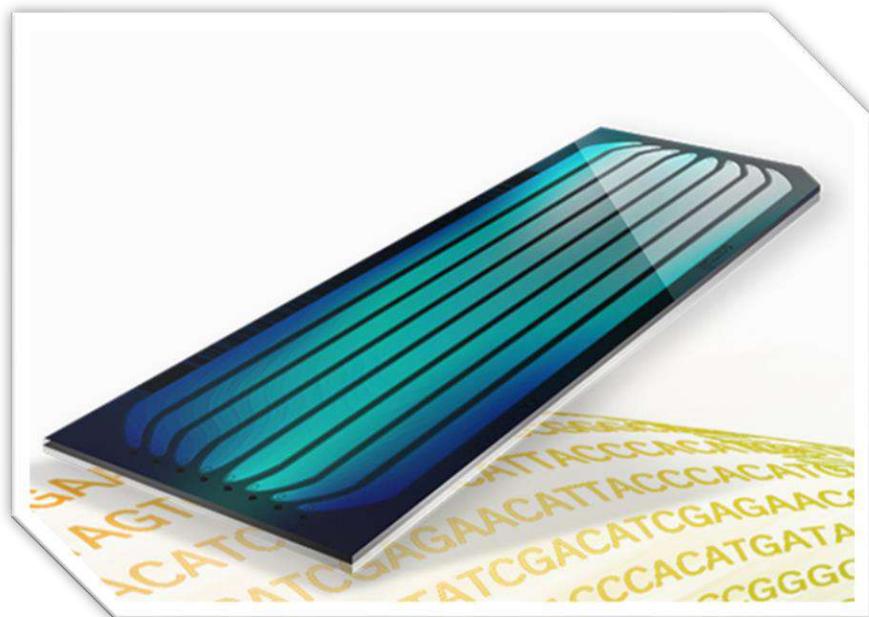
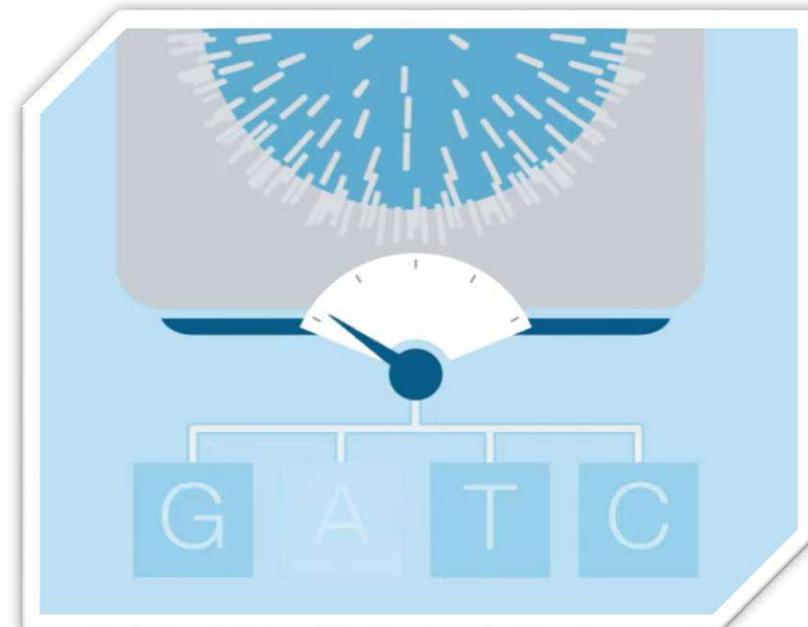


- Специфичность участка
- Необходимость в подборе праймеров
- Невозможность покрыть много участков разом



NGS

- Сниженная надежность отдельного прочтения
- Множественность прочтений
- Доступность геномного покрытия
- Различная цена



Особенности разных видов NGS

Ion-Torrent

Относительно дешевый, быстрый, проблемы с участками с повтором одного основания

Nanoball sequencing

В большой степени устраняет ошибки копирования ДНК, формирует короткие прочтения

Nanopore

Дешевый, возможно прочтение крайне протяженных участков, точность несколько ниже других методов

Illumina

Относительно дешевый, многокopiesность прочтений, формирует короткие прочтения

Общая схема работы с Illumina

- Получение ДНК
- Подготовка ДНК к секвенированию
- Непосредственно секвенирование
- Обработка сырых данных
- Интерпретация результатов

