

ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ

© Коллектив авторов

МАСТЕР-РЕГУЛЯТОРЫ, СВЯЗАННЫЕ С ПЛОХИМ ПРОГНОЗОМ ПРИ ГЛИОБЛАСТОМЕ

М. Калия^{1,2}, Т. Бейсбарт¹, А.Э. Кель^{2,3*}

¹Department of Medical Bioinformatics, University Medical Center Göttingen, Göttingen, 37099 Germany

²geneXplain GmbH, Wolfenbüttel, 38302 Germany

³Институт химической биологии и фундаментальной медицины СО РАН,
630090, Новосибирск, пр. Ак. Лаврентьева, 8; *эл. почта: alexander.kel@genexplain.com

Глиобластома (GBM, glioblastoma multiforme) является злокачественной опухолью головного мозга со средним сроком выживания 15 месяцев. Менее 2% пациентов выживают после 36 месяцев. Для понимания молекулярного механизма, ответственного за плохой прогноз, мы проанализировали 560 образцов GBM транскриптомных данных на основе микрочипов из базы данных TCGA (The Cancer Genome Atlas). Анализ с помощью одномерной регрессии Кокса позволил идентифицировать 720 генов, которые оказывают значительное влияние на продолжительность выживания. Для установления потенциальных механизмов регуляции активности этих генов и построения сетей генной регуляции мы использовали программное обеспечение Genome Enhancer (<https://genexplain.com/genome-enhancer/>). В результате мы выявили 12 транскрипционных факторов, сайты которых обогащены в промоторах этих генов, включая ключевой фактор регуляции многих патологических процессов в GBM — *STAT3*. Мы установили, что ген *STAT3* дифференциально экспрессируется в экстремальных группах выживания (кратковременное выживание — выживаемость менее 12 месяцев и длительное выживание — выживаемость более 36 месяцев), и его экспрессия коррелирует с продолжительностью жизни у пациентов при глиобластоме. На следующем этапе мы определили мастер-регуляторы в сети передачи сигналов в клетках, которые регулируют активность этих транскрипционных факторов. Под мастер-регуляторами мы подразумеваем такие управляющие белки в регуляторной сигнальной сети, которые играют ключевую роль в управлении исследуемым молекулярным процессом. Мастер-регуляторы были отфильтрованы на основе данных об их дифференциальной экспрессии в экстремальных группах выживания и корреляции с продолжительностью жизни у пациентов с GBM. Проведённое исследование подтверждает полученные нами ранее результаты, согласно которым мастер-регуляторы *IGFBP2*, *PDGFA*, *OSMR* и *AEBP1* связаны с уменьшением времени выживания пациентов с GBM. Кроме того, мы выявили гены *CD14*, *CD44*, *DUSP6*, *GRB10*, *ILIRAP*, *FGFR3* и *POSTN* в качестве мастер-регуляторов выживания при глиобластоме. Эти мастер-регуляторы предложены в качестве перспективных мишеней для терапевтического воздействия при неблагоприятном прогнозе при GBM. На основе их анализа с использованием программ Genome Enhancer и PASS мы отобрали несколько перспективных лекарственных препаратов в качестве потенциальных средств для борьбы с агрессивными формами GBM и повышения выживаемости пациентов.

Ключевые слова: глиобластома; GBM; регуляторные сигнальные сети; мастер-регуляторы; *STAT3*; промоторы; Genome Enhancer; PASS

DOI: 10.18097/PBMC20216703201

ВВЕДЕНИЕ

Глиобластома (GBM, glioblastoma multiforme) — наиболее распространённая злокачественная первичная опухоль головного мозга [1]. Несмотря на существенные изменения в стратегиях лечения, лишь 2% пациентов фактически реагируют на стандартную терапию и выживают после 36 месяцев; они известны как долгоживущие пациенты (ДВЖ — длительная выживаемость) [2]. Пациентов, которые выживают менее 12 месяцев, называют короткоживущими пациентами (КВЖ — краткосрочная выживаемость). Мы также выделили группу пациентов, выживших в период от 12 месяцев до 36 месяцев (СВЖ — среднесрочная выживаемость).

Анализ дифференциально экспрессированных генов (ДЭГ) является важной и широко применяемой в настоящее время стратегией для идентификации биологических макромолекул, характеризующих переходные клеточные состояния. Снижение экспрессии таких генов, как *CH3L1*, *FBLN4*, *EMP3*,

IGFBP2, *IGFBP3*, *LGALS3*, *MAOB*, *PDPN*, *SERPING1* и *TIMP1* связано с продолжительностью выживания, что было показано на основе анализа экспрессии генов экстремальных групп выживания (ДВЖ и КВЖ) [3-6].

В проведённой нами ранее работе на основе анализа общедоступных наборов данных мы выявили примерно 200 генов, которые дифференциально экспрессируются в случаях ДВЖ и КВЖ [7]. При этом мы исключали группу СВЖ, которая составляет большинство в популяции пациентов. В настоящей работе, используя одномерный регрессионный анализ Кокса по всему набору данных TCGA (n=560), мы идентифицировали 720 генов, связанных с продолжительностью выживания пациентов при GBM.

Реконструкция генных регуляторных сетей по конкретным заболеваниям позволяет выявить мастер-регуляторы соответствующего патологического процесса, под которыми подразумеваются белки в сигнальной регуляторной сети, играющие ключевую



роль в управлении конкретным молекулярным процессом. Мы использовали программу Genome Enhancer (<https://genexplain.com/genome-enhancer/>), инструмент для мультиомиксного анализа, и применили его для реконструкции регуляторной сети с использованием 720 генов, отобранных по критерию коэффициента ложного обнаружения $FDR < 0,05$ (False Discovery Rate), которые были идентифицированы с помощью регрессионного анализа Кокса. В этом подходе первым шагом является анализ промоторов и энхансеров генов и поиск в них сайтов связывания транскрипционных факторов (ТФ), вовлечённых в их регуляцию и важных для исследуемого процесса. На втором шаге исследования сигнальных путей, активирующих эти ТФ, проводится идентификация мастер-регуляторов в управляющей части сигнальных регуляторных путей [8-10].

Мы использовали программу Genome Enhancer для анализа 720 найденных генов, которые оказали максимальное влияние на выживание пациентов с GBM. На первом этапе мы определили наиболее важные транскрипционные факторы, сайты которых обогащены в промоторах исследуемых генов. Из них *NR3C1* (GR — глюкокортикоидный рецептор) и *STAT3* имели максимальные значения оценок влияния на регуляцию генной активности, что свидетельствует об их роли в контроле экспрессии генов, кодирующих мастер-регуляторы. *STAT3* является регуляторным белком, играющим важную роль в патологии GBM. Известно, что он индуцирует пролиферацию клеток, обеспечивает поддержание стволовых клеток глиомы, связан с инвазией опухоли, с ангиогенезом и с уходом опухоли от иммунного ответа [11]. На следующем этапе анализа мы определили мастер-регуляторные белки системы передачи сигналов в клетках, которые регулируют активность найденных транскрипционных факторов. Среди них 4 регуляторных белка, кодируемых генами *IGFBP2*, *PDGFA*, *AEBP1*, *OSMR*, совпали с выявленными нами в предыдущей работе [7], где было показана их ассоциация с плохим прогнозом в случае GBM [7]. В настоящей работе мы показали, что активность нескольких дополнительных мастер-регуляторов, кодируемых генами *POSTN*, *CD14*, *CD44*, *DUSP6*, *FGFR3*, *GRB10* и *ILIRAP*, также играет критическую роль в выживании пациентов при глиобластоме.

Настоящая работа направлена на анализ сети патологической генной регуляции в GBM, приводящей к низкому уровню выживания пациентов. Идентифицированные мастер-регуляторы могут указать вероятные пути блокирования компонентов регуляторной сети патологического процесса, подавление которых будет способствовать его нормализации.

МЕТОДИКА

Сбор данных

Профили экспрессии генов у пациентов с GBM и соответствующая клиническая информация для таких пациентов были получены из базы данных TCGA [12].

