International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 05, 2020
ISSN: 1475-7192

Catalogue of Cancer Genes – The First Step to Winning the Fight against Cancer

Dmitry V. Karpov, Lyubov I. Karpova, Ilya I. Vinogradov and
Sergey V. Dmitrienko

Abstract--- *Performance of an important task of creating a detailed catalogue of cancer genes will allow to select an optimal cancer therapy for each patient. To catalogue cancer genes, mutating with high (>20%) and medium (2–20%) frequency, 200 "tumour-norm" pairs on average for each gene should be analysed, i.e. about 100000 pairs for the most common types of cancer. Nowadays, the problem is no longer insoluble, since the cost of DNA sequencing has decreased a million fold over the last 10 years and will continue to decrease more.*

Keywords--- *Sequencing, Cancer, Proto-oncogene, Anti-oncogene, Cancer Genome Atlas.*

I. INTRODUCTION

Everyone remembers that cancer begins from mutations, disorders in the DNA sequence [1-3], after which the cell starts to divide uncontrollably. However, there are multiple types of cancer, and each one has its specific mutations, distinguishing only this particular type. But there are mutations that, vice versa, prove to be fairly universal and are found in many tumours. Furthermore, there are just mutations, which are perhaps detrimental, but have nothing to do with cancer.

To date, cancer ranks second among the reasons of human death, following only cardiovascular diseases [15,16] (every year, approximately 8 mln. people die worldwide from cancer), and in some developed countries, for example, in Denmark, cancer has already taken the first place. Cancer is a complex, dynamically developing disease of more than 200 known types and forms. Each of them requires an individual approach, individual treatment strategy. At the genetic level different cancers are characterised by various "architecture": sets of somatic mutations, chromosome rearrangements, and such epigenetic anomalies, as alteration of gene methylation profile. As a result, the activity of genes and (or) their products changes. Detailed data on the anomalies, related to occurrence and development of cancerous tumours, are needed to diagnose and effectively treat cancer, to define an optimal therapy, and develop new anticancer agents. The objects of studies involve anomalies in molecular DNA, RNA, and protein structure and epigenetic anomalies (methylation, in particular). It has been this approach, which has been adopted as a general strategy, implemented by several national and international consortia, incorporating dozens of institutions, universities, and clinics. It should finally lead to creation of more effective treatment agents, which will

Dmitry V. Karpov, Associate Professor of the Department of Surgery, Ryazan State Medical University named after Academician I.P. Pavlov, Ryazan, Russian Federation.
State Budgetary Institution of the Ryazan Region "Regional Clinical Hospital", Ryazan, Russian Federation.
Lyubov I. Karpova, 2-Year Postgraduate Student, Ryazan State Medical University named after Academician I.P. Pavlov, Ryazan, Russian Federation.
State Budgetary Institution of the Ryazan Region "City Clinical Hospital № 4", Ryazan, Russian Federation.
Ilya I. Vinogradov, Associate Professor, Department of Histology, State Budgetary Institution of the Ryazan Region "City Clinical Hospital № 4", Ryazan, Russian Federation.
Sergey V. Dmitrienko, Urologist, State Budgetary Institution of the Ryazan Region "Regional Clinical Hospital", Ryazan, Russian Federation.

DOI: 10.37200/IJPR/V24I5/PR2020200
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

accommodate, for example, not only genetic effect of mutation, but its influence on epigenetic processes.

II. MATERIALS AND METHODS

The paper specifies the classification of genes, involved in carcinogenesis, pathways of their activation; "The Cancer Genome Atlas" Consortium review.

