# Оглавление

[Введение 3](#_TOC_250016)

[Глава 1. Обзор литературы 5](#_TOC_250015)

* 1. [Эпидемиология злокачественных новообразований (ЗНО) 5](#_TOC_250014)
  2. [Человеческие рецепторы эпидермального фактора роста 7](#_TOC_250013)
  3. [Таргетное лечение ЗНО 8](#_TOC_250012)
     1. Моноклональные антитела 9
     2. Низкомолекулярные ингибиторы 11
     3. Недостатки таргетной терапии 12
  4. [Транскрипция и технологии, помогающие ее изучать 12](#_TOC_250011)
     1. Экстракция РНК 14
     2. Микрочипы на основе ДНК 15
     3. РНК-секвенирование 17

[1.5. Генные подписи как биомаркеры, предсказывающие результаты таргетной терапии ЗНО 19](#_TOC_250010)

[Глава 2. Материалы и методы 21](#_TOC_250009)

* 1. [Транскриптомные и клинические данные когорт пациентов 21](#_TOC_250008)
     1. Экспериментально полученные экспрессионные профили пациентов 23
     2. Экспрессионные профили пациентов из общедоступных наборов данных 24
  2. Анализ дифференциальной экспрессии генов РНК-секвенирования 26
  3. [Алгоритм, предсказывающий эффективность противоопухолевой терапии на основе сформированных генных подписей 27](#_TOC_250007)

[Глава 3. Результаты и обсуждение 29](#_TOC_250006)

* 1. [Поиск дифференциально экспрессированных данных 29](#_TOC_250005)
  2. [Тестирование генных подписей на общедоступных наборах данных для предсказания ответа на терапию 31](#_TOC_250004)
  3. [Модификация генной подписи для увеличения значимости предсказательной модели 35](#_TOC_250003)

[Заключение 36](#_TOC_250002)

[Выводы 37](#_TOC_250001)

[Список литературы 38](#_TOC_250000)

**Аннотация**

Метастазирование злокачественных новообразований является основной причиной смертности от опухолей. Действие противоопухолевых препаратов направлено на ингибирование роста и дифференцировки опухолевых клеток, однако пациенты могут по-разному реагировать на предложенную терапию. Предсказание ответа на терапию с помощью генных подписей может значительно улучшить подбор противоопухолевых препаратов и поможет индивидуализировать лечение. Целью исследования стоит выделение множества дифференциальных генов, высоко- или низко экспрессированных при различных состояниях клетки, для двух клеточных линий плоскоклеточной карциномы А431 и рака молочной железы BT474. Были сформированы генные подписи, которые лучше всего предсказывают ответ опухоли на терапию. Выбранные генные подписи были протестированы на публичных наборах данных, была проанализирована значимость полученных результатов. Далее генная подпись модифицировалась для повышения ее предсказательного потенциала. Как результат, была выделена генная подпись, которая хорошо предсказывает ответ пациентов на таргетную терапию цетуксимабом. Она была модифицирована для улучшения эффективности предсказания противоопухолевой терапии для каждого публичного набора экспрессионных данных. Таким образом, данная подпись может быть использована для прогнозирования ответа пациента на предложенную терапию.

# Введение

Злокачественные новообразования являются одной из наиболее серьезных и широко распространенных проблем в медицине современности.

Сигнальные пути EGFR (эпидермального фактора роста) и ERBB2 (рецептора ростового фактора эпидермиса человеческого) играют ключевую роль в регуляции различных процессов в клетке, включая пролиферацию, выживание, миграцию и дифференцировку. Аномалии в этих сигнальных путях часто наблюдаются в различных типах злокачественных новообразований и могут способствовать его развитию и прогрессированию. Поэтому противоопухолевые препараты, специфичные к EGFR и ERBB2, представляют собой важные инструменты в лечении опухолей, поскольку блокирование этих сигнальных путей может привести к подавлению роста опухоли и угнетению ее выживаемости. Однако, известно, что человеческая сыворотка крови и эпидермальный фактор роста способны модулировать действие препаратов и влиять на их эффективность.

В данном исследовании проведен анализ воздействия сыворотки крови и эпидермального фактора роста на рост опухолевых клеточных линий с использованием методов биоинформатики и генетики. Цель работы заключалась в поиске генных подписей, которые могут предсказать реакцию пациента на противоопухолевую терапию с препаратами, специфичными к EGFR и ERBB2.

Особое внимание уделено оптимизации генных подписей путем удаления лишних генов с целью увеличения точности прогнозирования реакции на различные препараты. Исходя из полученных результатов, были сделаны выводы о влиянии определенных генов на рост клеток опухолей, что может предоставить новые возможности для индивидуализации лечения и повышения эффективности терапии опухолей EGFR и ERBB2.

Исследование акцентирует важность генетических маркеров в прогнозировании реакции на терапию опухолей для развития

персонализированных подходов к лечению опухолей.

**Цель работы:** исследовать влияние сыворотки крови и эпидермального фактора роста на эффективность противоопухолевых препаратов, специфичных к EGFR и ERBB2.

# Задачи работы:

1. Выделить множества дифференциальных генов, высоко- или низко экспрессированных при различных состояниях клетки, для двух клеточных линий плоскоклеточной карциномы А431 и рака молочной железы BT474.
2. Сформировать пересечениями, объединениями и разностями множеств, которые отвечают различным состояниям клетки, генные подписи, выделить те из них, которые лучше всего предсказывают ответ опухоли на терапию.
3. Протестировать выбранные генные подписи на публичных наборах данных с наиболее близкими характеристиками опухоли и процесса лечения. Проанализировать значимость полученных результатов.
4. Оптимизация генной подписи путем удаления генов для повышения значимости определения набора генов, модулирующих действие различных препаратов.
5. На основании полученных результатов сделаны выводы о влиянии выделенных генов на рост и пролиферацию клеток злокачественных новообразований, предложена их потенциальная возможность применения для достижения необходимого ответа на терапию.

# Глава 1. Обзор литературы

# Эпидемиология злокачественных новообразований (ЗНО)

Злокачественные новообразования (ЗНО) являются значительным общественно-здравоохранительным вызовом на глобальном уровне. Исследования в области эпидемиологии опухолей имеют огромное значение для анализа распространенности этого заболевания, его особенностей и тенденций. Эти исследования необходимы для понимания масштабов проблемы ЗНО и разработки эффективных стратегий предотвращения и лечения этого заболевания. С увеличением продолжительности жизни людей в разных странах мира, особенно в странах с высоким уровнем экономического развития, риск развития опухоли также возрастает [[Wynder, Gori, 1977]](https://www.zotero.org/google-docs/?vLhu4j). Это связано с факторами старения населения, изменением образа жизни, в том числе диеты, физической активности, потреблением табака, алкоголя, воздействием окружающей среды, включая загрязнение и облучение [[Ferguson и др., 2015]](https://www.zotero.org/google-docs/?jfUtwn).

Статистика ЗНО представляет собой важный инструмент для понимания его распространенности, влияния и тенденций в различных странах и среди различных групп населения. Данные о новых случаях возникновения ЗНО, смертности от ЗНО, выживаемости и факторах риска помогают в разработке стратегий профилактики, диагностики и лечения этого заболевания. Согласно информации, представленной Всемирной организацией здравоохранения (ВОЗ) [[Global Cancer Observatory, ]](https://www.zotero.org/google-docs/?MhoFLO), на конец 2022 года по подсчетам зарегистрировано более 19 миллионов случаев заболевания ЗНО, при этом смертность от болезни достигла 50%, по сравнению с по сравнению с почти 14 миллионами случаев ЗНО и 8,2 миллионами смертей в 2012 году. Среди распространенных типов ЗНО у мужчин выделяются рак легких (15,2%), предстательной железы (14,2%) и толстой кишки (10,4%), а у женщин - молочной железы (23,8%), легких (9,4%) и толстой кишки (8,9%). Наиболее смертельными же являются рак легких (22,7%), печени (9,6%) и толстой кишки (9,2%) у мужчин и рак молочной железы (15,4%), легких (13,5%) и толстой кишки (9,4%) у женщин. За жизнь

каждого пятого человека на Земле диагностируется ЗНО, причиной смерти от него становится каждый восьмой мужчина и каждая одиннадцатая женщина. Общее количество людей, у которых диагностирована опухоль и которые живы на протяжении последних пяти лет, составляет 53,5 миллиона. В 2050 году ожидается увеличение количества новых случаев заболевания ЗНО примерно на 77% по сравнению с 2022 годом, то есть 35,3 миллиона случаев [[WHO: Global](https://www.zotero.org/google-docs/?SymZeZ) [Cancer Burden Rising Amid Striking Inequities, ]](https://www.zotero.org/google-docs/?SymZeZ).

Уровень заболеваемости и смертности от ЗНО варьироваться от региона к региону. Наибольшие значения достигаются для азиатских стран (почти половина всех случаев заболеваемости и смертности), европейские страны стоят на втором месте. Далее по мере убывания уровней заболеваемости и смертности идут страны Америки, Африки, Австралии и Океании. Многие регионы планеты не обладают необходимой точностью представления данных об уровнях распространения ЗНО среди населения, что затрудняет мониторинг и скрининг этой болезни по всему миру.

Распространенность злокачественных новообразований (ЗНО) в обществе зависит от различных факторов, которые влияют на качество жизни и благополучие населения. Эти переменные существенны для определения Индекса Развития Человеческого Потенциала (ИРЧП). Наблюдается значительное неравенство в распространении опухолей в странах с разным ИРЧП. В странах с высоким уровнем экономического развития 1 из 12 женщин в течение своей жизни столкнется с диагнозом рака груди, причем 1 из 71 женщины умрет от этого заболевания. В отличие от этого, в странах с низким уровнем экономического развития лишь у 1 из 27 женщин будет диагностирован рак молочной железы, но при этом 1 из 48 женщин умирает от него. Женщины в странах с низким ИРЧП имеют в два раза меньшую вероятность получить диагноз рака молочной железы, нежели их сверстницы в странах с высоким уровнем экономического развития, однако сталкиваются с гораздо более высоким риском смерти от этого заболевания из-за отсроченной

диагностики и ограниченного доступа к качественному лечению [[Global Cancer](https://www.zotero.org/google-docs/?ztNxvO) [Observatory, ]](https://www.zotero.org/google-docs/?ztNxvO).

# Человеческие рецепторы эпидермального фактора роста

Человеческое семейство эпидермальных факторов роста (HER) включает в себя 4 мембранных белка (ErbB1, ErbB2, ErbB3, ErbB4), играющих важную роль в регуляции роста, выживаемости и дифференцировки клеток [[Olayioye и](https://www.zotero.org/google-docs/?Ae5pyn) [др., 2000]](https://www.zotero.org/google-docs/?Ae5pyn). Все белки имеют одинаковый внеклеточный домен, который при связывании лиганда запускает процесс гомо- или гетеродимеризации белка [[Holbro, Hynes, 2004]](https://www.zotero.org/google-docs/?S8OSxC). После димеризации активируется тирозинкиназная активность внутриклеточного домена HER-белка, вследствие чего происходит аутофосфорилирование остатков аминокислоты, что запускает сигнальный каскад реакций, отвечающих определенным генетическим программам. На рецепторы эпидермальных факторов роста действуют различные наборы лигандов, обуславливая разные сигналы, подаваемые белками внутрь клетки.

Избыточная или недостаточная экспрессия генов и нарушения регуляции внутриклеточных сигнальных путей у членов семейства HER играют важное значение в развитии ЗНО, и гиперактивирующие мутации EGFR были обнаружены при различных формах онкологических заболеваний [[Brockhoff:](https://www.zotero.org/google-docs/?XQfaDI) [Differential impact of Cetuximab, Pertuzumab... - Академия Google, ]](https://www.zotero.org/google-docs/?XQfaDI). Так, например, гиперэкспрессия некоторых рецепторов (например, HER2-рецептора, то есть ErbB2 белка) является причиной появления до 20% опухолей молочной железы, что указывает на необходимость исследования дифференциально экспрессированных генов и важность анализа вклада этих генов в развитие опухоли. Для терапии гиперэкспрессированного HER2 рака молочной железы (сокращенно: HER2+ BC) одобрено применение моноклональных антител, таких как трастузумаб, пертузумаб, лапатиниб. Но проблема состоит в том, что пациенты с повышенным уровнем HER2 реагируют на эти препараты по-разному, и это лечение не всегда является эффективным. У некоторых

пациентов развивается резистентность к терапии, некоторые образцы опухолей становятся нечувствительными к предложенной терапии.

