Сравнительный анализ профиля экспрессии генов в опухолевой и здоровой ткани у больных колоректальным раком

С. А. Мерцалов1, Е. П. Куликов1, В. В. Стрельников2, А. И. Калинкин2, Е. И. Шумская1, Р. О. Пискунов1

1 Рязанский государственный медицинский университет имени академика И. П. Павлова, Рязань, Российская Федерация;
2 Медико-генетический научный центр имени академика Н. П. Бочкова, Москва, Российская Федерация

АННОТАЦИЯ

Введение. В структуре заболеваемости злокачественными новообразованиями колоректальный рак (КРР) без учета пола уверенно лидирует (12,3%). Пятилетняя выживаемость при КРР I стадии — 91%, IV стадии — 14%. Существующие на сегодняшний день методы лечения не помогают существенно снизить смертность — подходы необходимо персонифицировать, в т. ч. с помощью молекулярно-генетических методов.

Цель. Произвести сравнительную оценку экспрессионного профиля образцов опухолевой и здоровой ткани толстой кишки при КРР.

Материалы и методы. Материалом для исследования послужили 19 образцов опухолевой ткани, взятых из патологически измененной ткани слизистой толстого отдела кишечника у 19 пациентов с КРР, и 7 образцов «здоровой» ткани, отобранной на расстоянии 10-12 см дистальнее или проксимальнее от визуальной границы опухоли. Гомогенезация биоптатов выполнена механическим методом. Качество и количество рибонуклеиновой кислоты в элюированном растворе оценивались с помощью наноспектрофотометра IMPLEN (Германия). Для оценки экспрессии генов использовался набор микрочипов SurePrint G3 HumanGeneExpv3 ArrayKit (Agilent, США). Сканирование микрочипов производилось на аппарате InnoScan 1100 AL (США) с последующей обработкой изображения на программном обеспечении Mapix Software (США).

Результаты. Анализ экспрессионного профиля продемонстрировал 505 дифференциально экспрессируемых генов, среди них 337 проявили сниженную экспрессию в опухолевом материале и 168 — повышенную. Наиболее высокую экспрессию продемонстрировали гены, связанные с миРНК: hsa-miR-29b-3p и hsa-miR-1-5p, а также гены H19, FOXQ1, INHBA, MMP1, CDH3, CXCL2, MDFI, THBS2. Напротив, гены TMIGD1, GUCA2B, ZG16, AQP8, SLC4A4, CDKN2B-AS1, CA4, CA1 продемонстрировали низкую экспрессию в опухолевом материале. Экспрессия генов, ответственных за функционирование сигнальных путей: IL-17, NF-κappa B, TNF, — увеличена в опухолевых образцах. Гены, ответственные за сигнальные пути «Fatty acid degradation», «Drug metabolism — cytochrome P450», «Metabolic pathways», «Fatty acid metabolism» и «Steroid hormone biosynthesis», показали сниженную экспрессию.

Заключение. Выявлены значительные различия экспрессионного профиля между опухолевой и здоровой тканью у пациентов с КРР. Сравнительный анализ обогащения генов с данными международных баз данных позволил выявить ряд терминов, генов, кластеров, которые в дальнейшем могут быть использованы в поиске предикторов прогноза и ответа на лечение.

Ключевые слова: экспрессионные микрочипы; экспрессионный профиль; колоректальный рак

Comparative Analysis of Gene Expression Profile in Tumor and Healthy Tissue in Patients with Colorectal Cancer

Sergey A. Mertsalov¹, Evgeniy P. Kulikov¹, Vladimir V. Strel’nikov², Aleksey I. Kalinkin², Evgeniya I. Shumskaya¹, Roman O. Piskunov¹

¹ Ryazan State Medical University, Ryazan, Russian Federation; ² Research Centre for Medical Genetics, Moscow, Russian Federation

ABSTRACT

INTRODUCTION: Colorectal cancer (CRC) is a sure leader among malignant neoplasms (12.3%), without taking into account gender. Five-year survival rate in stage I CRC is 91%, in stage IV — 14%. The currently existing treatment methods are helpless to significantly reduce mortality the approaches should be personalized and include the use of molecular genetic methods.

