ISSN 0003-6838, Applied Biochemistry and Microbiology, 2022, Vol. 58, No. 9, pp. 992–996. © Pleiades Publishing, Inc., 2022.
Russian Text © The Author(s), 2021, published in Biotekhnologiya, 2021, Vol. 37, No. 6, pp. 95–100.

MEDICAL BIOTECHNOLOGY

Expression of CD44 Isoforms in Human Colorectal Cancer Cell Lines

M. P. Raigorodskaya[a], V. O. Novosad[b], S. A. Tonevitskaya[a, c], and D. V. Maltseva[a, *]

[a] *Bioclinicum Scientific Research Center, Moscow, 115088 Russia*
[b] *Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, 117997 Russia*
[c] *Faculty of Biology and Biotechnology, Higher School of Economics, Moscow National Research University, Moscow, 117312 Russia*
**e-mail: dmaltseva@gmail.com*

Received September 30, 2021; revised October 5, 2021; accepted November 10, 2021

Abstract—The expression profile of CD44 isoforms in 55 colorectal cancer (CRC) cell lines has been evaluated based on mRNA sequencing data from the Cancer Cell Line Encyclopedia database. The distribution of CD44 mRNA isoforms differs significantly between CRC lines. In 13 studied lines, including Caco-2 and RKO, CD44 expression was not detected. In most other lines, isoform 3 of CD44 mRNA was the most abundant; however, the level of its expression varied and it was absent in the MDST8 and SNU503 lines. The highest level of isoform 3 was observed in CW2 and T84 lines. The next most frequent was isoform 4 with its expression level being lower than that of isoform 3, except for the HCT116, SNU81, NCIH508 and SNUC4 lines. The highest expression of isoform 4 was detected in the MDST8 line, the only line in which isoform 6 was also expressed. Isoform 2 was also present in CRC cell lines; its highest expression level was found in the SNU503 line. Isoforms 1, 5, and 7 were not expressed in any of the studied lines. It is necessary to take into account the mRNA expression profile of specific CD44 isoforms when choosing a cell model to study its role in CRC.

Keywords: colorectal cancer, CD44 isoforms, mRNA sequencing, CCLE
DOI: 10.1134/S0003683822090071

INTRODUCTION

In recent years, the effectiveness of therapy of colorectal cancer (CRC) has significantly increased in the world; nevertheless, this oncological disease still ranks third in the frequency of deaths [1]. More than 70% of lethal outcomes in patients with CRC are associated with liver metastases, which are already detected in 25% of patients during primary diagnosis [2]. Approximately 50% of patients suffer from subsequently developed metastases. Currently, the prognosis of treatment and the choice of therapeutic regimen are based on the TNM classification [3], which does not accurately predict the CRC course in a patient at an early stage of the disease. In this regard, the identification of biomarkers of malignant tumors of the colon and rectum is a relevant task for the diagnosis and prognosis of CRC.

More data are accumulating on the crucial role of cancer stem cells (CSC) in the development and progress of oncological diseases [4]. This is the basis for the assumption that the study of CSC markers is promising for CRC diagnosis and prognosis. The transmembrane glycoprotein CD44 belongs to the proven and most significant CSC markers in CRC [4]. Increased CD44 expression has been found in various oncological diseases, including CRC [5–10].

The *CD44* gene in the human body includes 19 exons, which undergo alternative splicing to form lots of CD 44 mRNA isoforms [11, 12]. *In silico* analysis predicts the existence of 27 isoforms; however, the presence of only eight of them was experimentally confirmed (Fig. 1). The shortest isoform, consisting of the first five (1–5) and last four (15–17, 19) exons, is expressed in most tissue types and is regarded as the CD44 standard isoform (CD44s). The insertion of variable exons v2–v10 (corresponding to exons 6–14 in DNA) leads to the formation of isoforms with a higher molecular weight (CD44v); they are only expressed is some epithelial tissues and several types of tumor cells [13]. CD44 variable isoforms are often designated in accordance with numbers of variable exons included in their mRNA (e.g., CD44v2–10). Expression of specific CD44 isoforms is still poorly studied; however, new data are emerging, indicating the different functional load of CD44 isoforms [14–17].

It should be noted that the research is complicated by the lack of the standard nomenclature of this protein's isoforms. In other publications, researchers often use nomenclature based either on the commer-

Abbreviations: CCLE, Cancer Cell Line Encyclopedia; CD44s, CD44 standard isoform; CRC, colorectal cancer; CSC, cancer stem cells.