Суммарный набор данных содержал 560 образцов, из которых для 540 образцов, принадлежащих 526 пациентам, имелась информация о выживаемости. В данном исследовании мы рассматривали образцы от 271 пациента с GBM с краткосрочной выживаемостью (КВЖ — выживаемость менее 12 месяцев), 240 пациентов со среднесрочной выживаемостью (СВЖ — выживаемость от 12 до 36 месяцев), и 49 пациентов с долгосрочной выживаемостью (ДВЖ — выживаемость свыше 36 месяцев). Информация об образцах и наборы данных приведены в дополнительных материалах (табл. S1-A, S1-B).

Предварительная обработка микрочиповых данных (Affymetrix)

Файлы исходных данных (формата CEL) микрочиповой платформы U133 Affymetrix (Human Genome U133 Plus 2.0 Array) были предварительно обработаны с использованием алгоритма RMA из пакета R (affy) для коррекции влияния неизвестных факторов, проверки качества данных, нормализации данных и преобразования значений в логарифмическую шкалу (\log_2) [13]. Затем, с использованием эмпирического подхода Байеса, проводили коррекцию данных, учитывающую различные источники данных [14]. Полученные нормализованные значения экспрессии использовали для дальнейшего анализа (дополнительные материалы, рис. S1). В случаях множественных вариантов соответствия между пробами Affymetrix и генами для каждого гена брали максимальное из значений интенсивностей проб, принадлежащих к одному гену (после фильтрации статистических выбросов в соответствии с используемым ранее подходом с применением реализованной в пакете Bioconductor функции (<https://bioconductor.org/packages/devel/workflows/vignettes/maEndToEnd/inst/doc/MA-Workflow.html>)). Окончательная матрица экспрессии состояла из 13914 проб и 560 образцов.

Идентификация дифференциально экспрессированных генов

Анализ дифференциальной экспрессии генов между КВЖ и ДВЖ группами GBM был выполнен с использованием метода Limma [15] с отсечением FDR 5%. Анализ выявил 191 ген, значительно отличающийся по экспрессии более чем на 0,5 в логарифмической шкале (по основанию логарифма 2) ($adj.p < 0,05$).

Влияние на выживание

Для анализа выживаемости мы использовали библиотеку “Survminer” из пакета Bioconductor (<https://cran.r-project.org/web/packages/survminer/survminer.pdf>). Одномерная регрессия Кокса для анализа выживаемости была проведена с использованием функции *coxph* для вычисления отношения рисков (hazard ratio, HR) и значения FDR, скорректированного на эффект многократного тестирования. Мы идентифицировали 720 генов с $FDR < 0,05$, которые были использованы в дальнейшем анализе

в программе Genome Enhancer. Графики выживаемости Каплана-Мейера (КМ) генерировали с использованием 50% не перекрывающихся верхних и нижних квантилей, основанных на медиане значений экспрессии.

Базы данных, использованные в исследовании

Анализ сайтов связывания транскрипционных факторов в промоторах исследуемых генов проводили с использованием известных мотивов связывания ДНК, описанных в библиотеке TRANSFAC®, релиз 2019.3 (“geneXplain GmbH”, Германия) (<https://genexplain.com/transfac>) [16]. Для поиска мастер-регуляторов использовали базу данных TRANSPATH®, релиз 2019.3 (“geneXplain GmbH”) (<https://genexplain.com/transpath>) [17]. Сеть передачи сигнала в человеческих клетках строили с помощью программного обеспечения Genome Enhancer [10] на основе реакций, аннотированных в базе данных TRANSPATH®. Информацию об идентификаторах и названиях генов брали из базы данных Ensembl build 99.38 (<http://www.ensembl.org>) [18].

Анализ сигнальных путей

Для изучения биологической значимости генных сигнатур проводили анализ обогащения сигнальных путей с использованием биномиального распределения для расчёта статистической значимости (p) и процедуры Бенджамини-Хохберга для расчёта $adj.p$, скорректированного на множественное тестирование значения p . Анализ обогащения сигнальных путей проводили путём картирования белковых продуктов, кодируемых входными 720 генами, на канонические сигнальные и метаболические пути в базах данных TRANSPATH® и Reactome (<https://reactome.org>).

Программа Genome Enhancer

С применением вышеперечисленных подходов мы оценили влияние изучаемых генов на процессы возникновения и развития GBM. Чтобы понять причину возникновения патологии, обусловленной нарушениями в нормальной регуляции экспрессии генов в GBM, мы использовали программу Genome Enhancer платформы geneXplain. Она включает в себя автоматизированный конвейер в соответствии с ранее опубликованным подходом — “анализ восходящих связей” (upstream analysis) [8, 9] и его развитием — подходом “шагающих путей” (walking pathways) [10]. При этом были использованы гены, найденные нами на предыдущем этапе анализа, экспрессия которых оказывает статистически значимое ($FDR < 0,05$) влияние на длительность выживания при GBM. Конвейер включает в себя 2 этапа: 1) анализ промоторов и энхансеров генов на обогащённость сайтами связывания транскрипционных факторов (ТФ), участвующих в их регуляции и, таким образом, важных для изучаемого процесса; 2) анализ сигнальных путей, активирующих эти ТФ, и определение мастер-регуляторов с максимальными оценками влияния на активность этих транскрипционных факторов из-за ключевого положения в сети передачи

сигналов. На первом этапе анализа была использована база данных TRANSFAC® вместе с алгоритмами идентификации сайтов связывания транскрипционных факторов (ТФ) — MATCH™ и CMA [19, 20]. На втором этапе — база данных передачи сигналов TRANSPATH® и специальные алгоритмы анализа графов. Программа Genome Enhancer также генерирует визуализацию полученных сетей передачи сигналов и генной регуляции и выделяет в ней найденные мастер-регуляторы, а также использует значения \log_2FC и p для раскраски узлов созданной регуляторной сети.

Приоритизация лекарств

В настоящей работе мы провели поиск оптимального сочетания молекулярных мишеней (ключевых элементов регуляторной сети клетки), которые взаимодействуют с различными фармацевтическими субстанциями из более чем 9200 лекарственных препаратов, использующихся в медицинской практике и проходящих клинические испытания (включая лекарства и созданные на основе низкомолекулярных веществ биопрепараты) из базы данных HumanPSD™ [21]. Из них для 2507 низкомолекулярных фармакологических веществ был выполнен прогноз спектров биологической активности и потенциальных мишеней с помощью программы PASS [22].

Сначала были отобраны лекарства из базы данных HumanPSD™, которые имеют, по крайней мере, одну мишень из полученного нами в результате проведенного анализа списка. Далее осуществляли приоритизацию этих препаратов на основе интеграции оценок двух параметров: ранжирование по показателю активности мишеней ($T\text{-score}_{PSD}$) и ранжирование по показателю связи с болезнью ($D\text{-score}_{PSD}$).

$T\text{-score}_{PSD}$ рассчитывали следующим образом:

$$T\text{-score}_{PSD} = \frac{|T|}{|T| + w(|AT| - |T|)} \sum_{t \in T} \log_{10} \left(\frac{\text{rank}(t)}{1 + \max\text{Rank}(T)} \right),$$

где T — набор всех мишеней, связанных с соединением, пересекающимся с исходным списком; $|T|$ — количество элементов в T ; AT и $|AT|$ — набор всех мишеней, связанных с соединением и количество элементов; w — весовой коэффициент; $\text{rank}(t)$ — ранг данной мишени; $\max\text{Rank}(T)$ равен $\max(\text{rank}(t))$ для всех мишеней t в T .

Мы использовали следующую формулу для подсчёта $D\text{-score}_{PSD}$:

$$D\text{-score}_{PSD} = \begin{cases} \sum_{d \in D} \sum_{p \in P} \text{phase}(p, d) \\ 0, D = \emptyset \end{cases},$$

где D — набор выбранных заболеваний, а если D — пустой набор, то $D\text{-score}_{PSD} = 0$. P — набор всех известных клинических испытаний для каждого заболевания, $\text{phase}(p, d)$ равен номеру фазы клинического испытания p , для выбранного заболевания d .