Тем не менее, противоопухолевые препараты, целенаправленно действуя на HER-рецепторы, решают важную проблему онкологии. Они специфически воздействуют на HER-рецепторы и могут модулировать сигнальные пути, инициируемые этими белками, что может привести к подавлению роста опухоли или даже к индукции апоптоза клеток опухоли. Понимание механизмов взаимодействия между противоопухолевыми препаратами и HER-рецепторами имеет критическое значение для разработки эффективных стратегий лечения ЗНО, особенно тех опухолей, где гиперактивация HER-сигнального пути играет существенную роль в онкогенезе [[HER2-targeted therapies — a role beyond breast](https://www.zotero.org/google-docs/?fP6iCb) [cancer | Nature Reviews Clinical Oncology, ]](https://www.zotero.org/google-docs/?fP6iCb).

# Таргетное лечение ЗНО

К основным методам лечения ЗНО относят использование хирургического вмешательства, химиотерапии и лучевой терапии. Говоря о химиотерапии, она сталкивается с недостатками, связанными с низкой узнаваемостью опухолевых клеток по сравнению со здоровыми клетками пациента, вызывая ряд нежелательных реакций и развитие резистентности опухолевых клеток. В настоящее время таргетная терапия, которая направлена на конкретные мишени в опухолевых клетках, становится все более важным методом лечения [[Padma, 2015]](https://www.zotero.org/google-docs/?ziIK9a). Поэтому этот вид терапии обладает большей точностью в действии на опухолевые клетки, минимизируя при этом воздействие на обычные клетки.

Таргетная терапия для лечения ЗНО включает применение различных препаратов, таких как низкомолекулярные ингибиторы и моноклональные антитела, которые направлены на определенные молекулярные мишени в опухолевых клетках [[Gerber, 2008]](https://www.zotero.org/google-docs/?4FY0lO) [[Yewale и др., 2013]](https://www.zotero.org/google-docs/?zIzRHz). Но выбор самого препарата зависит от многих факторов, включая его специфику и рекомендации лечащего врача. С недавними колоссальными прорывами в области медицины и

биологии, человечество узнало намного больше о ДНК, белках и особенностях опухолей, поэтому лечение становится более точным и корректируется от пациента к пациенту Частота и продолжительность таргетной терапии зависят от конкретного типа опухоли, его стадии, а также индивидуальных особенностей пациента и его реакции на лечение. Для некоторых видов таргетной терапии характерно проведение циклов с периодами отдыха для восстановления организма.

Реакция на таргетную терапию может варьироваться у разных пациентов, что во многом зависит от их общего состояния здоровья до начала лечения, вида опухоли, степени ее распространенности, а также конкретного препарата. Поэтому невозможно точно предсказать, как именно пациент будет чувствовать себя во время проведения терапии. Важным аспектом проведения таргетной терапии является регулярное наблюдение врачей и проведение медицинских обследований для оценки эффективности лечения и корректировки стратегии при необходимости. Таргетная терапия уникальна своей доставкой с помощью активирующих лекарство генов, вирусов или антител (GDEPT, VDEP, ADEP соответственно) внутрь опухоли с последующим введением лекарства, которое за счет первого этапа будет приобретать специфичность к клеткам ЗНО [[Padma,](https://www.zotero.org/google-docs/?alcYZF) [2015]](https://www.zotero.org/google-docs/?alcYZF). Моноклональные антитела и низкомолекулярные ингибиторы являются основными семействами препаратов, использующихся для таргетной терапии.

* + 1. Моноклональные антитела

Моноклональные антитела (МАТ) - это белки, которые относятся к семейству гликопротеинов суперсемейства иммуноглобулинов, которые участвуют в иммунном ответе B-клеток на чужеродные вещества. Они состоят из двух тяжелых и двух легких цепей, в состав которых входит вариабельный домен. Именно этот компонент моноклональных антител помогает им достигать специфичности [[Stern, Herrmann, 2005]](https://www.zotero.org/google-docs/?aDCiJv). Первые моноклональные антитела были получены из мыши в 1975 году, однако они вызывали иммунный ответ. С того времени велись многочисленные эксперименты по улучшению качества

препарата для достижения желаемого ответа пациента. Вскоре появились химерные МАТ, в которых постоянную невариабельную область цепи составляют человеческие иммуноглобулины. У них удлиненный срок существования в организме человека и уменьшенная способность вызывать иммунный ответ, однако они все еще способны опосредовать индукцию антинаркотических антител. Гуманизированные МАТ в некоторой степени обходят эту проблему, так как он мышиного антитела извлекается только гипервариабельный домен, а 95% антитела является человеческим. С появлением фагового дисплея и с помощью трансгенных технологий для внедрения вариабельного домена человека в мышиный геном получилось сделать МАТ на полностью человеческими, что делает их максимально близкими по действию и эффективности к иммуноглобулинам человека IgG [[Buss и др., 2012]](https://www.zotero.org/google-docs/?vMoea6).

Моноклональные антитела способны блокировать молекулы, необходимые клеткам ЗНО для роста, помечать опухолевые клетки для их дальнейшего уничтожения иммунной системой организма или доставлять препарат к опухолевым клеткам. Например, трастузумаб присоединяется на поверхности некоторых опухолевых клеток к HER2 - рецептору и блокирует сигнал, направленный на рост и пролиферацию [[Stern, Herrmann, 2005]](https://www.zotero.org/google-docs/?CJcAZS). Другой пример - это ретуксимаб, который помечает опухолевые клетки, связываясь на их поверхности с рецепторами CD20, и подает сигнал для уничтожения таких клеток иммунной системой [[Buss и др., 2012]](https://www.zotero.org/google-docs/?XMQgNg) [[Sun и др., 2015]](https://www.zotero.org/google-docs/?93P2HE). VEGF - молекула, которая заставляет сосуды расти. МАТ бевацизумаб блокирует VEGF и вследствие восстанавливает рост кровеносных сосудов, необходимых для выживания опухоли [[Garcia и др., 2020]](https://www.zotero.org/google-docs/?qU50dJ). Исследователи ЗНО продолжают изучать новые способы использования высокоточных МАТ для лечения онкологически больных. Эти соединения достаточно крупные, поэтому, как правило, вводятся внутривенно.

* + 1. Низкомолекулярные ингибиторы

Открытие в 1960 году Филадельфийской хромосомы и дальнейшее изучение явления транслокации, в частности образование онкогена bcr-abl и соответствующего ему белка, которые являются причиной до 95% случаев хронического миелолейкоза из-за повышенной активности тирозинкиназы, стало началом изучения и синтеза неких веществ, направленных на борьбу с опухолевыми клетками [[Erp van, Gelderblom, Guchelaar, 2009]](https://www.zotero.org/google-docs/?PpRrRd).

Связывание молекулы с внеклеточной частью EGFR приводит к образованию димеров или олигомеров этого рецептора, что инициирует активацию цитоплазматической тирозинкиназы и добавление фосфорильных групп к тирозиновым остаткам. Однажды фосфорилированные, эти тирозиновые остатки служат платформой для привлечения специфических белков, что в свою очередь регулирует разнообразные клеточные сигнальные реакции, такие как пролиферация клеток, апоптоз, миграция клеток и активация транскрипции генов [[Schlessinger, 2000]](https://www.zotero.org/google-docs/?T9rNZ7) [[Ebrahimi и др., 2023]](https://www.zotero.org/google-docs/?vYadGK).

Другой класс терапевтических агентов, нацеленных на EGFR, представлен низкомолекулярными ингибиторами. Эти ингибиторы могут действовать на различные мишени внутри клеток, такие как ферменты, рецепторы и сигнальные молекулы, которые играют ключевую роль в прогрессировании опухоли. Примером низкомолекулярных ингибиторов, широко используемых в лечении ЗНО, являются ингибиторы тирозинкиназы (TKI). Их механизм действия отличается от МАТ тем, что TKI воздействуют на тирозинкиназный домен EGFR рецептора внутри клеток, затрудняя таким образом передачу сигналов и активацию роста клеток, ангиогенеза и метастазирования.

Несколько ингибиторов тирозинкиназы, таких как эрлотиниб, гефитиниб, афатиниб, осимертиниб и момелотиниб, получили свое применение для терапии метастатической аденокарциномы легких. Ингибитор алмонертиниб способен подавлять пролиферацию и миграцию, а также стимулировать апоптоз в клетках

немелкоклеточного рака легкого.

Кроме того, низкомолекулярные ингибиторы могут использоваться в комбинации с другими средствами лечения ЗНО, такими как химиотерапия, иммунотерапия и радиотерапия, чтобы усилить эффект и повысить эффективность терапии.

* + 1. Недостатки таргетной терапии

Существуют некоторые ограничения у таргетной терапии, которые важно учитывать. Во-первых, клетки ЗНО могут приобретать устойчивость к таргетной терапии, например, когда целевая мишень изменяется, меняет свои пути активации или блокирования, и препарат перестает воздействовать на нее, или когда клетки приобретают устойчивость, находят новые пути для роста и пролиферации [[Corcoran, O’Driscoll, 2015]](https://www.zotero.org/google-docs/?5IGBP0) [[Lamorte, Park, 2001]](https://www.zotero.org/google-docs/?aS6SfO). В таких случаях эффективность таргетной терапии может быть улучшена комбинированным применением нескольких видов терапий или совместно с химиотерапией и облучением. Еще одним вызовом является сложность разработки препаратов для некоторых мишеней из-за структурно-функциональных особенностей и трудностей с доставкой определенных веществ.

# Транскрипция и технологии, помогающие ее изучать

Транскрипция представляет собой процесс передачи генетической информации, закодированной в ДНК, к РНК, чаще всего в форме молекул матричной РНК, путем взаимодействия с определенными белковыми ферментами. Фермент РНК-полимераза связывается с конкретным участком генома в последовательности ДНК, разделяет двойную спираль ДНК, чтобы начать транскрипцию с одной из цепей ДНК. На этой цепи РНК-полимераза синтезирует новую, дочернюю РНК-цепь, используя комплементарные нуклеотиды матричной цепи в качестве шаблона. Как результат, получается цепь РНК, где вместо тимина ДНК присутствует нуклеотид урацил, при этом цепь является копией матричной цепи. Эта РНК кодирует необходимую для

дальнейшей трансляции и других процессов информацию, которая надежно хранится и в дальнейшем может быть использована [[Cooper, 2000]](https://www.zotero.org/google-docs/?RFBAeT). В этом процессе мРНК действует как промежуточная молекула в информационной цепи, а некодирующие РНК выполняют функции регуляции транскрипции и трансляции генов, контроль за стабильностью мРНК, модуляцию хроматина и вовлечение в клеточные сигнальные пути. Транскриптом представляет мгновенный снимок общего количества транскриптов, присутствующих в клетке в определенный момент времени.

Транскрипцию можно исследовать с помощью микрочипов и высокопроизводительного секвенирования. Использование технологии микрочипов позволяет одновременно проводить массовое параллельное измерение различных связанных событий. Микрочипы обычно содержат множество маленьких точек, каждая из которых содержит одинаковые зонды. Эти зонды представляют собой олигонуклеотиды или комплементарные ДНК, что позволяет им связываться с целевыми молекулами путем гибридизации комплементарных цепей [[Lowe и др., 2017]](https://www.zotero.org/google-docs/?69Zhv2) [[Bier и др., 2008]](https://www.zotero.org/google-docs/?ErlJTZ). Для идентификации событий связывания используют флуоресцентные метки, которые навешивают на образец до его инкубации и введения на подложку микрочипа. После инкубации чип промывают, затем данные считываются с помощью сканирующего устройства или устройства для создания изображения, чаще всего в сухом виде, что позволяет получить мгновенное изображение интенсивности флуоресценции и таким образом определить интересующие нас последовательности.

РНК-секвенирование - второй, но не менее популярный вид анализа транскриптомных данных, который заключается в применении передовой технологии секвенирования для быстрого и всестороннего определения транскрипционного состояния биологических образцов в конкретный момент времени [[Zhang, He, Cai, 2018]](https://www.zotero.org/google-docs/?eRyrZc) РНК-секвенирование имеет преимущество над микрочипами при изучении транскриптома в том смысле, что позволяет анализировать всю РНК-цепь полностью, а также дает возможность изучать

новые РНК-молекулы, идентифицировать альтернативный сплайсинг и тому подобные уникальные фрагменты цепи.

