Биоинженерия

Лекция 4

ПЦР в случаях, когда известна неполная последовательность фрагмента ДНК

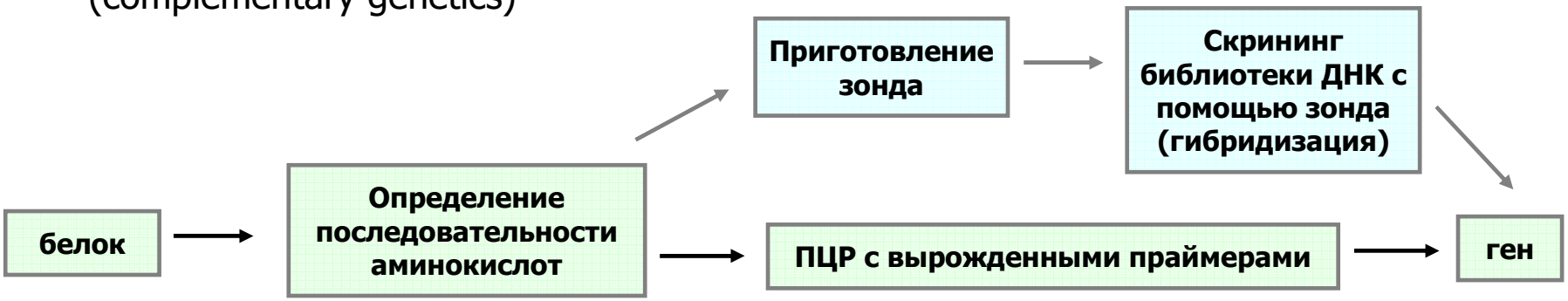
Обратная транскрипция, приготовление кДНК из малых количеств материала, технологии Genome Walking

Сергей Лукьянов

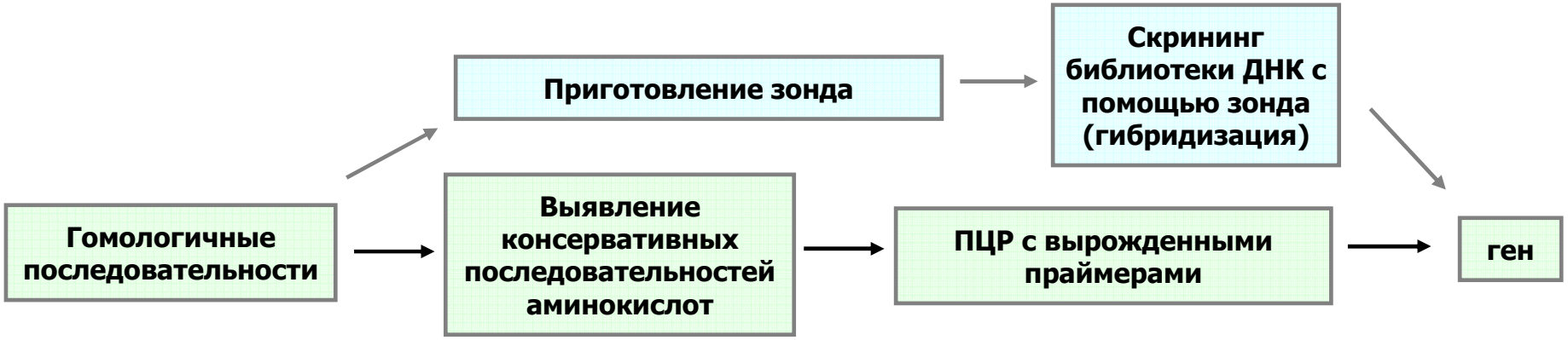
Институт биоорганической химии имени академиков М.М. Шемякина и Ю.А. Овчинникова РАН

ПЦР и клонирование генов

- Поиск генов, для которых известна частичная последовательность белка (complementary genetics)



- Поиск генов, гомологичных уже известным (homology cloning)



Перекрестная гибридизация ДНК возможна только для близких видов, гибридизация с вырожденным зондом часто низкоэффективна,

ПЦР с вырожденным праймером

НОМЕОБОХ – эволюционно консервативный участок большой группы транскрипционных факторов

```
AL_DROME        QRRYRTTFTSFQLEELEKAFSRTHYPDVFTREELAMKIGLTEARIQVWFQNRRAKWRKQEKVG
HMPB_DROME      PRRLRTAYTNTQLLELEKEFHFNKYLCRPRRIEIAASLDLTERQVKVWFQNRRMKHKRQTLSK
MOX2_XENLA      PRKERTAFTKEQIRELEAEFAHHNYLTRLRRYEIAVNLDLTERQVKVWFQNRRMKWKRVKGGQ
O62548          PNTGRTNFTNKQLTELEKEFHFNKYLTRARRIEIAAALGLNETQVKIWFQNRRMKQKKRMKEG
ANF1_CHICK      GRRPRTAFTRNQIEVLENVFKMNSYPGIDIREELARKLDLEEDRIQIWFQNRRAKLKRSHRES
AAD15943        GRKKRCPYSKHQILELEKEFLFNMYLTRERRLEISRSINLTDRQVKIWFQNRRMKLKKMTREH
CDX4_HUMAN      KEKYRVVYTDHQRLELEKEFHCNRYITIQRKSELAVNLGLSERQVKIWFQNRRAKERKMIKKK
Q23818          SKRIRTAYTSIQLLELEKEFQNNRYLSRLRRIQIAAILDLTEKQVKIWFQNRRVKWKKDKKGY
CAA10306        GRRLRTAFTSDQISTLEKTFQKHRYLGASERRKLAAKLQLSEVQIKTWFQNRRMKHKREIQDG
MAB5_CAEEL      SKRTRQTYSRSQTLELEKEFHYHKYLTRKRRQEISETLHLTERQVKIWFQNRRMKHKKEAKGE
```

E L E K/N E F W F Q N R R

ПЦР с вырожденным праймером

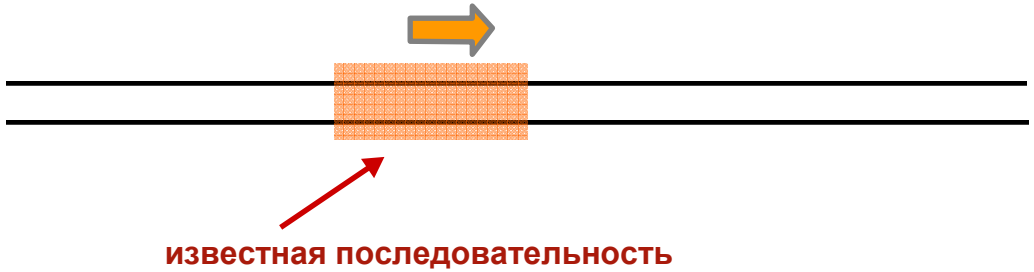
Second base

First base	U	C	A	G
U	UUU, UUC — Phenylalanine F UUA, UUG — Leucine L	UCU, UCC, UCA, UCG — Serine S	UAU, UAC — Tyrosine Y UAA Stop codon UAG Stop codon	UGU, UGC — Cysteine C UGA Stop codon UGG Tryptophan W
C	CUU, CUC, CUA, CUG — Leucine L	CCU, CCC, CCA, CCG — Proline P	CAU, CAC — Histidine H CAA, CAG — Glutamine Q	CGU, CGC, CGA, CGG — Arginine R
A	AUU, AUC, AUA — Isoleucine I AUG Methionine start codon M	ACU, ACC, ACA, ACG — Threonine T	AAU, AAC — Asparagine N AAA, AAG — Lysine K	AGU, AGC — Serine S AGA, AGG — Arginine R
G	GUU, GUC, GUA, GUG — Valine V	GCU, GCC, GCA, GCG — Alanine A	GAU, GAC — Aspartic acid D GAA, GAG — Glutamic acid E	GGU, GGC, GGA, GGG — Glycine G

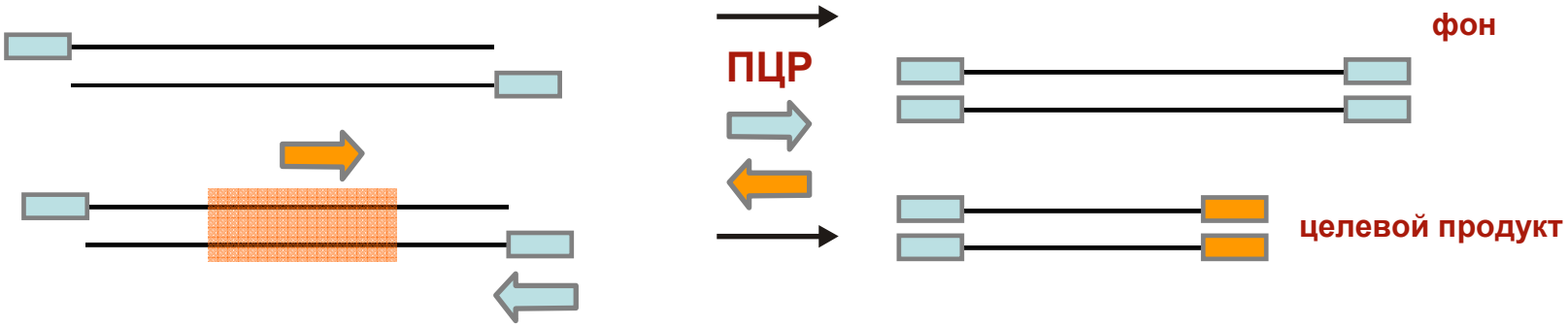
```
                          E          L   E        K/N        E      F
5'- gatcgaattc  ga(a/g)   ctx ga(a/g) aa(x)    ga(a/g) tt-3'

      W     F          Q        N        R        R
     tgg tt(t/c)  ca(a/g)  aa(t/c) (c/a)gx  (c/a)gx
3'-  acc aa(a/g)  gt(t/c)  tt(g/a) (t/g)ct  (t/g)ct cgagatgc -5'
```

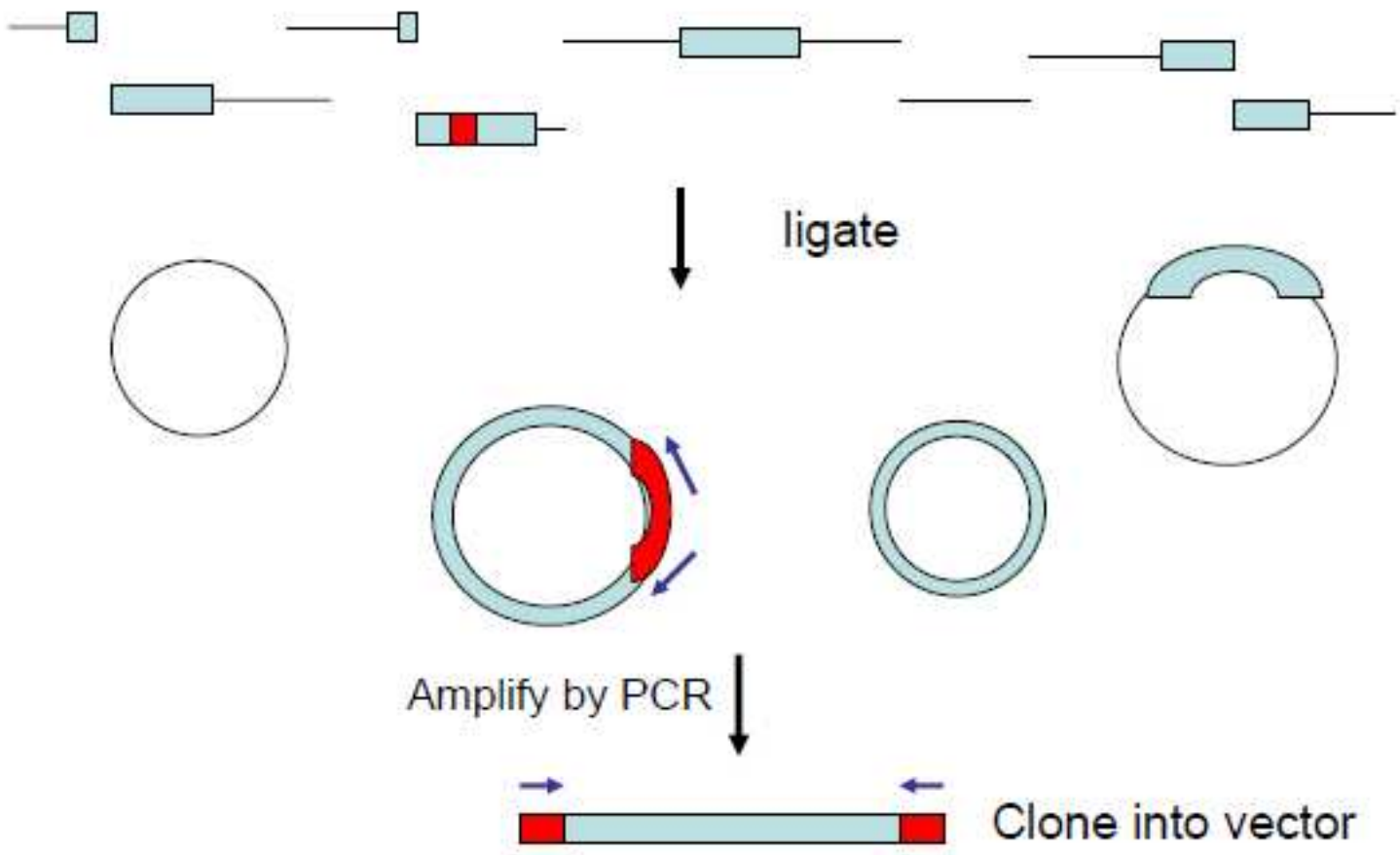
ПЦР амплификация фрагментов ДНК фланкирующих участок с известной последовательностью (клонирование промотеров, выявление точки интеграции трансгена, анализ геномных перестроек)



- Ассиметричная амплификация (с одного праймера) - линейное увеличение количества целевой последовательности + «линейный» ПЦР - крайне неэффективный метод
- Лигирование адаптеров позволяет вести амплификацию с двух праймеров, но не решает проблемы фоновой амплификации