С помощью программы PASS мы предсказали спектры биологической активности 2507 структур фармакологических веществ из HumanPSD, включая их возможные побочные и токсические эффекты, а также возможные механизмы действия (мишени). PASS вычисляет вероятности наличия конкретных видов биологической активности, обозначаемые как Pa .

Были отобраны лекарственные препараты, для которых с вероятностями $Pa > 0,3$ прогнозировалось не менее двух из отобранного списка мишеней. Далее мы рассчитали приоритеты для этих соединений аналогично приведённому выше методу. Мы использовали рассчитанную на основе прогноза PASS оценку “Ранг лекарств”, который представляет собой сумму двух других рангов: ранга по “PASS показатель активности мишени” (T -Score) и ранга по “PASS показатель связи с болезнью” (D -Score), которые рассчитываются следующим образом.

PASS показатель активности мишеней:

$$T\text{-score}(s) = \frac{|T|}{|T| + w(|AT| - |T|)} \sum_{m \in M(s)} Pa(m) \sum_{g \in G(m)} IAP(g) \times optWeight(g),$$

где $M(s)$ — набор активностей, рассчитанных PASS для данной структуры (который преодолел выбранный порог $Pa > 0,3$); $G(m)$ — набор мишеней (преобразованных в гены), соответствующий данной активности (m) для данного соединения; $Pa(m)$ — вероятность активности (m), $IAP(g)$ — инвариантная точность предсказания для гена из $G(m)$; $optWeight(g)$ — дополнительный весовой множитель для гена, например, величина дифференциальной экспрессии этого гена (в данной работе этот множитель равен 1,0 для всех генов); T — множество всех мишеней, связанных с соединением, пересекающимся с входным списком; $|T|$ — количество элементов в T ; AT и $|AT|$ — множество всех мишеней, связанных с соединением и количество элементов в нём; w — весовой множитель, задаваемый пользователем.

Рассчитанный на основе прогноза PASS показатель связи с болезнью:

$$D\text{-score}(s) = \max_{m \in M(s)} (Pa(m)),$$

где $M(s)$ — набор активностей соответствующих изучаемой болезни и предсказанных программой PASS для данного соединения s ; $Pa(m)$ — вероятность для данного соединения иметь активность m .

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Идентификация генов, влияющих на выживание

Регрессионный анализ Кокса выявил 720 генов, которые оказались статистически значимо связаны с продолжительностью выживания при GBM. Среди генов с наиболее высоким отношением рисков были выявлены такие гены, как *PDCD1LG2* (HR=2,1), *PPA3* (HR=1,8), *SIGLEC9* (HR=1,7) и с наименьшим отношением рисков такие гены, как *MLNR* (HR=0,28), *ZNF208* (HR=0,35) и *NEUROG1* (HR=0,37).

Значение отношений рисков и статистическая значимость связи 720 генов с продолжительностью выживания при GBM приведены в дополнительных материалах (табл. S2-A).

Анализ сигнальных путей

Анализ сигнальных путей был выполнен путём картирования 720 найденных генов на канонические пути в базах данных TRANSPATH® и Reactome и оценки статистически значимого обогащения этих путей нашими генами. Мы выявили 35 коротких цепочек и полных сигнальных путей из базы данных TRANSPATH® и 22 сигнальных пути из базы данных Reactome, обогащённых белками, кодируемыми анализируемыми нами генами. Среди выявленных сигнальных путей мы обнаружили: сеть бета-катенина, сигнальный путь гипоксии, сигнальные пути и цепочки цитокина IL-6, сигнальные пути регуляции межклеточных взаимодействий, сигнальный путь интерферона. Известно, что многие из этих сигнальных путей играют важную роль на нескольких стадиях прогрессии опухоли. Полный список обогащённых сигнальных путей приведён в дополнительных материалах (табл. S2-B, S2-C). На рисунке 1 показаны обращающие на себя внимание две наиболее статистически значимые короткие цепи, вовлечённые в передачу сигналов к ТФ *STAT3* и в регуляцию генов при гипоксии, которые оказались обогащены несколькими генами из нашего списка, характеризующимися экстремальными значениями отношения риска выживания при GBM.

Анализ сайтов связывания транскрипционных факторов и композиционных модулей

На следующем этапе с помощью библиотеки мотивов связывания ТФ из базы данных TRANSFAC® мы выполнили анализ сайтов связывания транскрипционных факторов в промоторах (-1000 пар оснований до места старта транскрипции) 720-ти генов. Анализ был проведен при помощи метода СМА [19], который применяется для выявления транскрипционных факторов, взаимодействие которых приводит к синергичному влиянию на процесс генной регуляции. Используя СМА, мы определили два композиционных модуля факторов транскрипции, контролирующих экспрессию генов: модуль 1: *MYOGNF1*, *JUN*, *NFATC2*, *AP2*, *LEF1*, *TFAP2A*, *HOXA10*; модуль 2: *E2A*, *CEBPA*, *GR*, *MYOGNF1*, *MZF1*, *IK*, *NFIC*, *STAT3* (дополнительные материалы, рис. S2, табл. S2-D). Эти два композиционных модуля совместно обеспечивают достаточно хорошую дискриминацию промоторов 720 генов от промоторов генов домашнего хозяйства (p -тест Уилкоксона = $3,7 \times 10^{-33}$; AUC = 0,74, что достоверно выше, чем ожидалось для случайного множества регуляторных областей, Z-скор = 4,44).

Обращают на себя внимание три транскрипционных фактора, которые участвуют в контроле экспрессии выявленных генов: *JUN*, *STAT3* и *GR*. *JUN* является протоонкогеном, играющим критическую роль в пролиферации клеток и злокачественной трансформации, уровень которого, по имеющимся данным, повышен в GBM. *GR* — рецептор