AIM: To perform a comparative evaluation of expression profile of samples of tumor and healthy colon tissue in CRC.

MATERIALS AND METHODS: The material for the study was 19 samples of tumor tissue taken from the pathologically altered colonic mucosa of 19 patients with CRC, and 7 samples of ‘healthy’ tissue taken 10 cm–12 cm distally or proximally from the visual boundary of the tumor. Biopsy materials were homogenized using a mechanical method. The quality and quantity of ribonucleic acid in the eluted solution were evaluated using IMPLEN nanospectrophotometer (Germany). Gene expression was evaluated using microchip kit SurePrint G3 HumanGeneExpv3 ArrayKit (Agilent, USA) and subsequent image processing in Mapix Software program (USA).

RESULTS: The analysis of expression profile demonstrated 505 differentially expressed genes, 337 of which showed reduced expression and 168 — enhanced expression in the tumor material. The highest expression was demonstrated by genes bound with miRNA: hsa-miR-29b-3p and hsa-miR-1-5p, and also genes H19, FOXQ1, INHBA, MMP1, CDH3, CXCL2, MDFI, THBS2. On the contrary, genes TMIGD1, GUCA2B, ZG16, AQP8, SLC4A4, CDKN2B-AS1, CA4, CA1 demonstrated a low expression in the tumor material. Expression of genes responsible for functioning of signal pathways: IL-17, NF-κappa B, TNF, was increased in tumor samples. Genes responsible for signal pathways Fatty acid degradation, Drug metabolism — cytochrome P450, Metabolic pathways, Fatty acid metabolism and Steroid hormone biosynthesis, showed reduced expression.

CONCLUSION: Significant differences were found in the expression profile of tumor and healthy tissue in patients with CRC. A comparative analysis of gene enrichment and the data of the international databases permitted to identify a number of terms, genes, clusters that can be used in future in search for predictors of prognosis and of response to treatment.

Keywords: expression microchips; expression profile; colorectal cancer


Received: 25.01.2023 Accepted: 08.02.2023 Published: 30.06.2023
INTRODUCTION

According to A. D. Kaprin, et al. CRC is a confident leader in the structure of morbidity with malignant neoplasms, without taking into account gender, and accounts for 12.3%. With this, the annual growth rate is 4.24% [1], five-year survival in stage I — 91%, in stage IV — 14% [2, 3].

Development of CRC is associated with environmental and genetic factors. A great progress has been made in the understanding of the molecular pathogenesis of colon cancer, which includes four main mechanisms: adenoma–carcinoma sequence, hereditary forms, deficit of deoxyribonucleic acid (DNA) repair enzymes and microsatellite instability [5]. However, the exact molecular mechanisms are not completely clear, and the currently existing methods of treatment do not help any significantly reduce mortality. All this, in turn, gives grounds to state the need to personalize the approaches to treatment [4].

The approach can be individualized using molecular genetic methods that allow not only to detect genetic anomalies, mutations, but also to identify their transcription products — the expression profile, which is a genetic ‘portrait’ of the neoplasm. One of the promising directions is expression microchip analysis, which has already proved effective in predicting the course of breast cancer and response to treatment [6], in tumors of the hematopoietic system and lymphoproliferative diseases [7]. Since the peculiarities of the course of the tumor process, the response to the therapy and the prognosis are largely determined by the genetic profile, a large number of scientific studies are aimed at finding a connection between the molecular genetic ‘portrait’ of cells and oncopathological processes and their course in different cancer localities.

The aim of this study to perform a comparative evaluation of the expression profile of samples of tumor and healthy colon tissue in colorectal cancer.

MATERIALS AND METHODS

The study was conducted on the base of Ryazan State Medical University and Regional Clinical Oncologic Dispensary (Ryazan), Research Centre for Medical Genetics (Moscow) and was approved by the Local ethics committee at Ryazan State Medical University (Protocol No. 5 of 2020, December 09). All the included patients signed the Informed consent.