Расшифровка транскриптома помогает провести сравнение между референсным человеческим транскриптомом и транскриптомом больного и выделить интересующие исследователя участки, которые могут провоцировать возникновение болезни. Сборка базы транскриптомных данных является важной задачей, так как помогает выделить паттерны, характерные для того или иного заболевания в той или иной ткани человека.

* + 1. Экстракция РНК

Перед анализом транскриптома необходимо провести соответствующие приготовления по извлечению РНК из организма, что можно делать различными методами в зависимости от типа ткани и требуемого количества выделенной РНК [[Suntsova и др., 2019]](https://www.zotero.org/google-docs/?TsNz8O). Выделенные хирургическим вмешательством клетки подвергаются фиксации формалином и заливаются парафином для сохранения структуры ткани [[Chomczynski, Sacchi, 2006]](https://www.zotero.org/google-docs/?cHTEjW). Потом клетки подвергаются механическому воздействию для разрушения целостности мембраны, далее происходит угнетение активности РНКазы путем добавления хаотропных солей для предотвращения разрушения РНК. После лизирования РНК с помощью специальных веществ очищали от разных ненужных макромолекул и загрязнений, далее осаждали РНК центрифугированием, потом промывали раствор и ресуспендировали полученный образец для дальнейшего исследования [[Chomczynski, Sacchi, 1987]](https://www.zotero.org/google-docs/?UKPdOT). Чтобы избавиться от остатков ДНК и тем самым повысить выход РНК, дополнительно обрабатывают полученный образец ДНКазой для расщепления оставшейся ДНК [[Green, Sambrook, 2019]](https://www.zotero.org/google-docs/?9Jwctn). Если необходимо удалить рибосомальную РНК, которая по массе занимает большую часть всей выделенной РНК, можно использовать метод аффинной очистки полиаденилированной матричной РНК с помощью олиго-дезокситиминов. Однако таким образом теряются и другие неполиаденилированные РНК. Чтобы этого избежать, используют метод

связывания целевой РНК с модифицированными нуклеотидами с биотиновой меткой для отделения нужной РНК от других видов РНК. Далее происходит выделение целевой РНК, связанной с биотиновыми метками, с помощью магнитных шариков со стрептавидином, который специфично связывается с биотином [[Bryant, Priest, Mockler, 2012]](https://www.zotero.org/google-docs/?yNUtVa).

* + 1. Микрочипы на основе ДНК

Матрицы ДНК представляют собой плоские поверхности, на которых расположены нуклеиновые кислоты в форме точечных образцов. Каждая точка содержит специфическую нуклеиновую кислоту, представленную цепью длиною в 400-1000 нуклеотидов, отличающейся от других нуклеиновых кислот в матрице. Эти нуклеотиды обычно включают аденин (A), цитозин (C), гуанин

(G) и тимин (T). Такие нуклеиновые кислоты, прикрепленные к положительно заряженной подложке, называются “зондами”. Выделенные ранее РНК-образцы взаимодействуют с твердой поверхностью чипа из стекла, кремния или полимера, например, олигонуклеотиды или комплементарная ДНК связываются с целевыми молекулами через гибридизацию. Детектировать количество связанной РНК на каждом из зондов можно путем сравнения интенсивностей флуоресценции каждого участка, для чего необходимо флуоресцентно пометить транскрипты перед инкубацией на подложке [[Bilitewski, 2009]](https://www.zotero.org/google-docs/?MjvhyO). Основная проблема данного метода заключается в том, что для определения генной экспрессии необходимо иметь некоторые представления о транскриптоме исследуемого организма для соответствующего дизайна зондов.

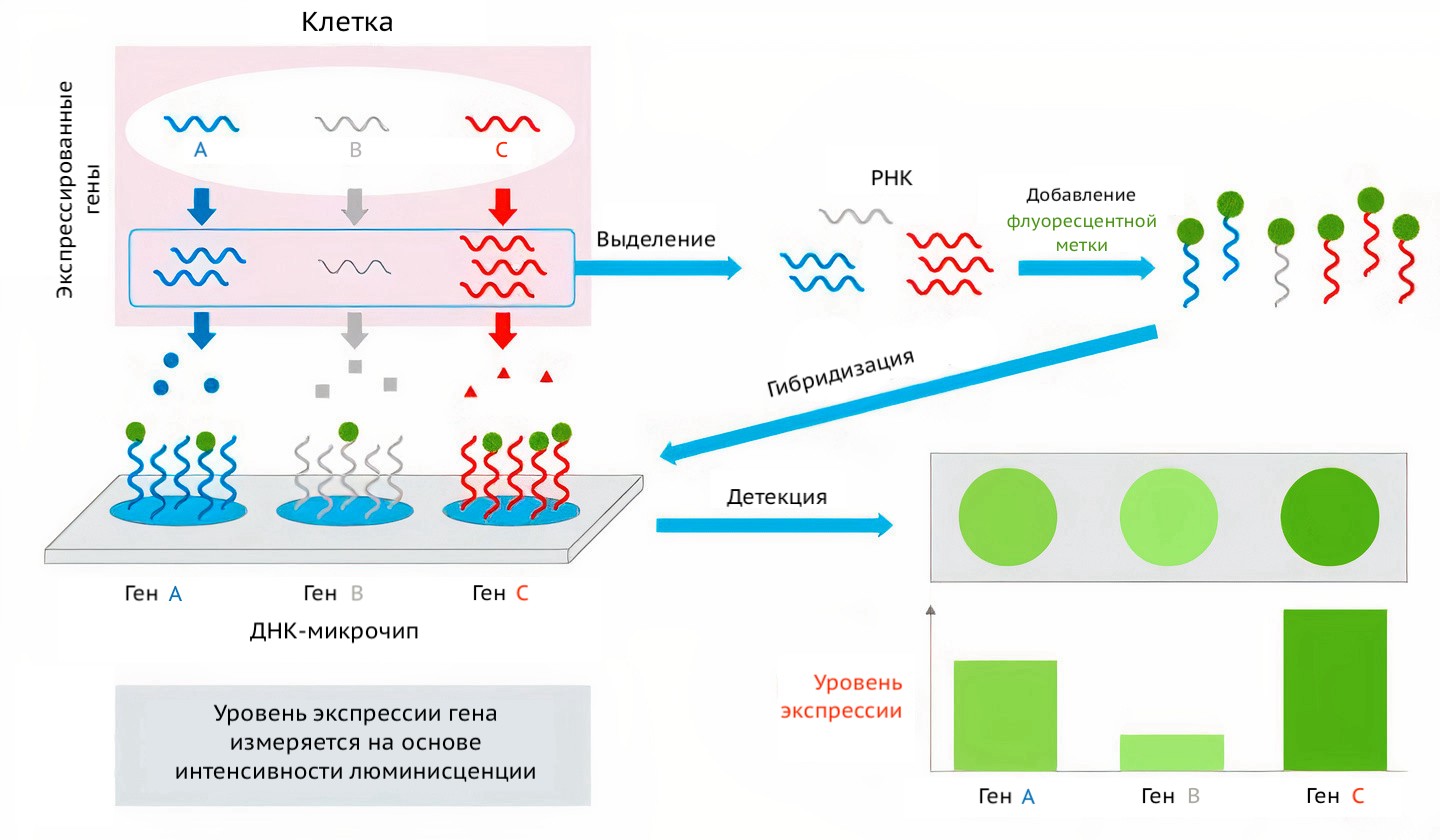


Рисунок 1. Схема метода обнаружения генов с помощью матриц ДНК [[Outline of detection method of genes by DNA microarrays | 3D-Gene® [Toray DNA](https://www.zotero.org/google-docs/?vIYw67) [Chips] | TORAY, ]](https://www.zotero.org/google-docs/?vIYw67).

В наше время используются микрочипы высокой плотности, например компании Affymetrix. Они основаны на параллельном проведения гибридизации большого количества олигонуклеотидов. Они начали разрабатываться в начале 1990-х годов и объединяют фотолитографию с синтезом ДНК для анализа транскриптома [[Heller, 2002]](https://www.zotero.org/google-docs/?NVPMjS). Другой вид матриц, точечные матрицы низкой плотности, содержат капли очищенных кДНК на прозрачной подложке. Зонды в таких матрицах длиннее, но обычно недостаточно разрешения для полного детектирования последовательностей. После гибридизации с флуоресцентными зондами, которые содержали обратно транскрибированные мРНК человека и контрольные арабидопсисы (модельный организм как стандарт для сравнения и оценки точности результатов), происходит сканирование сигналов и определение их интенсивности. Дифференциальная экспрессия определяется как среднее значение соотношений флуоресценции для двух независимых гибридизаций [[Parallel human genome analysis: microarray-based expression](https://www.zotero.org/google-docs/?d80F10)

[monitoring of 1000 genes., ]](https://www.zotero.org/google-docs/?d80F10).

* + 1. РНК-секвенирование

Секвенирование ДНК нового поколения дало возможность секвенировать и РНК с помощью обратной транскрипции в кДНК. Эта технология решила проблему ДНК-чипов, если необходимо обработать сигналы с широкими разбросами по интенсивности [[Ozsolak, Milos, 2011]](https://www.zotero.org/google-docs/?Zfg6bH). Созданные нуклеотидные последовательности или прочтения обычно составляют около 100 пар нуклеотидов в длину, но это значение может варьироваться от 30 до более чем 10 000 пар нуклеотидов в зависимости от конкретного метода секвенирования. Эти последовательности играют важную роль в молекулярной биологии и генетике, позволяя исследователям анализировать геномы организмов, выявлять мутации, идентифицировать гены и многое другое.

Первым методом РНК-секвенирования было секвенирование по Сэнгеру. При использовании метода секвенирования по Сэнгеру происходит сложная последовательность манипуляций с ДНК. Сначала амплифицированную ДНК или кДНК подвергают отжигу с использованием олигонуклеотидного праймера. После этого происходит удлинение с помощью фермента ДНК-полимеразы, который включает в себя либо все четыре дезоксинуклеотидтрифосфата либо дидезоксинуклеотидтрифосфаты. Добавление последних ограничивает скорость реакции удлинения за счет особенной структуры нуклеотидов, что приводит к остановке удлинения по мере их включения. В результате такой манипуляции формируются фрагменты ДНК различной длины, которые могут быть однозначно идентифицированы [[Ozsolak, Milos, 2011]](https://www.zotero.org/google-docs/?eohrKG).

Новое поколение секвенирования РНК заключается в синтезе кДНК на основе РНК, фрагментации и добавлении адаптеров к каждому фрагменту, амплификации и синтезе на матрице этих фрагментов комплементарных им участков, которые впоследствие детектируются [[Wang, Gerstein, Snyder, 2009]](https://www.zotero.org/google-docs/?m1Etgh). Полученные прочтения выравниваются на референсную последовательность, а количество прочтений указывает на интенсивность экспрессии того или иного участка РНК. В РНК-секвенировании используются нанограммы образца, в

отличие от микрограмм в ДНК-микрочипах, что является одним из

преимуществ использования первой технологии для изучения транскриптомного профиля организма. Также РНК-секвенирование отличается тем, что может быть использовано для поиска событий альтернативного сплайсинга [[Ozsolak, Milos, 2011]](https://www.zotero.org/google-docs/?r9qnyA), идентификации уникальных перестроек в геноме, однонуклеотидных полиморфизмов и зависящих от аллели различий в генах. Именно из-за этих преимуществ появившееся в 2006-2008 годах РНК-секвенирование вытеснило в начале 2010 годов метод микрочипов [[Hu и](https://www.zotero.org/google-docs/?mxfEn3) [др., 2021]](https://www.zotero.org/google-docs/?mxfEn3).