Сырые данные

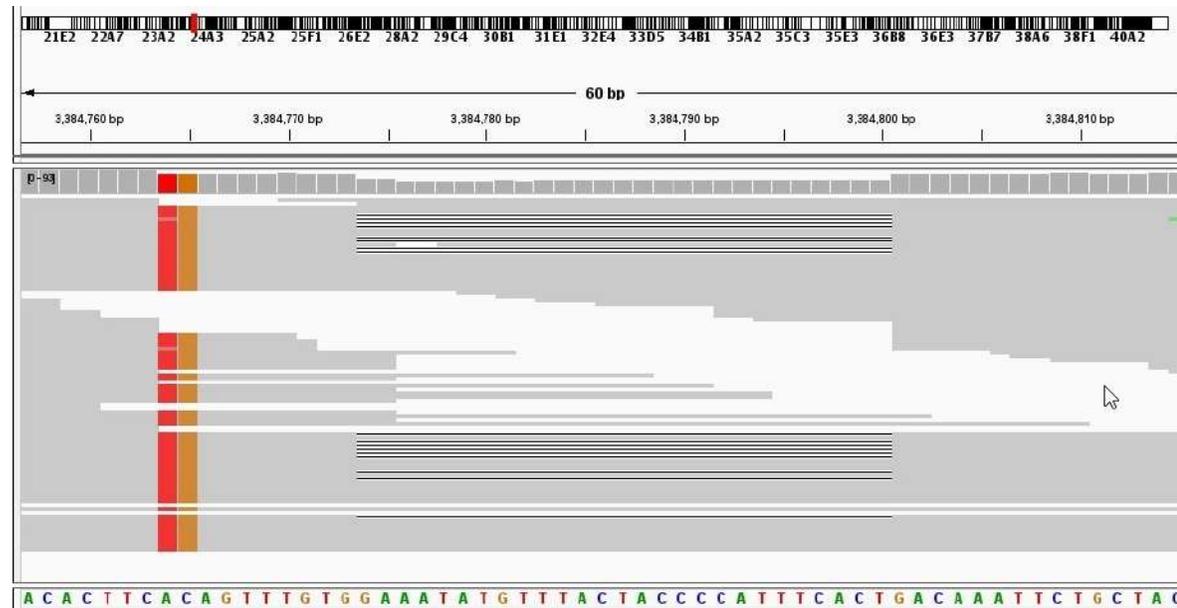
```
Identifier — ● @SRR566546.970 HWUSI-EAS1673_11067_FC7070M:4:1:2299:1109 length=50
Sequence — ● TTGCCTGCCTATCATTTTAGTGCCTGTGAGGTGGAGATGTGAGGATCAGT
'+' sign — ● +
Quality scores — ● hhhhhhhhhghghghhhhhfhhhhhfffffe'ee['X]b[d[ed'[Y[~Y
Identifier — ● @SRR566546.971 HWUSI-EAS1673_11067_FC7070M:4:1:2374:1108 length=50
Sequence — ● GATTTGTATGAAAGTATACAACATAAACTGCAGGTGGATCAGAGTAAGTC
'+' sign — ● +
Quality scores — ● hhhhgfhhcghghggfcffdhfehhhhcehdchhdhahehffffde'bVd
```