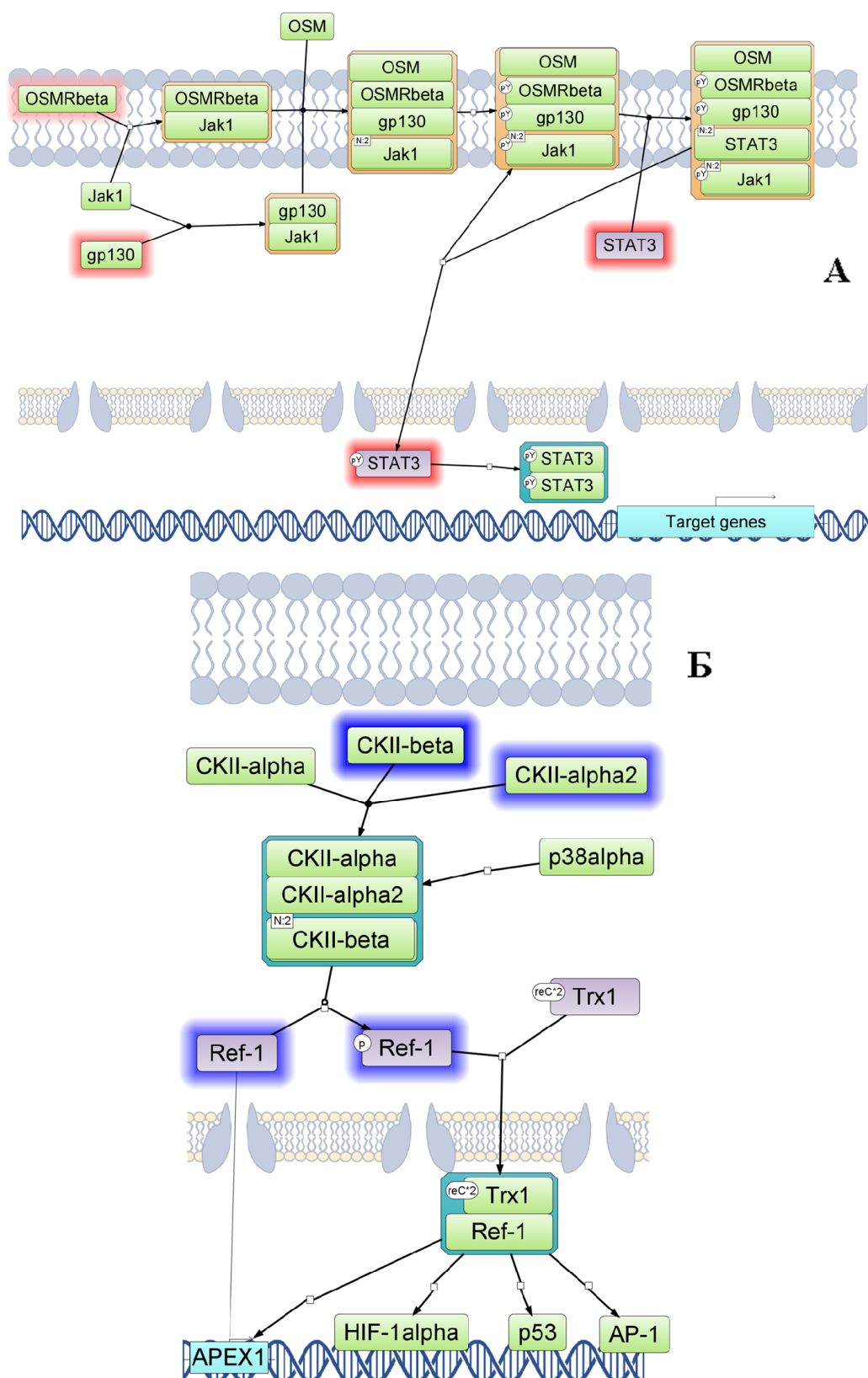


Рисунок 1. Две наиболее примечательные цепи передачи сигнала (из базы данных TRANSPATH®), статистически значимо обогащённые генами с экстремальными значениями отношения рисков выживания при GBM. А) Цепь *OSM* → *STAT3*, на которую картировались продукты трёх генов из нашего списка с высокими отношениями рисков плохой выживаемости при GBM (красные внешние овалы вокруг элементов диаграммы) ($p < 6,5 \times 10^{-4}$). Б) Цепь *CHN* → *AP-1*, на которую картировались продукты трёх генов из нашего списка с уменьшенными значениями отношения рисков плохой выживаемости при GBM (синие овалы вокруг элементов диаграммы) ($p < 6,5 \times 10^{-4}$). Интенсивность окраски овалов вокруг элементов диаграммы соответствовала значениям регрессии Кокса.

глюкокортикоидов — по некоторым сообщениям, способствует развитию фенотипа, подобного фенотипу стволовых клеток, и повышению резистентности к химиотерапии [23]. *STAT3* является очень важным транскрипционным фактором, связанным с GBM. Постоянная активация *STAT3* индуцирует пролиферацию клеток, блокирование апоптоза, поддержание стволовых клеток глиомы, инвазию опухоли, ангиогенез и избегание иммунного ответа [11, 24]. Мы выявили, что *STAT3* достоверно дифференцированно выражен у пациентов с краткосрочной выживаемостью ($\log_2FC=0,403$, $adj.p=0,00129$). *STAT3* также показал достоверное превышение отношения рисков выживаемости, $HR=1,4$ ($p=0,0015$ при $FDR=0,009$ — см. рис. 2). Таким образом, мы рассматриваем *STAT3* в качестве одной из перспективных терапевтических мишеней для контроля опухолевого процесса путём формирования иммунного микроокружения опухоли [11].

Поиск мастер-регуляторов в сигнальных сетях

С помощью базы данных сигнальной трансдукции TRANSPATH® и алгоритма Genome Enhancer мы идентифицировали мастер-регуляторы, контролирующие активность выявленных факторов транскрипции. Алгоритм Genome Enhancer осуществляет поиск ключевых узлов в глобальной сети трансдукции сигналов в активирующие транскрипционные факторы, как это описано ранее [8, 9], и фильтрует идентифицированные ключевые узлы по критериям наличия петель положительной обратной связи [10] (подход “шагающих путей”). Это соответствует требованию, чтобы белки ключевых узлов кодировались генами, которые, в свою очередь, находятся под регуляторным контролем этих ключевых узлов.

В результате анализа таких мастер-регуляторов на основе исследования 720 генов, связанных с выживаемостью при GBM, нами было выявлено 43 потенциальных мастер-регулятора, имеющих петли положительной обратной связи. Полный

список выявленных мастер-регуляторов приведён в дополнительных материалах (табл. S2-E). Эти мастер-регуляторы в основном находятся в таких сигнальных путях, как сигнальная сеть бета-катенина, EGF сигнальный путь, стресс-ассоциированный путь и путь активации цитокинов. Мы построили тепловую карту значений экспрессии этих 43 генов-мастер-регуляторов (дополнительные материалы, рис. S3). Она разделяет их на две подгруппы генов — в среднем с более высокой экспрессией в группе КВЖ (*POSTN*, *IGFBP2*, *FGFR3* и др.) и в среднем более высокой экспрессией в группе ДВЖ (*CASP9*, *PARD3*, *APEX1*). Тем не менее, мы видим очень высокую вариабельность экспрессии этих генов в каждой из этих групп. Это указывает на то, что мастер-регуляторы из обеих этих групп могут служить перспективными мишенями для воздействия на продолжительность выживаемости пациентов с GBM.

Определение перспективных мишеней для воздействия лекарствами

Выявленные мастер-регуляторы можно рассматривать в качестве ключевых кандидатов для терапевтических мишеней, так как они оказывают существенное воздействие на регуляцию внутриклеточных путей, активизирующих патологические процессы в GBM, которые были выявлены в ходе нашего исследования. Для выбора наиболее перспективных мишеней лекарственных средств мы, прежде всего, отфильтровали все найденные мастер-регуляторы, основываясь на их дифференциальной экспрессии ($\log_2FC>0,5$) между группами ДВЖ и КВЖ. Полные результаты анализа Limma приведены в дополнительных материалах (табл. S2-F). Мы выделили несколько мастер-регуляторов, для которых наблюдалось статистически значимое различие между крайними группами выживания (полный список см. в дополнительных материалах, табл. S2-G). Из них 11 основных регуляторов, имеющих более

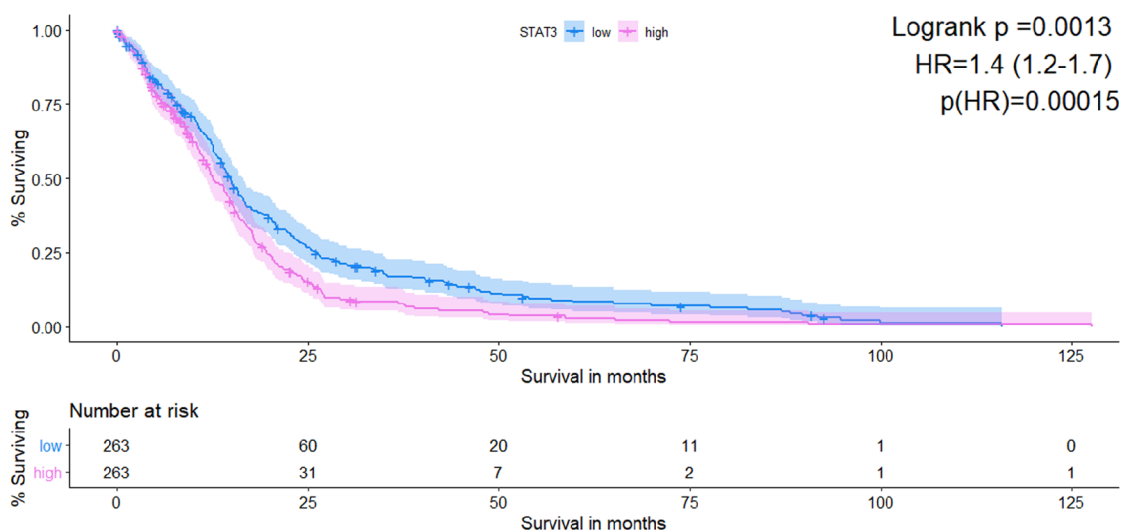


Рисунок 2. График кривых Каплана-Мейера для отображения связи экспрессии гена *STAT3* с выживаемостью с использованием 526 образцов данных из базы данных TCGA-GBM. На диаграмме также указана рассчитанная величина отношения рисков (HR) и статистическая значимость ($p(HR)$) в соответствии с оценками выживаемости Кокса.