Inverse PCR: amplification "out" from known sequence



Vectorette PCR

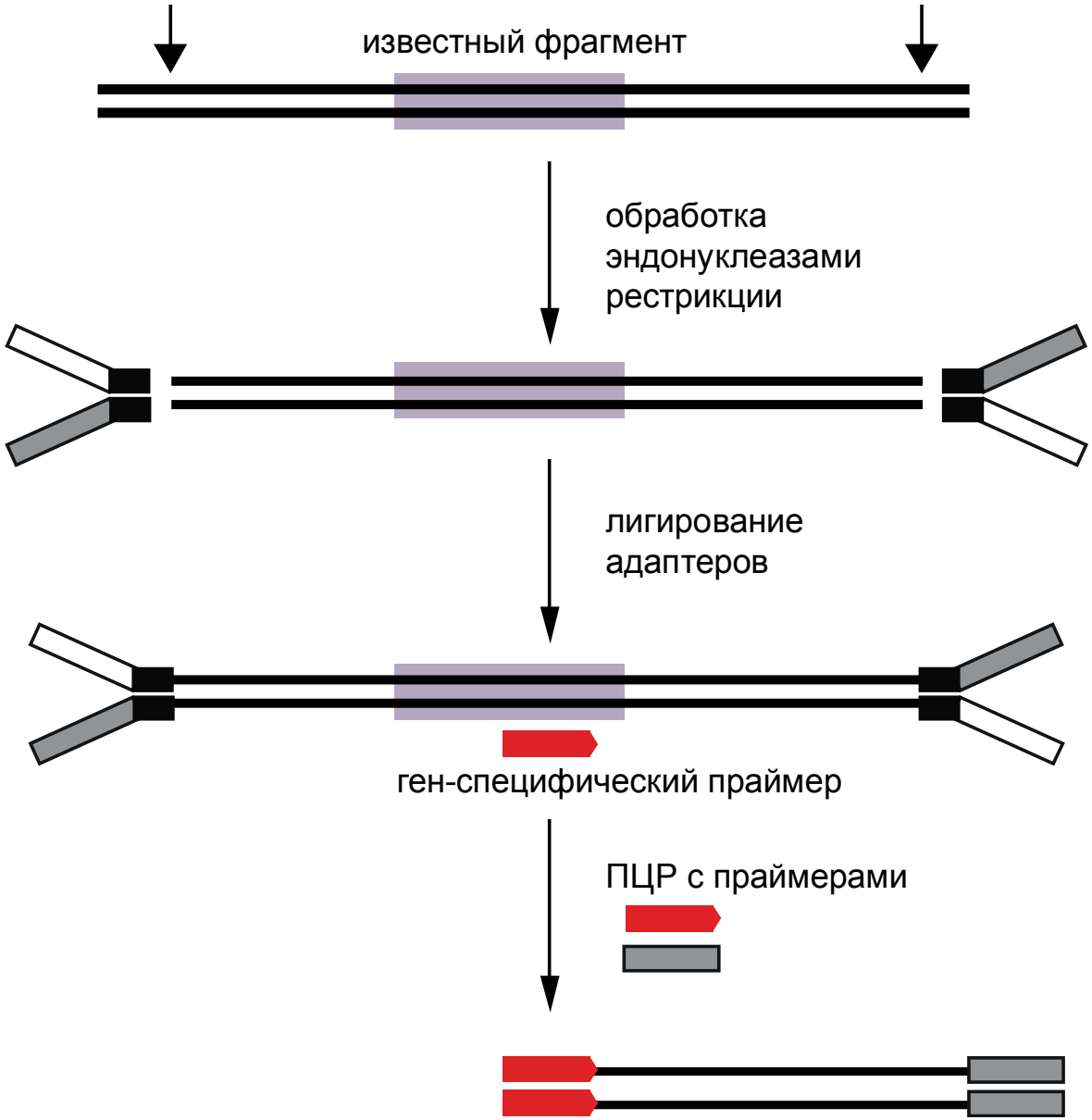
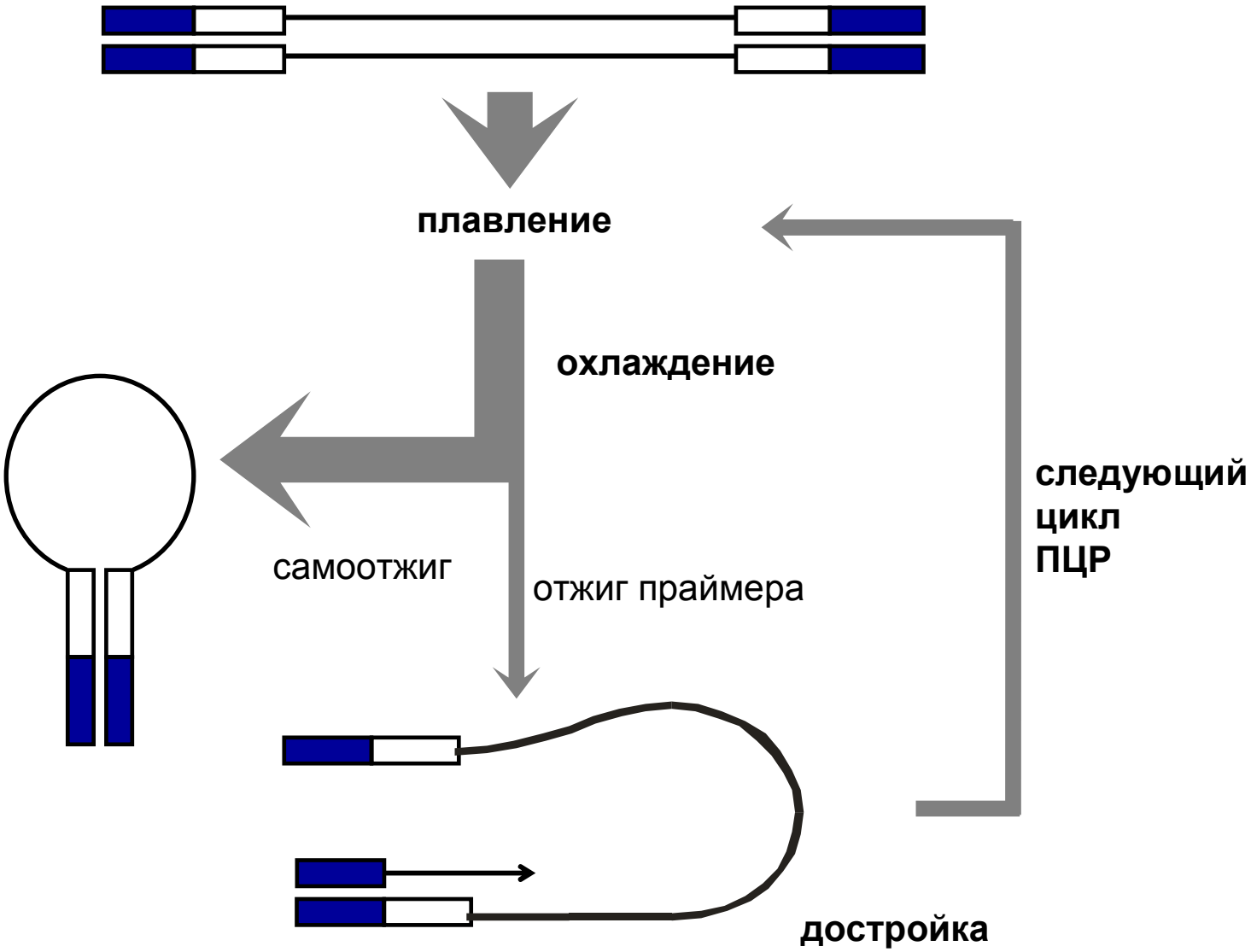
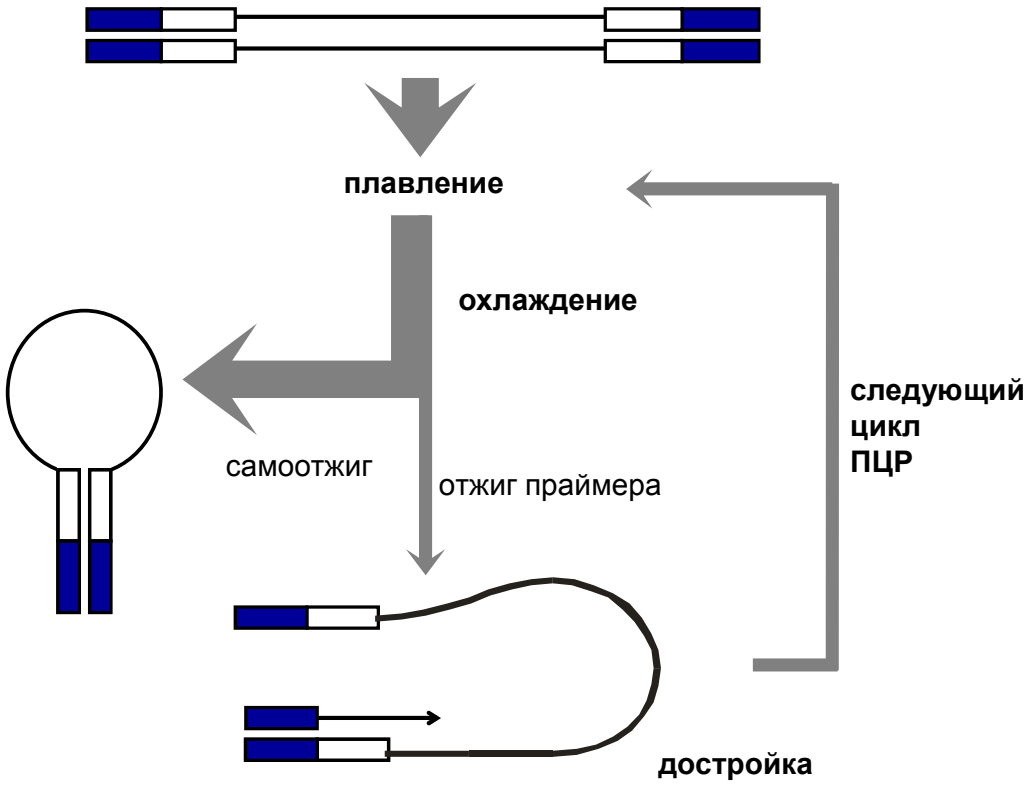


Схема селективной супрессии ПЦР



Параметры, влияющие на ССП-эффект

1. Различия в температурах отжига ИКП и праймера, используемого для амплификации.
   - Соотношение длины ИКП и последовательности, соответствующей структуре праймера.
   - GC-состав ИКП и праймера
2. Длина молекулы ДНК, несущей ИКП
3. Концентрация праймера в ПЦР

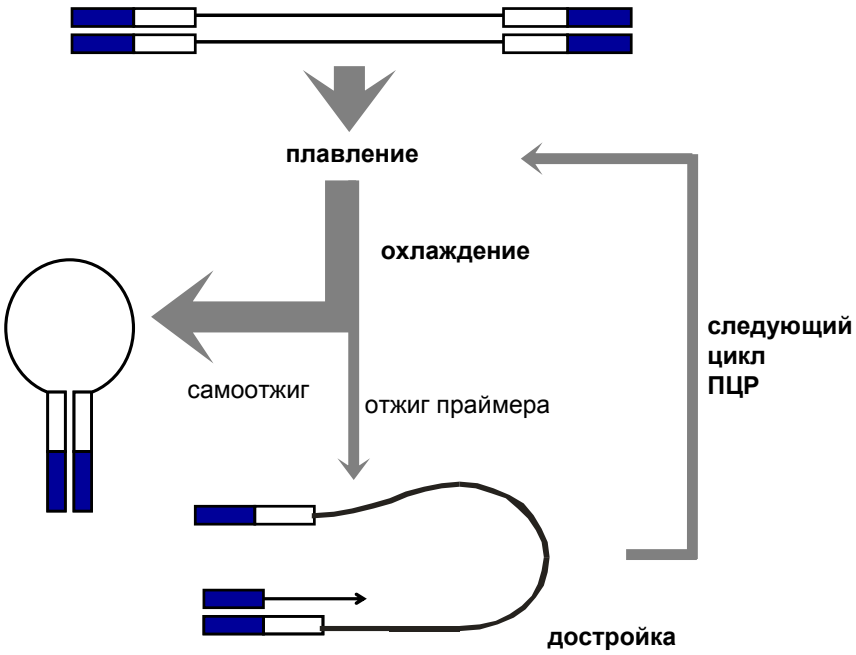


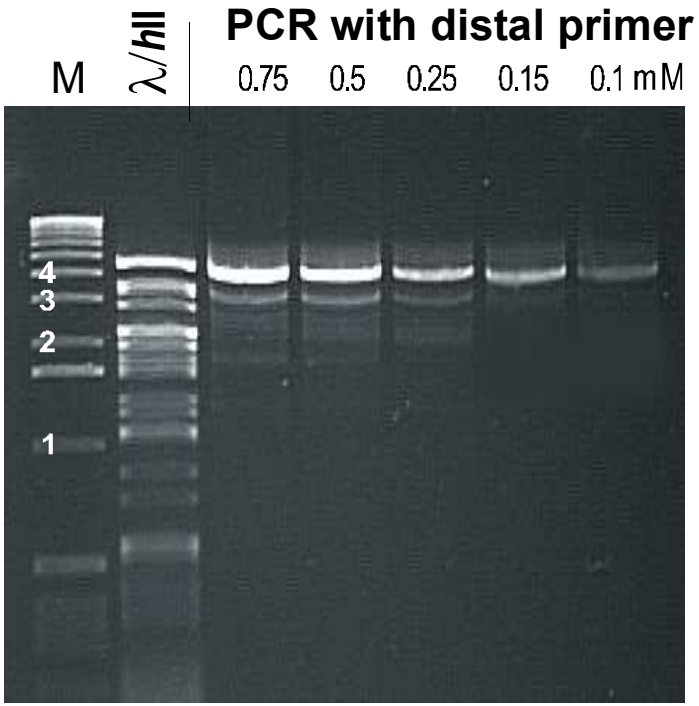
Регуляция средней длины сложных продуктов ПЦР с помощью ССП

adapter 5’-GTAATACGACTCACTATAGGGCACGGTTGCACGACGGCCCGGGCTGGT-3’
3’-CCCGACCA-5’

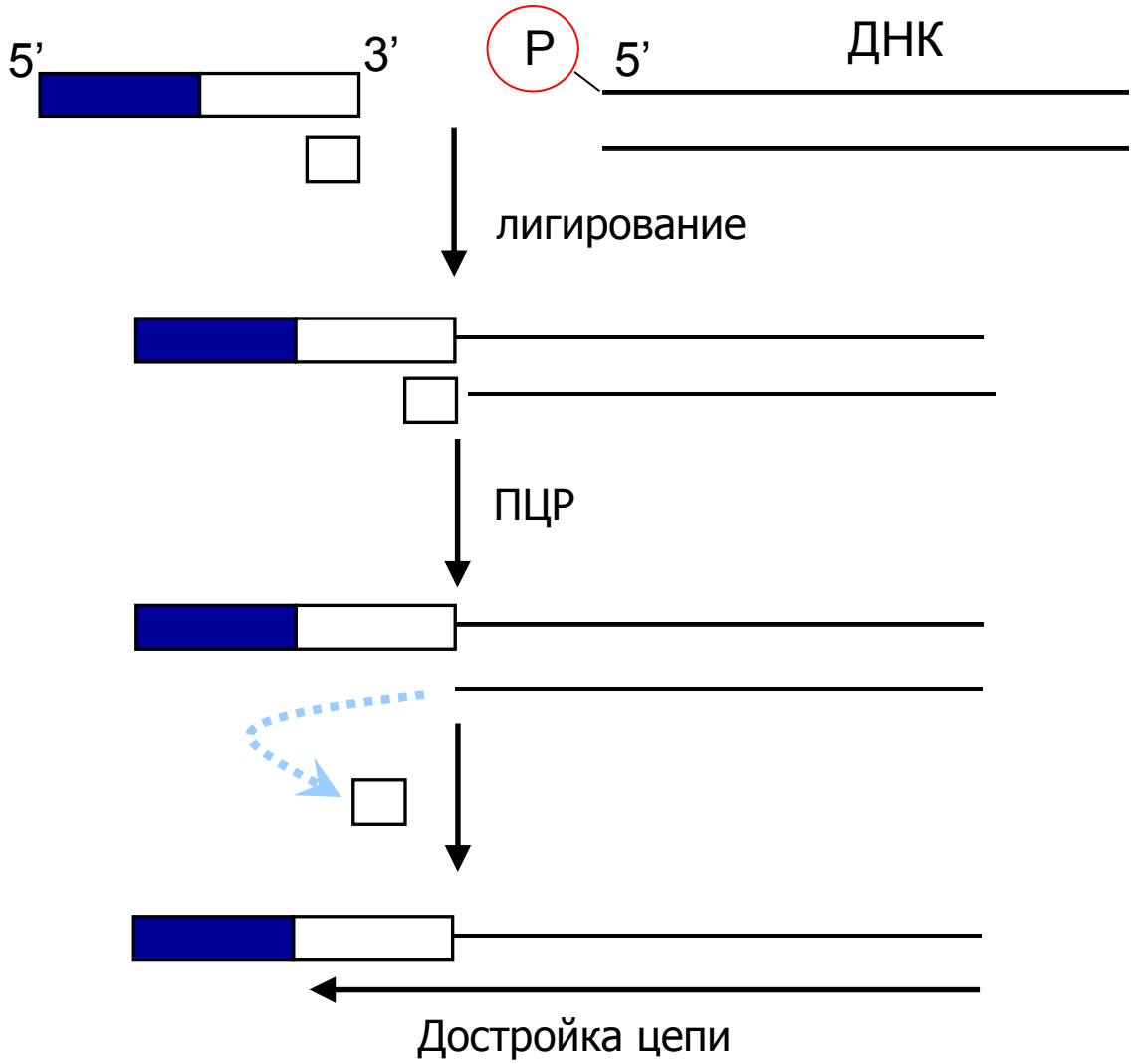
distal primer 5’-GTAATACGACTCACTATAGGGC-3’

proximal primer 5’-ACGGTTGCACGACGGCCCGGGCTGGT-3’





Введение ИКП в структуру ДНК с помощью лигирования псевдо-двухцепочечного адаптера