Обработка сырых данных

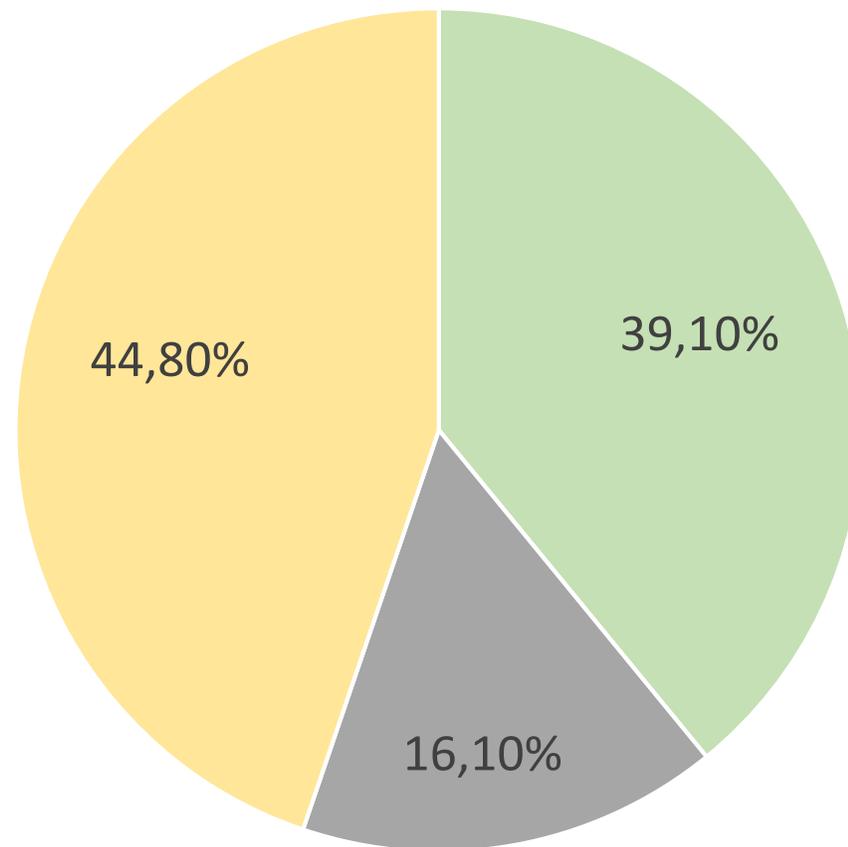
- Картирование прочтений
- Постобработка картирования
- Маркировка ПЦР дубликатов
- Анализ покрытия
- Поиск генетических вариантов
- Анализ найденных вариантов

Анализ обработанных данных

- Определяется экспериментом
- Определяется использованной платформой
- IGV
- Excel



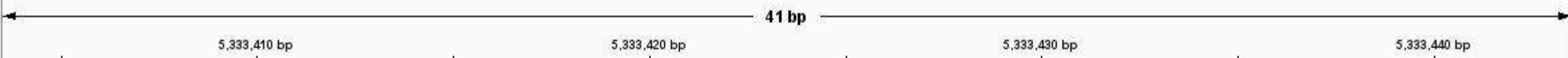
296629	A		80	OK
1.9E+07	G		543	OK
5333407	G		35	OK
1.9E+07	C		101	Вероятен артефакт (слонный для сб)
1.5E+07	T		478	OK
2.4E+07	T	<DUP>	456	OK
2.9E+07	G	[2R:12257854]		Реципрокная транслокация (двойная)
2.9E+07	C	C[2R:12257852]		Реципрокная транслокация (двойная)
9374390	G	<DUP>	####	Полиморфный участок или артефакт
1728330	A	<INV>	358	Мозаицизм или артефакт
5194197	A	<INV>	460	Вероятен артефакт (слонный для сб)
1E+07	G		1731	OK, возможен мозаицизм
2242025	C	C[2R:11959807]		Реципрокная транслокация (двойная)
2242041	T	[2R:11959524]		Реципрокная транслокация (двойная)
7204686	A		106	OK
7204767	G		36	OK
8763945	T	<INS>	28	Вероятен артефакт (слонный для сб)
8844066	A		335	OK
8861575	T		40	Вероятен артефакт (слонный для сб)
8875238	G		40	Полиморфный участок или артефакт
8902080	T	<INS>	38	Вероятен артефакт (слонный для сб)
9009997	A		53	Полиморфный участок или артефакт
9011175	A		36	Полиморфный участок или артефакт
2.4E+07	C	<INV>	302	Вероятен артефакт (слонный для сб)
2.5E+07	C	<INV>	454	Мозаицизм или артефакт
1455277	C	<INV>	823	Мозаицизм или артефакт
3023892	T	<INV>	627	Мозаицизм или артефакт
2.1E+07	C		115	OK
2.1E+07	C		37	OK
357117	A	<INV>	384	Полиморфный участок или артефакт
1221506	G	<INV>	485	Мозаицизм или артефакт
2.3E+07	T	<INV>	668	Мозаицизм или артефакт
5786664	G		30	Мозаицизм или артефакт