The study materials were obtained in video-colonoscopy with biopsy in 19 patients with the verified diagnosis of colon cancer: 19 samples of tumor tissue taken from the pathologically altered tissue of the colon mucosa, and 7 samples of ‘healthy’ tissue.

The area of ‘healthy’ tissue was chosen visually 10 cm–12 cm distally or proximally of the visible tumor boundary. For transportation and storage of the biomaterial, RNA later stabilization solution (Thermo Fisher Scientific Inc., USA) was used.
The next stage was performed in the Central Research Laboratory of Ryazan State Medical University. Tissue was homogenized by a mechanical method, in 500 µl of lysis solution with addition of mercaptoethanol. Then ribonucleic acid (RNA) was isolated on spin columns of RNA easy Plus Mini Kit (Qiagen, USA) using filtering DNA and RNA-membranes according to the manufacturer’s instruction. The quality and quantity of RNA in the eluted solution were evaluated with IMPLEN nanospectrophotometer (Germany). An important parameter of RNA quality is the RNA Integrity Number (RIN). Three classes of RIN are distinguished: high degree of RNA integrity (RIN = 10), partially degraded RNA (RIN below 5), fully degraded RNA (RIN ≈ 3) [8].

The analysis of bands corresponding to an individual RNA sample and to the marker of molecular size, was performed automatically. In result of the analysis, a profile was generated, and the quantitative and qualitative characteristics of RNA sample were calculated. In the analysis of RNA samples, the results obtained imitated electrophoretic separation. Major bands in the 49 and 42 regions corresponded to 28S and 18S rRNA. In the work, samples were used that strictly satisfied the optimal parameters of purity (with RIN not less than 7.5, with the most part having 8.0 and higher index).

For evaluation of gene expression, a set of SurePrint G3 HumanGeneExpv3 Array Kit microchips (Agilent, USA) was used. Microchips were scanned on InnoScan 1100 AL apparatus (USA).

The data from the microchip analyzer were sent for bioinformatic processing to Research Centre for Medical Genetics (cooperation agreement of 2022, May 30). The data were obtained from 19 samples of colorectal cancer and 7 samples of normal tissue (colon mucosa). All the patients were with the verified diagnosis. Histologically, the tumors were adenocarcinomas, mainly of moderate differentiation. The stage of the disease was determined according to the International Classification TNM 8th edition (2018). The Limma package was used to import, control the quality of the data obtained, and search for differentially expressed genes. In order to identify signaling pathways and molecular functions by analyzing the enrichment of signaling pathways, the Cluster Profiler package was used. The R ver. 3.6.3 programming language was used for all calculations and packages.

For the procedure of searching for differentially expressed genes, moderated t-statistics method was used that was realized in Limma software package. In the multiple testing, type I errors can occur (a large number of false-positive results) which was adjusted by us by correction for multiple testing set up at 0.09 (because of a relatively small number of samples, but with the aim to control and prevent omission of a large number of false-positive results). The basis of the analysis of enrichment of signal pathways was a hypergeometric test; to cut off the most significant signal pathways, the correction for multiple testing was also used, set up at 0.05.

RESULTS

After the import of the raw data obtained from expression chips, the check was performed for the difference between samples, and normalization of the input data. Normalization of the data was followed by noise correction, after which the data were cleared of bad and technical microchip probes. As can be seen in Figure 1, the data required additional normalization and subsequent filtration.

There were found 505 differentially expressed genes, 337 of which demonstrated reduced expression in the tumor material and 168 — enhanced expression. The CA1 gene was found to be least expressed with an average negative expression index (log FC = -4.31), the lowest expression values were also identified in the TMIGD1, GUCA2B, ZG16, AQP8, SLC4A4, CDKN2B-AS1, CA4 genes. On the contrary, the H19 gene demonstrates the highest average positive expression (log FC = 4.23). Besides this gene, high expression is demonstrated by the FOXQ1, INHBA, MMP1, CDH3, CXCL2, MDFI, THBS2 genes. Figure 2 graphically shows the results of search for differentially expressed genes.