Введение ИКП в структуру ДНК с помощью лигирования и достройки концов

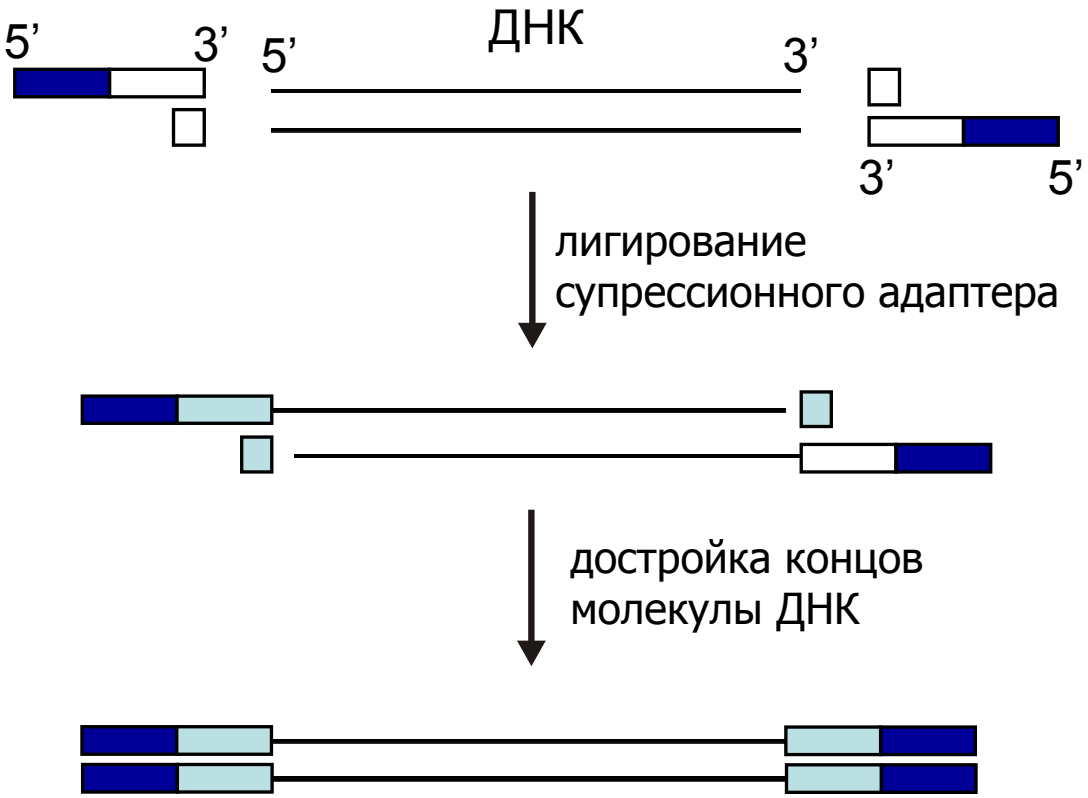
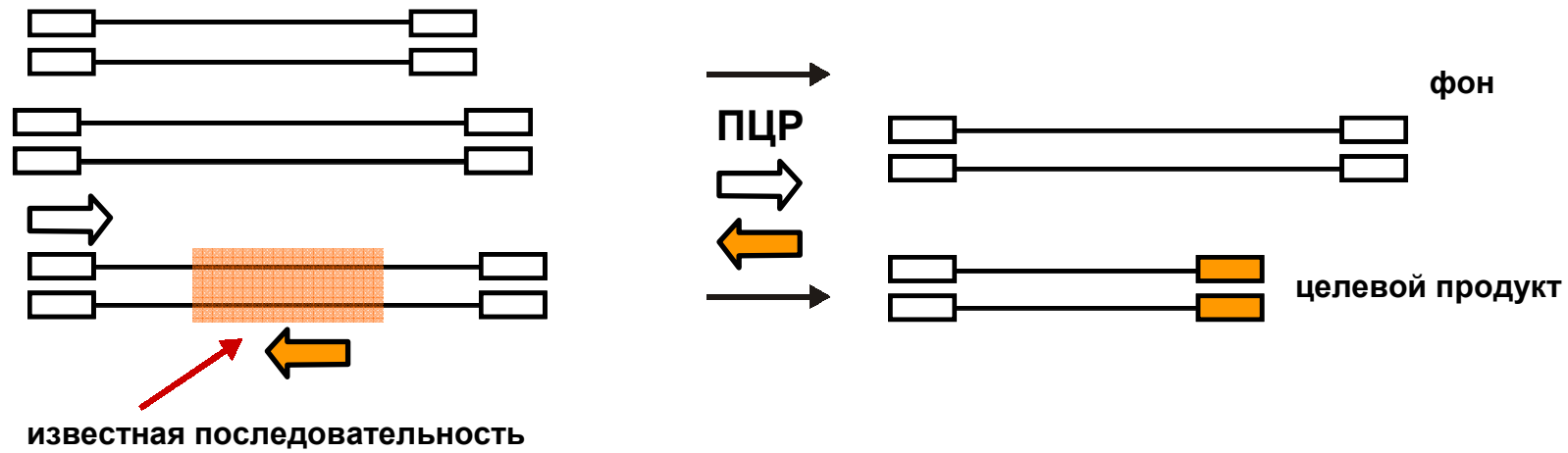
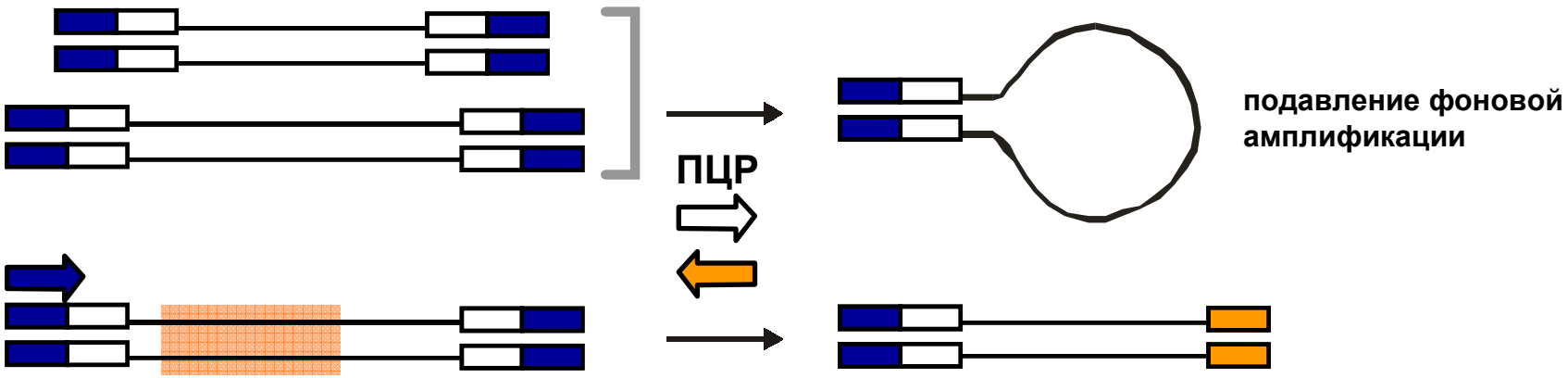
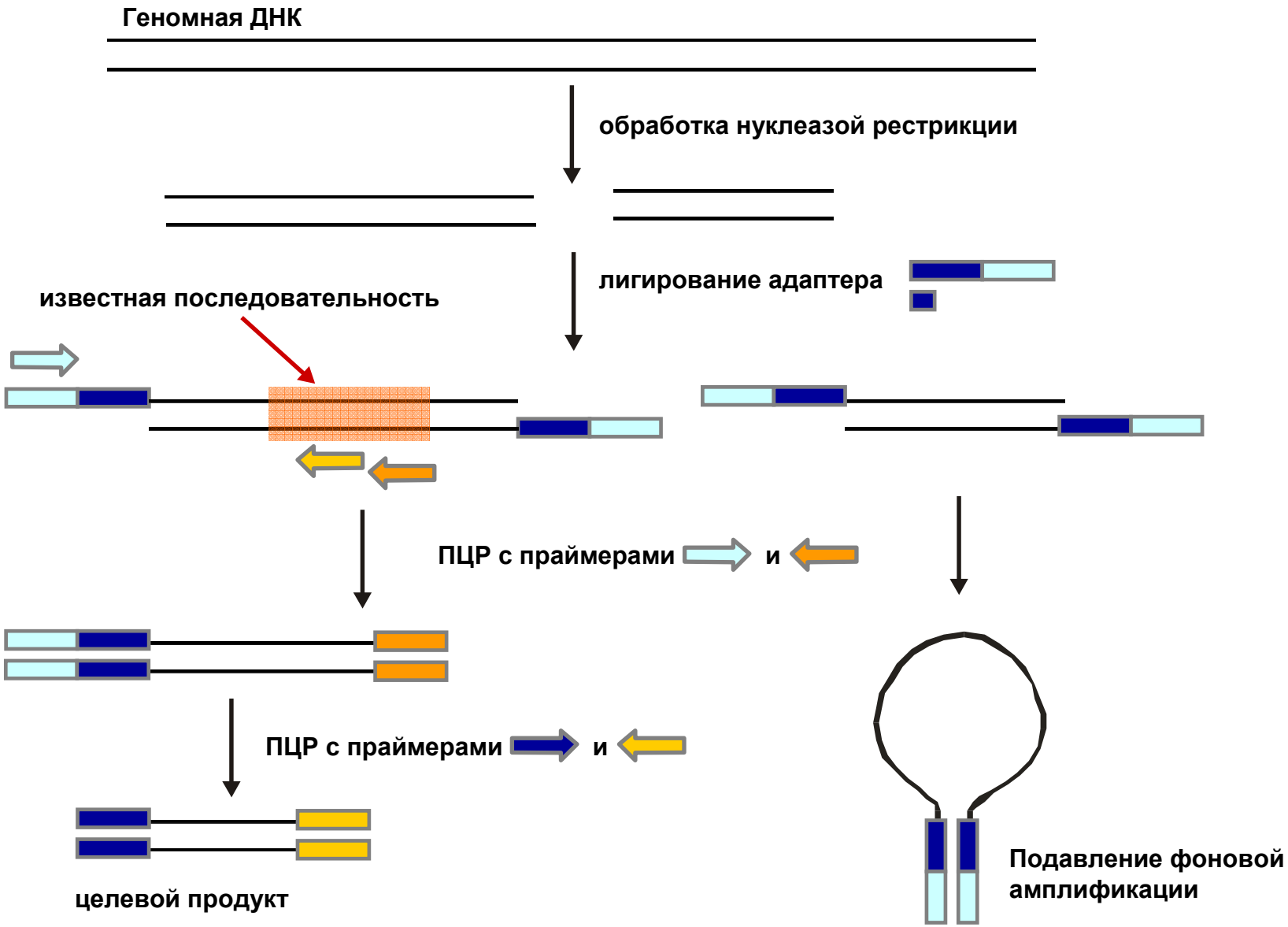


Схема метода быстрой амплификации концевых фрагментов ДНК без и с использованием ССП

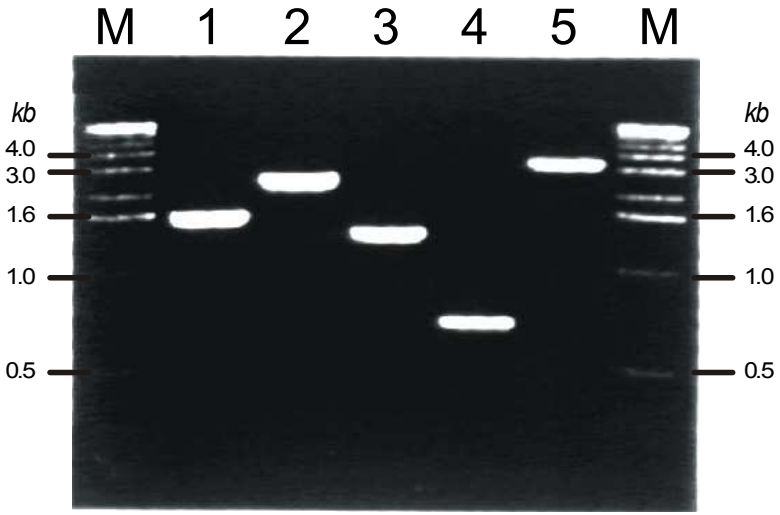




Genome Walking с использованием ССП-эффекта

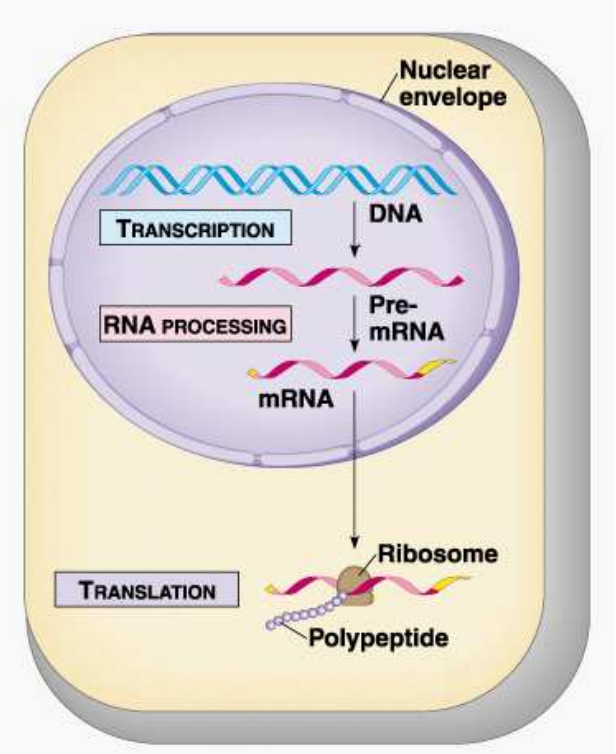


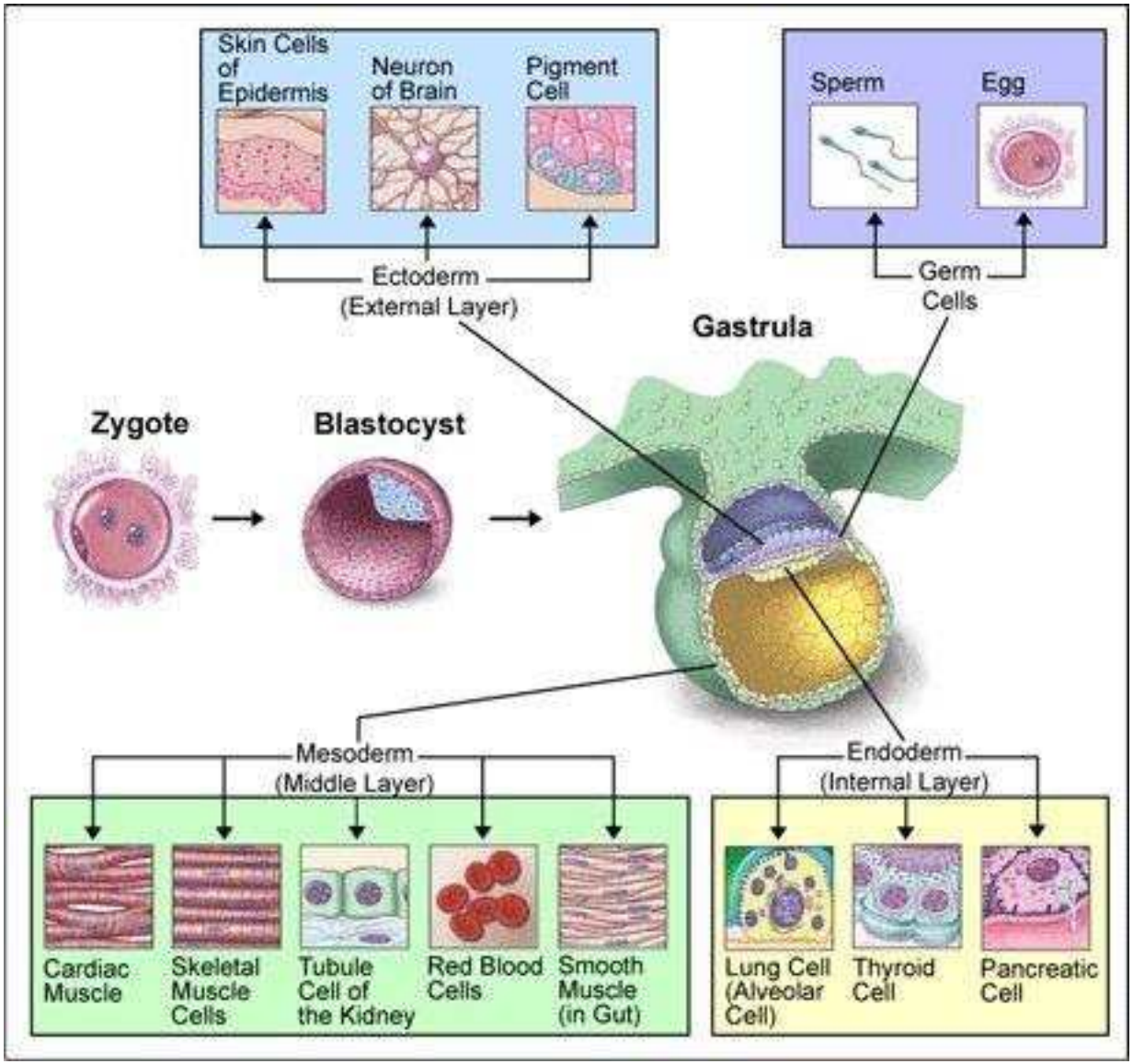
Результат амплификации промоторной зоны гена IL-1beta мыши



Амплифицированный продукт был получен в ходе двух последовательных “long and accurate” ПЦР (34 и 20 циклов) с использованием внешнего супрессионного праймера и двух последовательных специфических праймеров. В качестве матрицы использовалась геномная ДНК, обработанная различными эндонуклеазами рестрикции: 1 - EcoR V; 2 - ScaI; 3 - DraI; 4 - PvuII; 5 - SspI; M - маркер (1 kb ladder).

