

Современная оценка нкРНК как биомаркеров прогноза при светлоклеточном почечно-клеточном раке

Логинов В.И.^{1, 2*}, Бурденный А.М.^{1, 3}, Пронина И.В.¹, Лукина С.С.¹, Брага Э.А.^{1, 2}

¹ Научно-исследовательский институт общей патологии и патофизиологии, Москва, Россия

² Медико-генетический научный центр им. академика Н.П. Бочкова, Москва, Россия

³ Институт биохимической физики им. Н.М. Эмануэля РАН, Москва, Россия

* loginov7w@gmail.com

Ключевые слова: мРНК; миРНК; днРНК; светлоклеточный почечноклеточный рак; скПКР

Мотивация и цель: Ключевую роль в развитии и прогрессии опухолей играют эпигенетические изменения. В частности, результаты многочисленных исследований однозначно подтверждают вовлеченность в онкогенез аномального метилирования ДНК. Региональное гиперметиличивание промоторных CpG-островков ряда генов (белок-кодирующих, микроРНК, длинных некодирующих РНК и др.) – еще одно распространенное явление при опухолевом процессе. Результатом таких эпигенетических изменений может стать ингибирование транскрипции генов-супрессоров, обеспечивающих в нормальных клетках защиту от злокачественной трансформации [1]. Катализаторами добавления метильных групп служат ферменты из семейства ДНК-метилтрансфераз (DNMTs), в то время как за удаление модификации главным образом ответственны диоксигеназы ТЕТ [2]. Нарушения в работе этих ферментов, обусловленные мутациями, эпигенетическими событиями или иными причинами, могут быть ассоциированы с развитием злокачественных новообразований.

Целью настоящей работы был поиск возможных осей днРНК/миРНК/мРНК, участвующих в регуляции некоторых генов, ответственных за механизмы метилирования и деметилирования ДНК, при скПКР.

Методы и алгоритмы: Биоинформатический анализ проводился с использованием открытых баз данных входящих в базу NCBI (<https://www.ncbi.nlm.nih.gov/>): PubMed, Gene и др. В работе также были использованы Lnc2Cancer 2.0 (www.bio-bigdata.net/lnc2cancer), GEPIA2.0 (<http://gepia2.cancer-pku.cn/#index>), miRCode (<http://www.mircode.org/>), miRWalk 2.0 (<http://mirwalk.umm.uni-heidelberg.de/>), TargetScan 8.0. (https://www.targetscan.org/vert_80/) и The Genotype-Tissue Expression (GTEx, <https://www.gtexportal.org/home/>).

Результаты: Отбор днРНК, участвующих в прогрессии эпителиальных форм рака, в том числе скПКР, проводился на основании анализа данных литературы PubMed с использованием ключевых слов по формуле: “human” AND (“long non-coding RNA”) OR (“lncRNA”) OR (“lincRNA”) OR (“long intergenic non-protein coding RNA”) OR (“long non-protein-coding RNA”)) AND “cancer”. Критерием отбора являлось то, что исследования должны были быть оригинальными и в них оценивалась экспрессия днРНК в опухолевой ткани эпителиальных опухолей, в том числе и при скПКР. Критерии исключения были следующими: публикации написаны на языке, отличном от английского; статьи относятся к категориям: отчеты о случаях заболевания, комментарии, письма, обзоры или метаанализы;

другие заболевания; исследования не оценивали роль днРНК в диагностике, прогнозе и ответе на лечение. Связь днРНК с онкозаболеваниями также была оценена с использованием баз данных LncRNADisease 2.0 (www.cuilab.cn/lncrnadisease) и lncATLAS (<https://lncatlas.crg.eu/>). Транскриптомный ландшафт днРНК был идентифицирован с помощью Lnc2Cancer 2.0. По результатам анализа было отобрано около 400 днРНК, изменяющих экспрессию в эпителиальных опухолях разных локализаций.

С привлечением базы данных GEPID2.0 [3] был проведен интерактивный анализ профиля экспрессии отобранных генов при скПКР. Были отобраны гены с дифференциальной экспрессией в опухолевой ткани. В результате было отобрано около 150 днРНК, изменяющих экспрессию в опухолевой ткани более чем в 2 раза. Следующим этапом анализа было выявление наличия CpG островков в области не далее 2000 п.н. от 5' конца гена или старта транскрипции по базе Epigenomics, входящей в пакет NCBI. В результате поиска было отобрано 50 днРНК. Характеристики представленности днРНК и мРНК в нормальной ткани получены с помощью ресурса GTEx.