высокие значения дифференциальной экспрессии ($\log_2FC > 0,5$) в экстремальных группах выживания, были рассмотрены для реконструкции сети геной регуляции. Ниже мы охарактеризуем эти гены с точки зрения их вовлечённости в канцерогенез.

Периостин (POSTN) — секретируемый белок внеклеточного матрикса — играет важную роль в прогрессии GBM, а также в увеличении инвазивности опухоли и играет важную роль в клиническом ответе на антиангиогенную терапию [25]. Следующий мастер-регулятор — рецептор-связывающий белок 10 (GRB10), субстрат mTOR — известен в качестве участника PI3K-AKT пути передачи сигнала в клетке и связан с прогрессированием роста опухоли при раке простаты [26]. Было показано, что он имеет высокую экспрессию в мезенхимальном подтипе опухолей, но более низкую экспрессию в подтипах опухолей G-CIMP GBM [27]. Различные aberrации *FGFR* — гена, кодирующего рецептор фактора роста фибробластов, — широко вовлечены в развитие и прогрессирование опухолей различных типов и включают в различных типах рака такие изменения, как повышенная экспрессия этого гена, различные амплификации, мутации, вариации белковых изоформ, возникающих в результате альтернативного сплайсинга, транслокации *FGFR* [28]. Кроме того, изменения экспрессии гена *FGFR* в астроцитах могут привести к злокачественной трансформации и прогрессии GBM в результате активации процессов деления, миграции клеток и подавления апоптоза [28]. *IGFBP2* считается одним из наиболее важных биомаркеров агрессивного развития GBM [29, 30], а также известен как прогностический маркер низкой продолжительности выживания при GBM [30, 31]. *IGFBP2* наряду с *AEBP1* (*ACLP*), *PDGFA* и *OSMR* являются мастер-регуляторами, связанными с короткой выживаемостью при GBM, что было показано нами ранее [7].

Ген *CD14* модулирует клеточный и гуморальный иммунный ответ, взаимодействуя непосредственно с Т- и В-клетками и играет важную роль в подавлении иммунного ответа, что коррелирует с плохим прогнозом при GBM [32]. Как известно из литературы, повышенная экспрессия *CD44* предсказывает плохую выживаемость при GBM и играет важную роль в прогрессировании GBM; этот ген рассматривается как перспективная терапевтическая мишень [33, 34]. Установлено, что при GBM повышенная экспрессия гена фосфатазы 6 двойной специфичности — *DUSP6* — играет жизненно важную роль в эпителиально-мезенхимальном переходе при прогрессии опухоли [35].

Необходимо отметить, что, как это видно на тепловой карте (дополнительные материалы, рис. S3), мы наблюдали очень высокую вариабельность значений экспрессии одних и тех же генов между различными образцами GBM. Хотя наблюдается слегка пониженная экспрессия ряда генов в среднем во всей группе КВЖ, однако для достаточно большой подгруппы образцов из этой группы видны высокие значения экспрессии этих генов (см. тепловую карту значений экспрессии 43 генов — рис. S3

в дополнительных материалах). Поэтому, некоторые важные мишени лекарств могли быть найдены среди таких генов, экспрессия которых особенно высока только лишь в подмножестве образцов КВЖ. Такие мишени могут иметь узкую специфичность для подгруппы пациентов с плохим прогнозом по выживаемости. Поэтому мы расширили круг поиска перспективных мишеней и рассмотрели также все другие потенциальные мастер-регуляторы с обратной связью, обращая особое внимание на их известную или предсказанную с помощью программы PASS роль в качестве мишеней для препаратов, связанных с терапией рака либо других заболеваний.

Все выявленные мастер-регуляторы, были оценены с точки зрения их потенциальной лекарственной пригодности (drugability score) с использованием базы данных биомаркеров и лекарственных препаратов HumanPSD™ [21] и программного обеспечения PASS [22] для прогнозирования биологической активности. Оценка лекарственной пригодности представлена количеством известных препаратов, которые действуют на соответствующую мишень, либо на основе информации из базы данных HumanPSD™, либо по данным прогноза активности соединений в отношении исследуемой мишени с помощью программы PASS. Таким образом, мы провели дальнейший отбор мишеней для лекарств с использованием параметра лекарственной пригодности и добавили три дополнительные мишени: *APEX1*, *MAPK8* и *PTK2B*. Эти три мишени, хотя и имеют относительно низкое значение отношения рисков (HR), но характеризуются достаточно высокой потенциальной лекарственной пригодностью.

APEX Nuclease (фермент многофункционального восстановления ДНК) 1 — это фермент репарации ДНК, который, как известно, положительно коррелирует с изменённым статусом гена *MGMT*, с признаками резистентности к лечению тимозоломидом, с рецидивами GBM и с поляризацией к иммуносупрессивному микроокружению в GBM [36]. Считается, что сниженная экспрессия *APEX1* повышает чувствительность к обработке тимозоломидом в резистентных линиях клеток GBM [37]. Установлено, что гены *MAPK8* и *MAPK* активируются в линиях клеток GBM, устойчивых к тимозоломиду. Сигнальные пути, идущие от *MAPK8*, также усиливают пролиферацию клеток и тормозят апоптоз [38]. Экспрессия *PTK2B* (PYK2, белковая тирозинкиназа 2 бета) предположительно играет критическую роль в миграционном поведении опухолевых клеток, что приводит к большей агрессивности в GBM [39]. Помимо этих известных мишеней, в качестве новых потенциальных мишеней, роль которых в терапии агрессивных форм GBM ещё предстоит подтвердить, можно выделить два гена, являющихся компонентами сигнального пути бета-катенина — *CSNK2A2* и *CSNK2B* (рис. 1B) [40].

Схема сети мастер-регуляторов с положительными контурами обратной связи представлена на рисунке 3. Как было отмечено выше, выявленные нами мастер-регуляторы потенциально могут выступать в качестве мишеней для терапии GBM.