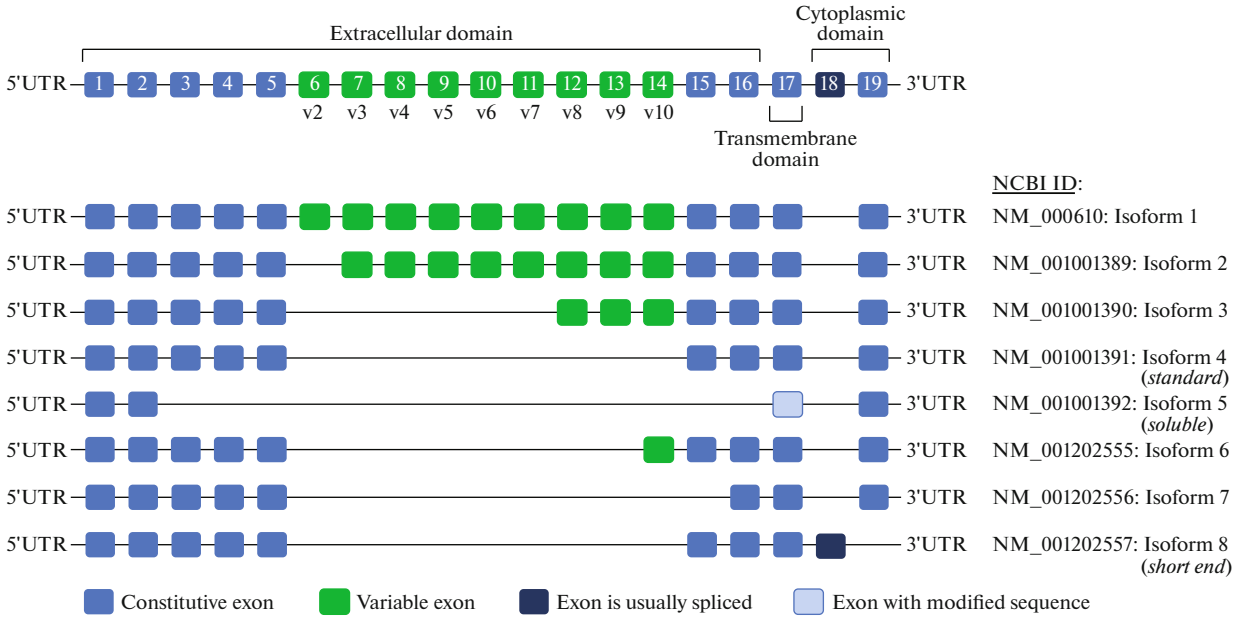

Fig. 1. The structural scheme of CD44 mRNA isoforms, whose existence has been confirmed experimentally.

cial names of monoclonal antibodies, or of the variable exon for which the antibodies are specific, for instance CD44v3, CD44v6, or CD44v9. In these cases, it is often not taken into account that the identification of a specific variable exon can lead to detection of all isoforms containing it and not just one specific protein. As an example, exon v9 may be part of various isoforms: CD44v2–10, CD44v3–10, and CD44v8–10, which correspond to isoforms 1, 2, and 3 in the NCBI database (Fig. 1). In view of this, total mRNA sequencing is seen as a powerful tool for analyzing the expression profiles of CD44 isoforms. It should also be noted that there are differences in the designation of CD44 isoforms in the UniProt and NCBI databases. In this work, we will adhere to the NCBI nomenclature.

According to a recently published meta-analysis [5], variants of CD44 mRNAs containing variable exons v2, v3, v6 and v9 are expressed in tumors of patients with CRC. However, the functional significance of these exons has not yet been understood. In this regard, studies on the role of specific CD44 isoforms (or a specific isoform profile) in CRC remain relevant. To solve this problem, model cell lines with a well-annotated expression profile of CD44 isoforms are needed.

In this work, we have analyzed expression profiles of CD44 mRNA isoforms in CRC cell lines using publicly available total mRNA sequencing data from the Cancer Cell Line Encyclopedia (CCLE) [18]. As a result, more potentially relevant lines for studying the role of CD44 isoforms in CRC were identified.

MATERIALS AND METHODS

Processing of Sequencing Results

Public data on mRNA sequencing for 55 CRC lines were downloaded from the CCLE portal (https://sites.broadinstitute.org/ccle) as TPM expression tables. TPM units were obtained by normalizing the reads first by the length of the splice variants (in thousands of nucleotides), and then by the depth of the library (the sum of the obtained values in millions reads). The data set contained information on isoforms from 1 to 7; the data on isoform 8 (in accordance with the NCBI nomenclature) were absent.

To analyze and visualize the distribution of expression levels of CD44 mRNA isoforms in CRC lines, we built a heat map and carried out hierarchical clustering using the Ward method [19]. The graphical representation of the obtained data was made using the seaborn module of the Python programming language [20].