Одним из основных этапов биоинформатического анализа был поиск миРНК, взаимодействующих как с отобранными днРНК, так и белок-кодирующими генами (*TET1*, *TET2*, *TET3*, *TDG*, *DNMT3A*, *DNMT3B* и *DNMT1*). Скрининг был основан на анализе данных полногеномного секвенирования NCBI GEO и определении процента комплементарных пар с помощью локального выравнивания с использованием алгоритма Смита-Ватермана. Затем проверяли отобранные миРНК по известным базам данных (miRCode, miRWalk 2.0 и TargetScan 8.0). Отбирали миРНК минимум с двумя сайтами взаимодействия. Минимальная свободная энергия (ΔG) < -15 ккал/моль была установлена в качестве порога взаимодействия. Рассматривались дуплексы миРНК/мРНК и миРНК/днРНК со спариванием оснований по меньшей мере из 7-8 нуклеотидов в затравочной последовательности миРНК. В результате, для дальнейшего использования были отобраны miR-124-3p, miR-129-5p и miR-125b-5p, негативно коррелирующие как с экспрессией мРНК белок-кодирующих генов-мишеней, так и днРНК.

В дальнейшем потребуется проверка полученных данных на репрезентативной выборке клинических образцов с последующей валидацией на культурах клеток.

Выводы: В результате комплексного анализа отобрано 17 днРНК (*OIP5-AS1*, *FAM222A-AS1*, *GAS5*, *HAND2-AS1*, *LINC00339*, *LINC00886*, *MALAT1*, *MEG3*, *NCK1-DT*, *NEAT1*, *SNHG1*, *SNHG12*, *TINCR*, *TP73-AS1*, *XIST*, *SNHG6*, *ADAMTS9-AS2*), 3 миРНК (miR-124-3p, miR-129-5p, miR-125b-5p) и 7 белок-кодирующих генов (*TET1*, *TET2*, *TET3*, *TDG*, *DNMT3A*, *DNMT3B* и *DNMT1*). Проведена оценка вероятности их взаимодействия. Получены теоретические данные о комбинации триплетов (осей), возможно участвующих в развитии скПКР.

Финансирование: Исследование поддержано государственным заданием FGFU-2022-0007.

Modern assessment of ncRNAs as biomarkers of prognosis in clear cell renal cell carcinoma

Loginov V.I.^{1,2*}, Burdenny A.M.^{1,3}, Pronina I.V.¹, Lukina S.S.¹, Braga E.A.^{1,2}

¹ Institute of General Pathology and Pathophysiology, Moscow, Russia

² Bochkov Research Centre for Medical Genetics, Moscow, Russia

³ Emanuel Institute for Biochemical Physics, Moscow, Russia

* loginov7w@gmail.com

Key words: mRNA; miRNA; lncRNA; clear cell renal cell carcinoma; ccRCC

Motivation and Aim: Epigenetic changes play a key role in the development and progression of tumors. In particular, the results of numerous studies unequivocally confirm the involvement of abnormal DNA methylation in oncogenesis. Regional CpG islands promoter hypermethylation of a number of genes (including genes coding proteins, microRNAs, long non-coding RNAs, etc.) is another common phenomenon in the tumor process. The result of these epigenetic changes may be the inhibition of the transcription of suppressor genes that protect normal cells from malignant transformation [1]. Enzymes from the DNA methyltransferase (DNMT) family catalyze the transfer of methyl groups, while TET dioxygenases remove this modification [2]. Disorders in the functioning of these enzymes induced by mutations, epigenetic events or other causes may be associated with the development of malignant neoplasms.

The aim of this work was to search for possible lncRNA/miRNA/mRNA axes involved in the regulation of some genes responsible for the mechanisms of DNA methylation and demethylation in ccRCC.

Methods and Algorithms: Bioinformatic analysis was carried out using open-access databases included in the NCBI database (<https://www.ncbi.nlm.nih.gov/>): PubMed, Gene, etc. Lnc2Cancer 2.0 (www.bio-bigdata.net/lnc2cancer), GEPIA2.0 (<http://gepia2.cancer-pku.cn/#index>), miRCode (<http://www.mircode.org/>), miRWalk 2.0 (<http://mirwalk.umm.uni-heidelberg.de/>), TargetScan 8.0. (https://www.targetscan.org/vert_80/) and The Genotype-Tissue Expression (GTEx, <https://www.gtexportal.org/home/>) was also used in the work.