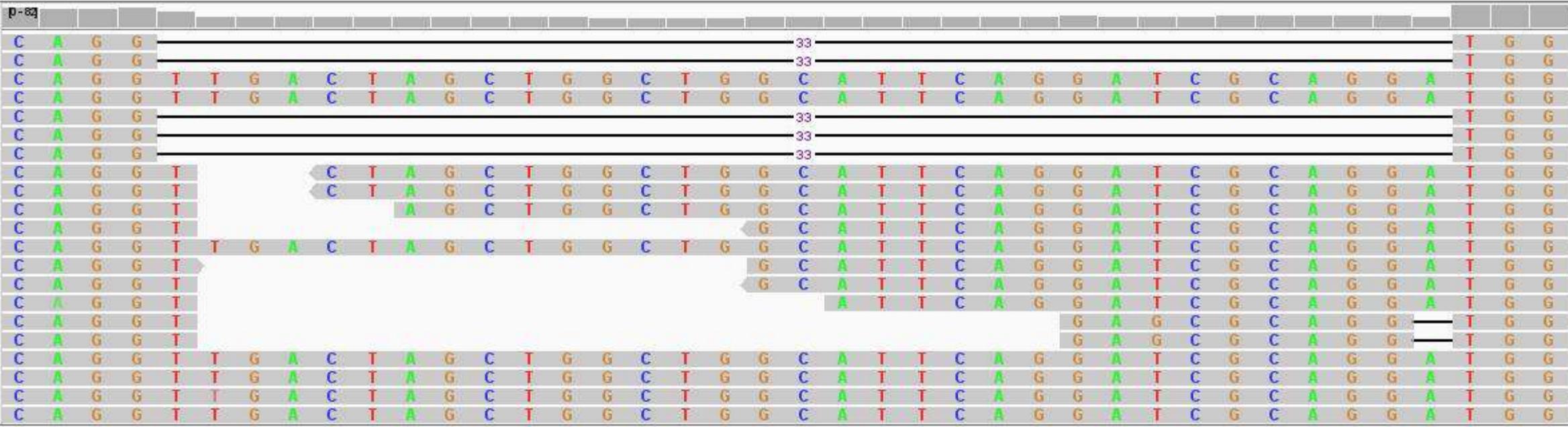
Полные (■), мозаичные (■) и не идентифицированные (■) изменения ДНК выявляемые биоинформационным анализом после полногеномного секвенирования.

Примеры нюансов анализа

61C6 62A4 62E2 63D3 64B2 64C13 65B1 65F1 66C1 67A1 67D8 68A3 68F1 69F2 70C5 70F7 71F2 73B2 74D5 75D2 76B4 77B7 78C6 79D3 80E1