РНК-секвенирование нового поколения позволяет генерировать последовательности размером более 10 килобайт прямо из выделенной РНК. Так как РНК трансформируется в кДНК перед началом фрагментации, репарации концов, лигирования адаптеров и амплификации фрагментов, то платформа для обоих видов данных является одинаковой [[Hu и др., 2021]](https://www.zotero.org/google-docs/?X9hDoM). Фрагментация является важным этапом при приготовлении библиотеки последовательностей, так как наличие большого числа правильных фрагментов обеспечивает высокий выход секвенирования и высокую точность прочтений. Фрагментирование ДНК достигается через сдвиг с помощью сжатого азота или воздуха, ультразвуковое воздействие и ферментативное расщепление [[Knierim и](https://www.zotero.org/google-docs/?9AW6Tg) [др., 2011]](https://www.zotero.org/google-docs/?9AW6Tg). Амплификация с помощью ПЦР фрагментов ДНК нужна для увеличения количества фрагментов, чтобы можно было получать высокий выход секвенирования при малом количестве образца [[Parekh и др., 2016]](https://www.zotero.org/google-docs/?UA7WHn). Существует парноконцевое и непарное прочтение ридов, то есть секвенирование в одном или двух направлениях; они выбираются в зависимости от задач, необходимой скорости и цены. Первый вариант является более предпочтительным для тщательного и точного изучения транскриптомных данных, так как минимизирует количество ошибок и более достоверно предсказывают гены у немодельных, малоизученных организмов [[Ozsolak, Milos, 2011]](https://www.zotero.org/google-docs/?1Q5pLp).

# 1.5. Генные подписи как биомаркеры, предсказывающие результаты таргетной терапии ЗНО.

Прогностические показатели эффективности таргетной терапии при лечении злокачественных новообразований (ЗНО) играют ключевую роль в оценке терапевтических результатов у пациентов. Таргетированная терапия, направленная на конкретные молекулярные цели в опухоли, стала одним из важных методов борьбы с ЗНО. Однако не все люди реагируют на такие препараты одинаково, поэтому важно выявить прогностические индикаторы, способные предсказать эффективность данного вида лечения.

Исследования продемонстрировали, что конкретные биомаркеры способны предсказывать реакцию пациента на таргетную терапию, что в свою очередь помогает персонализировать лечение [[Guo и др., 2022]](https://www.zotero.org/google-docs/?v1Ixyd). Примерами таких биомаркеров могут служить мутации в генах, ответственных за клеточный рост и деление, уровень экспрессии определенных белков, амплификация генов и другие характеристики опухолевых клеток.

Научные исследования показывают, что использование прогностических биомаркеров способствует выбору подходящих кандидатов для таргетированной терапии, предсказывает возможные побочные эффекты и помогает оценить вероятность успешного лечения. Это в свою очередь помогает избежать назначения неэффективных терапий и повышает эффективность лечения у пациентов с ЗНО.

Систематизация данных секвенирования является важной задачей на пути к пониманию паттернов, которые отвечают на изменения в активности определенных типов клеток. Благодаря использованию полнотранскриптомных профилей для анализа генных комплексов вместо отдельных генов возможно создание математических моделей, называемых генными подписями. Эти подписи позволяют более точно описывать характеристики злокачественных опухолей и предсказывать клинический исход. Проекты The Cancer Genome Atlas и International Cancer Genome Consortium играют ключевую роль в

исследованиях онкологии, предоставляя ученым информацию о молекулярной карте опухолей [[Lawrence и др., 2014]](https://www.zotero.org/google-docs/?EDAr85).

Рак молочной железы, толстой кишки, легких и другие типы ЗНО имеют сложные механизмы возникновения, которые часто связаны с повышенной или пониженной экспрессией некоторых генов. Именно такие гены формируют генную подпись, которая в дальнейшем может быть использована для прогнозирования состояния пациента и его ответа на терапию. Известно, что таргетные препараты, такие как цетуксимаб и трастузумаб, негативно влияют на рост и пролиферацию опухоли из-за блокирующего действия EGFR рецепторов, вследствие чего прекращается передача активирующего сигнала внутрь клетки [[Kamashev и др., 2023]](https://www.zotero.org/google-docs/?tD57CS) [[M и др., 2020]](https://www.zotero.org/google-docs/?W0eZNV).

Например, была создана подпись для рака молочной железы, которая помогает предсказывать сильно варьирующиеся ответы на противоопухолевую терапию и предсказывать возникновение метастазов. Сильно выбивающиеся из общей экспрессионной картины гены были выделены с помощью методов машинного обучения, в частности, обученной классификации, и таким образом позволили сформировать подпись генов, потенциально влияющих на развитие определенного типа ЗНО [[Van ’T Veer и др., 2002]](https://www.zotero.org/google-docs/?t3yOYn). А также была установлена подпись, которая определяет опухоли у носителей мутаций в гене BRCA1. Это ген, который часто отвечает за возникновение рака молочной железы у пациентов, но его роль в образовании ЗНО может быть точно предсказана только после формирования соответствующей генной подписи. Другой пример - новая генетическая подпись, состоящая из генов, имеющих отношение к раку желудка [[Zhao и др., 2019]](https://www.zotero.org/google-docs/?1DbjBM). Был проведен анализ генов, которые дифференциально экспрессировались в контрольном пункте перед входом в фазу митоза. Вследствие было выделено 5 генов, связанных с выживаемостью пациентов в ходе исследования. Эта генная сигнатура помогает разделить пациентов на подгруппы высокого и низкого риска развития ЗНО для персонализации терапии и прогнозирования ответа на нее.

# Глава 2. Материалы и методы

# Транскриптомные и клинические данные когорт пациентов

В работе использовались экспрессионные профили когорт пациентов, которые страдают опухолевыми заболеваниями и получают лечение с помощью таргетной терапии цетуксимабом или трастузумабом. Данные когорт пациентов были получены либо экспериментально с помощью РНК-секвенирования [[M и](https://www.zotero.org/google-docs/?CigNXy) [др., 2020]](https://www.zotero.org/google-docs/?CigNXy) [[D и др., 2022]](https://www.zotero.org/google-docs/?p3PYJg), либо из общедоступной базы данных генной экспрессии Gene Expression Omnibus (GEO). Данные состоят из данных транскрипции и информации о заболевании и процессе и особенностях лечения пациентов.

GEO - это база данных, созданная Национальным Центром биотехнологической информации (NCBI), которая предоставляет ученым возможность получить доступ к геномным, транскриптомным и протеомным данным, полученным из различных экспериментов по изучению уровней экспрессии генов в различных состояниях и условиях.

Информация из GEO для пациентов с ЗНО содержит ряд важных данных для каждого исследования, включая тип, местонахождение и особенности опухоли, стадии и специфику лечения, информацию о пациенте (возраст, пол, иногда - статус курения, перенесенные операции, хронические заболевания и тому подобное) и ответ на предложенную терапию. Эти клинические данные очень важны для правильного выбора таргетного препарата, а также для оценки набора генов, влияющих на эффективность терапии. Данные по ответу пациентов на терапию сформированы согласно критериям RECIST [[Schwartz и](https://www.zotero.org/google-docs/?kgaBMF) [др., 2016]](https://www.zotero.org/google-docs/?kgaBMF) [[Eisenhauer и др., 2009]](https://www.zotero.org/google-docs/?MZXYkr).

RECIST (Response Evaluation Criteria in Solid Tumors) — это значимый инструмент для оценки объема опухолевой нагрузки в рамках клинических исследований в онкологии. Эти критерии имеют важное значение при оценке эффективности и активности новых противоопухолевых препаратов для солидных опухолей. Для того, чтобы предсказывать ответ на терапию,

необходимо пользоваться категориальной переменной, которая может принимать только 2 значения. Исходя из этого, пациенты в зависимости от их реакции на лечение были разделены на 2 непересекающиеся группы: те, которые ответили на терапию и те, которые не ответили на терапию. К пациентам с положительным ответом на терапию относились группы пациентов, которые показали:

* + - “полный ответ” - исчезновение всех пораженных мишеней, все патологические лимфатические узлы (как целевые, так и нецелевые) должны быть уменьшены по короткой оси до менее чем 10 мм;
    - “частичный ответ” - уменьшение суммы диаметров пораженных областей на не менее 30% по сравнению с исходными значениями;
    - “стабильное заболевание” - отсутствие достаточного уменьшения суммы диаметров пораженных областей для классификации как "частичный ответ" или "прогрессирующее заболевание", при этом не наблюдается увеличения суммы диаметров поражений.

К пациентам с отрицательным ответом на терапию относилась группа пациентов, которая показала только “прогрессирующее заболевание” - увеличение суммы диаметров целевых поражений не менее чем на 20% от наименьшей суммы в исследовании, появление новых очагов заболевания [[Eisenhauer и др., 2009]](https://www.zotero.org/google-docs/?TDo9HU).

Транскриптомные профили были получены с помощью технологии ДНК-микрочипов или РНК-секвенирования с высоким выходом. Отбирались только те образцы пациентов, которые исследовались до начала терапии. Изучение такого претритмента в будущем может быть полезным для предсказания эффективности лечения, ведь задача состоит в определении наиболее подходящих препаратов исходя из экспрессионного профиля пациента. Генные символы в данном исследовании были представлены в соответствии с номенклатурой HUGO (Human Genome Organization) [[Seal и др.,](https://www.zotero.org/google-docs/?J6Gt9R) [2022]](https://www.zotero.org/google-docs/?J6Gt9R), что обеспечило единообразное и точное наименование генов, облегчая правильное понимание результатов исследования.

* + 1. Экспериментально полученные экспрессионные профили пациентов

Для предиктивного исследования ответа пациентов на таргетную терапию цетуксимабом была проанализирована клеточная линия А431, в которой была обнаружена экспрессия EGFR рецептора [[D.Kamashev и др., 2022]](https://www.zotero.org/google-docs/?baxLL9). Эта клеточная линия подвегралась воздействию цетуксимаба (3 биологические копии), цетуксимаба + сыворотки крови (3), цетуксимаба + эпидермального фактора роста (4), сыворотку крови (5), эпидермальный фактор роста (4), либо которая не подвергалась воздействию никаких веществ (5).

Для предиктивного исследования ответа пациентов на таргетную терапию трастузумабом была проанализирована клеточная линия BT474, подвергавшаяся воздействию трастузумаба (1 биологических образцов), трастузумаба + сыворотки крови, трастузумаб + эпидермального фактора роста (1), сыворотки крови (5), эпидермального фактора роста (1), либо которая не подвергалась воздействию никаких веществ (2). Образцы клеточной линии отображали рак молочной железы со сверхэкспрессией HER2 [[M.Sorokin и др.,](https://www.zotero.org/google-docs/?m0Vnlw) [2020]](https://www.zotero.org/google-docs/?m0Vnlw). Перед секвенированием образцы обрабатывали соответствующими препаратами и впоследствие сравнивались их экспрессионные профили для выделения генных подписей.

Для извлечения РНК в обоих случаях использовали набор RecoverAll™ (Invitrogen). Концентрацию РНК измеряли с помощью набора Qubit для анализа РНК, а для оценки целостности РНК (RIN) применяли биоанализатор Agilent 2100. Для деградации рибосомальной РНК использовали набор RNA Hyper (Roche), а затем анализировали концентрацию библиотек и длину фрагментов при помощи Qubit (Life Technologies) и Agilent Tapestation (Agilent) соответственно. Секвенирование РНК осуществляли на платформе Illumina NextSeq 550 для однонитевых в соответствии с общепринятым протоколом [[Suntsova и др., 2019]](https://www.zotero.org/google-docs/?1Tm3xi).

* + 1. Экспрессионные профили пациентов из общедоступных наборов данных

Экспрессионные данные наборов пациентов из общедоступного сайта GEO можно разделить на две группы: страдающие от колоректального рака и от рака молочной железы. К когортам пациентов, страдающих от колоректального рака и получающих терапию цетуксимабом, относятся транскрипционные данные GSE104645 [[Okita и др., 2018]](https://www.zotero.org/google-docs/?QUzAwy) и GSE5851 [[Khambata-Ford и др., 2007]](https://www.zotero.org/google-docs/?csZCik). Когорты пациентов, страдающих от рака молочной железы с активным HER2 рецептором и получающих терапию трастузумабом, относятся к наборам данных GSE50948 [[Prat и др., 2014]](https://www.zotero.org/google-docs/?ZaQcpF), GSE66305 [[Guarneri и др., 2015]](https://www.zotero.org/google-docs/?yQXcN2), GSE37946 [[Liu и др., 2012]](https://www.zotero.org/google-docs/?BrlE3x). Для получения экспрессионных профилей этих данных использовался метод ДНК-микрочипов компании Affymetrix. Молекулярные профили были получены до получения таргетной терапии.