Results: The selection of lncRNAs involved in the progression of epithelial forms of cancer, including ccRCC, was carried out based on the analysis of PubMed publication data using keywords according to the formula: “human” AND ((“long non-coding RNA”) OR (“lncRNA”) OR (“lincRNA”) OR (“long intergenic non-protein coding RNA”) OR (“long non-protein-coding RNA”)) AND “cancer”. The selection criteria were as follows: the studies had to be original and they evaluated the expression of lncRNA in the cancer tissue of epithelial tumors, including in ccRCC. The exclusion criteria were as follows: publications were written in a language other than English; articles belong to the categories: case reports, comments, letters, reviews or meta-analyses; other diseases; studies did not evaluate the role of lncRNA in diagnosis, prognosis and response to treatment. The association of lncRNA with cancer was also evaluated using LncRNADisease 2.0 databases (www.cuilab.cn/lncrnadisease) and lncATLAS (<https://lncatlas.crg.eu/>). The transcriptomic landscape of lncRNA was identified using Lnc2Cancer 2.0. According to the results of the analysis, about

400 lncRNAs were selected that change expression in epithelial tumors of different localizations.

Using the GEPIA2.0 database [3] an interactive analysis of the expression profile of the selected genes was performed in ccRCC. Genes with differential expression in tumor tissue were selected. As a result, about 150 lncRNAs were selected that change expression in tumor tissue by more than 2 times.

The next stage of the analysis was to identify the presence of CpG islands in the region no further than 2000 bp from the 5' end of the gene or the start of transcription according to the Epigenomics database included in the NCBI package. As a result of the search, 50 lncRNAs were selected. The characteristics of the representation of lncRNA and mRNA in normal tissue were obtained using the GTEx resource.

One of the main stages of bioinformatic analysis was the search for miRNAs interacting with both selected lncRNAs and protein-coding genes (*TET1*, *TET2*, *TET3*, *TDG*, *DNMT3A*, *DNMT3B* and *DNMT1*). The screening was based on the analysis of NCBI GEO full-system sequencing data and the determination of the percentage of complementary pairs using local alignment using the Smith-Waterman algorithm. Then the selected miRNAs were checked against well-known databases (miRCode, miRWalk 2.0 and TargetScan 8.0). miRNAs with at least two interaction sites were selected. The minimum free energy (ΔG) < -15 kcal/mol was set as the interaction threshold. Duplexes of miRNA/mRNA and miRNA/lncRNA with base pairing of at least 7-8 nucleotides in the miRNA seed sequence were considered. As a result, miR-124-3p, miR-129-5p and miR-125b-5p were selected for further use, negatively correlating with both mRNA expression of protein-coding target genes and lncRNA.

In the future, it will be necessary to verify the data obtained on a representative sample of clinical samples with subsequent validation on cell cultures.

Conclusion: As a result of the complex analysis, 16 lncRNAs (OIP5-AS1, FAM222A-AS1, GAS5, HAND2-AS1, LINC00339, LINC00886, MALAT1, MEG3, NCK1-DT, NEAT1, SNHG1, SNHG12, TINCR, TP73-AS1, XIST, SNHG6, ADAMTS9-AS2), 3 miRNAs (miR-124-3p, miR-129-5p, miR-125b-5p) and 7 protein-coding genes (*TET1*, *TET2*, *TET3*, *TDG*, *DNMT3A*, *DNMT3B* and *DNMT1*). The probability of their interaction has been estimated. Theoretical data have been obtained on the combination of triplets (axes) possibly involved in the development of ccRCC.

Funding: The study is supported by the FGFU-2022-0007 state assignment.

Список литературы/References

1. Клаг У.С., Каммингс М.Р., Спенсер Ш.А., Палладино М.А. Основы генетики. М.: Техносфера, 2021
[Klug U.S., Cummings M.R., Spencer S.A., Palladino M.A. Fundamentals of Genetics. Moscow: Technosphere, 2021 (in Russian)]
2. Максимова В.П., Бугаева П.Е., Жидкова Е.М. и др. Современные подходы к выявлению и изучению эпигенетически активных ксенобиотиков. *Успехи молекулярной онкологии*. 2019;6(3):8-27
[Maksimova V.P., Bugaeva P.E., Zhidkova E.M. et al. Modern approaches to the identification and study of epigenetically active xenobiotics. *Uspehi Molekulârnj Onkologii = Successes of Molecular Oncology*. 2019;6(3):8-27 (in Russian)]
3. Tang Z., Li C., Kang B., Gao G., Li C., Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 2017;45(W1):W98-W102. doi 10.1093/nar/gkx247