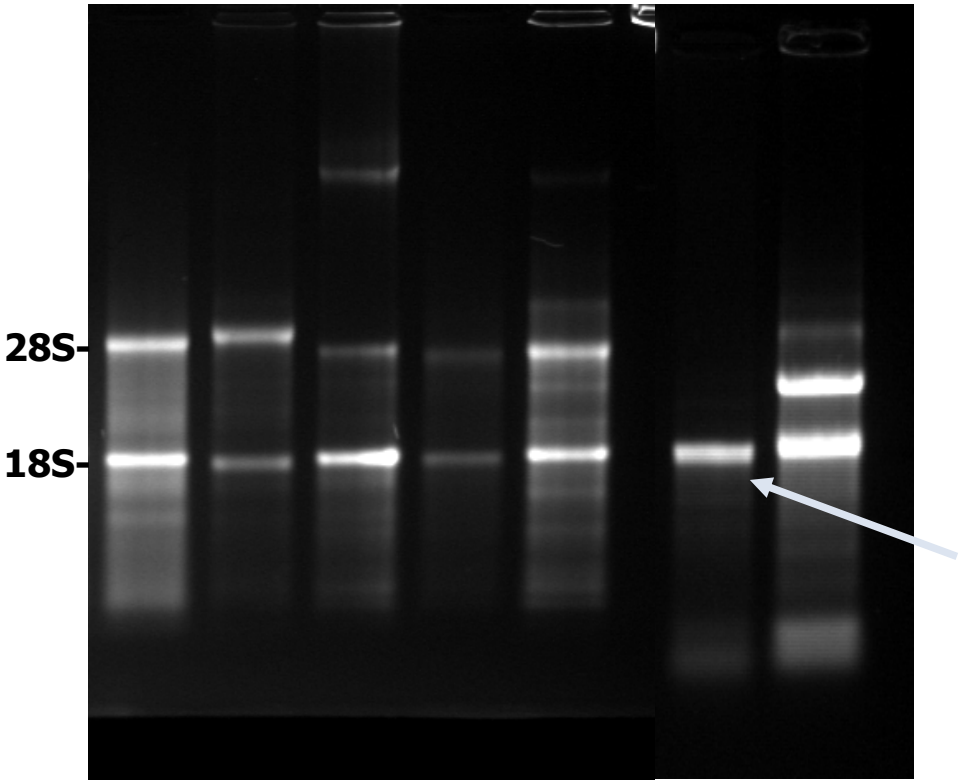
Экспрессия генов на уровне РНК





Приготовление комплементарной кДНК

Визуализация РНК на гель-электрофорезе



Тотальная РНК из тканей млекопитающих выглядит как шмер, содержащий две интенсивные полосы 28S и 18S рибосомальной
РНК. Соотношение интенсивностей полос 28S к 18S РНК должно быть не менее 1:1. Изменение соотношения в пользу 18S РНК свидетельствует о частичной деградации образца.

РНК из других организмов может иметь иные характеристики: например у насекомых может быть видна только одна полоса рибосомальной РНК на уровне 18S (Ishikawa, 1977) или шмер может быть длиной до 2-3 kb.

Выделение мРНК эукариот

мРНК составляет всего 2,5-5% тотальной РНК

мРНК имеет кэп на 5’-конце

мРНК (кроме гистонов) имеет поли(А) последовательность на 3’-конце (полиА+ РНК). Эту последовательность можно использовать при выделении фракции мРНК

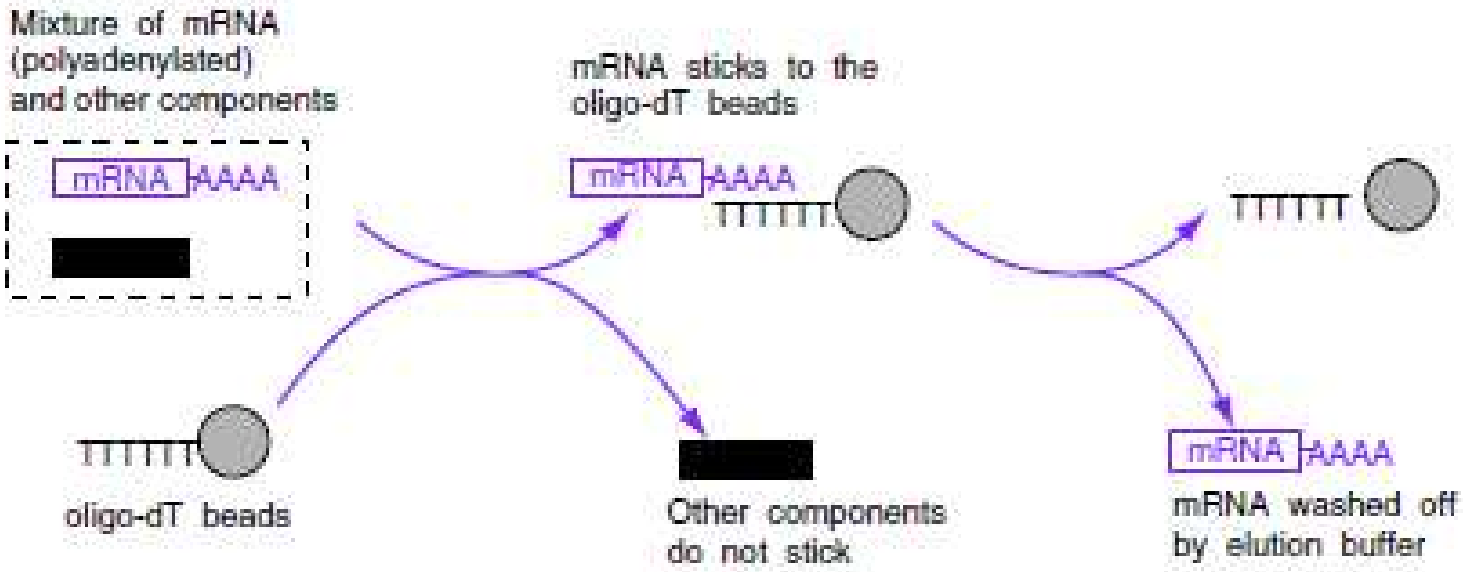
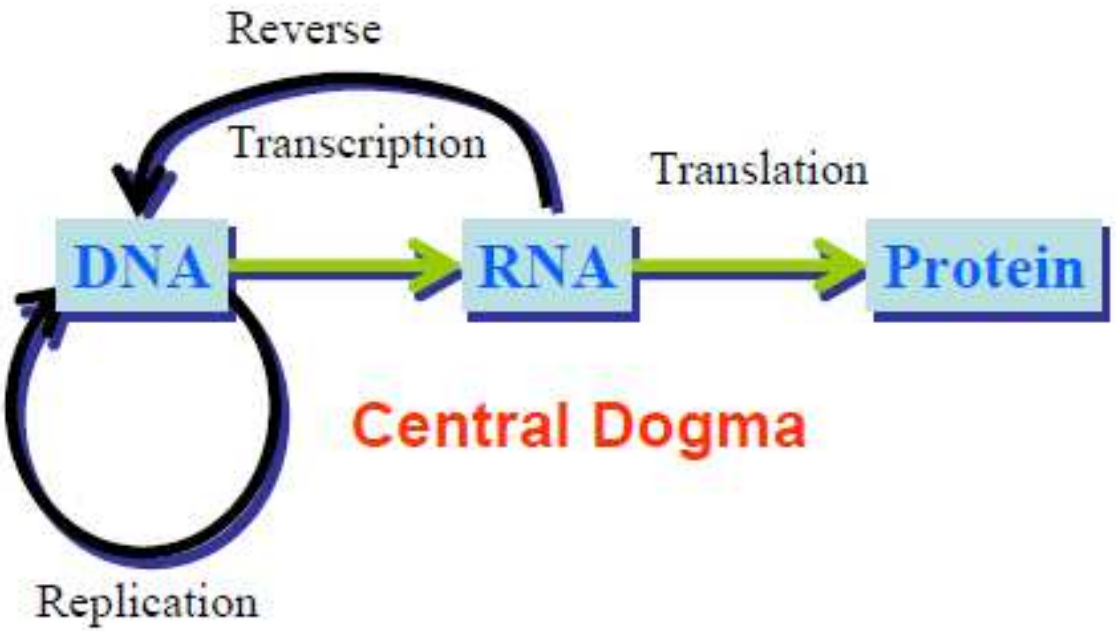


Figure 7.5 Principle of oligo-dT purification of mRNA

Центральная догма молекулярной биологии

Retroviruses add a new "loop" to central dogma making "cDNA" from RNA.

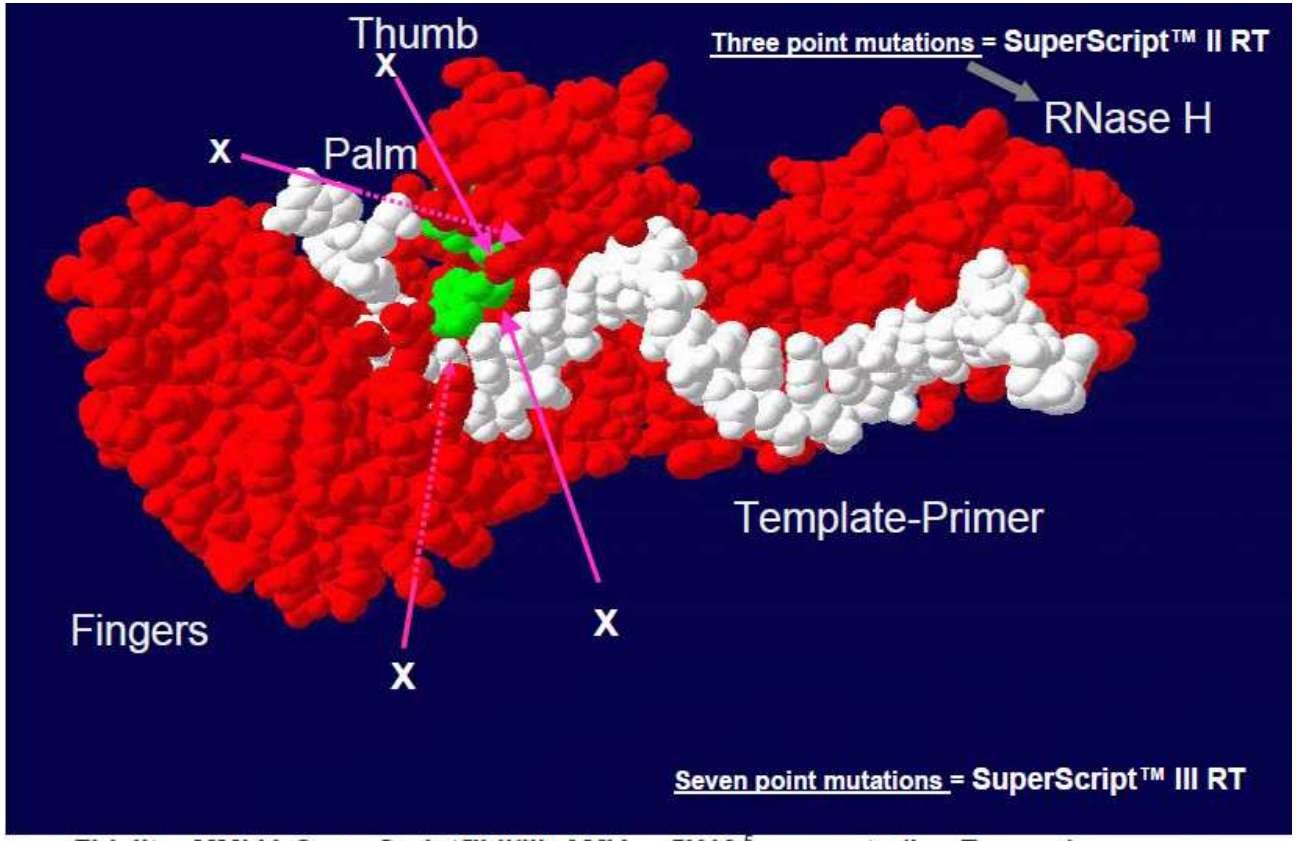


Обратные транскриптазы

- РНК-зависимые ДНК полимеразы ретровирусов
  - вируса миелобластоза птиц (Avian Myeloblastosis Virus RT, AMV RT)
  - вируса мышиного лейкоза Молони (Moloney Murine Leukemia Virus RT, M-MLV RT)
- Описаны в 60-е годы прошлого столетия (Dubelco, Temin, and Baltimore, 1975 - Noble Prize)
- Нуждаются в ДНК или РНК затравке (праймере)
- Обладают активностью РНКазы Н, деградируют РНК-матрицу в процессе синтеза первой цепи кДНК
- AMV RT и M-MLV RT:
  - AMV RT – гетеродимер, состоящий из двух субъединиц
  - M-MLV RT мономер, имеет более слабую активность РНКазы Н

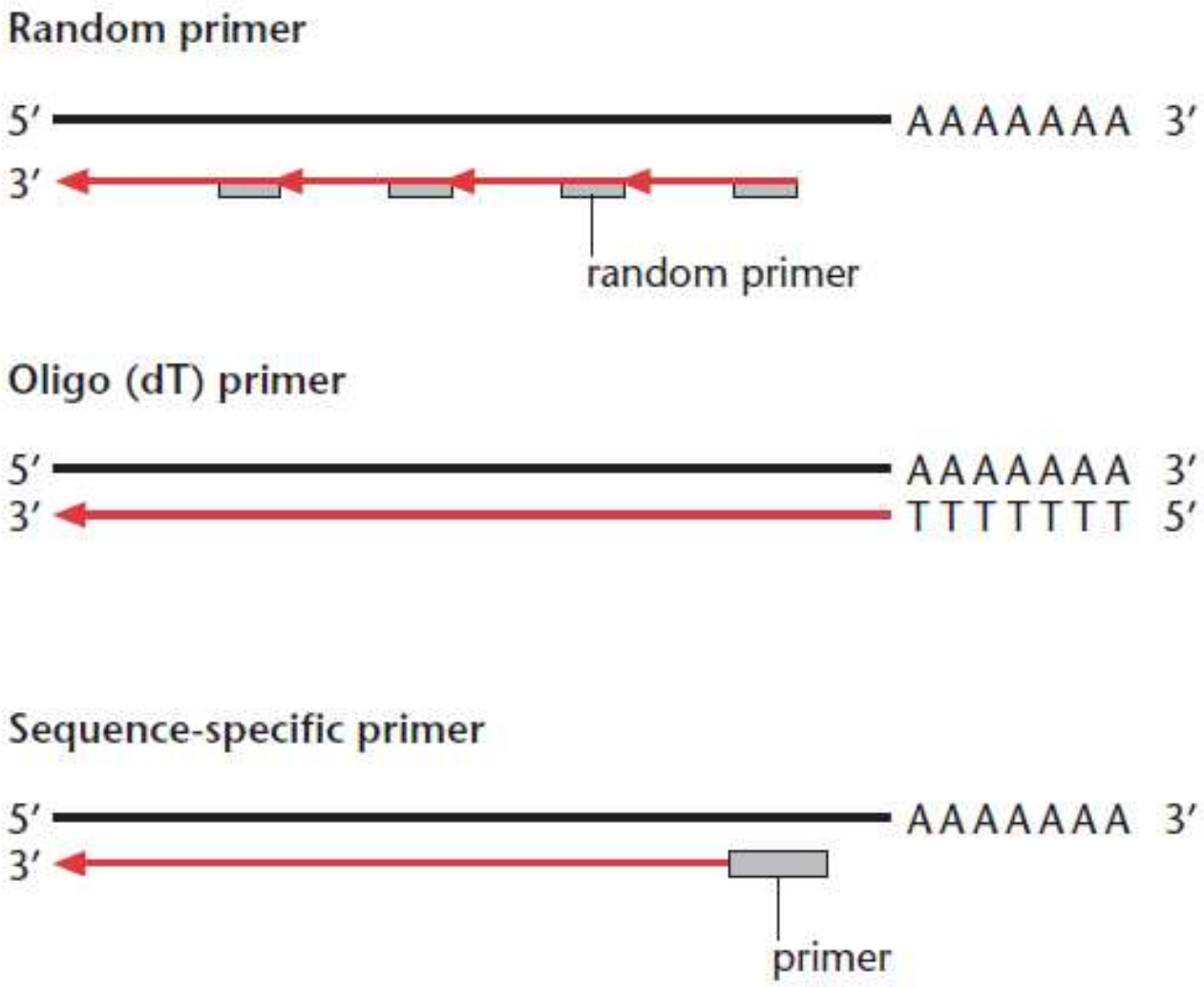
Обратные транскриптазы

- Мутанты рекомбинантной M-MLV RT с инактивированным РНКаза H доменом более эффективны в синтезе первой цепи кДНК
- M-MLV RT мутанты с подавленной активностью РНКазы H, внематрично добавляют несколько нуклеотидов на 5’-конец новосинтезированной кДНК (достигнув конца РНК-матрицы).