Таблица 1. Основные характеристики наборов данных для когорт пациентов.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Источник данных | Кол-во пациентов с положительным ответом на терапию | Кол-во пациентов с отрциательным ответом на терапию | Терапия | Патология | Тип данных |
| GSE104645 | 40 | 25 | Цетуксимаб | Колоректаль ный рак IV стадия | экспрессионные микрочипы, log2 нормализация |
| GSE5851 | 35 | 56 | Цетуксимаб | Колоректаль ный рак | экспрессионные микрочипы, log2 нормализация |
| GSE50948 | 41 | 70 | Трастузумаб | Рак молочной железы | экспрессионные микрочипы, log2 нормализация |
| GSE66305 | 19 | 38 | Трастузумаб | Рак молочной железы | экспрессионные микрочипы, log2 нормализация |
| GSE37946 | 24 | 23 | Трастузумаб | Рак молочной железы | экспрессионные микрочипы, log2 нормализация |

Набор данных GSE104645 состоит из 193 образцов опухолей колоректального рака, находящегося на разных стадиях прогрессирования, из которых стадией интереса является третья стадия. Цетуксимаб добавлялся к лечению как вторая линия терапии, после лечения комбинациями из препаратов FOLFOX (фолинат кальция, оксалиплатин, фторурацил), SOX (оксалиплатин, тегафур, гимерацил, отерацил), Bev (бевацизумаб) и IRIS (иринотекан, тегафур, гимерацил, отерацил) [[Okita и др., 2018]](https://www.zotero.org/google-docs/?lfeoml).

Для когорты пациентов GSE5851 применялась монотерапия цетуксимабом для лечения метастатического колоректального рака. Ответ пациентов на терапию основывался на наличии у пациентов определенных биомаркеров (экспрессия эпирегулина и амфирегулина, статус мутации конкретного белка), что позволяло предсказать эффективность лечения и контроль над заболеванием при применении цетуксимаба. [[Khambata-Ford и др.,](https://www.zotero.org/google-docs/?9mW8el) [2007]](https://www.zotero.org/google-docs/?9mW8el).

В исследовании для набора данных GSE50948 анализировалась экспрессия генов у пациентов с опухолями молочной железы, которые были заранее обработаны терапией и имели HER2-положительный статус. У 114 пациентов проводили неоадъювантную терапию доксорубицином и паклитакселом, а затем им назначали циклофосфамид, метотрексат и фторурацил либо ту же схему в сочетании с трастузумабом в течение года [[Prat](https://www.zotero.org/google-docs/?MYjCOd) [и др., 2014]](https://www.zotero.org/google-docs/?MYjCOd). Событием интереса считались те пациенты, которые подвергались второму варианту терапии.

Общее количество участниц исследования GSE66305 составило 121 женщина с HER2-положительным раком молочной железы, которые были случайным образом распределены на группы для проведения неоадъювантной химиотерапии с трастузумабом, лапатинибом или их комбинацией [[Guarneri и](https://www.zotero.org/google-docs/?KfOYLU) [др., 2015]](https://www.zotero.org/google-docs/?KfOYLU). Интересом обладали первый и третий вариант терапии.

В исследовании для набора данных GSE37946 участвовали 50 пациентов с HER2-положительным раком молочной железы, лечившихся химиотерапией на основе трастузумаба перед операцией. Из первичных опухолей были взяты

образцы до начала лечения, и РНК была извлечена и гибридизирована с помощью микрочипов Affymetrix. Патологический ответ был оценен в конце периода неоадъювантного лечения [[Liu и др., 2012]](https://www.zotero.org/google-docs/?HH5H4L).

# Анализ дифференциальной экспрессии генов полученных с помощью РНК-секвенирования

Язык программирования R часто используется для статистический расчетов и анализа данных. Он предоставляет широкий ассортимент необходимых инструментов для обработки больших объемом биоинформатических данных, поэтому нашел применение в этой области. Дифференциальный анализ экспрессии генов - это эффективный подход для выявления различий в уровнях экспрессии генов между двумя или более группами образцов. Пакет DESeq2 языка программирования R использовался для моделирования количества генов и определения групп генов, экспрессия которых выше или ниже ожидаемого среднего на значимую величину [[Love,](https://www.zotero.org/google-docs/?Z0snDn) [Huber, Anders, 2014]](https://www.zotero.org/google-docs/?Z0snDn). DESeq2 работает на модели отрицательного биномиального распределения, так как существуют различия между экспрессией генов у биологических копий. Модель использует первоначальную матрицу подсчета и метаданные, которые включают информацию о лечении пациентов, а также формулу дизайна для создания объекта DESeq2. Приведенная расчетная формула должна содержать основные ожидаемые источники изменений, которые необходимо контролировать, и интересующее условие в качестве последнего члена формулы. Первым этапом обработки данных с помощью этого пакета является выравнивание подаваемых генов на референсный геном, подсчет экспрессии и их нормализация. Дальше идет построение модели анализа: DESeq2 применяет обобщенные линейные модели для учёта различий в экспрессии генов между образцами и условиями эксперимента, учитывая такие факторы, как размер выборки, дисперсия и другие характеристики. Конечным этапом является оценка дифференциальной экспрессии генов. Система вычисляет статистическую значимость изменений и

корректирует значения p-значений для обеспечения контроля над ошибками. На вход программы подавались данные из экспериментальных наборов.

# Алгоритм, предсказывающий эффективность противоопухолевой терапии на основе сформированных генных подписей

В данной работе на основе экспериментально полученных наборов данных РНК-секвенирования были сформированы генные подписи после применения к данным отрицательной биномиальной модели из пакета DESeq2 языка программирования R. Были применены фильтры p-значения и логарифмического отношения среднего для подтверждения статистической значимости различий в экспрессии между разными условиями каждого гена из набора, а также для определения направления дифференциальной экспрессии.

Сформированные генные подписи далее тестировались на предварительно обработанных общедоступных наборах данных, которые имели максимальную схожесть по дизайну лечения и выборе когорт пациентов. Алгоритм предсказания эффективности таргетного препарата основывался на экспрессии генов-мишеней из генных подписей, а оценка качества предсказания зависела от метрики AUC [[Hajian-Tilaki, 2013]](https://www.zotero.org/google-docs/?U1LnfL). Для когорт пациентов, имеющих данные о результатах ответа на терапию, можно оценить точность предсказания таргетной терапии на основе метрики “матрица ошибок”. Эта матрица представляет собой таблицу с четырьмя комбинациями предсказанных и реальных величин: ложно положительные, ложно отрицательные, истинно положительные и истинно отрицательные. Для создания этой таблицы необходимо, чтобы предсказываемые и фактические значения были бинарными величинами (присутствует или отсутствует результат). Поскольку предсказываемая величина является непрерывной числовой величиной и не дает однозначного ответа о том, будет ли эффективно использовать данный препарат или нет, для построения матрицы ошибок необходимо выбрать пороговое значение. Ниже этого значения будут предполагаемые пациенты с отрицательным ответом на терапию, а выше - пациенты с положительынм

ответом на терапию. На основе матрицы ошибок рассчитываются метрики чувствительности (насколько правильно предсказывается истинный результат) и специфичности (насколько правильно предсказывается ложный результат).

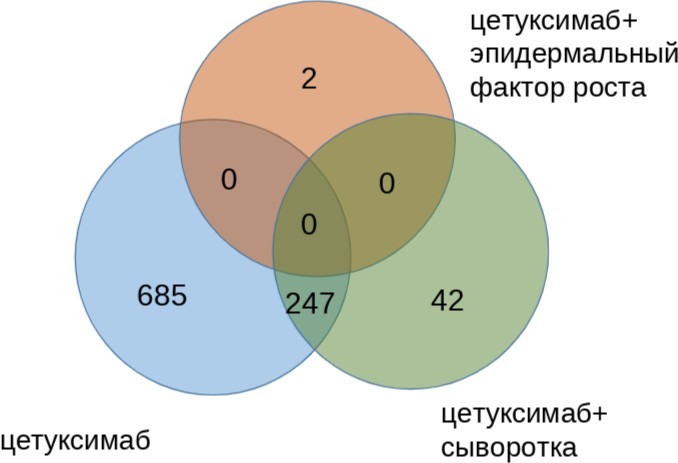
Выбор порогового значения предсказательной величины может быть сложной задачей, поэтому более удобно использовать метрику AUC (площадь под ROC-кривой), которая не требует выбора конкретного порогового значения. ROC-кривая (receiver operating characteristic) показывает зависимость чувствительности от вероятности неправильного предсказания положительного результата (1 - специфичность) при изменении порогового значения. Метрика AUC оценивается как площадь под ROC-кривой. Если значение AUC близкое к 1, то это указывает на идеальное качество прогноза, а значение 0.5 означает предсказание случайным образом. Важно, что значения AUC около 0.7 и выше считаются приемлемыми для биомедицинских тест-систем.

# Глава 3. Результаты и обсуждение

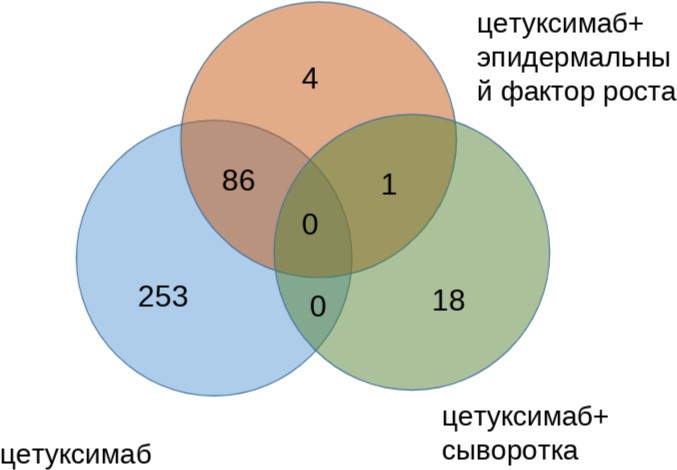
В данном исследовании анализировалась дифференциальная экспрессия генов и на ее основе формировались генные подписи для опухолей 2 клеточных линий экспериментальных данных при лечении цетуксимабом и трастузумабом. Сформированные наборы генов тестировались на общедоступных наборах данных пациентов с колоректальным раком и раком молочной железы для предсказания эффективности лечения до начала таргетной терапии. Эффективность терапии оценивалась с помощью метрики AUC. При недостаточной точности оценки генные подписи модифицировались для улучшения предсказательной модели и вновь тестировались на данных общедоступных когорт пациентов. Для 5 наборов данных клинических испытаний, на которых проводилось тестирование для предсказания ответа на терапию до начала лечения, предсказательные величины рассчитывались только для таргетных препаратов, не беря во внимание дополнительное лечение или первую линию лечения (как в случае с данными GSE104645).

# Поиск дифференциально экспрессированных данных

Для клеточной линии плоскоклеточного рака A431 был проведен анализ с помощью пакета DESeq2 и найдены дифференциально экспрессируемые гены для разных состояний образцов по сравнению с контрольным состоянием, когда ничего не добавляли: терапия цетуксимабом, добавление сыворотки крови, эпидермального фактора роста, сыворотки крови и цетуксимаба, сыворотки крови и эпидермального фактора роста (Рисунок 2, 3).



А



Б

Рисунок 2. Пересечения дифференциально экспрессированных генов для трех сравнений на клеточной линии А431. (А) гены с пониженной экспрессией (логарифмическое значение среднего < -1, поправка на множественное сравнение < 0.1). (Б) гены с повышенной экспрессией (логарифмическое значение среднего > 1, поправка на множественное сравнение < 0.1).