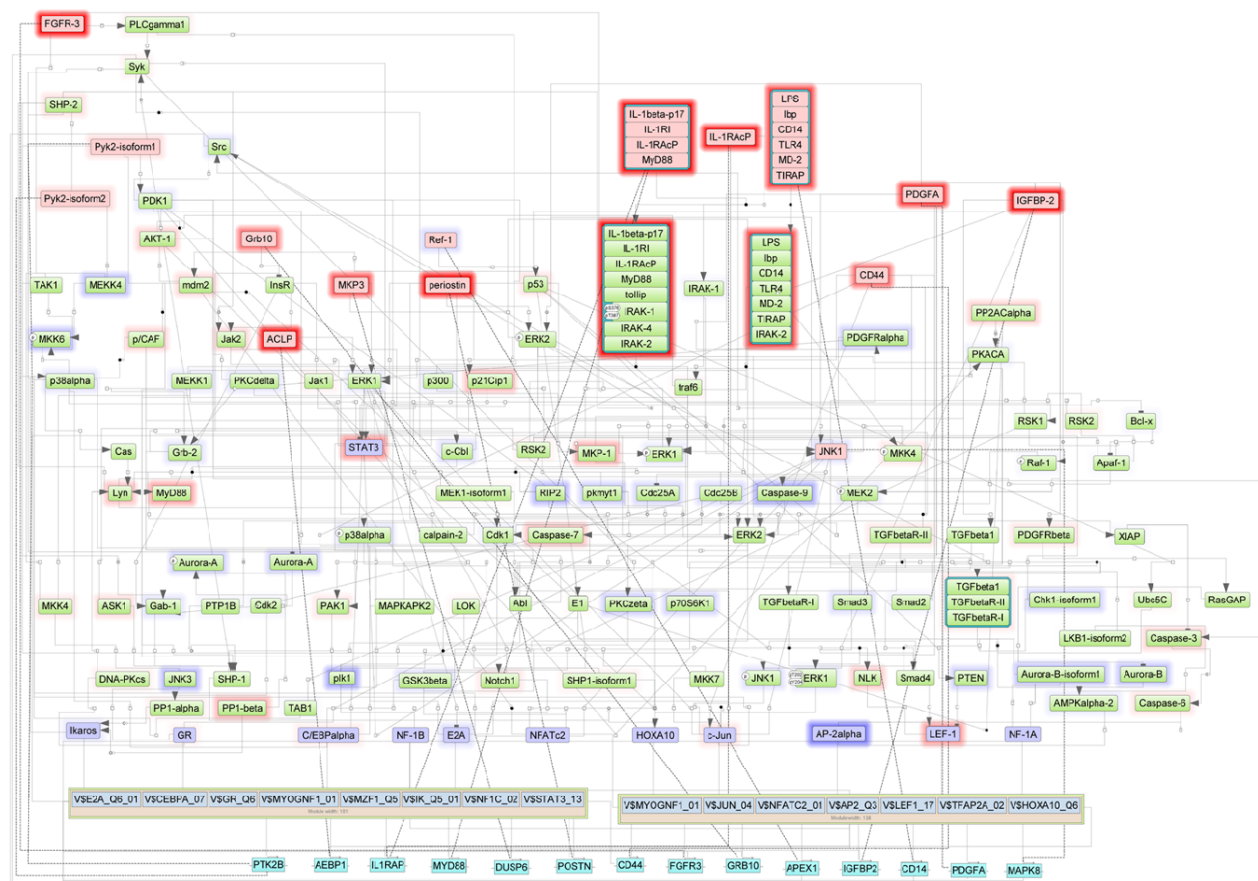


Рисунок 3. Сеть сигнальной трансдукции и генной регуляции найденных нами мастер-регуляторов (красные узлы), регулирующих два модуля транскрипционных факторов (фиолетовые узлы), которые были нами обнаружены как обогащённые в промоторах исследуемых генов. Пунктирные линии от генов до кодируемых ими сигнальных белков представляют собой процессы транскрипции и трансляции (петли положительной обратной связи). Сеть построена с использованием мастер-регуляторов, которые имеют значения $\log_2FC > 0,5$. OSMR включён для проверки ранее показанного драйвера короткой выживаемости при GBM 7. На диаграмму также добавлены три дополнительных мастер-регулятора (APEX1, MAPK8, PTK2B) в качестве известных мишеней лекарств. Расцветка овалов вокруг элементов диаграммы основана на дифференциальной экспрессии генов в текущем исследовании и заполняется красным цветом при повышенном уровне экспрессии в группе КВЖ ($\log_2FC > 0,1$) и синим при пониженном уровне экспрессии ($\log_2FC < -0,1$).

Приоритизация потенциальных препаратов

Мы ранжировали лекарства, известные или предсказанные с помощью программы PASS, на основе ранга препаратов (drug rank), который представляет собой сумму частичных рангов, вычисленных, во-первых, на основе роли мишеней этого лекарства в качестве мастер-регуляторов исследуемой сети регуляции генов (target activity score), во-вторых — на основе оценки возможности этой мишени взаимодействовать с лекарствами, а также на основе показателя связи с болезнью (disease activity score) (см. раздел “Методика”). Результаты представлены в таблицах 1 и 2.

Среди препаратов с наивысшим приоритетом имеется несколько фармакологических веществ, которые используются или проходят клинические и доклинические испытания при глиоме, GBM, неоплазмах центральной нервной системы и других неоплазмах. Среди них можно выделить такие препараты, как лефлуномид (клиническое исследование: NCT00003293), нинтаниб [41], памидронат [42],

палифермин (в клеточных линиях) [43]. В другом исследовании мета-анализ показал, что этопозид и тенипозид улучшают выживаемость при глиоме [44].

Применение нашего алгоритма также позволило выбрать несколько дополнительных препаратов, известных при лечении других заболеваний, которые могут быть предложены для дальнейших исследований с целью репозиционирования лекарств для лечения GBM. В частности, в последнее время сообщалось о повышении эффективности лечения GBM доксорубицином (широко известным противоопухолевым антибиотиком) в сочетании со стандартной терапией [45].

ЗАКЛЮЧЕНИЕ И ВЫВОДЫ

Настоящая работа посвящена анализу данных, полученных из архива TCGA по GBM ($n=560$), с целью идентификации мастер-регуляторов, определяющих экспрессию генов, которые оказывают существенное влияние на выживание при GBM.

Таблица 1. Наиболее перспективные лекарственные препараты, выбранные для воздействия на найденные нами мишени на основе информации по различным литературным источникам собранной в базе данных HumanPSD™

Препарат	Идентификатор мишени	Статус (согласно Drugbank)	Клинические испытания (Фаза)	Оценочная функция для мишени	Оценочная функция для болезни	Ранг
Palifermin	FGFR3	biotech, approved	Brain Abscess (1,3,4); Neoplasms (1,2,3); Leukemia (1,2,3); Multiple Myeloma (1,2,3,4), Mucosis (3)	0,1115	0	16
Pazopanib	FGFR3	small molecule, approved	Glioma (1,2); Neoplasms (1,2,3,4), Central Nervous System Neoplasms (2,3)	0,0820	9	17
Leflunomide	PTK2B	small molecule, approved, investigational	Arthritis (1,2,3,4); Psoriatic (1); Central Nervous System Neoplasms (2,3)	0,0874	0	20
Lenvatinib	FGFR3	small molecule, approved	Adenocarcinoma (1,2), Neoplasms (1,2,3)	0,0899	3	21
Nintedanib	FGFR3	small molecule, approved	Adenocarcinoma (1,2); Neoplasms (1,2,3), Pulmonary Fibrosis (3,4)	0,0753	0	23
XL999	FGFR3	small molecule, investigational	Lung neoplasms (1,2); Neoplasms (1,2); Brain Abscess (2)	0,1267	0	24
Ponatinib	FGFR3	small molecule, approved	Leukemia (1,2,3); Neoplasms (1,2,3)	0,0569	2	27
Hyaluronic acid	CD44	small molecule, approved	Arthritis (1,2,3,4); Osteoarthritis (1,2,3,4); Glaucoma (4)	0,0009	0	28
Lucanthone	APEX1	small molecule, approved, investigational	Glioblastoma (2), Neoplasms (2)	0,0755	2	29
Genistein	PTK2B	small molecule, investigational	Carcinoma (1); Neoplasms (1,2,3); Bone Diseases (3,4)	0,0559	0	30
Pyrazolanthrone	MAPK8	small molecule, experimental		0,1050	0	30
Flavopiridol	CDK8	small molecule, experimental, investigational	Carcinoma (1), Lymphoma (1,2)	0,0488	0	35
Adenosine triphosphate	NAE1	small molecule, approved, nutraceutical	Neoplasms (1), Pain (1); Alzheimer Disease (2)	0,0027	0	36

Таблица 2. Перспективные лекарственные препараты, которые, согласно прогнозам программы PASS, являются активными в отношении выявленных нами мишеней для лечения GBM