To obtain the data about the role of differentially expressed genes in the biological processes, molecular functions and signaling pathways, we used GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes) analyses, as well as mirTarBase analysis for genes that are regulated with small interfering (micro-) RNA (miRNA).

In the analysis, 44 and 75 biological processes were identified that are regulated by reduced and enhanced expression of genes, respectively (Figure 3).

These biological, molecular processes, as well as signaling pathways play an important role in the regulation, transport, adhesion of molecules, in catalytic activity, drug metabolism, biosynthesis, degradation and metabolism of fatty acids. We have obtained the data on the ‘NF-κappa B’, ‘TNF’ signaling pathways and on the regulation of different miRNAs.

DISCUSSION

In the analysis of the data about gene expression obtained from microchips, 337 genes with decreased expression and 168 with enhanced expression were identified. The CA1 gene (Carbonic Anhydrase 1) shows the lowest expression in the tumor material, which is confirmed by other studies of the expression profile [9]. CA1 is a potential oncogene and promotes abnormal calcification, apoptosis and migration of cells in breast cancer [10]. Carbonic anhydride isoenzymes may play
Fig. 1. Box plots for raw data (A) obtained from expression chips and data (B) obtained after filtration and normalization.

Fig. 2. Identification of differentially expressed genes in colon cancer.

Note: Red and blue dots show enhanced and reduced expression of genes in the tumor material of the patients.
an important role in development of cancer, since they control pH homeostasis in tumors, which, probably, models behavior of tumor cells [11].

Expression of the H19 gene (gene for a long non-coding RNA) is increased in the tumor material, which can lead to activation of epithelial-mesenchymal transition and further to metastasis and invasion [12, 13]. Further studies confirmed expression of H19 leading to activation of Raf-ERK signaling pathway and to induction of epithelial-mesenchymal transition, which, probably, correlates with metastasis in patients with colorectal cancer [14]. Besides, the genes, involved in the epithelial-mesenchymal transition, participate not only in migration and invasion of tumor cells, but also in suppression of cell death, regulation of cell cycle and also in the processes responsible for resistance to radiation therapy and chemotherapy [15, 16].

For a more extensive analysis and understanding of the significance of the obtained profiles of expressed genes in the biological processes, we used the Gene Set Enrichment Analysis (GSEA) on the GO database. For example, analysis of enrichment of a set of genes that are activated under certain conditions (in particular, in cancer) based on statistical methods, revealed the sets of genes of this database that are present in excess (upregulated) or in deficit compared to genes in healthy tissue. It is believed that the differences found in this way are reflected at the morpho-functional level.

binding’, ‘small molecule binding’, ‘ion binding’ and ‘anion binding’ change their activity in reduced gene expression [17–21]. However, the main mechanisms by which the corresponding genes in these biological and molecular processes contribute to the oncogenesis, still remain undescribed. Consequently, further investigation of these identified biological and molecular processes can help both in elucidating the main mechanisms of CRC carcinogenesis and in personalization of approaches to therapy of these patients.

Analysis of signaling pathways using the KEGG database revealed that the signaling pathways ‘Fatty acid degradation’, ‘Drug metabolism — cytochrome P450’, ‘Metabolic pathways’, ‘Fatty acid metabolism’ and ‘Steroid hormone biosynthesis’ change in reduced gene expression, while the signaling pathways ‘IL—17’, ‘NF-kappa B’, ‘TNF’ change in enhanced gene expression. The signaling pathways ‘Fatty acid degradation’ and ‘Fatty acid metabolism’ play an important role in the pathogenesis of different oncological diseases. In a recent study by C. Ding, et al. a signature of fatty acid genes was obtained that effectively predict survival for patients with colorectal cancer, as well as resistance to 5-fluorouracil [21]. The ‘Drug metabolism — cytochrome P450’ signaling pathway and reduced expression of genes associated with it may indicate acquired resistance to 5-fluorouracil [22]. Genes associated with ‘Metabolic pathways’ have a high potential for therapeutic effects [23]. Genes with high expression in the ‘IL—17 signaling pathway’ can lead to the oncogenesis through stimulating the production of growth factors, and to glycolysis, angiogenesis, and metastasis of colon cancer [24]. High expression of the genes of the ‘NF-kappa B’ signaling pathway leads to the progression of colorectal cancer, while therapeutic influence on the genes of this signaling pathway leads to a decrease in proliferation, metastasis and angiogenesis and increases the level of apoptosis and sensitivity to chemotherapeutic drugs [25]. Up regulated genes belonging to the ‘TNF’ signaling pathway, contribute to the tumor microenvironment, and their overexpression can lead to epithelial-mesenchymal transition and subsequent metastasis [26].