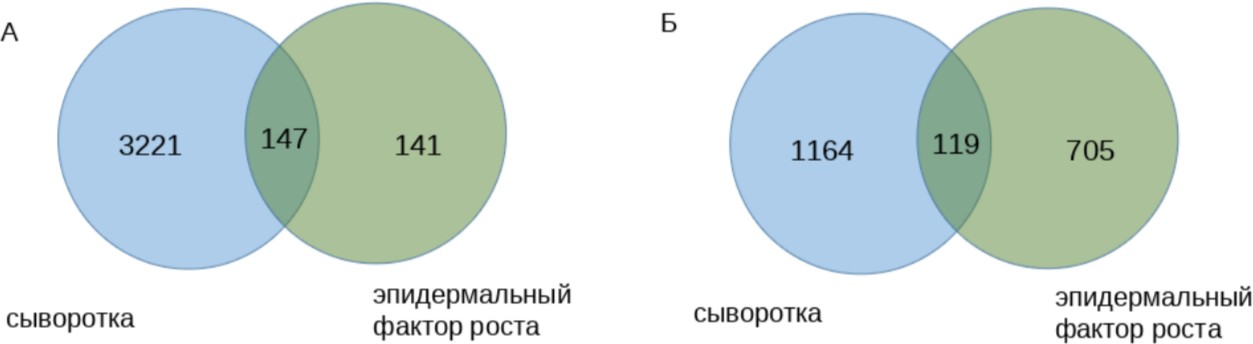


Рисунок 3. Пересечения дифференциально экспрессированных генов для двух сравнений на клеточной линии А431. (А) гены с пониженной экспрессией (логарифмическое значение среднего < -1, поправка на множественное сравнение < 0.1). (Б) гены с повышенной экспрессией (логарифмическое значение среднего > 1, поправка на множественное сравнение < 0.1).

Для клеточной линии рака молочной железы BT474 были найдены дифференциально экспрессируемые гены в сравнении с терапией трастузумабом и в сравнении с контролем для случаев лечения трастузумабом, добавления сыворотки крови, эпидермального фактора роста, сыворотки крови и трастузумаба, эпидермального фактора роста и трастузумаба. Наиболее значимыми результатами обладали дифференциально экспрессируемые гены образцов при добавлении эпидермального фактора роста, который замедляет действие трастузумаба и сыворотки крови [[D и др., 2022]](https://www.zotero.org/google-docs/?6aTdmT).

# Тестирование генных подписей на общедоступных наборах данных для предсказания ответа на терапию

Сформированные из экспериментальных данных генные подписи для клеточных линий A431 и BT474 на публичных наборах данных, содержащих экспрессионные профили пациентов, которые получали цетуксимаб и трастузумаб с известным статусом ответа на терапию. Протестировано более

150 генных подписей, которые для цетуксимаба включали следующие множества дифференциально экспрессируемых генов: цетуксимаб против контроля, цетуксимаб с сывороткой крови против контроля, цетуксимаб с эпидермальным фактором роста против контроля, сыворотка крови против контроля, эпидермальный фактор роста против контроля, разные пересечения и объединения вышеперечисленных множеств. Также были протестированы все вышеуказанные модели для полного датасета GSE104645, а не только для метастатических образцов, не учитывая состояние “стабильное заболевание” как ответ на терапию, однако значимых результатов не наблюдалось. Для каждого множества генов были проверены варианты формирования генной подписи, а именно сформированная на основе только оверэкспрессированных генов подпись, на основе только недоэкспрессированных генов, а также разность и отношение. Наилучшими показателями предсказательных величин AUC и поправки на множественное сравнение обладали генные подписи, которые формировались на основе пересечений дифференциально экспрессированных генов образцов, которые подвергались воздействию сыворотки крови или эпидермального фактора роста (ефр) (Таблица 2).

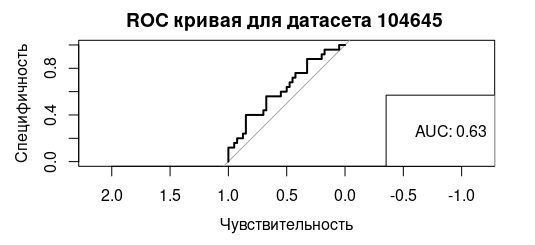
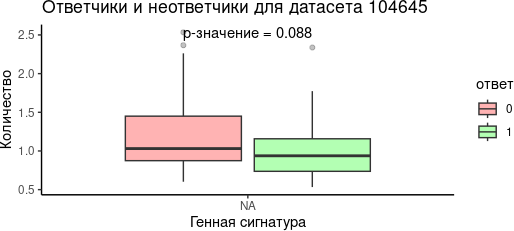
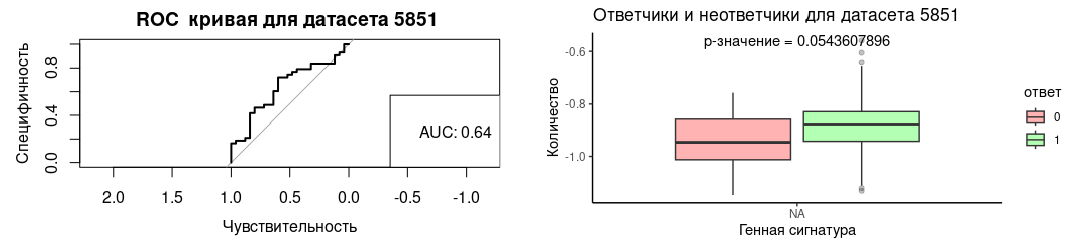
Таблица 2. Результаты тестирования генных подписей на датасетах, подвергшихся лечению цетуксимабом.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | GSE104645 | | GSE5851 | |
|  | Генная подпись | AUC | p-значение | AUC | p-значение |
| дифференциально экспрессированные гены ефр против контроля | Up - down | 0.57 | 0.33 | 0.67 | 0.02 |
| Up / down | 0.57 | 0.37 | 0.66 | 0.03 |
| up | 0.53 | 0.74 | 0.65 | 0.05 |
| down | 0.55 | 0.54 | 0.54 | 0.59 |
| дифференциально | Up - down | 0.56 | 0.45 | 0.56 | 0.42 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| экспрессирвоаные гены сыворотки крови против контроля | Up / down | 0.56 | 0.43 | 0.54 | 0.56 |
| up | 0.53 | 0.68 | 0.59 | 0.29 |
| down | 0.60 | 0.21 | 0.54 | 0.56 |
| объединение дифференциальных генов ефр против контроля и дифференциальных генов сыворотки крови против контроля | Up - down | 0.59 | 0.22 | 0.64 | 0.06 |
| Up / down | 0.59 | 0.25 | 0.63 | 0.08 |
| up | 0.54 | 0.62 | 0.62 | 0.12 |
| down | 0.59 | 0.25 | 0.55 | 0.51 |
| пересечение дифференциальных генов ефр против контроля с дифференциальными генами сыворотки крови против контроля | Up - down | 0.63 | 0.09 | 0.64 | 0.07 |
| Up / down | 0.62 | 0.12 | 0.63 | 0.08 |
| up | 0.56 | 0.43 | 0.66 | 0.03 |
| down | 0.6 | 0.19 | 0.52 | 0.83 |

Наилучшей оценкой предсказательной модели эффективности таргетной терапии на клетки плоскоклеточного рака клеточной линии A431 обладает генная подпись, сформированная на основе пересечения множеств генов, дифференциально экспрессированных при воздействии сыворотки крови с множеством генов, дифференциально экспрессированных при воздействии эпидермального фактора роста, при этом взята разность между экспрессией оверэкспрессированных и недоэкспрессированных генов.

Для этих двух наборов данных построены соответствующие ROC кривые и боксплоты для определения и визуализации p-значения, чтобы наглядно показать эффективность данных наборов генов в предсказательной модели (Рисунок 4).



А

Б

В

Г

32

Рисунок 4. Величины, отражающие эффективность применения генной сигнатуры до ее модификации для клеточной линии А431 как возможного предиктора ответа на проотивоопухолевую терапию цетуксимабом до начала лечения. (А) ROC-кривая для датасета GSE5851. (Б) боксплот распределения вероятности получить тот или иной ответ на терапию для датасета GSE5851. (В) ROC-кривая для датасета GSE104645. (Г) боксплот распределения вероятности получить тот или иной ответ на терапию для датасета GSE104645.

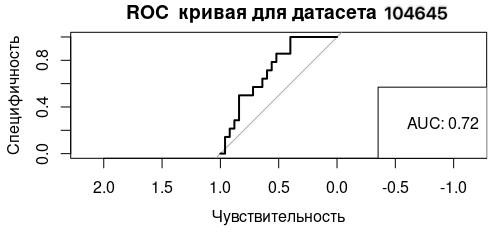
Для клеточной линии BT474 с терапией трастузумабом также были выделены различные генные подписи, учитывая аналогичное с наблюдаемым на клеточной линии A431 действие эпидермального фактора роста (замедление действия противоопухолевого препарата) и противоположное действие сыворотки крови (усиление действия противоопухолевого препарата). Однако, ни одна из протестированных генных подписей не обладала достаточной точностью для предсказания ответа на терапию до начала лечения. Дальнейшие модификации не привели к существенному улучшению качества предсказательной модели (Таблица 3).

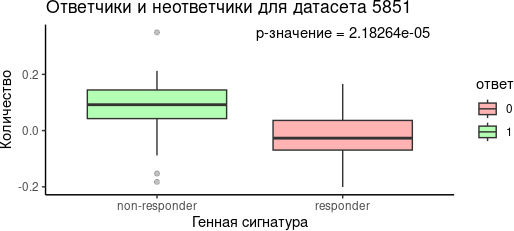
Таблица 3. Результаты тестирования генных подписей на датасетах, подвергшихся лечению трастузумабом.

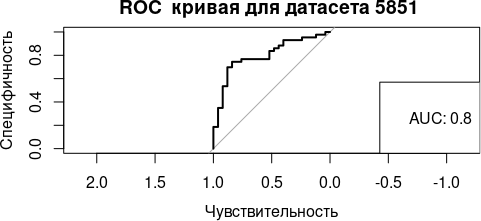
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | GSE50948 | | GSE66305 | | GSE37946 | |
|  |  | AUC | p-значение | AUC | p-значение | AUC | p-значение |
| дифференциально  экспрессированные гены ефр против контроля (1) | Up - down | 0.54 | 0.37 | 0.54 | 0.68 | 0.56 | 0.49 |
| Up / down | 0.55 | 0.37 | 0.47 | 0.69 | 0.56 | 0.52 |
| up | 0.52 | 0.74 | 0.48 | 0.85 | 0.52 | 0.86 |
| down | 0.53 | 0.64 | 0.55 | 0.57 | 0.58 | 0.38 |
| дифференциально экспрессированные гены сыворотки крови против  контроля (2) | Up - down | - | - | 0.61 | 0.16 | 0.60 | 0.28 |
| Up / down | - | - | 0.61 | 0.16 | 0.60 | 0.26 |
| up | - | - | 0.58 | 0.37 | 0.55 | 0.58 |
| down | - | - | 0.57 | 0.39 | 0.55 | 0.58 |
| (1) - (2) | up | 0.48 | 0.73 | 0.47 | 0.73 | 0.58 | 0.38 |
| down | 0.62 | 0.47 | 0.51 | 0.87 | 0.47 | 0.73 |

# Модификация генной подписи для увеличения значимости предсказательной модели

После тестирования генной сигнатуры, сформированной из дифференциально экспрессируемых генов в образцах клеточной линии А431, на

публичных наборах данных, получили значения AUC и p-значения для оценки качества предиктивного ответа на терапию цетуксимабом до начала лечения относительно публичных наборов экспрессионных данных. Однако для увеличения значимости подписи ее необходимо модифицировать путем поочередного удаления генов и проверки прироста значения AUC. На каждой итерации удаляется ген, который наибольше увеличивает значение AUC, и так до тех пор, пока величина не будет постоянной или не будет начинать уменьшаться при удалении всех последующих генов. Генная подпись формировалась при тестировании top-down подхода по удалению генов сначала на одном датасете, а потом на втором для валидации результатов. Лучшие значения AUC достигаются при выборе набора данных GSE5851 как тестового, а GSE5851 - как валидационного. С помощью этого принципа удалось увеличить значения AUC для датасетов GSE104645 и GSE5851, чтобы они стали > 0.7, что считается приемлемым для биоинформатических систем. При этом p-значение, которое отвечает за то, насколько достоверно можно отделить генную экспрессию в опухолях пациентов с положительным ответом на терапию от пациентов с отрицательным ответом на терапию, также улучшилось (Рисунок 5).