Препарат	Идентификаторы мишеней	Оценочная функция для мишени	Оценочная функция для болезни	Ранг
2,5,7-Trihydroxynaphthoquinone	MAPK8, DUSP5, DUSP6, DUSP3	0,0480	0,380	8
Teniposide	APEX1, CASP9	0,0443	0,378	12
Daunorubicin	STAT3, APEX1	0,0506	0,290	14
Doxorubicin	STAT3, APEX1	0,0423	0,308	18
Epirubicin	STAT3, APEX1	0,0423	0,308	18
Idarubicin	STAT3, APEX1	0,0426	0,293	18
Pyrazolanthrone	MAPK8, PTK2B, FES	0,0359	0,420	19
Etoposide	APEX1, CASP9	0,0343	0,422	19
Alendronate	DUSP5, FGFR3, DUSP6, DUSP3	0,0445	0,248	22
Pamidronate	DUSP5, FGFR3, DUSP6, DUSP3	0,0411	0,255	26
Fluorouracil	FGFR3, PTK2B, FES	0,0226	0,537	31
Oxybenzone	MAPK8, DUSP5, DUSP6, DUSP3	0,0463	0,202	36

Мы выявили такие гены и построили сеть регуляции генов, обуславливающую плохой прогноз при GBM, используя компьютерную программу Genome Enhancer. Мы определили транскрипционные факторы, потенциально участвующие в этом процессе. Среди них был выявлен транскрипционный фактор *STAT3*, который показал максимальные значения показателей участия в регуляции (как мастер-регулятор), дифференциальной экспрессии в экстремальных группах выживания, а также в оценках отношения рисков выживания при GBM. Наряду с генами, выявленными нами в предыдущей работе — *IGFBP2*, *PDGFA*, *OSMR* и *AEBP1*, мы предложили ещё 7 мастер-регуляторов, которые потенциально могут выступать в качестве терапевтических мишеней. Мы показали, что *STAT3*, а также некоторые другие транскрипционные факторы находятся в петле положительной обратной связи с этими мастер-регуляторами, в результате возникают патологически самоусиливающиеся процессы в опухолевых клетках, ведущие к плохой выживаемости при GBM. На основе реконструированной мастер-регуляторной сети генной регуляторной и трансдукции сигнала в клетках, а также с помощью компьютерной программы PASS мы определили наиболее перспективные лекарственные мишени и приоритетные препараты, которые потенциально могут быть использованы для лечения агрессивных форм GBM.

БЛАГОДАРНОСТИ

Авторы выражают признательность Э. Вингендеру за помощь в обсуждении работы.

ФИНАНСИРОВАНИЕ

Данный проект финансировался из средств программы Европейского Союза в области исследований и инноваций “Горизонт 2020” в рамках грантового соглашения Марии Склодовской-Кюри № 766069.

СОБЛЮДЕНИЕ ЭТИЧЕСКИХ СТАНДАРТОВ

Настоящая статья не содержит каких-либо исследований с участием людей или с использованием животных в качестве объектов.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы М. Каля и А.Е. Кель являются сотрудниками компании geneXplain.

Доступность программного обеспечения, данных и материалов

Анализируемый в настоящем исследовании набор данных, дополнительные файлы и графики доступны в проекте GitHub по следующей ссылке: https://github.com/genexplain/Manasa_KP_et_al_Master_regulators_of_poor_prognosis_in_Glioblastoma

Дополнительные материалы доступны в электронной версии статьи на сайте журнала (pbmc.ibmc.msk.ru).

ЛИТЕРАТУРА

1. Wen P.Y., Kesari S. (2008) N. Engl. J. Med., **359**, 492-507.
2. Krex D., Klink B., Hartmann C., von Deimling A., Pietsch T., Simon M., Sabel M., Steinbach J.P., Heese O., Reifenberger G., Weller M., Schackert G., German Glioma Network (2007) Brain, **130**(10), 2596-2606.
3. de Vega S., Iwamoto T., Yamada Y. (2009) Cell. Mol. Life Sci., **66**, 1890-1902.
4. Bi W.L., Beroukhi R. (2014) Neuro-Oncology, **16**(9), 1159-1160.
5. Reifenberger G., Weber R.G., Riehm V., Kaulich K., Willscher E., Wirth H., Gietzelt J., Hentschel B., Westphal M., Simon M., Schackert G., Schramm J., Matschke J., Sabel M.C., Gramatzki D., Felsberg J., Hartmann C., Steinbach J.P., Schlegel U., Wick W., Radlwimmer B., Pietsch T., Tonn J.C., von Deimling A., Binder H., Weller M., Loeffler M., German Glioma Network (2014) Int. J. Cancer, **135**(8), 1822-1831.
6. Franceschi S., Mazzanti C.M., Lessi F., Aretini P., Carbone F.G., la Ferla M., Scatena C., Ortenzi V., Vannozzi R., Fanelli G., Pasqualetti F., Bevilacqua G., Zavaglia K., Naccarato A.G. (2015) Oncology Lett., **10**(6), 3599-3606.
7. Kalya M.P., Kel A., Wlochowitz D., Wingender E., Beißbarth T. (2021) Front. Genet., **12**, DOI: 10.3389/fgene.2021.670240.
8. Koschmann J., Bhar A., Stegmaier P., Kel A.E., Wingender E. (2015) Microarrays (Basel), **4**(2), 270-286.
9. Boyarskikh U., Pintus S., Mandrik N., Stelmashenko D., Kiselev I., Evshin I., Sharipov R., Stegmaier P., Kolpakov F., Filipenko M., Kel A. (2018) BMC Med Genomics, **11**, 12. DOI: 10.1186/s12920-018-0330-5.
10. Kel A., Boyarskikh U., Stegmaier P., Leskov L.S., Sokolov A.V., Yevshin I., Mandrik N., Stelmashenko D., Koschmann J., Kel-Margoulis O., Krull M., Martínez-Cardús A., Moran S., Esteller M., Kolpakov F., Filipenko M., Wingender E. (2019) BMC Bioinformatics, **20**, 119. DOI: 10.1186/s12859-019-2687-7.
11. Chang N., Ahn S.H., Kong D.-S., Lee H.W., Nam D.-H. (2017) Mol. Cell. Endocrinol., **451**, 53-65.
12. Grossman R.L., Heath A.P., Ferretti V., Varmus H.E., Lowy D.R., Kibbe W.A., Staudt L.M. (2016) N. Engl. J. Med., **375**, 1109-1112.
13. Gautier L., Cope L., Bolstad B.M., Irizarry R.A. (2004) Bioinformatics, **20**(3), 307-315.
14. Leek J.T., Johnson W.E., Parker H.S., Jaffe A.E., Storey J.D. (2012) Bioinformatics, **28**(6), 882-883.
15. Ritchie M.E., Phipson B., Wu D., Hu Y., Law C.W., Shi W., Smyth G.K. (2015) Nucleic Acids Res., **43**(7), e47. DOI: 10.1093/nar/gkv007.
16. Wingender E., Dietze P., Karas H., Knüppel R. (1996) Nucleic Acids Res., **24**(1), 238-241.
17. Krull M., Voss N., Choi C., Pistor S., Potapov A., Wingender E. (2003) Nucleic Acids Res., **31**(1), 97-100.
18. Aken B.L., Ayling S., Barrell D., Clarke L., Curwen V., Fairley S., Fernandez Banet J., Billis K., García Girón C., Hourlier T., Howe K., Kähäri A., Kokocinski F., Martin F.J., Murphy D.N., Nag R., Ruffier M., Schuster M., Tang Y.A., Vogel J.-H., White S., Zadiša A., Flicek P., Searle S.M.J. (2016) Database, **2016**, baw093. DOI: 10.1093/database/baw093.
19. Kel A.E., Gössling E., Reuter I., Cheremushkin E., Kel-Margoulis O.V., Wingender E. (2003) Nucleic Acids Res., **31**(13), 3576-3579.
20. Waleev T., Shtokalo D., Konovalova T., Voss N., Cheremushkin E., Stegmaier P., Kel-Margoulis O., Wingender E., Kel A. (2006) Nucleic Acids Res., **34**(Suppl_2), W541-W545.