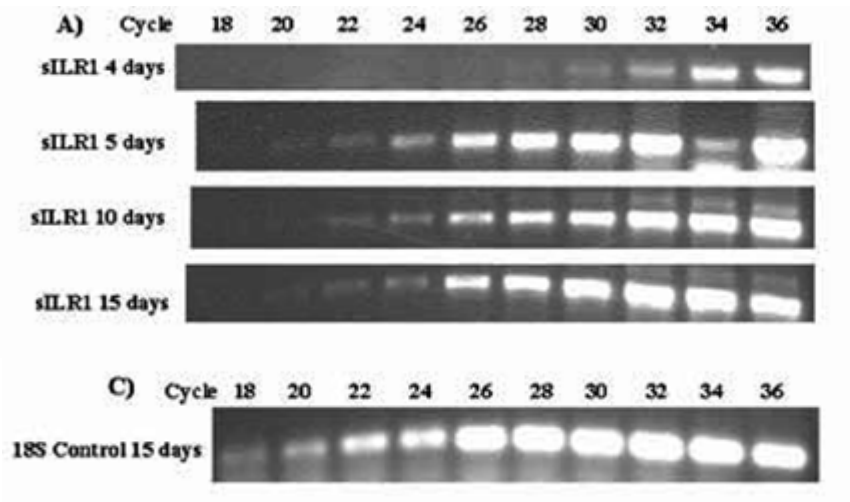
Fidelity: MMLV, SuperScript™ II/III, AMV: ~ $5X10^{-5}$ error rate (LacZ assay)

Три стратегии затравки синтеза кДНК обратной транскриптазой



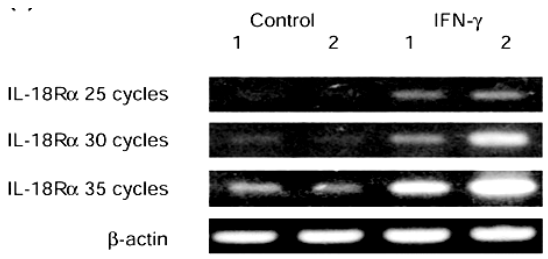
Обратная транскрипция-ПЦР (ОТ-ПЦР \ RT-PCR)

- **Синтез первой цепи с помощью обратной транскриптазы на матрице РНК**
- **Амплификация кДНК в ПЦР на матрице первой цепи**



sILR1 expression examined by semi-quantitative RT-PCR from 4, 5, 10, and 15 day old seedlings. 5 µl aliquots of the RT-PCR reaction mixtures were taken for 10 alternating cycles starting at cycle #18.

c) Amplification of the 18S control. The control was expressed at the same level for all samples examined. All samples were analyzed on 1% agarose gels stained with ethidium bromide.



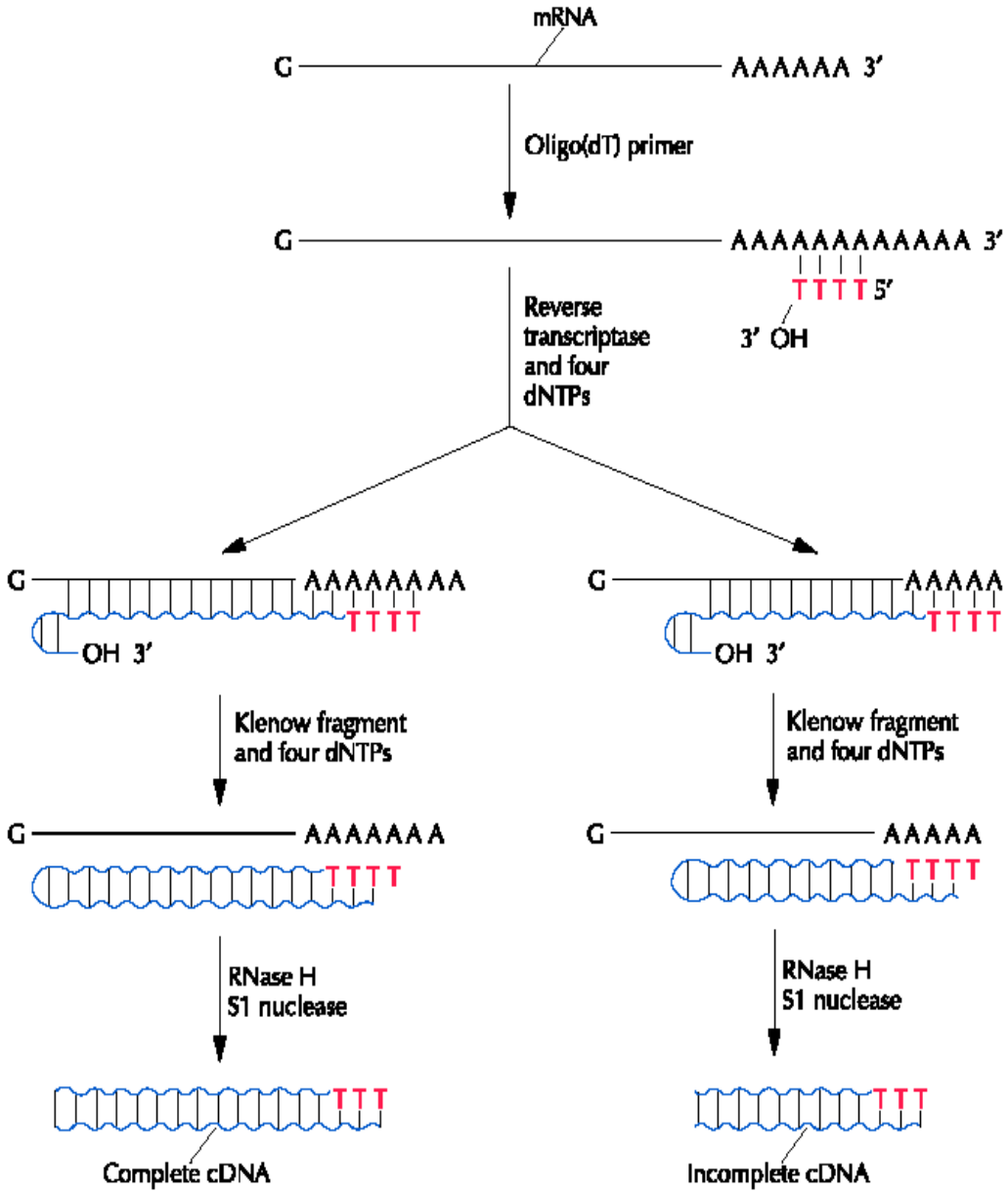
IFN-γ induces IL-18Rα mRNA expression in RA-FLS. IL-18Rα RT-PCR products of two representative cultures after 25, 30, and 35 cycles, and the corresponding β-actin RT-PCR products (bottom lane).

FLS = fibroblast-like synoviocytes; IL-18Rα = IL-18 receptor α; RA = rheumatoid arthritis; RT-PCR= reverse transcriptase polymerase chain reaction.

Приготовление образцов суммарной двухцепочечной кДНК

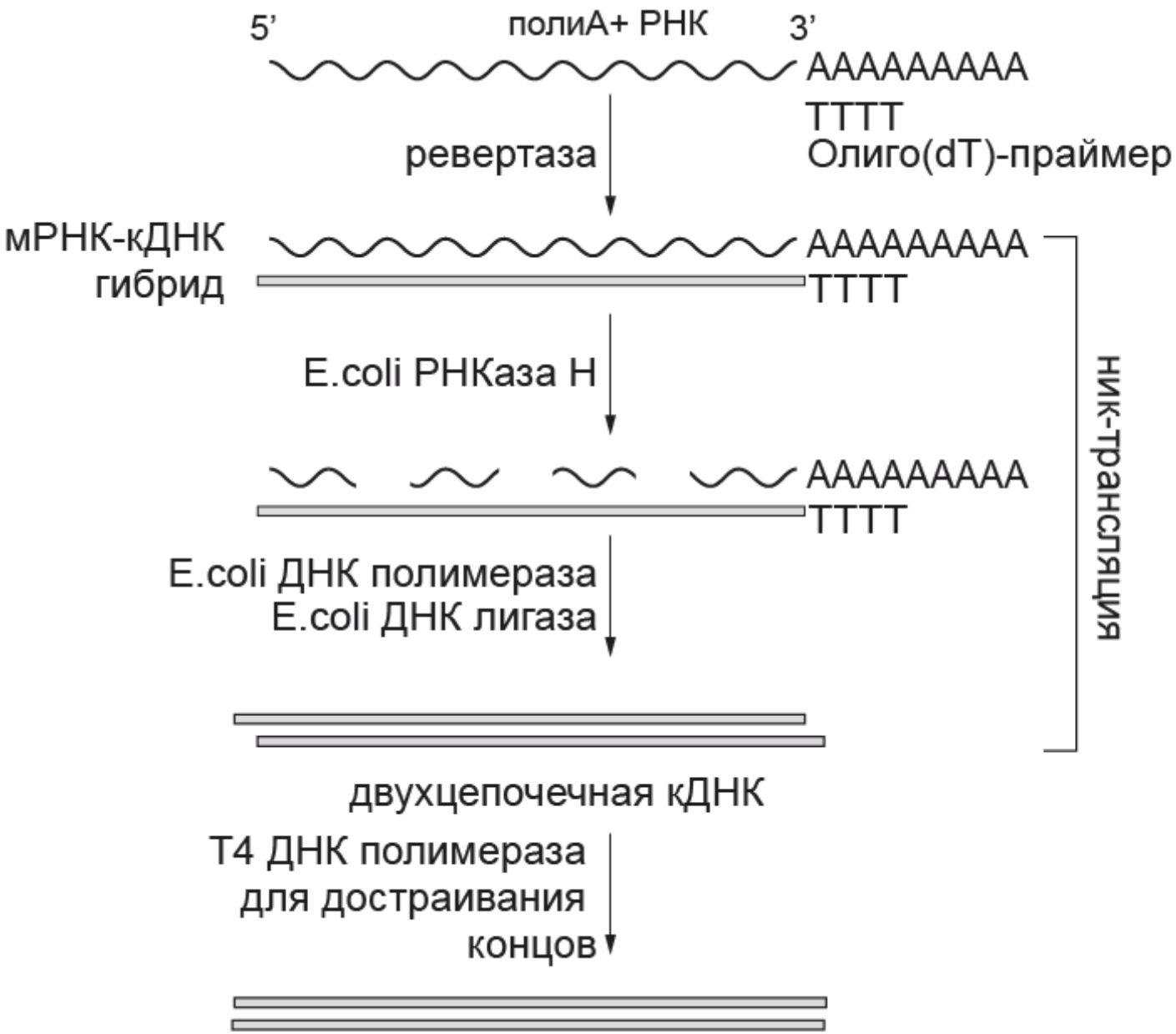
- **Приготовление библиотек кДНК**
- **Приготовление экспессионных библиотек кДНК**
- **Выявление 3′ и 5′ –концевых фрагментов РНК (RACE) и клонирование полноразмерных молекул кДНК**
- **Сравнение популяций РНК**
- **Широкомасштабный анализ транскриптомов**

Схема приготовления двухцепочечной кДНК - 1



Происходит потеря 5′-концов РНК при обработке S1 нуклеазой

Gubler–Hoffman метод приготовления двухцепочечной кДНК



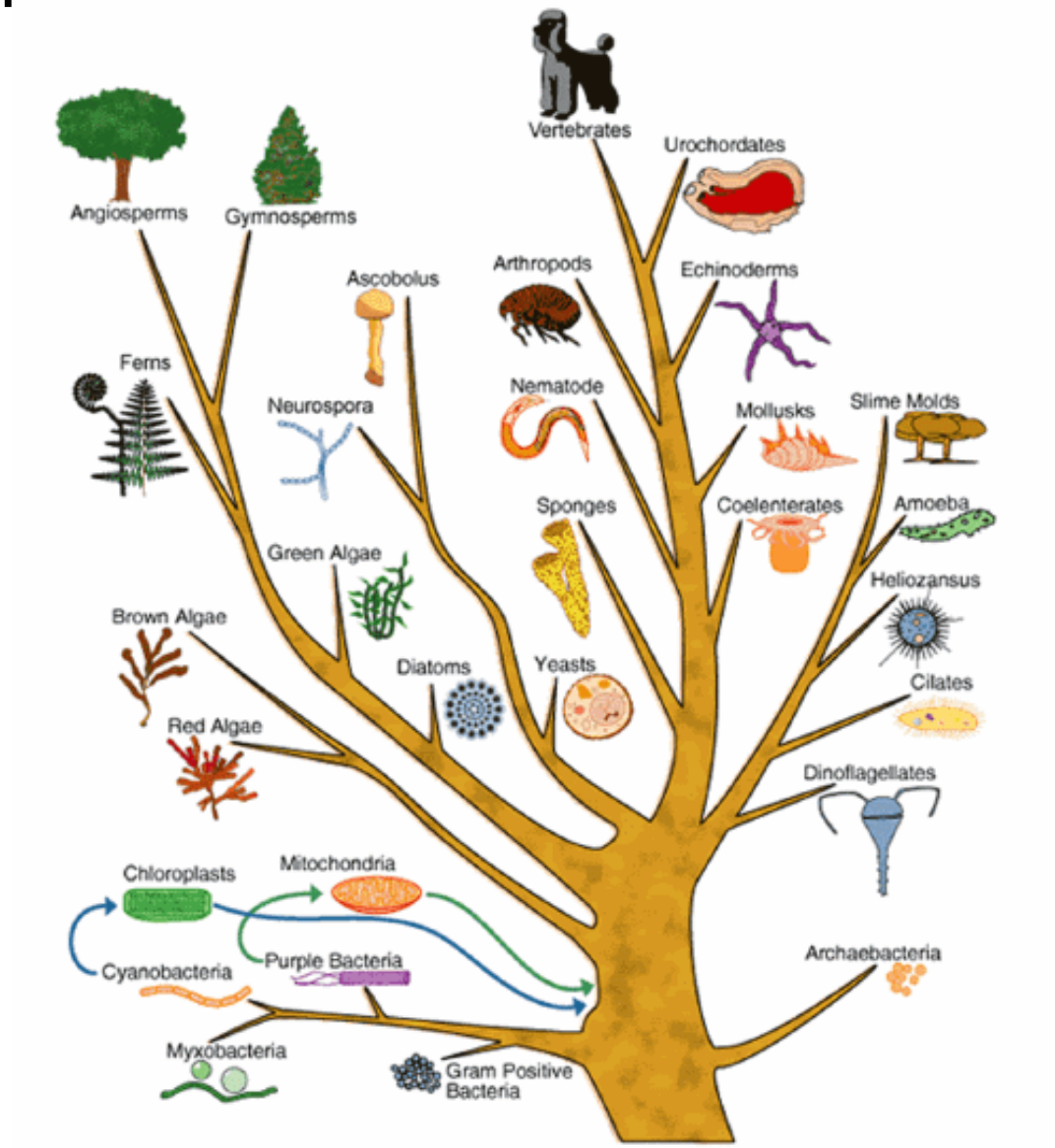
ПРОБЛЕМЫ СИНТЕЗА кДНК

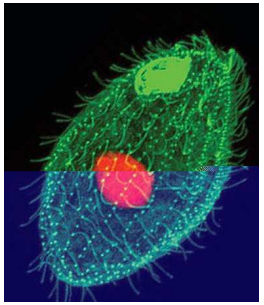
- Для выделения полиА+ фракции нужно значительное колличество РНК (>100 mkg)

- Популяция кДНК обеднена 5’ –последовательностями - обратная транскриптаза часто не доводит синтез до конца