А Б



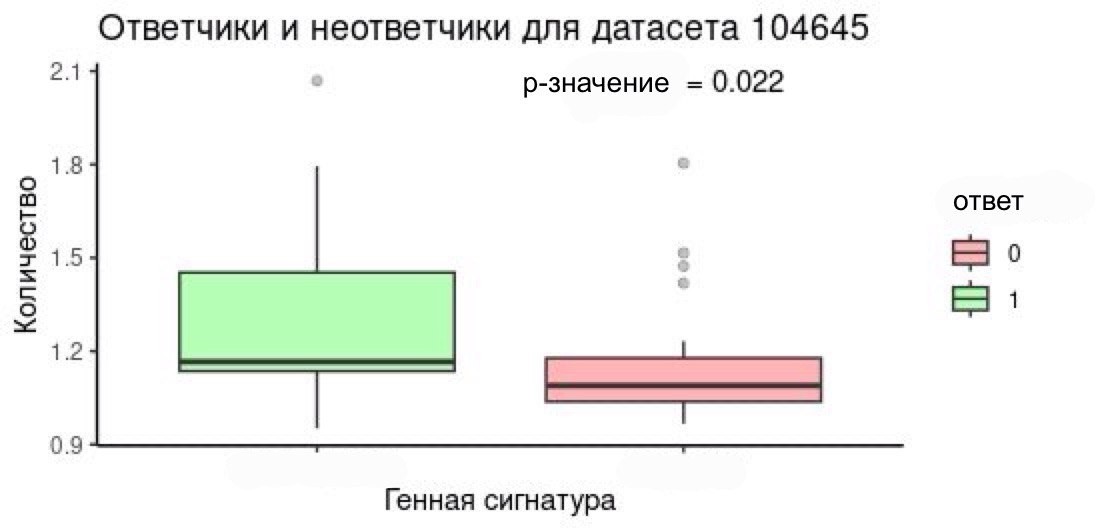
В Г

Рисунок 5. Величины, отражающие эффективность применения генной сигнатуры после модификации для клеточной линии А431 как возможного предиктора ответа на проотивоопухолевую терапию цетуксимабом до начала лечения. (А) ROC-кривая для датасета GSE5851. (Б) боксплот распределения вероятности получить тот или иной ответ на терапию для датасета GSE5851. (В) ROC-кривая для датасета GSE104645 после модификации генной подписи. (Г) боксплот распределения вероятности получить тот или иной ответ на терапию для датасета GSE104645.

Модифицировання генная подпись, выделенная из дифференциально экспрессированных генов клеточной линии плоскоклеточного рака А431 состоит из 48 генов, из которых 29 недоэкспрессированных: ADAMTS12, ADH1C, CALB2, CD177, CENPS.CORT, CHRNA3, CLCA2, CLU, DBP, DTNA, FOXP3, IFT74, IGFLR1, IPCEF1, IRAK3, LBH, LSP1, MMP13, MMP7, PAOX, PLAT, PTGES, PTPRN, SRPX2, TBX6, TLR1, TMPRSS4, TXK, WARS1, и 19

оверэкспрессированных генов: AEBP1, COL18A1, COL6A1, CYP1B1, DPYSL3, EGR3, FST, GLDC, IL32, ITK, JAK3, MUC5B, MYO1A, PIGR, PRM2, PRUNE2, SLC15A3, TGM2, WIPF1.

# Заключение

В ходе данной работы был проведен анализ эффективности различных генных подписей, сформированных на основе экспериментальных экспрессионных данных, на публичных когортах пациентов, для идентификации тех подмножеств генов, которые бы лучше всего предсказывали ответ на терапию цетуксимабом или трастузумабом до начала лечения. Проанализирована дифференциальная экспрессия генов из опухолей клеточных линий А431 и BT474, применены методы нормализации и фильтрации данных на основе логарифмического изменения относительно среднего экспрессии и поправки на множественное сравнение. Данные наборы генов протестированы на 5 набор публичных транскриптомных данных для выделения подписей, которые будут лежать в основе предсказательной модели. Выделенные генные подписи улучшены путем поочередного удаления генов до максимизации значения AUC при обходе всех генов до конца.

Было обнаружено, что лучшим предсказательным потенциалом обладает множество генов, модифицированное на основе датасета GSE5851, а потом валидированное на датасете GSE104645. Генная подпись для предсказания ответа на терапию цетуксимабом обладает большим потенциалом вследствие высоких показателей AUC и p-значения для обеих когорт пациентов.

Стоит отметить, что результаты для построения предсказательной модели ответа на терапию трастузумабом не были значимыми, и поэтому дальнейшая модификация подписи не рассматривалась. Необходимы дальнейшие исследования датасетов пациентов, подвергшихся действию трастузумаба, для более глубокого изучения того, как разные генные подписи предсказывают ответ на терапию.

# Выводы

* + 1. Генные подписи, в зависимости от способа формирования и входящих в них гены, имеют разный потенциал предсказания ответа на терапию на других наборах данных РНК-секвенирования. Лучше всего на пациентах из публичных наборов данных, подвергавшихся воздействию цетуксимаба, работает генная подпись, сформированная на основе пересечения генов, дифференциально экспрессированных при действии сыворотки крови, и генов, дифференциально экспрессированных при действии эпидермального фактора роста, выделенных из данных РНК-секвенирования плоскоклеточного рака клеточной линии А431.
    2. Эффективность предсказания противоопухолевой терапии для каждого публичного набора экспрессионных данных варьируется:
       1. для наборов данных, подвергшихся лечению цетуксимабом, построенная успешная предиктивная модель ответа на химиотерапию на основе дифференциально экспрессируемых генов экспериментальных наборов данных позволяет использовать множество генов на датасетах, содержащие экспрессионные профили до начала лечения, для предсказания резистентности пациента в ответ на лечение;
       2. для наборов данных, подвергшихся лечению трастузумабом, протестированные генные подписи как модель предиктивного ответа на терапию до начала лечения не дала значимых результатов, что может быть связано с несоответствием процесса лечения противоопухолевыми препаратами между пациентами из экспериментальных наборов данных и публичными когортами пациентов.

# Список литературы

1. [Bier F. F. и др. DNA Microarrays // Biosensing for the 21st Century / под ред. R.](https://www.zotero.org/google-docs/?ArRkV1) [Renneberg, F. Lisdat. Berlin, Heidelberg: Springer, 2008. С. 433–453.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Bilitewski U. DNA Microarrays: An Introduction to the Technology // Microchip](https://www.zotero.org/google-docs/?ArRkV1) [Methods in Diagnostics / под ред. U. Bilitewski. Totowa, NJ: Humana Press, 2009.](https://www.zotero.org/google-docs/?ArRkV1) [С. 1–14.](https://www.zotero.org/google-docs/?ArRkV1)
3. [Bryant D. W., Priest H. D., Mockler T. C. Detection and Quantification of](https://www.zotero.org/google-docs/?ArRkV1) [Alternative Splicing Variants Using RNA-seq // RNA Abundance Analysis: Methods](https://www.zotero.org/google-docs/?ArRkV1) [and Protocols / под ред. H. Jin, W. Gassmann. Totowa, NJ: Humana Press, 2012. С.](https://www.zotero.org/google-docs/?ArRkV1) [97–110.](https://www.zotero.org/google-docs/?ArRkV1)
4. [Buss N. A. и др. Monoclonal antibody therapeutics: history and future // Curr.](https://www.zotero.org/google-docs/?ArRkV1) [Opin. Pharmacol. 2012. Т. 12. № 5. С. 615–622.](https://www.zotero.org/google-docs/?ArRkV1)
5. [Chomczynski P., Sacchi N. Single-step method of RNA isolation by acid](https://www.zotero.org/google-docs/?ArRkV1) [guanidinium thiocyanate-phenol-chloroform extraction // Anal. Biochem. 1987. Т.](https://www.zotero.org/google-docs/?ArRkV1) [162. № 1. С. 156–159.](https://www.zotero.org/google-docs/?ArRkV1)
6. [Chomczynski P., Sacchi N. The single-step method of RNA isolation by acid](https://www.zotero.org/google-docs/?ArRkV1) [guanidinium thiocyanate–phenol–chloroform extraction: twenty-something years on](https://www.zotero.org/google-docs/?ArRkV1)

[// Nat. Protoc. 2006. Т. 1. № 2. С. 581–585.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Cooper G. M. Regulation of Transcription in Eukaryotes // The Cell: A Molecular](https://www.zotero.org/google-docs/?ArRkV1) [Approach. 2nd edition. : Sinauer Associates, 2000.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Corcoran C., O’Driscoll L. Receptor Tyrosine Kinases and Drug Resistance:](https://www.zotero.org/google-docs/?ArRkV1) [Development and Characterization of In Vitro Models of Resistance to RTK](https://www.zotero.org/google-docs/?ArRkV1) [Inhibitors // Receptor Tyrosine Kinases Methods in Molecular Biology. / под ред. S.](https://www.zotero.org/google-docs/?ArRkV1) [Germano. New York, NY: Springer New York, 2015. С. 169–180.](https://www.zotero.org/google-docs/?ArRkV1)
3. [D K. и др. Human Blood Serum Inhibits Ductal Carcinoma Cells BT474 Growth](https://www.zotero.org/google-docs/?ArRkV1) [and Modulates Effect of HER2 Inhibition // Biomedicines. 2022. Т. 10. № 8.](https://www.zotero.org/google-docs/?ArRkV1)
4. [Ebrahimi N. и др. Receptor tyrosine kinase inhibitors in cancer // Cell. Mol. Life](https://www.zotero.org/google-docs/?ArRkV1) [Sci. 2023. Т. 80. № 4. С. 104.](https://www.zotero.org/google-docs/?ArRkV1)
5. [Eisenhauer E. A. и др. New response evaluation criteria in solid tumours: Revised](https://www.zotero.org/google-docs/?ArRkV1)

[RECIST guideline (version 1.1) // Eur. J. Cancer. 2009. Т. 45. № 2. С. 228–247.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Erp N. P. van, Gelderblom H., Guchelaar H.-J. Clinical pharmacokinetics of](https://www.zotero.org/google-docs/?ArRkV1) [tyrosine kinase inhibitors // Cancer Treat. Rev. 2009. Т. 35. № 8. С. 692–706.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Ferguson L. R. и др. Genomic instability in human cancer: Molecular insights](https://www.zotero.org/google-docs/?ArRkV1) [and opportunities for therapeutic attack and prevention through diet and nutrition //](https://www.zotero.org/google-docs/?ArRkV1) [Semin. Cancer Biol. 2015. Т. 35 Suppl. № Suppl. С. S5–S24.](https://www.zotero.org/google-docs/?ArRkV1)
3. [Garcia J. и др. Bevacizumab (Avastin®) in cancer treatment: A review of](https://www.zotero.org/google-docs/?ArRkV1)

[15 years of clinical experience and future outlook // Cancer Treat. Rev. 2020. Т. 86.](https://www.zotero.org/google-docs/?ArRkV1) [С. 102017.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Gerber D. E. Targeted therapies: a new generation of cancer treatments // Am.](https://www.zotero.org/google-docs/?ArRkV1) [Fam. Physician. 2008. Т. 77. № 3. С. 311–319.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Green M. R., Sambrook J. Removing DNA Contamination from RNA Samples](https://www.zotero.org/google-docs/?ArRkV1) [by Treatment with RNase-Free DNase I // Cold Spring Harb. Protoc. 2019. Т. 2019.](https://www.zotero.org/google-docs/?ArRkV1)

[№ 10. С. pdb.prot101725.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Guarneri V. и др. Prospective Biomarker Analysis of the Randomized](https://www.zotero.org/google-docs/?ArRkV1)

[CHER-LOB Study Evaluating the Dual Anti-HER2 Treatment With Trastuzumab and](https://www.zotero.org/google-docs/?ArRkV1) [Lapatinib Plus Chemotherapy as Neoadjuvant Therapy for HER2-Positive Breast](https://www.zotero.org/google-docs/?ArRkV1) [Cancer // The Oncologist. 2015. Т. 20. № 9. С. 1001–1010.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Guo H. и др. Biomarker-Targeted Therapies in Non–Small Cell Lung Cancer:](https://www.zotero.org/google-docs/?ArRkV1) [Current Status and Perspectives // Cells. 2022. Т. 11. № 20. С. 3200.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for](https://www.zotero.org/google-docs/?ArRkV1) [Medical Diagnostic Test Evaluation // Casp. J. Intern. Med. 2013. Т. 4. № 2. С.](https://www.zotero.org/google-docs/?ArRkV1) [627–635.](https://www.zotero.org/google-docs/?ArRkV1)
3. [Heller M. J. DNA Microarray Technology: Devices, Systems, and Applications //](https://www.zotero.org/google-docs/?ArRkV1) [Annu. Rev. Biomed. Eng. 2002. Т. 4. № 1. С. 129–153.](https://www.zotero.org/google-docs/?ArRkV1)
4. [Holbro T., Hynes N. E. ErbB receptors: directing key signaling networks](https://www.zotero.org/google-docs/?ArRkV1) [throughout life // Annu. Rev. Pharmacol. Toxicol. 2004. Т. 44. С. 195–217.](https://www.zotero.org/google-docs/?ArRkV1)
5. [Hu T. и др. Next-generation sequencing technologies: An overview // Hum.](https://www.zotero.org/google-docs/?ArRkV1) [Immunol. 2021. Т. 82. № 11. С. 801–811.](https://www.zotero.org/google-docs/?ArRkV1)
6. [Kamashev D. и др. Human Blood Serum Can Diminish EGFR-Targeted](https://www.zotero.org/google-docs/?ArRkV1) [Inhibition of Squamous Carcinoma Cell Growth through Reactivation of MAPK and](https://www.zotero.org/google-docs/?ArRkV1)