RESULTS AND DISCUSSION

We have analyzed the expression of CD44 mRNA isoforms in 55 CRC cell lines using total mRNA sequencing data from the CCLE database. As a result, it was established that the distribution of isoforms significantly depends on the cell line. Figure 2 shows a heat map with hierarchical clustering of the CRC lines by expression of CD44 mRNA isoforms.

The analyzed lines can be conditionally divided into five groups by the expression levels of isoforms 2, 3, and 4: (1), 13 lines (from Colo320 to RKO), which do not express any CD44 mRNA isoform, practically or completely; (2), 17 lines (from CL11 to SW1463),

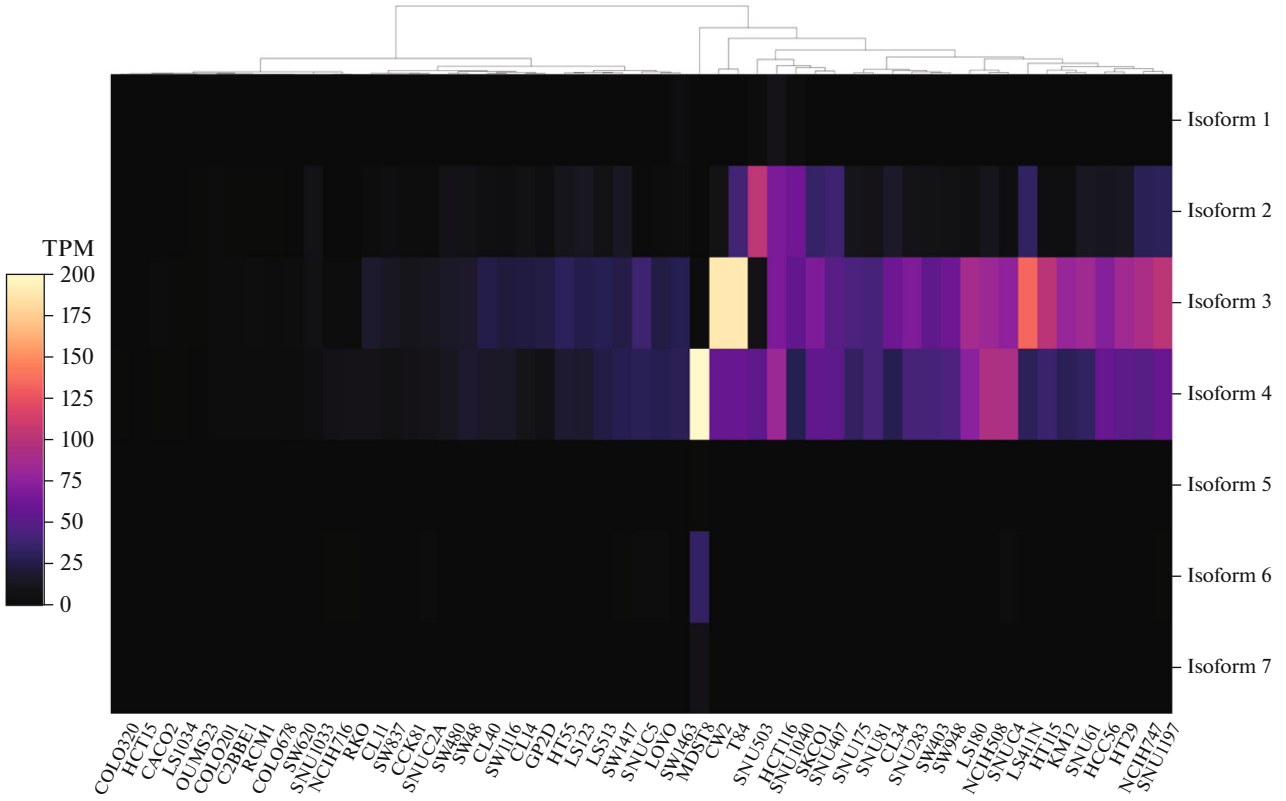

Fig. 2. A heat map with hierarchical clustering of CRC cell lines by expression of CD44 mRNA isoforms, built based on total mRNA sequencing data from the CCLE database.

which are capable of weak expression of isoform 3, and of which about half also poorly express isoform 4; (3), 17 lines (from SNU175 to SNU1197), which are able to express isoform 3 well, and some of which also express isoform 4, but in an amount much lower than isoform 3 (except for the SNU81, NCIH508, and SNUC4 lines); (4), 4 lines (from HCT116 to SNU407), in which a similar moderate expression level of isoforms 2,3 and 4 is observed; and (5), the remaining 4 lines, in which the expression pattern of CD44 mRNA isoforms significantly differs from the characteristics of the above listed lines.