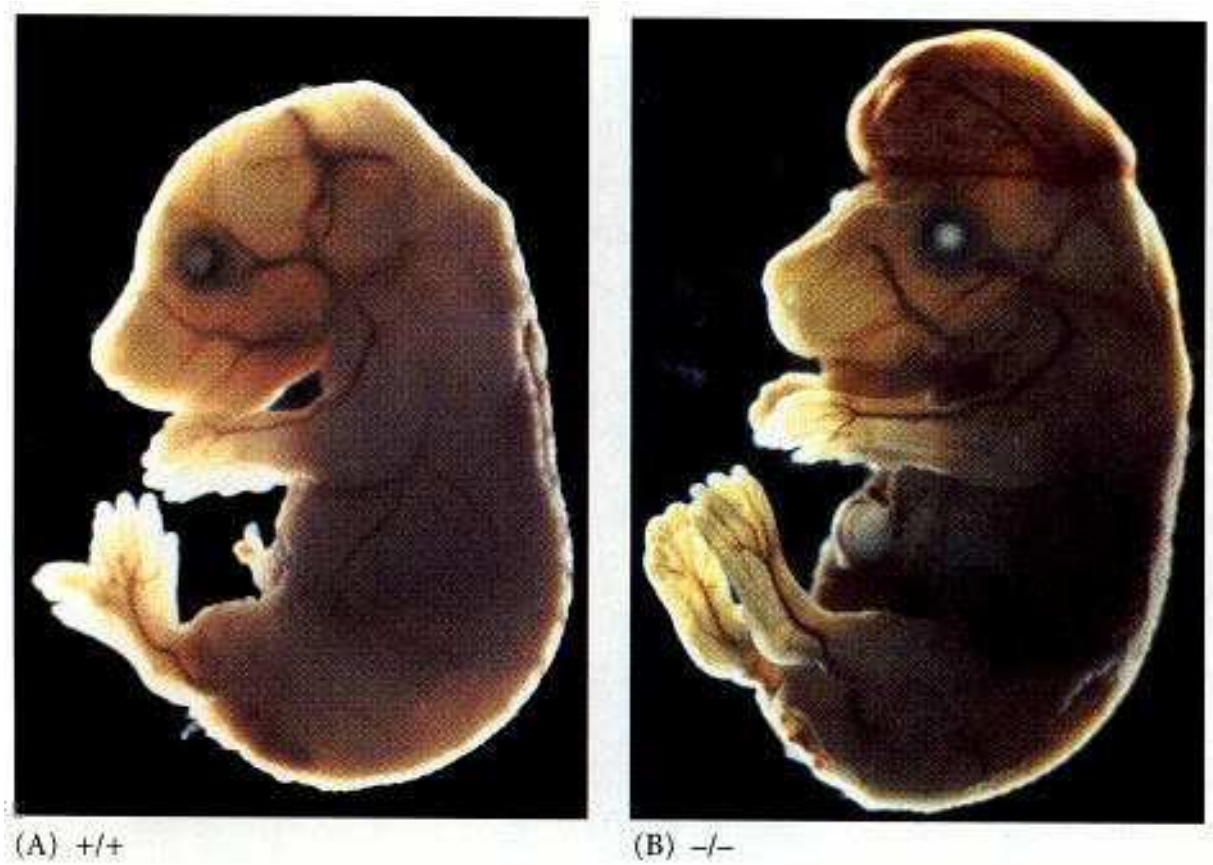
- Часть информации о структуре 5’ –концевой последовательности РНК теряется при синтезе 2-ой цепи кДНК

Многие модельные объекты имеют небольшие размеры



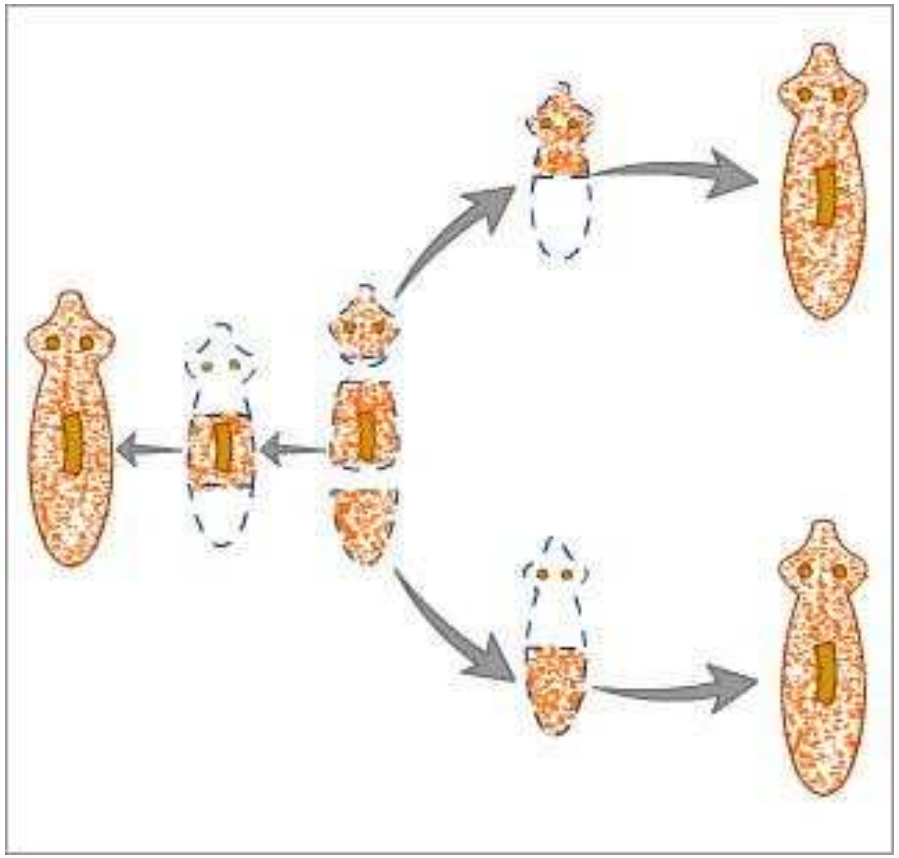


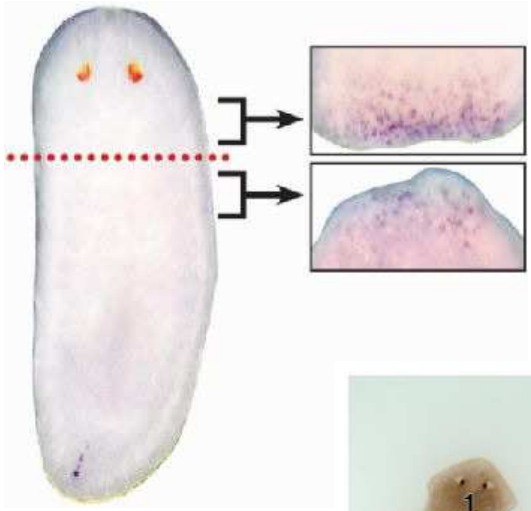
Ткани или популяции клеток, вовлеченные в биологический процесс, имеют малые размеры



Патологии развития мозга (мышь)

Ткани или популяции клеток, вовлеченные в биологический процесс, имеют малые размеры





Регенерация планарий

- Для выделения полиА+ фракции нужно значительное количество РНК (>100 mkg)

- Необходимы методы приготовления кДНК из небольших количеств биологического материала

ОСНОВНАЯ ПРОБЛЕММА АМПЛИФИКАЦИИ СУММАРНОЙ кДНК – отсутствие на 5′-конце мРНК (соответствует 3′-концу 1-ой цепи кДНК) определенной последовательности, которая могла бы использоваться для отжига праймера.

Использование терминальной трансферазы для приготовления амплифицированной кДНК (схема 1)

мРНК AAAAA

Синтез первой цепи кДНК

Т-праймер: TTTTT

обратная транскриптаза

AAAAA

TTTTT

TTTTT

избыток Т-праймера

Олиго-dG-тэйлинг

терминальная трансфераза,

dGTP

GGGG AAAAA

TTTTT

GGGG TTTTT

ПЦР

Т-праймер: TTTTT

С-праймер: CCCC

CCCC AAAAA

GGGG TTTTT

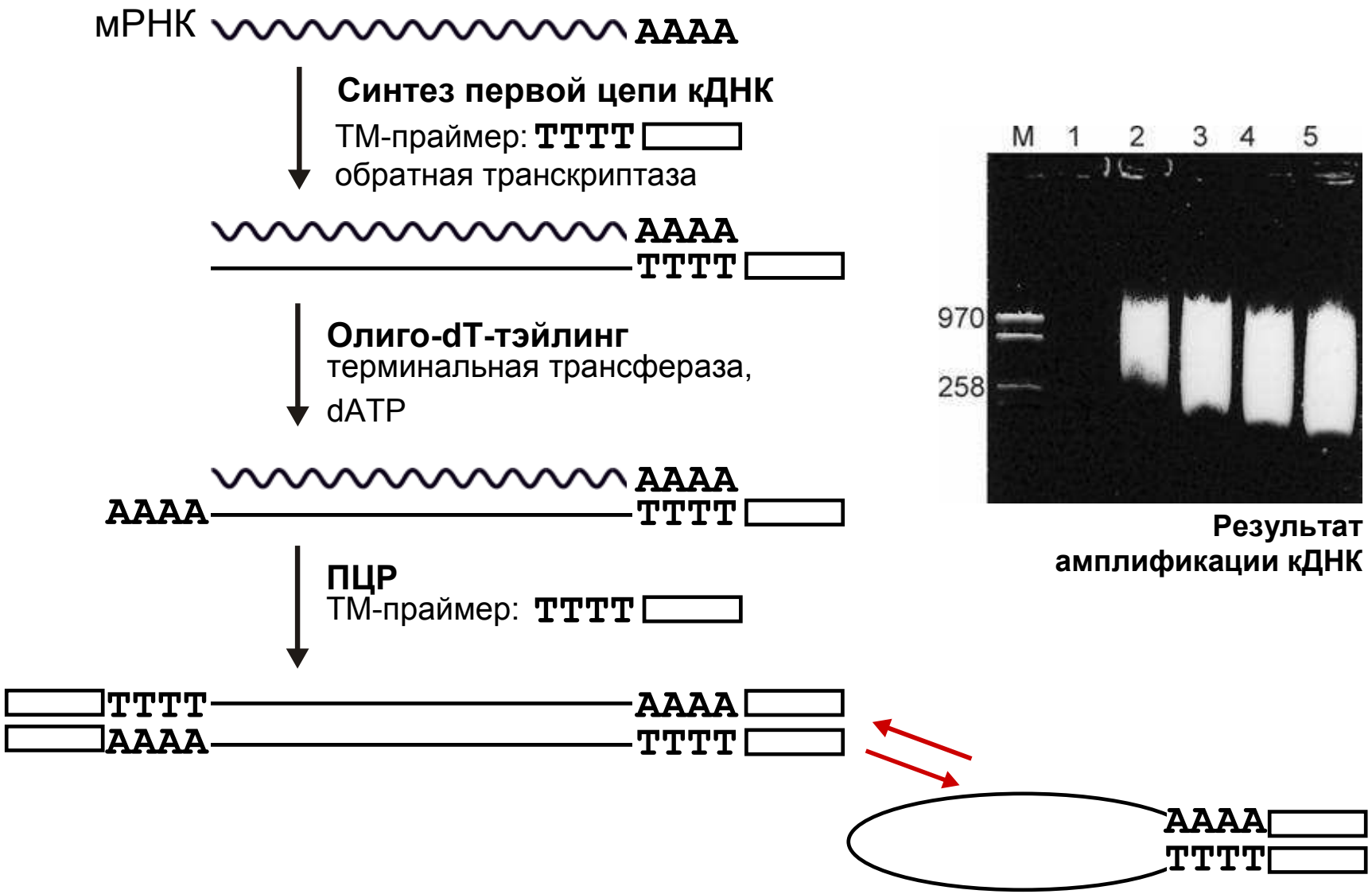
5′-CCCC AAAAA-3′

GGGG TTTTT

димер праймеров

Результат амплификации кДНК

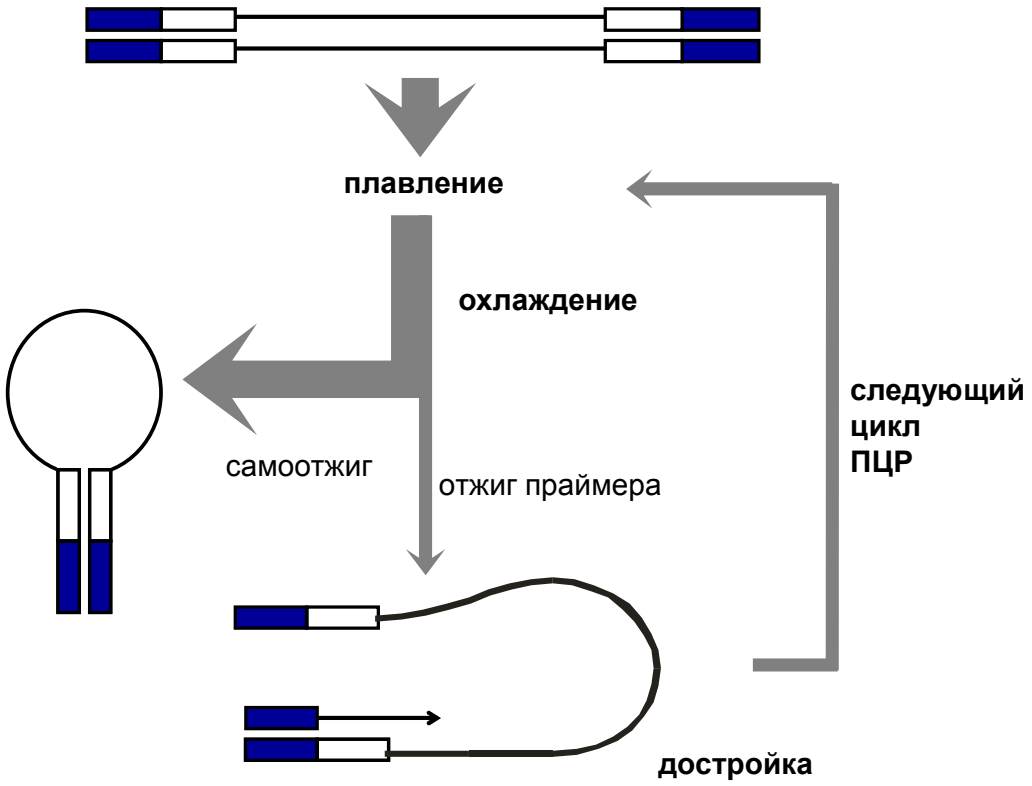
Использование терминальной трансферазы для приготовления кДНК (схема 2)



Результат амплификации кДНК

Параметры, влияющие на ССП-эффект

1. Различия в температурах отжига ИКП и праймера, используемого для амплификации.
   - Соотношение длины ИКП и последовательности, соответствующей структуре праймера.
   - GC-состав ИКП и праймера
2. Длина молекулы ДНК, несущей ИКП
3. Концентрация праймера в ПЦР