Enhanced expression of genes associated with hsa-miR-29b-3p, leads to enhancement of angiogenesis and epithelial-mesenchymal transition, which is associated with poor survival of patients with colorectal cancer [27]. The role of hsa-miR-29a-3p as an oncogene or onco-suppressor and of associated genes, remains unspecified until now, however, it is considered as a diagnostic and prognostic marker [28]. According to A. Safa, et al. hsa-miR-1-5p can suppress progression of colorectal cancer, while the active expression of genes associated with this miRNA, may interfere with this process [29].

CONCLUSION
Thus, the analysis of expression profile of patients with colorectal cancer was performed, the data of 505 differentially expressed genes were obtained, among them 337 showed reduced expression in the tumor material and 168 — enhanced expression. Highest expression was demonstrated by genes associated with miRNA (hsa-miR-29b-3p and hsa-miR-1-5p), and also by the H19, FOXQ1, INHBA, MMP1, CDH3, CXCL2, MDFI, THBS2 genes. On the contrary, the TMIGD1, GUCA2B, ZG16, AQP8, SLC4A4, CDKN2B—AS1, CA4 genes, as well as the CA1 gene showed low expression in the tumor material.

Expression of genes responsible for functioning of signaling pathways ‘IL—17’, ‘NF-kappa B’, ‘TNF’ was enhanced in the tumor samples. Genes responsible for signaling pathways ‘Fatty acid degradation’, ‘Drug metabolism — cytochrome P450’, ‘Metabolic pathways’, ‘Fatty acid metabolism’ and ‘Steroid hormone biosynthesis’, showed reduced expression in the tumor material. The obtained results of the expression profile can be extrapolated to clinical data and can be used in search for predictor of prognosis and of response to treatment of patients of this group. In future, these markers will permit personalization of the approach to treatment of each patient.

ADDITIONALLY
Funding. The work was performed within the framework of the state assignment of the Ministry of Education and Science of Russia No. FGFF—2022—0004 for Research Center for Medical Genetics; Moscow to carry out research in 2023.
Conflict of interests. The authors declare no conflicts of interests.

Contribution of the authors: S. A. Mertsalov — setting goals and objectives, working with patients, collecting laboratory material, working with primary material, conducting statistical analysis, analyzing the literature, drawing conclusions; E. P. Kulikov — setting goals and objectives, working with patients, collecting laboratory material, working with primary material, conducting statistical analysis, analyzing the literature, drawing conclusions; V. V. Strel’nikov — conducting statistical analysis, analyzing the literature, drawing conclusions, construction of graphs and charts, editing; A. I. Kalinkin — article concept, bioinformatic and statistical information processing, construction of graphs and charts, editing; E. I. Shumskaya — sample preparation and microarray analysis, editing, work with primary material, conducting laboratory research; R. O. Piskunov — working with patients, analyzing the literature, completing the paper, translating the text, and drawing conclusions. The authors confirm the correspondence of their authorship to the ICMJE International Criteria. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work.
Список источников

8. Андреева Т.В., Кукинченко С.С. Анализ регуляторных РНК, связанных с развитием болезни мозга. М.: Цифровичок; 2012.


REFERENCES