As an example, line MDST8 demonstrated a very high level of the isoform 4 expression and a low level of isoform 6 (the latter was only detected in this line of all the others analyzed). An extremely high level of isoform 3 was found in lines CW2 and T84, while the level of isoform 4 was below average, as was isoform 2 in the T84 line. The SNU503 line was characterized by pronounced expression of isoform 2; isoform 4 was also expressed. Thus, isoforms 1, 5, and 7 are not expressed in any of the studied lines.

We note that isoform 4 corresponds to the so-called CD44 standard isoform (CD44s) that does not contain variable exons. CD44s is the most common isoform that is expressed in most vertebrate tissues [12, 15]. It has been found in almost all examined lines expressing the *CD44* gene (groups 2–5). Isoform 3 contains variable exons v8, v9, and v10 and corresponds to the CD44v8–10 (or CD44E) variant. This isoform was previously regarded only as an epithelial variant of the CD44 protein, since it is expressed in the normal epithelial tissues, including intestinal, and in the glandular epithelial tissues [13, 21, 22]. However, later it was shown that exon v9 included into CD44 can be considered as a CSC marker [23, 24], and the content of isoforms containing exon v9 increases during the epithelial–mesenchymal transition (EMT) [25].

Isoform 2 contains variable exons v3–v10 and includes all exons, v3, v6, and v9, that are most frequently used in research. Line SNU503 appears to be the most relevant model for studying the role of isoform 2 in CRC. Isoform 2 is expressed in lines assigned to groups 4 and 5, and also in several lines of group 3.

Since none of the CRC lines express CD44 isoform 1 containing exon v2, there are currently no models to study the role of this CD44 variant, whose expression was shown in clinical samples as a predictive marker of poor survival in patients with CRC [5].

CONCLUSIONS

Summarizing the obtained data, it should be concluded that the distribution of CD44 isoforms strongly

differs in CRC lines. As an example, the expression patterns of CD44 mRNA isoforms vary greatly between lines that are often used in CRC research, such as RKO, HCT116, and HT29. When selecting a model for the study of the CD44 role in CRC it is necessary to take into consideration the expression profile of its specific mRNA isoforms.

FUNDING

The study was funded by the Russian Science Foundation (project no. 17-14-01338).

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflicts of interest.

This article does not contain any studies involving animals performed by any of the authors.

This article does not contain any studies involving human participants performed by any of the authors outside the scope of their normal professional activities.