Введение ИКП в структуру ДНК с помощью лигирования

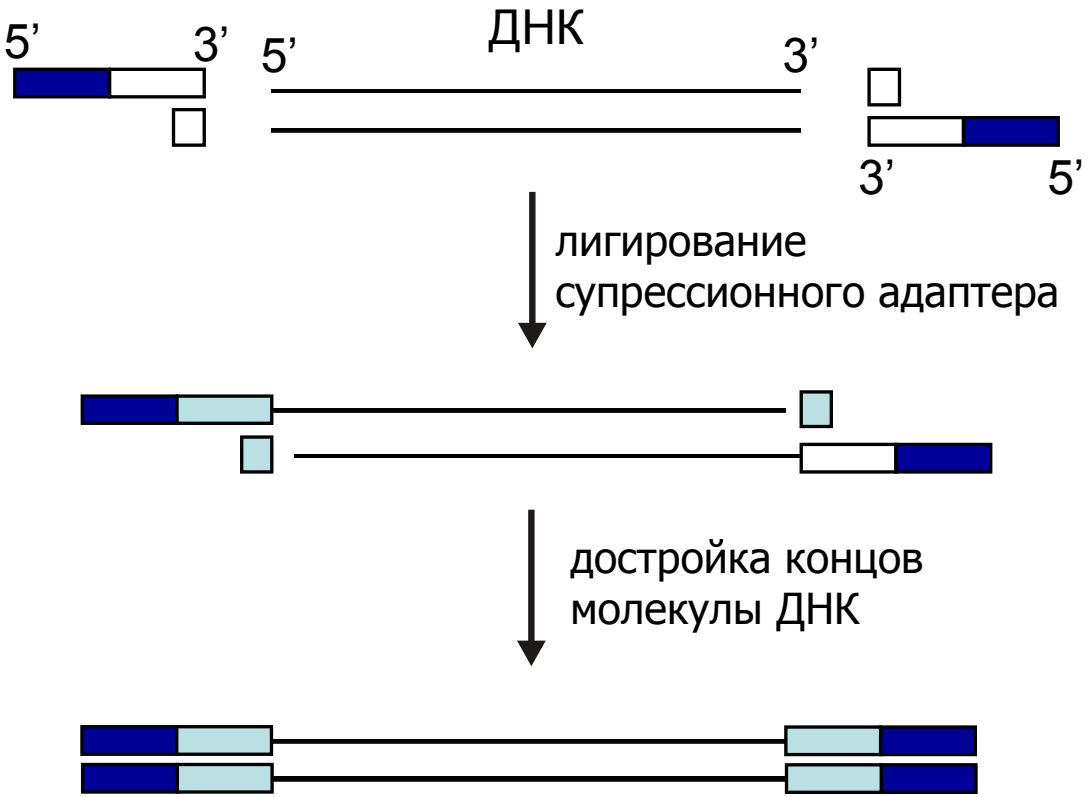
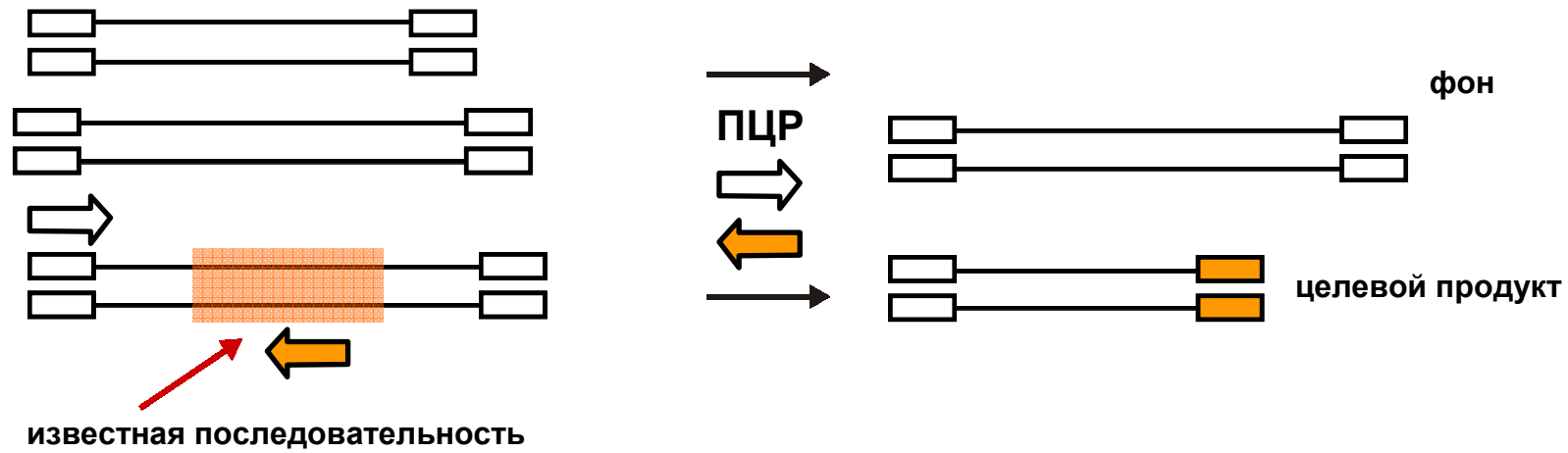
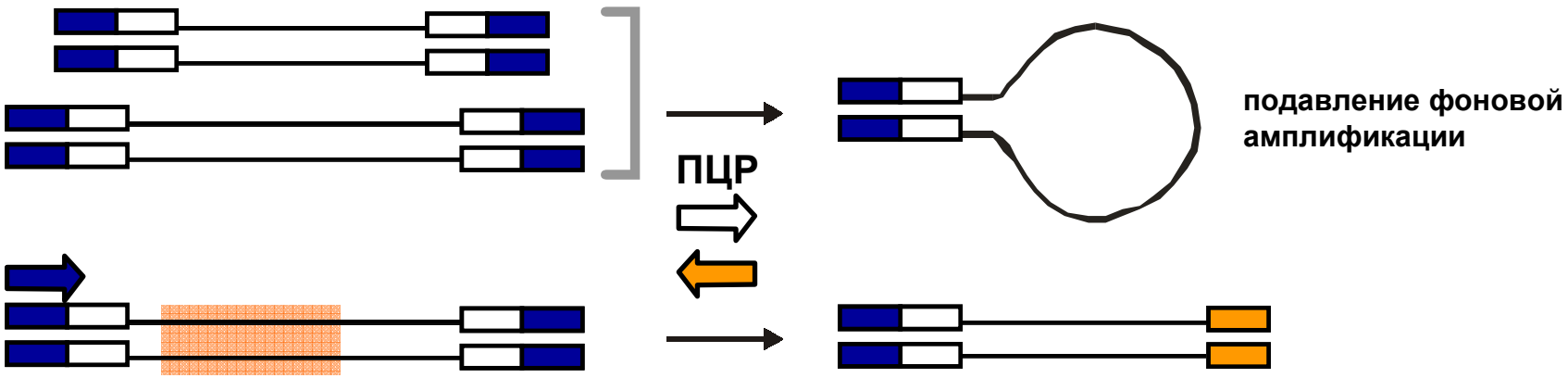
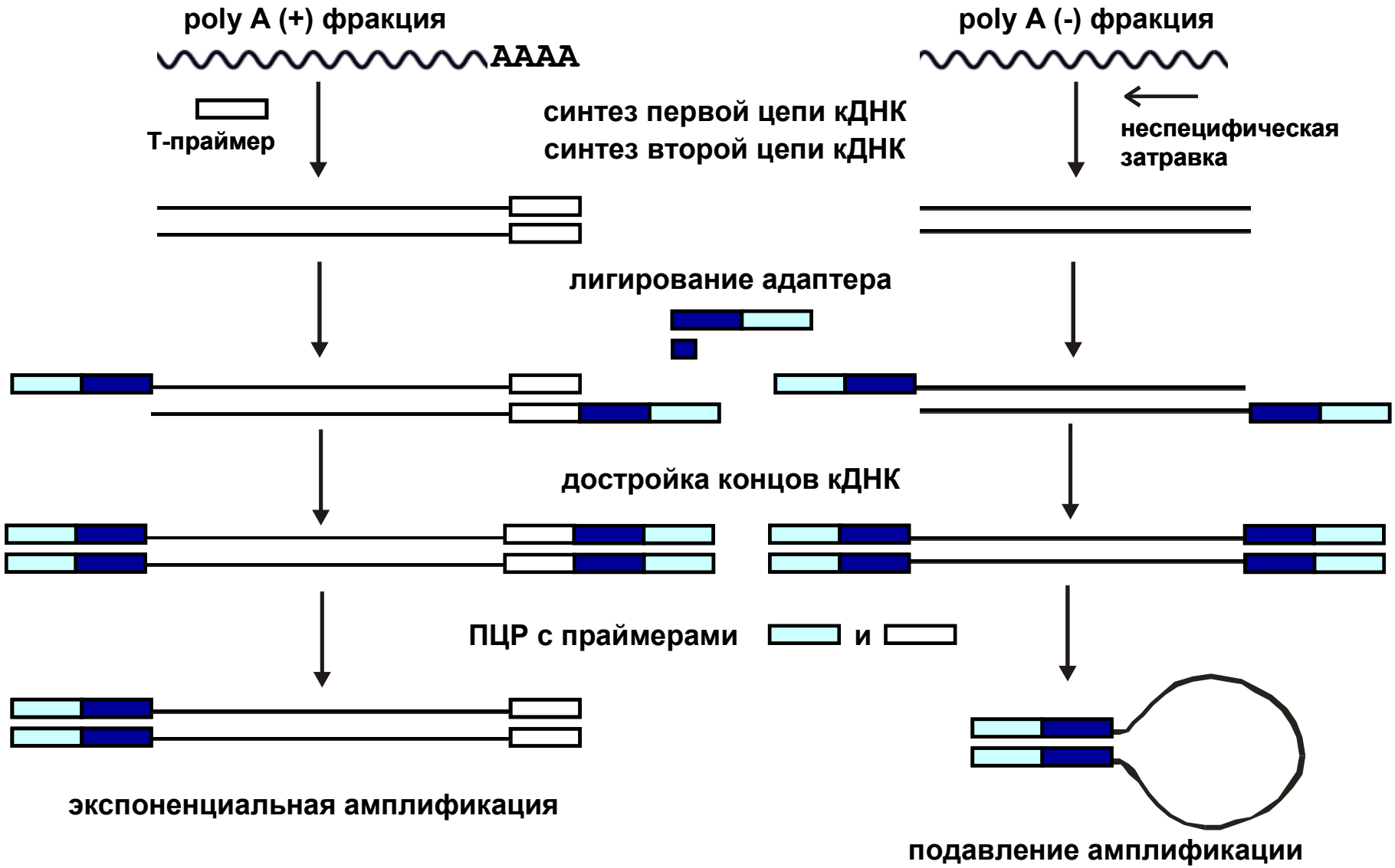


Схема метода быстрой амплификации концевых фрагментов ДНК без и с использованием ССП

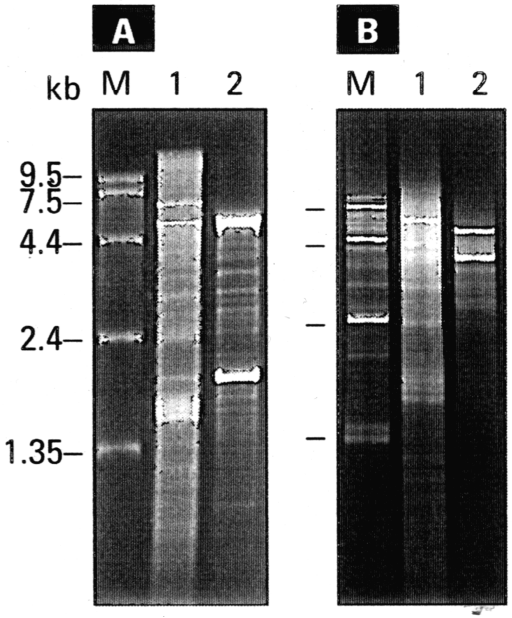




Marathon метод получения амплифицированной кДНК из малого количества биологического материала

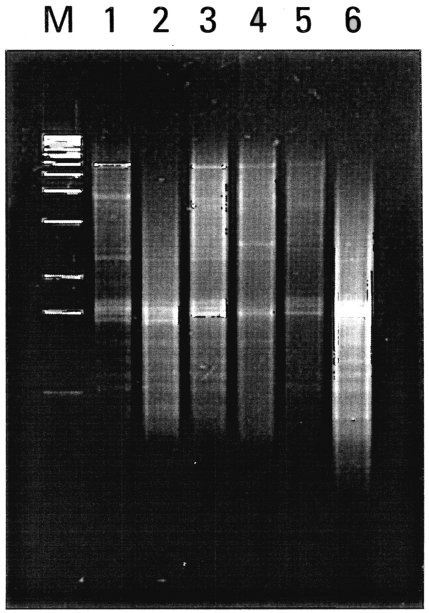


Гель-электрофорез образцов РНК скелетной мышцы человека и полученных на их основе кДНК библиотек



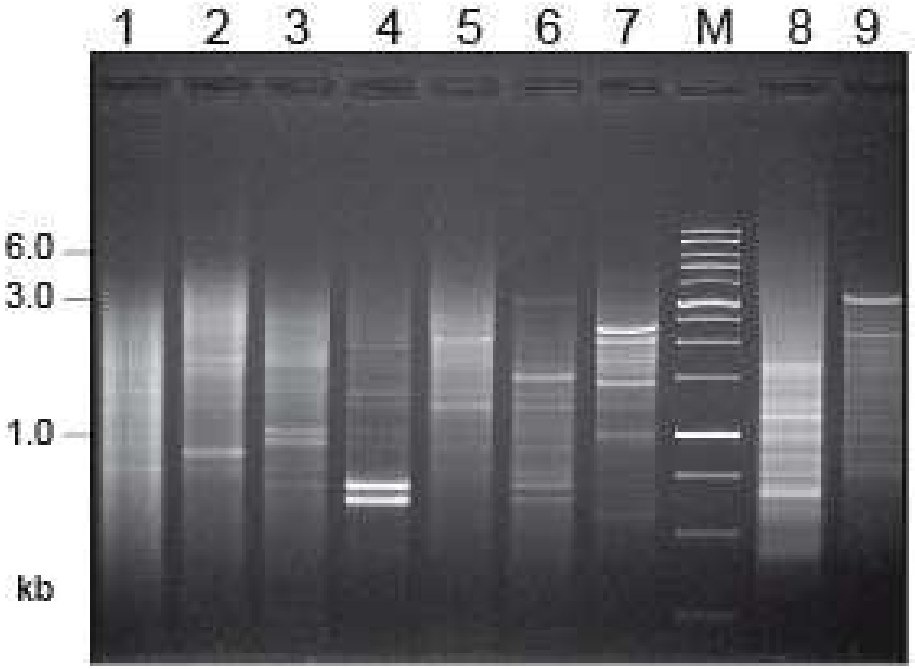
(A) Формальдегидный агарозный гель-электрофорез РНК. 1 - поли(A)+ РНК (1 мкг); 2 - суммарная РНК (1 мкг); M - маркер молекулярных длин, RNA ladder (Life Technologies; poly(A)+ RNA).

(B) Агарозный гель-электрофорез двухцепочечной кДНК, синтезированной по методу (Gubler & Hoffman 1983). Дорожка 1 - синтез с 1 мкг поли(A)+ РНК. Дорожка 2 - синтез с 1 мкг суммарной РНК. Дорожка M - синтез с 500 нг маркера RNA ladder.



Агарозный гель-электрофорез образцов кДНК после ПЦР-амплификации. Синтез исходной двухцепочечной кДНК проводили на основе различных количеств поли(A)+ или суммарной РНК. Дорожки 1, 2 - синтез с 1 мкг (10 циклов ПЦР) и 1 нг (20 циклов ПЦР) поли(A)+ РНК, соответственно. Дорожки 3-6 - синтез с 1 мкг (15 циклов ПЦР), 100 нг (17 циклов ПЦР), 10 нг (21 циклов ПЦР) и 1 нг (25 циклов ПЦР) суммарной РНК, соответственно. Дорожка M - маркер молекулярных длин 1 kb ladder (Gibco BRL).

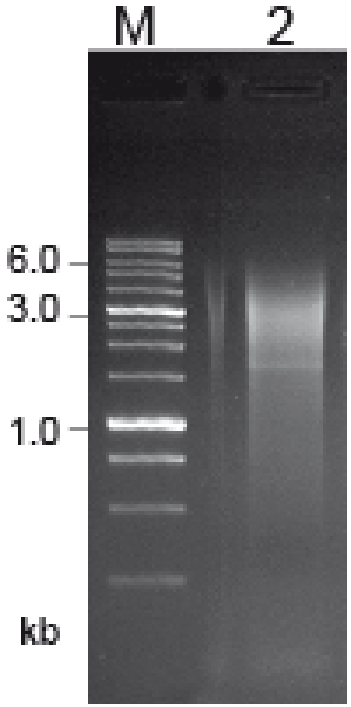
Амплификация образцов кДНК



Amplified cDNA from different sources:
1 - mouse liver; 2 - mouse skeletal muscle; 3 - mouse brain; 4 - human leucocytes; 5 - human lung; 6 - human skeletal muscle; 7 - mosquito grub; 8 - copepod Pontella sp.; 9 - tomato Lycopersicon esculentum. M - 1 kb DNA size marker, SibEnzyme, Russia.

кДНК, как правило, имеет несколько ярких полос, соответствующих высоко-представленным транскриптам.

Вследствие высокой сложности полиА+ фракции РНК из мозга, тимуса и селезенки – кДНК из этих тканей может не иметь ярких полос.