verage



→

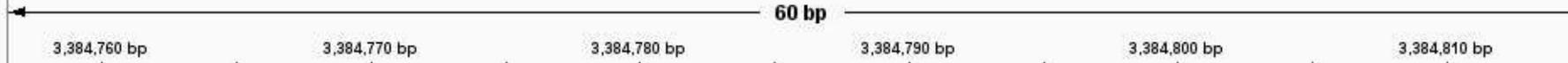
C A G G T T G A C T A G C T G G C A T T C A G G A T C G C A G G A T G G

D. melanogaster (dm3)

chr2L

chr2L:3,384,757-3,384,815

Go



2_GGACTCCT-...am Coverage



2_GGACTCCT-AGAGGATA_ALM



Sequence

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

UCSC Refseq Genes

D. melanogaster (dm3)

chr2L

chr2L:3,333,511-3,333,983

Go



473 bp

3,333,600 bp

3,333,700 bp

3,333,800 bp

3,333,900 bp

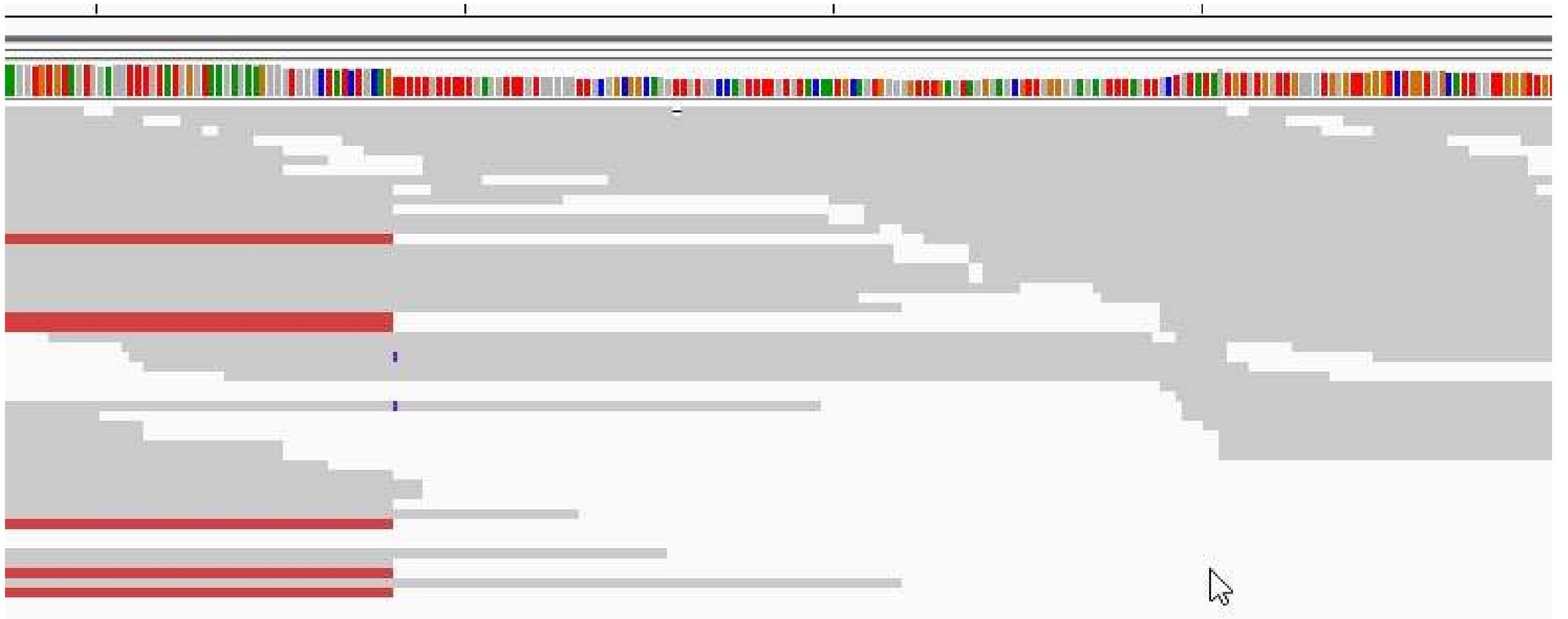
2_GGACTCCT-...am Coverage



Sequence

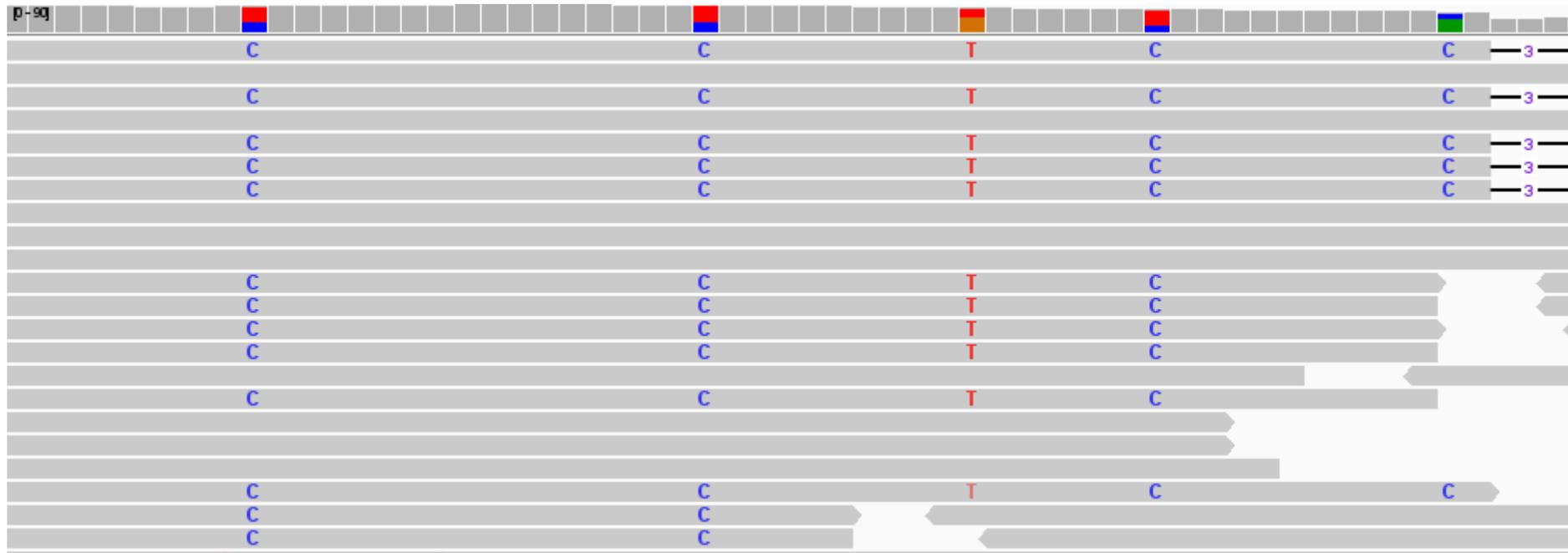
UCSC Refseq Genes

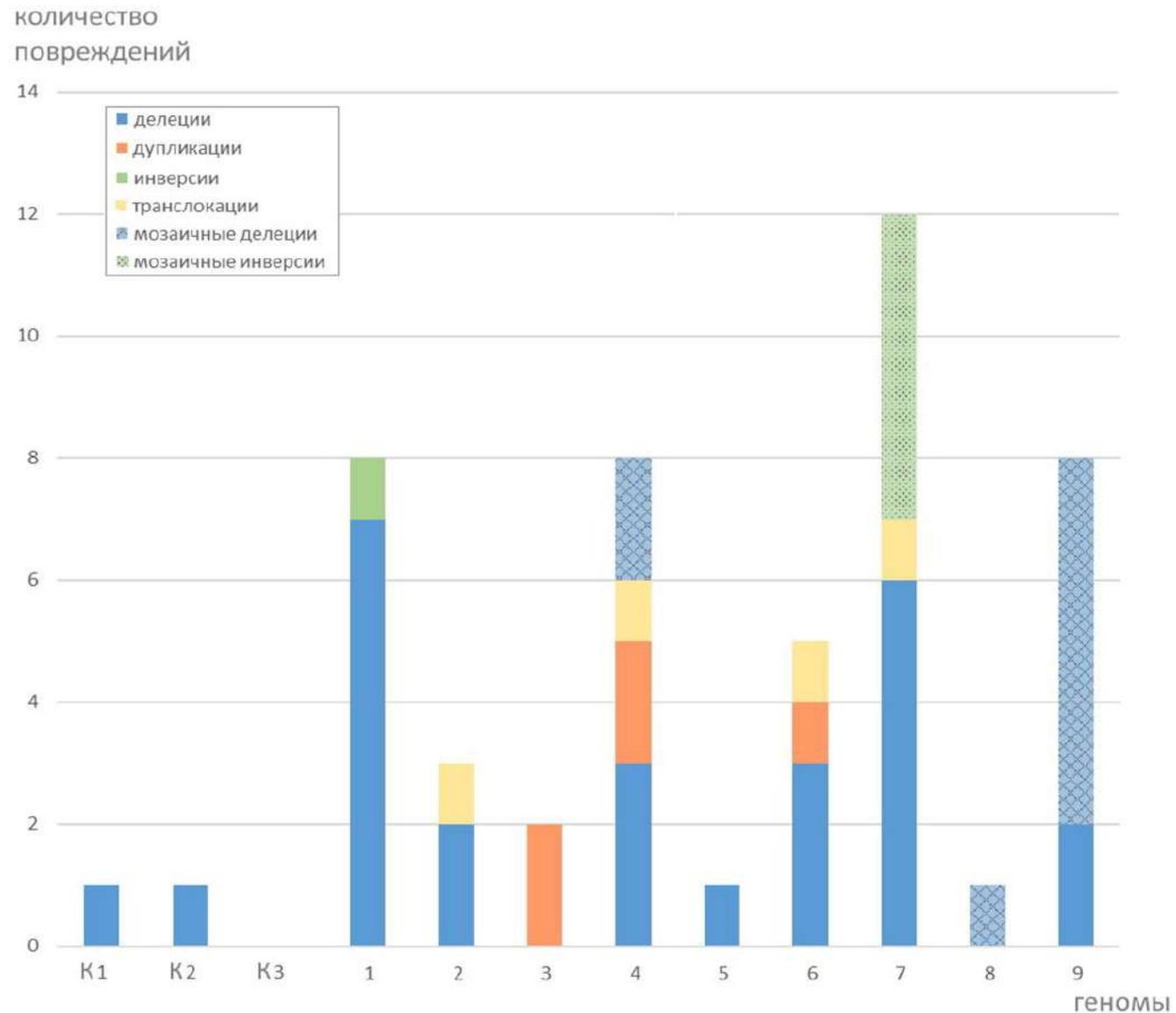
CG15408



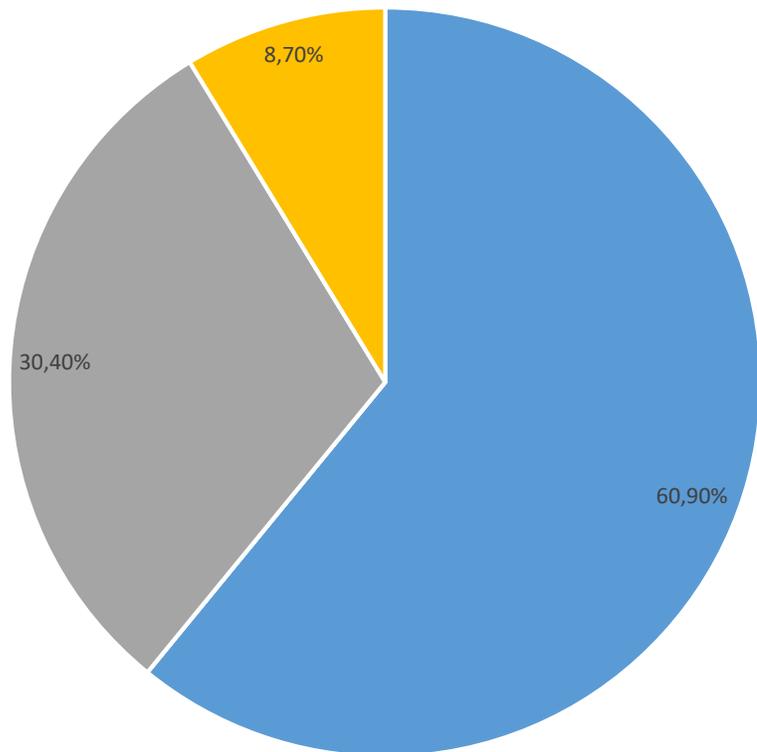
Дополнения к уже произведенному анализу

- Анализ однонуклеотидных изменений
- Другой подход к обработке сырых данных
- Другие методы секвенирования