REFERENCES

1. Bray, F., Ferlay, J., Soerjomataram, I., et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA Cancer J. Clin.*, 2018, vol. 68, pp. 394–424. https://doi.org/10.3322/caac.21492
2. Van Cutsem E., Cervantes A., Nordlinger B., Arnold D., ESMO Guidelines Working Group. Metastatic colorectal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.*, 2014, vol. 25, suppl. 3, pp. iii1–9. https://doi.org/10.1093/annonc/mdu260
3. Weiser, M.R., AJCC 8th ed.: Colorectal cancer. *Ann. Surg. Oncol.*, 2018, vol. 25, pp. 1454–1455. https://doi.org/10.1245/s10434-018-6462-1
4. Munro, M.J., Wickremesekera, S.K., Peng, L., et al., Cancer stem cells in colorectal cancer: a review, *J. Clin. Pathol.*, 2018, vol. 71, pp. 110–116. https://doi.org/10.1136/jclinpath-2017-204739
5. Wang, Z., Tang, Y., Xie, L., et al., The prognostic and clinical value of CD44 in colorectal cancer: a metaanalysis, *Front. Oncol.*, 2019, vol. 9, p. 309. https://doi.org/10.3389/fonc.2019.00309
6. Chen, J., Zhou, J., Lu, J., et al., Significance of CD44 expression in head and neck cancer: a systemic review and meta-analysis, *BMC Cancer*, 2014, vol. 14, pp. 15. https://doi.org/10.1186/1471-2407-14-15
7. Luo, Y. and Tan, Y., Prognostic value of CD44 expression in patients with hepatocellular carcinoma: meta-analysis, *Cancer Cell Int.*, 2016, vol. 16, p. 47. https://doi.org/10.1186/s12935-016-0325-2
8. Li, X., Ma, X., Chen, L., et al., Prognostic value of CD44 expression in renal cell carcinoma: a systematic review and meta-analysis, *Sci. Rep.*, 2015, vol. 5, p. 13157. https://doi.org/10.1038/srep13157
9. Lin, J. and Ding, D., The prognostic role of the cancer stem cell marker CD44 in ovarian cancer: a meta-analysis, *Cancer Cell Int.*, 2017, vol. 17, p. 8. https://doi.org/10.1186/s12935-016-0376-4
10. Chen, Y., Fu, Z., Xu, S., et al., The prognostic value of CD44 expression in gastric cancer: a meta-analysis, *Biomed. Pharmacother.*, 2014, vol. 68, pp. 693–697. https://doi.org/10.1016/j.biopha.2014.08.001
11. Screaton, G.R., Bell, M.V., Jackson, D.G., et al., Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, vol. 89, pp. 12160–12164. https://doi.org/10.1073/pnas.89.24.12160
12. Azevedo, R., Gaiteiro, C., Peixoto, A., et al., CD44 glycoprotein in cancer: a molecular conundrum hampering clinical applications, *Clin. Proteomics*, 2018, vol. 15, p. 22. https://doi.org/10.1186/s12014-018-9198-9
13. Ponta, H., Sherman, L., and Herrlich, P.A., CD44: from adhesion molecules to signalling regulators, *Nat. Rev. Mol. Cell Biol.*, 2003, vol. 4, pp. 33–45. https://doi.org/10.1038/nrm1004
14. Zöller, M., CD44: can a cancer-initiating cell profit from an abundantly expressed molecule?, *Nat. Rev. Cancer*, 2011, vol. 11, pp. 254–267. https://doi.org/10.1038/nrc3023
15. Xu, H., Niu, M., Yuan, X., et al., CD44 as a tumor biomarker and therapeutic target, *Exp. Hematol. Oncol.*, 2020, vol. 9, p. 36. https://doi.org/10.1186/s40164-020-00192-0
16. Bhattacharya, R., Mitra, T., Ray Chaudhuri, S., and Roy, S.S., Mesenchymal splice isoform of CD44 (CD44s) promotes EMT/invasion and imparts stem-like properties to ovarian cancer cells, *J. Cell. Biochem.*, 2018, vol. 119, pp. 3373–3383. https://doi.org/10.1002/jcb.26504
17. Skandalis, S.S., Karalis, T.T., Chatzopoulos, A., and Karamanos, N.K. Hyaluronan–CD44 axis orchestrates cancer stem cell functions, *Cell. Signal.*, 2019, vol. 63, pp. 109377. https://doi.org/10.1016/j.cellsig.2019.109377
18. Ghandi, M., Huang, F.W., Jané-Valbuena, J., et al., Next-generation characterization of the Cancer Cell Line Encyclopedia, *Nature*, 2019, vol. 569, pp. 503–508. https://doi.org/10.1038/s41586-019-1186-3
19. Ward, J.H., Hierarchical grouping to optimize an objective function, *J. Am. Stat. Assoc.*, 1963, vol. 58, pp. 236–244. https://doi.org/10.2307/2282967
20. Waskom, M.L., (n. d.) seaborn: statistical data visualization, *J. Open Source Software*, 2021, vol. 6, pp. 3021. https://doi.org/10.21105/joss.03021
21. Brown, T.A., Bouchard, T., St John, T., et al., Human keratinocytes express a new CD44 core protein (CD44E) as a heparan-sulfate intrinsic membrane proteoglycan with additional exons, *J. Cell Biol.*, 1991,

vol. 113, pp. 207–221.
https://doi.org/10.1083/jcb.113.1.207

22. Naor, D., Nedvetzki, S., Golan, I., et al., CD44 in cancer, *Crit. Rev. Clin. Lab. Sci.*, 2002, vol. 39, pp. 527–579. https://doi.org/10.1080/10408360290795574

23. Suwannakul, N., Ma, N., Thanan, R., et al., Overexpression of CD44 variant 9: a novel cancer stem cell marker in human cholangiocarcinoma in relation to inflammation, *Mediat. Inflamm.*, 2018, vol. 1–8. https://doi.org/10.1155/2018/4867234

24. Suwannakul, N., Ma, N., Midorikawa, K., et al., CD44v9 induces stem cell-like phenotypes in human cholangiocarcinoma, *Front. Cell Dev. Biol.*, 2020, vol. 8, p. 417. https://doi.org/doi.org/10.3389/fcell.2020.00417

25. Taniguchi, D., Saeki, H., Nakashima, Y., et al., CD44v9 is associated with epithelial-mesenchymal transition and poor outcomes in esophageal squamous cell carcinoma, *Cancer Med.*, 2018, vol. 7, pp. 6258–6268. https://doi.org/10.1002/cam4.1874

Translated by I. Gordon