21. Wingender E., Hogan J., Schacherer F., Potapov A.P., Kel-Margoulis O. (2007) *In Silico Biol.*, **7**, S17-S25.
22. Filimonov D.A., Druzhilovskiy D.S., Lagunin A.A., Glorizova T.A., Rudik A.V., Dmitriev A.V., Pogodin P.V., Poroikov V.V. (2018) *Biomedical Chemistry: Research and Methods*, **1**(1), e00004. DOI: 10.18097/BMCRM00004.
23. Kostopoulou O.N., Mohammad A.-A., Bartek J., Winter J., Jung M., Stragliotto G., Söderberg-Nauclér C., Landázuri N. (2018) *Int. J. Cancer*, **142**(6), 1266-1276.
24. Jahani-Asl A., Yin H., Soleimani V.D., Haque T., Luchman H.A., Chang N.C., Sincennes M.-C., Puram S.V., Scott A.M., Lorimer I.A.J., Perkins T.J., Ligon K.L., Weiss S., Rudnicki M.A., Bonni A. (2016) *Nat. Neurosci.*, **19**, 798-806.
25. Emini E., Ramos-Moreno T., Stefani R.F., Svensson A., Bengzon J. (2018) *J. Stem Cell Res. Ther.*, **8**, 414. DOI: 10.4172/2157-7633.1000414.
26. Khan M.I., Johani A.A., Hamid A., Ateeq B., Manzar N., Adhami V.M., Lall R.K., Rath S., Sechi M., Siddiqui I.A., Choudhry H., Zamzami M.A., Havighurst T.C., Huang W., Ntambi J.M., Mukhtar H. (2019) *FASEB J.*, **33**(3), 3198-3211.
27. Smith A.A., Huang Y.-T., Eliot M., Houseman E.A., Marsit C.J., Wiencke J.K., Kelsey K.T. (2014) *Epigenetics*, **9**(6), 873-883.
28. Jimenez-Pascual A., Siebzehnruhl F.A. (2019) *Cells*, **8**(7), 715. DOI: 10.3390/cells8070715.
29. Holmes K.M., Annala M., Chua C.Y.X., Dunlap S.M., Liu Y., Hugen N., Moore L.M., Cogdell D., Hu L., Nykter M., Hess K., Fuller G.N., Zhang W. (2012) *Proc. Natl. Acad. Sci. USA*, **109**(9), 3475-3480.
30. Phillips L.M., Zhou X., Cogdell D.E., Chua C.Y., Huisinga A., Hess K.R., Fuller G.N., Zhang W. (2016) *J. Pathol.*, **239**(6), 355-364.
31. McDonald K.L., O'Sullivan M.G., Parkinson J.F., Shaw J.M., Payne C.A., Brewer J.M., Young L., Reader D.J., Wheeler H.T., Cook R.J., Biggs M.T., Little N.S., Teo C., Stone G., Robinson B.G. (2007) *J. Neuropathol. Exper. Neurol.*, **66**(5), 405-417.
32. Deininger M.H., Meyermann R., Schluesener H.J. (2003) *Acta Neuropathol.*, **106**, 271-277.
33. Bradshaw A., Wickremsekera A., Tan S.T., Peng L., Davis P.F., Itinteang T. (2016) *Front. Surg.*, **3**, 21. DOI: 10.3389/fsurg.2016.00021.
34. Si D., Yin F., Peng J., Zhang G. (2020) *CMAR*, **2020**(12), 769-775. DOI: 10.2147/CMAR.S233423.
35. Zuchegna C., di Zazzo E., Moncharmont B., Messina S. (2020) *BMC Res. Notes*, **13**, 374. DOI: 10.1186/s13104-020-05214-y.
36. Hudson A.L., Parker N.R., Khong P., Parkinson J.F., Dwight T., Ikin R.J., Zhu Y., Chen J., Wheeler H.R., Howell V.M. (2018) *Front. Oncol.*, **8**, 314. DOI: 10.3389/fonc.2018.00314.
37. Montaldi A.P., Godoy P.R.D.V., Sakamoto-Hojo E.T. (2015) *Mutat. Res. Genet. Toxicol. Environ. Mutagen.*, **793**, 19-29. DOI: 10.1016/j.mrgentox.2015.06.001.
38. Xu P., Zhang G., Hou S., Sha L. (2018) *Biomedicine Pharmacotherapy*, **106**, 1419-1427.
39. Lipinski C.A., Tran N.L., Menashi E., Rohl C., Kloss J., Bay R.C., Berens M.E., Loftus J.C. (2005) *Neoplasia*, **7**(5), 435-445.
40. Zheng Y., McFarland B.C., Drygin D., Yu H., Bellis S.L., Kim H., Bredel M., Benveniste E.N. (2013) *Clin. Cancer Res.*, **19**(23), 6484-6494.
41. Muhic A., Poulsen H.S., Sorensen M., Grunnet K., Lassen U. (2013) *J. Neurooncol.*, **111**, 205-212.
42. Jarry U., Chauvin C., Joalland N., Léger A., Minault S., Robard M., Bonneville M., Oliver L., Vallette F.M., Vié H., Pecqueur C., Scotet E. (2016) *Oncol Immunology*, **5**(6), e1168554. DOI: 10.1080/2162402X.2016.1168554.
43. Brake R., Starnes C., Lu J., Chen D., Yang S., Radinsky R., Borges L. (2008) *Mol. Cancer Res.*, **6**(8), 1337-1346.
44. Leonard A., Wolff J.E. (2013) *Anticancer Res.*, **33**, 3307-3315.
45. Norouzi M., Yathindranath V., Thliveris J.A., Kopec B.M., Siahaan T.J., Miller D.W. (2020) *Sci. Rep.*, **10**, 11292. DOI: 10.1038/s41598-020-68017-y.

Поступила в редакцию: 25. 04. 2021.
После доработки: 05. 05. 2021.
Принята к печати: 11. 05. 2021.

MASTER REGULATORS ASSOCIATED WITH POOR PROGNOSIS IN GLIOBLASTOMA MULTIFORME

M.P. Kalya^{1,2}, T. Beißbarth¹, A. Kel^{2,3*}

¹Department of Medical Bioinformatics, University Medical Center Göttingen, Göttingen, 37099 Germany

²geneXplain GmbH, Wolfenbüttel, 38302 Germany

³Institute of Chemical Biology and Fundamental Medicine SBAS,

8 Ak. Lavrentiev ave., Novosibirsk, 630090 Russia; *e-mail: alexander.kel@genexplain.com

Glioblastoma multiforme (GBM) is a highly malignant brain tumor with average survival time of 15 months. Less than 2% of the patients survive beyond 36 months. To understand the molecular mechanism responsible for poor prognosis, we analyzed GBM samples of TCGA microarray (n=560) data. We have identified 720 genes that have a significant impact upon survival based on univariate cox regression. We applied the Genome Enhancer pipeline to analyze potential mechanisms of regulation of activity of these genes and to build gene regulatory networks. We identified 12 transcription factors enriched in the promoters of these genes including the key molecule of GBM — *STAT3*. We found that *STAT3* had significant differential expression across extreme survivor groups (short-term survivors— survival <12 months and long-term survivors – survival >36 months) and also had a significant impact on survival. In the next step, we identified master regulators in the signal transduction network that regulate the activity of these transcription factors. Master regulators are filtered based on their differential expression across extreme survivors groups and impact on survival. This work validates our earlier report on master regulators *IGFBP2*, *PDGFA*, *OSMR*, and *AEBP1* driving short survival. Additionally, we propose *CD14*, *CD44*, *DUSP6*, *GRB10*, *ILIRAP*, *FGFR3*, and *POSTN* as master regulators driving poor survival. These master regulators are proposed as promising therapeutic targets to counter poor prognosis in GBM. Finally, the algorithm has prioritized several drugs for the further study as potential remedies to conquer the aggressive forms of GBM and to extend survival of the patients.

Key words: glioblastoma; gene regulatory networks; master regulators; upstream analysis; STAT3; transcription factors

Funding. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement no. 766069.

Received: 25.04.2021, revised: 05.05.2021, accepted: 11.05.2021.