[EGFR Pathways // Cells. 2023. Т. 12. № 16. С. 2022.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Khambata-Ford S. и др. Expression of epiregulin and amphiregulin and K-ras](https://www.zotero.org/google-docs/?ArRkV1) [mutation status predict disease control in metastatic colorectal cancer patients treated](https://www.zotero.org/google-docs/?ArRkV1) [with cetuximab // J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 2007. Т. 25. № 22. С.](https://www.zotero.org/google-docs/?ArRkV1) [3230–3237.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Knierim E. и др. Systematic Comparison of Three Methods for Fragmentation of](https://www.zotero.org/google-docs/?ArRkV1) [Long-Range PCR Products for Next Generation Sequencing // PLOS ONE. 2011. Т.](https://www.zotero.org/google-docs/?ArRkV1) [6. № 11. С. e28240.](https://www.zotero.org/google-docs/?ArRkV1)
3. [Lamorte L., Park M. The Receptor Tyrosine Kinases: Role in Cancer Progression](https://www.zotero.org/google-docs/?ArRkV1)

[// Surg. Oncol. Clin. N. Am. 2001. Т. 10. № 2. С. 271–288.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Lawrence M. S. и др. Discovery and saturation analysis of cancer genes across 21](https://www.zotero.org/google-docs/?ArRkV1) [tumor types // Nature. 2014. Т. 505. № 7484. С. 495–501.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Liu J. C. и др. Seventeen-gene signature from enriched Her2/Neu mammary](https://www.zotero.org/google-docs/?ArRkV1) [tumor-initiating cells predicts clinical outcome for human HER2+:ERα- breast cancer](https://www.zotero.org/google-docs/?ArRkV1)

[// Proc. Natl. Acad. Sci. U. S. A. 2012. Т. 109. № 15. С. 5832–5837.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Love M. I., Huber W., Anders S. Moderated estimation of fold change and](https://www.zotero.org/google-docs/?ArRkV1) [dispersion for RNA-seq data with DESeq2 // Genome Biol. 2014. Т. 15. № 12. С.](https://www.zotero.org/google-docs/?ArRkV1) [550.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Lowe R. и др. Transcriptomics technologies // PLoS Comput. Biol. 2017. Т. 13.](https://www.zotero.org/google-docs/?ArRkV1)

[№ 5. С. e1005457.](https://www.zotero.org/google-docs/?ArRkV1)

1. [M S. и др. Molecular Pathway Activation Markers Are Associated with Efficacy](https://www.zotero.org/google-docs/?ArRkV1) [of Trastuzumab Therapy in Metastatic HER2-Positive Breast Cancer Better than](https://www.zotero.org/google-docs/?ArRkV1) [Individual Gene Expression Levels // Biochem. Biokhimiia. 2020. Т. 85. № 7.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Okita A. и др. Consensus molecular subtypes classification of colorectal cancer](https://www.zotero.org/google-docs/?ArRkV1) [as a predictive factor for chemotherapeutic efficacy against metastatic colorectal](https://www.zotero.org/google-docs/?ArRkV1) [cancer // Oncotarget. 2018. Т. 9. № 27. С. 18698–18711.](https://www.zotero.org/google-docs/?ArRkV1)
3. [Olayioye M. A. и др. The ErbB signaling network: receptor heterodimerization in](https://www.zotero.org/google-docs/?ArRkV1) [development and cancer // EMBO J. 2000. Т. 19. № 13. С. 3159–3167.](https://www.zotero.org/google-docs/?ArRkV1)
4. [Ozsolak F., Milos P. M. RNA sequencing: advances, challenges and opportunities](https://www.zotero.org/google-docs/?ArRkV1)

[// Nat. Rev. Genet. 2011. Т. 12. № 2. С. 87–98.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Padma V. V. An overview of targeted cancer therapy // BioMedicine. 2015. Т. 5.](https://www.zotero.org/google-docs/?ArRkV1)

[№ 4. С. 19.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Parekh S. и др. The impact of amplification on differential expression analyses by](https://www.zotero.org/google-docs/?ArRkV1) [RNA-seq // Sci. Rep. 2016. Т. 6. № 1. С. 25533.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Prat A. и др. Research-based PAM50 subtype predictor identifies higher](https://www.zotero.org/google-docs/?ArRkV1) [responses and improved survival outcomes in HER2-positive breast cancer in the](https://www.zotero.org/google-docs/?ArRkV1) [NOAH study // Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 2014. Т. 20. № 2.](https://www.zotero.org/google-docs/?ArRkV1) [С. 511–521.](https://www.zotero.org/google-docs/?ArRkV1)
3. [Schlessinger J. Cell Signaling by Receptor Tyrosine Kinases // Cell. 2000. Т. 103.](https://www.zotero.org/google-docs/?ArRkV1)

[№ 2. С. 211–225.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Schwartz L. H. и др. RECIST 1.1—Update and clarification: From the RECIST](https://www.zotero.org/google-docs/?ArRkV1) [committee // Eur. J. Cancer. 2016. Т. 62. С. 132–137.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Seal R. L. и др. Genenames.org: the HGNC resources in 2023 // Nucleic Acids](https://www.zotero.org/google-docs/?ArRkV1) [Res. 2022. Т. 51. № D1. С. D1003–D1009.](https://www.zotero.org/google-docs/?ArRkV1)
3. [Stern M., Herrmann R. Overview of monoclonal antibodies in cancer therapy:](https://www.zotero.org/google-docs/?ArRkV1) [present and promise // Crit. Rev. Oncol. Hematol. 2005. Т. 54. № 1. С. 11–29.](https://www.zotero.org/google-docs/?ArRkV1)
4. [Sun L. L. и др. Anti-CD20/CD3 T cell-dependent bispecific antibody for the](https://www.zotero.org/google-docs/?ArRkV1) [treatment of B cell malignancies // Sci. Transl. Med. 2015. Т. 7. № 287. С. 287ra70.](https://www.zotero.org/google-docs/?ArRkV1)
5. [Suntsova M. и др. Atlas of RNA sequencing profiles for normal human tissues //](https://www.zotero.org/google-docs/?ArRkV1) [Sci. Data. 2019. Т. 6. С. 36.](https://www.zotero.org/google-docs/?ArRkV1)
6. [Van ’T Veer L. J. и др. Gene expression profiling predicts clinical outcome of](https://www.zotero.org/google-docs/?ArRkV1) [breast cancer // Nature. 2002. Т. 415. № 6871. С. 530–536.](https://www.zotero.org/google-docs/?ArRkV1)
7. [Wang Z., Gerstein M., Snyder M. RNA-Seq: a revolutionary tool for](https://www.zotero.org/google-docs/?ArRkV1) [transcriptomics // Nat. Rev. Genet. 2009. Т. 10. № 1. С. 57–63.](https://www.zotero.org/google-docs/?ArRkV1)
8. [Wynder E. L., Gori G. B. Contribution of the environment to cancer incidence: an](https://www.zotero.org/google-docs/?ArRkV1) [epidemiologic exercise // J. Natl. Cancer Inst. 1977. Т. 58. № 4. С. 825–832.](https://www.zotero.org/google-docs/?ArRkV1)
9. [Yewale C. и др. Epidermal growth factor receptor targeting in cancer: A review](https://www.zotero.org/google-docs/?ArRkV1) [of trends and strategies // Biomaterials. 2013. Т. 34. № 34. С. 8690–8707.](https://www.zotero.org/google-docs/?ArRkV1)
10. [Zhang H., He L., Cai L. Transcriptome Sequencing: RNA-Seq // Computational](https://www.zotero.org/google-docs/?ArRkV1) [Systems Biology Methods in Molecular Biology. / под ред. T. Huang. New York,](https://www.zotero.org/google-docs/?ArRkV1)

[NY: Springer New York, 2018. С. 15–27.](https://www.zotero.org/google-docs/?ArRkV1)

1. [Zhao L. и др. Identification of a novel cell cycle-related gene signature predicting](https://www.zotero.org/google-docs/?ArRkV1) [survival in patients with gastric cancer // J. Cell. Physiol. 2019. Т. 234. № 5. С.](https://www.zotero.org/google-docs/?ArRkV1) [6350–6360.](https://www.zotero.org/google-docs/?ArRkV1)
2. [Global Cancer Observatory [Электронный ресурс]. URL:](https://www.zotero.org/google-docs/?ArRkV1) [https://gco.iarc.who.int/en (дата обращения: 16.06.2024).](https://www.zotero.org/google-docs/?ArRkV1)
3. [WHO: Global Cancer Burden Rising Amid Striking Inequities [Электронный](https://www.zotero.org/google-docs/?ArRkV1) [ресурс]. URL:](https://www.zotero.org/google-docs/?ArRkV1)

[https://www.medscape.com/viewarticle/global-cancer-burden-rising-amid-striking-in](https://www.zotero.org/google-docs/?ArRkV1) [equities-2024a10002an (дата обращения: 16.06.2024).](https://www.zotero.org/google-docs/?ArRkV1)

1. [Brockhoff: Differential impact of Cetuximab, Pertuzumab... - Академия Google](https://www.zotero.org/google-docs/?ArRkV1) [[Электронный ресурс]. URL:](https://www.zotero.org/google-docs/?ArRkV1) [https://scholar.google.com/scholar\_lookup?journal=Cell+Prolif.&title=Differential+i](https://www.zotero.org/google-docs/?ArRkV1) [mpact+of+Cetuximab,+Pertuzumab+and+Trastuzumab+on+BT474+and+SK-BR-3+](https://www.zotero.org/google-docs/?ArRkV1) [breast+cancer+cell+proliferation&author=G.+Brockhoff&author=B.+Heckel&author](https://www.zotero.org/google-docs/?ArRkV1)

[=E.+Schmidt-Bruecken&author=M.+Plander&author=F.+Hofstaedter&volume=40&](https://www.zotero.org/google-docs/?ArRkV1) [publication\_year=2007&pages=488-507&pmid=17635517&doi=10.1111/j.1365-218](https://www.zotero.org/google-docs/?ArRkV1) [4.2007.00449.x& (дата обращения: 16.06.2024).](https://www.zotero.org/google-docs/?ArRkV1)

1. [HER2-targeted therapies — a role beyond breast cancer | Nature Reviews Clinical](https://www.zotero.org/google-docs/?ArRkV1) [Oncology [Электронный ресурс]. URL:](https://www.zotero.org/google-docs/?ArRkV1)

[https://www.nature.com/articles/s41571-019-0268-3 (дата обращения: 16.06.2024).](https://www.zotero.org/google-docs/?ArRkV1)

1. [Outline of detection method of genes by DNA microarrays | 3D-Gene® [Toray](https://www.zotero.org/google-docs/?ArRkV1) [DNA Chips] | TORAY [Электронный ресурс]. URL:](https://www.zotero.org/google-docs/?ArRkV1)

[https://www.3d-gene.com/en/about/chip/chi\_003.html (дата обращения:](https://www.zotero.org/google-docs/?ArRkV1) [16.06.2024).](https://www.zotero.org/google-docs/?ArRkV1)

1. [Parallel human genome analysis: microarray-based expression monitoring of](https://www.zotero.org/google-docs/?ArRkV1) [1000 genes. [Электронный ресурс]. URL:](https://www.zotero.org/google-docs/?ArRkV1) [https://www.pnas.org/doi/epdf/10.1073/pnas.93.20.10614 (дата обращения:](https://www.zotero.org/google-docs/?ArRkV1) [16.06.2024).](https://www.zotero.org/google-docs/?ArRkV1)