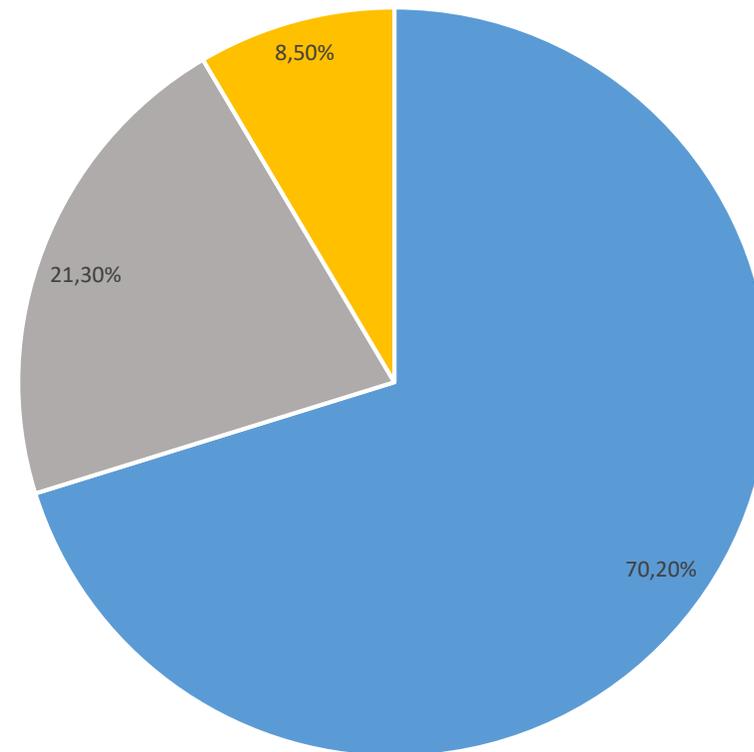




Количество и типы геномных изменений у потомков от контрольных (K1-K3) и γ -облученных (40 Гр, Co^{60}) самцов *D. melanogaster*



А



Б

Относительная частота идентичных изменений (делеций (■), инверсий (■), транслокаций (■)), выявляемая классическим методом анализа мутаций отдельных генов (А) и полногеномного секвенирования нового поколения (Б)



Спасибо за внимание!