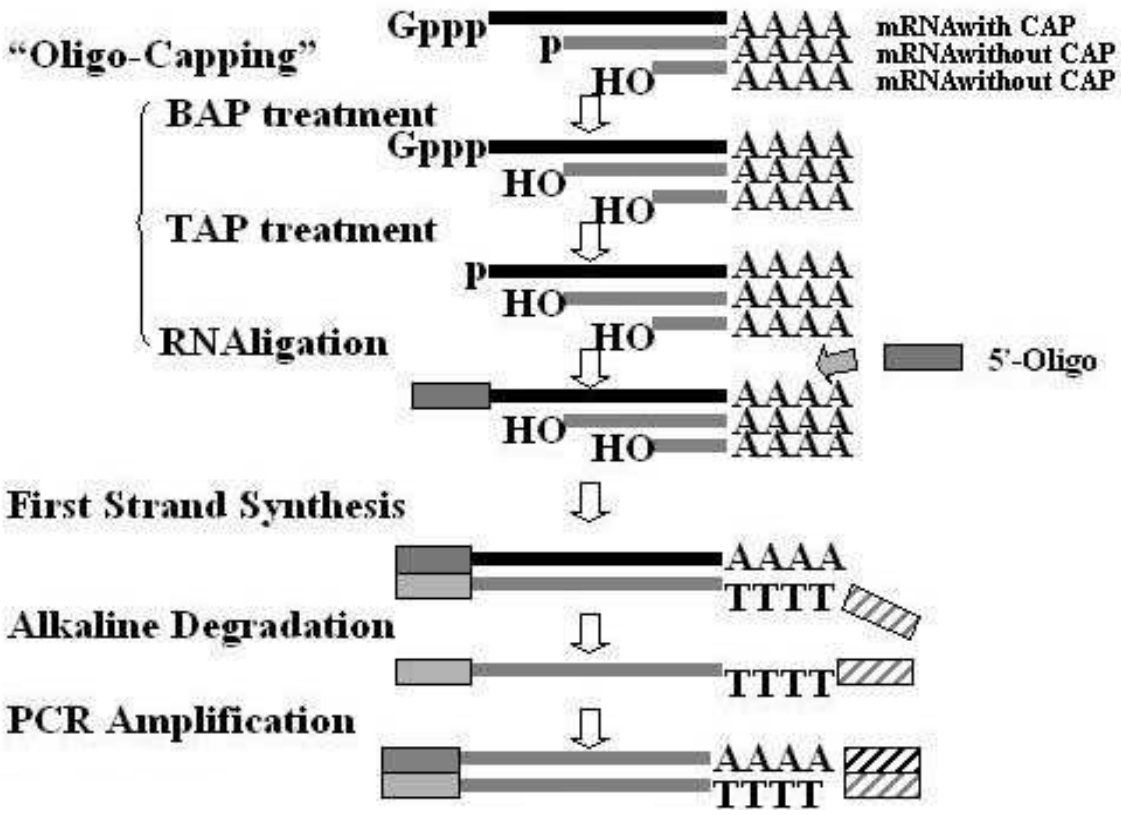


Amplified cDNA from human brain

Недостатки технологии Marathon

- Сложный многостадийный процесс, включающий ряд плохо контролируемых стадий – синтез двуцепочечной кДНК и лигирование адаптеров
- Гарантированная потеря 5’- концевой последовательности РНК

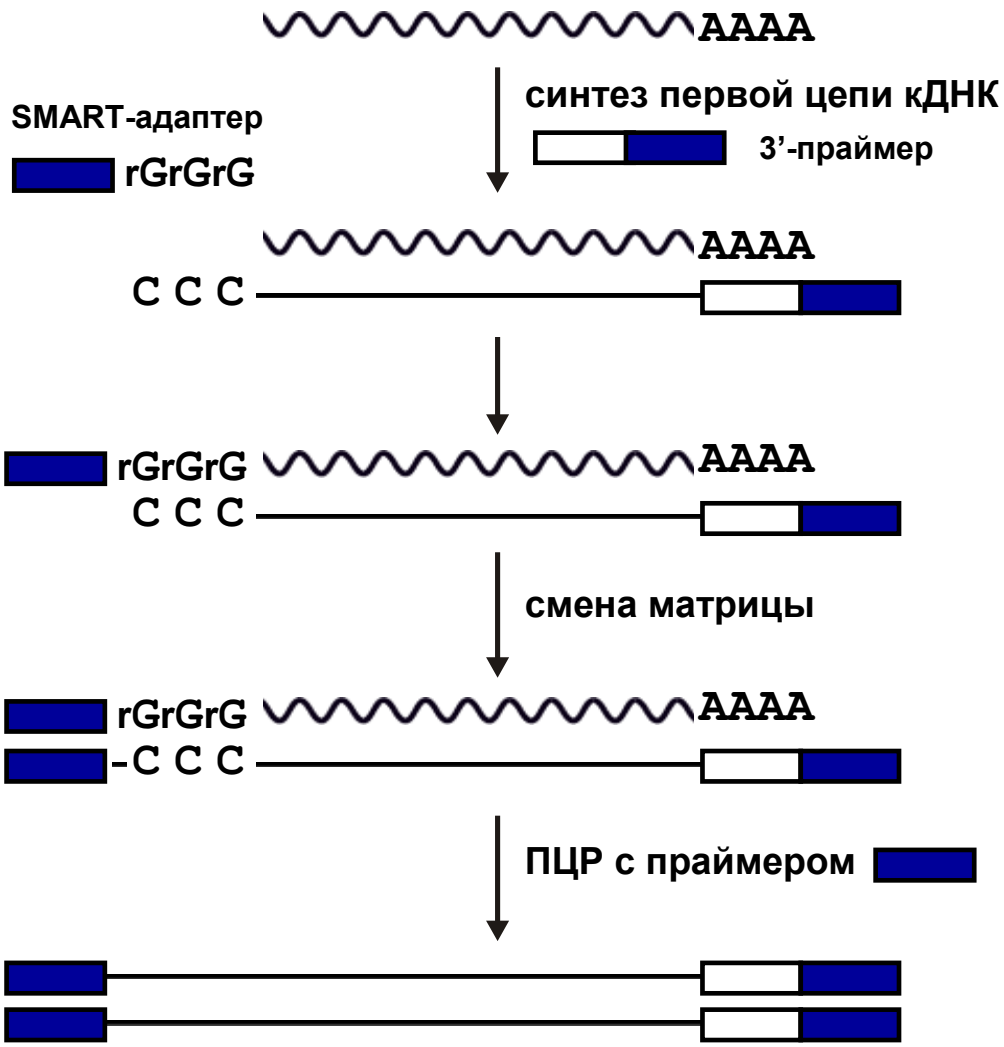
Обогащение полноразмерными последовательностями кДНК – Oligo-Capping метод



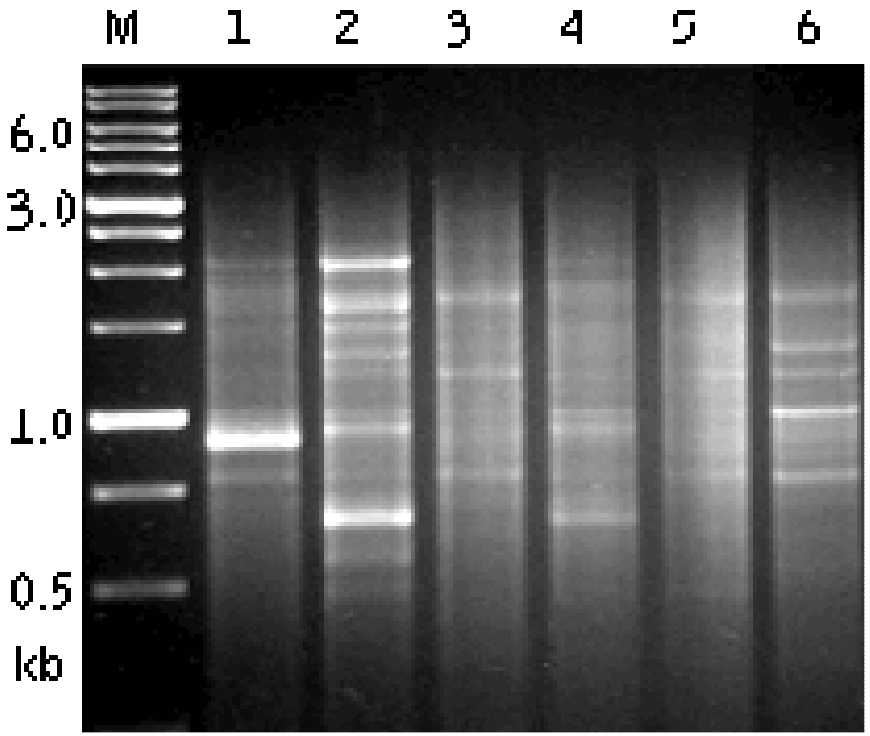
The basis of the method is that RNA is treated with alkaline phosphatase and acid pyrophosphatase:

- Bacterial alkaline phosphatase removes phosphate
- Tobacco acid pyrophosphatase removes the cap
- T4 RNA ligase ligates adapter

Схема приготовления кДНК по технологии SMART-
наиболее простой и эффективной технологии амплификации кДНК



Образцы кДНК, приготовленные методом SMART



ds cDNA synthetized on the basis of total RNA from different human tissues

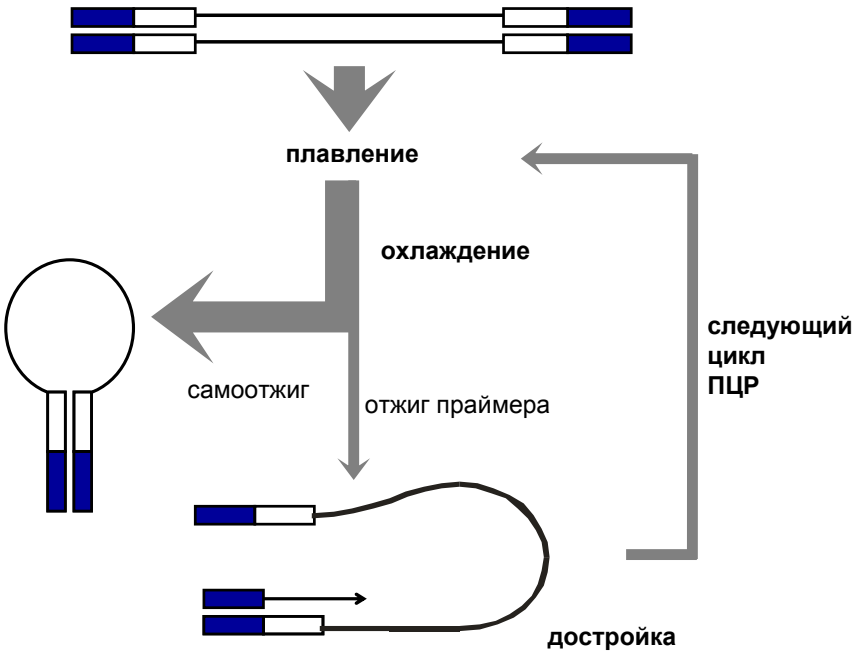
: M, 1-kb DNA size markers (SibEnzyme); lane 1 - placenta; lane 2 - fetal liver; lane 3 - brain; lane 4 - fetal lung; lane 5 - testis; lane 6 - small intestine.

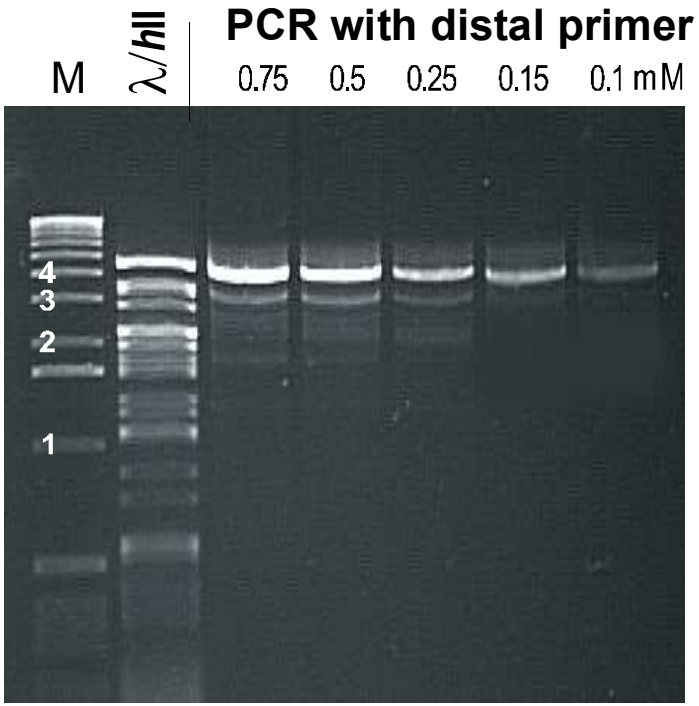
Регуляция средней длины сложных продуктов ПЦР с помощью ССП

adapter 5’-GTAATACGACTCACTATAGGGCACGGTTGCACGACGGCCCGGGCTGGT-3’
3’-CCCGACCA-5’

distal primer 5’-GTAATACGACTCACTATAGGGC-3’

proximal primer 5’-ACGGTTGCACGACGGCCCGGGCTGGT-3’



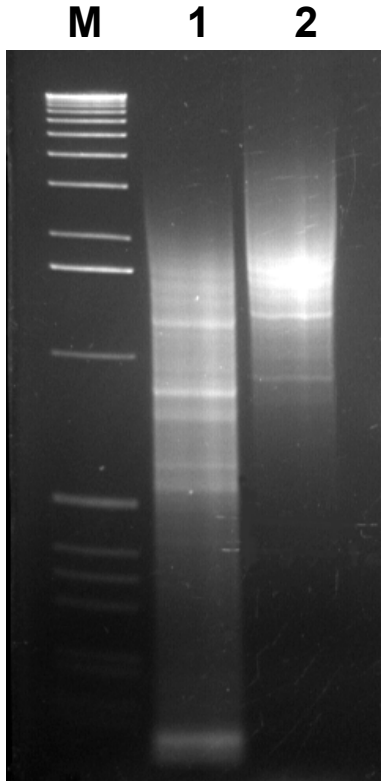


Регуляция длины амплифицированной кДНК

adapter 5’-GTAATACGACTCACTATAGGGCACGGTTGCACGACGGCCCGGGCTGGT-3’
3’-CCCGACCA-5’

distal primer 5’-GTAATACGACTCACTATAGGGC-3’

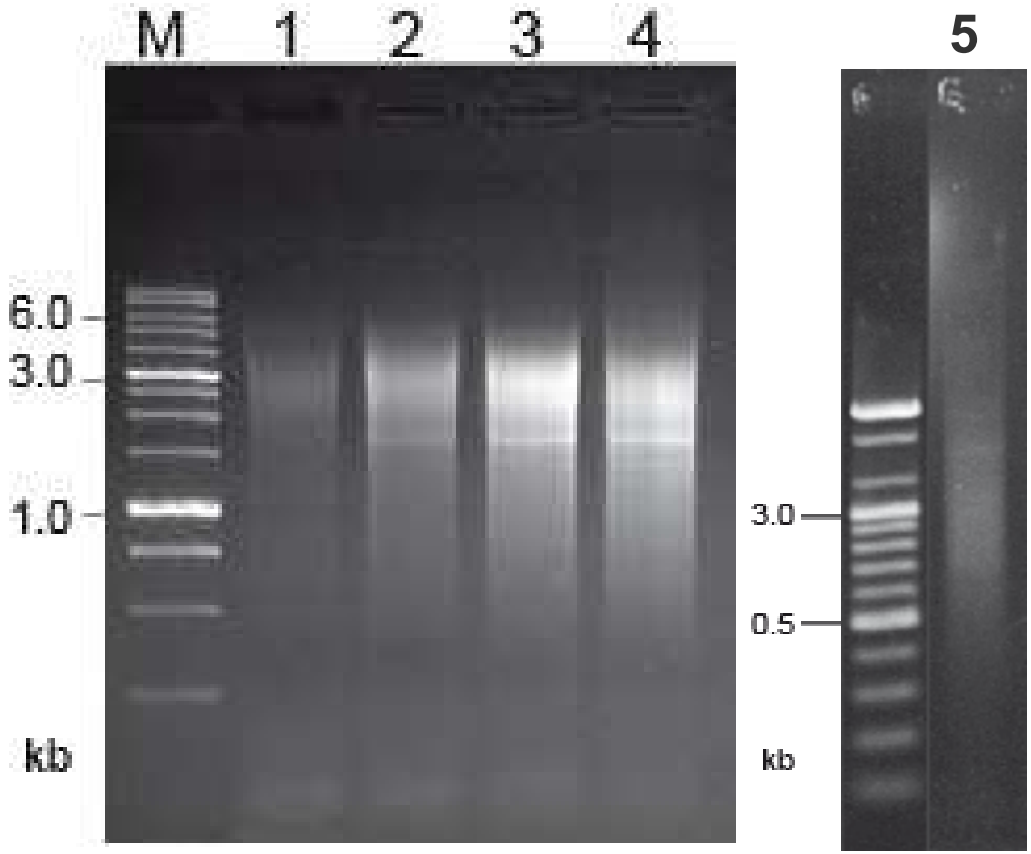
proximal primer 5’-ACGGTTGCACGACGGCCCGGGCTGGT-3’



1. SMART кДНК
2. SMART кДНК после лигирования супрессионного адаптера и амплификации с проксимального праймера

Амплификация образцов кДНК

Отсутствие изменений в концентрации продукта ПЦР при добавлении циклов указывает на то, что реакция вышла на плато. Оптимальное для амплификации образца количество циклов должно быть на 2-3 цикла меньше, чем то, которое необходимо для выхода реакции на плато.
Общая интесивность шмера должна соответствовать той, которую имеют образцы кДНК на дорожках 2-3 (сравнение относительно интенсивности маркера).
Если интенсивность шмера не нарастает с увеличением циклов ПЦР и смещена к лунке (дорожка 4), значит, образец кДНК подвергся избыточной амплификации.
Если интенсивность шмера существенно ниже (например, как на дорожке 1), значит, количество циклов ПЦР было недостаточно для амплификации этого образца.



Отсутствие ярких полос и шмер, начинающийся от лунки, – свидетельство чрезмерного числа циклов ПЦР (дорожка 4) или слишком длинной элонгации в ходе ПЦР (дорожка 5)