III. RESULTS AND DISCUSSION

Maximum progress has been made and the most valuable results have been obtained as a result of searching for genes, related to initiation, development, and maintenance of malignant transformation of cells [12,17]. A delicate equilibrium should be maintained in the organism between the activity of genes and their products, which, on the one hand, ensure growth and division of cells, and on the other hand, prevent them from growing and dividing with no limits. An excessive activity of the first or suppression of the function of the second result in uncontrollable growth of cells, occurrence and development of neoplasms, or cancerous tumours. Cancer-related genes can be classified as oncogenes and anti-oncogenes (tumour suppressors), the products of which can, respectively, promote or supress oncogenesis. MicroRNA (miRNA) - short (~22 nucleotides on average) non-coding RNAs hold a special place. To date, about 2000 different microRNAs have been identified. They are able to suppress mRNA translation, read from 30-60% of human genes. Certain microRNAs (oncomiR) facilitate malignant transformation of cells, others can function as anti-oncogenes. Normal gene-protooncogene, which promotes the growth of cells constantly or at the certain stages of the organism development, can be transformed into oncogene. Transformation of protooncogene into oncogene occurs due to relatively minor modification of its natural function. There are the following major ways of protooncogene activation:

1. Mutation inside protooncogene or in its regulatory elements, which changes the structure of protein and increases the activity of its coded protein (enzyme) or enhances the expression of the corresponding gene.
2. Increasing protein concentration due to an increase in its stability in the cell, an increase in the half-life and, respectively, an increase in the activity.
3. Gene duplication (an increase in the number of copies), which results in the increased concentration of protein in the cell.
4. Translocation of gene, which causes strengthening of its expression or emergence of aggressive hybrid gene.
5. Oncogenes are, e.g., Ras (*abbr.* from Rat sarcoma) family genes, GT phases, involved in transmitting the signals that promote division of cells. The function of anti-oncogenes is opposite to that of proto-oncogenes. Anti-oncogenes control different processes impeding malignant transformation of cells:
   1) Suppression of excessive expression of genes, ensuring proliferation of cells.
   2) DNA reparation (damages in DNA, when reparation is suppressed, intensify mutagenesis and, as a result, activation of proto-oncogenes and inactivation of anti-oncogenes).
   3) Coordination of proliferation of cells with DNA reparation. If DNA reparation is suppressed, they impede the cell division and trigger apoptosis.
   4) Controlling adhesion and mechanisms of contact inhibition of dividing cells.

DOI: 10.37200/IJPR/V24I5/PR2020200
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

Generally speaking, anti-oncogenes set a barrier to the unlimited growth of cells. Loss of anti-oncogene function destroys this barrier. The most well-known anti-oncogene, which frequently mutates in cancerous tumours of multiple types, is TP53 [4,5]. The TP53 product is a phosphoprotein, regulating transcription of a number of different genes. It is inactive in the normal cell. In case of extreme events, it becomes active and serves as a "genome guardian", fulfilling various anticancer functions:

1) Activation of DNA reparation system.
2) If DNA is damaged, TP53 impedes mitosis of dividing cells, blocking transition from G1-phase to S-phase and providing the reparation system with the time to eliminate damages.
3) Failing elimination of DNA damages, TP53 initiates apoptosis, the program of cell death.

If cancer is caused by proto-oncogene, to activate this proto-oncogene on one of the two paired cell chromosomes is usually enough. However, if cancer emerged due to loss of anti-oncogene function, mutations or loss of both its copies are mostly required.

By now, about 300 oncogenes and tumour suppressors have been identified [7,9,18-20]. Cancer genes have been sought using comparative sequencing, i.e. sequences of nucleotides in DNA of tumours are compared with those in normal tissues, and then somatic mutations, missing in DNA of the normal tissue, are identified, which are found more often than just random events. This strategy is implemented in several stages. First, it is required to obtain the samples of tumour tissue from patients with a definite diagnosis and thoroughly described disease course. Here, the samples should be, whenever possible, free from normal cells. Samples of a patient's normal tissue or blood are used for comparison. DNA is extracted from the tumour and normal tissue and studied using sequencing. In recent years, sequencing has often been performed using platforms of new generation, which enable sequencing human genome quickly and relatively inexpensively. The results of comparative sequencing are further analysed applying specially developed fairly complex mathematical and bioinformatic methodologies. The main purpose of these projects implies creation of the detailed catalogue of the genome structure anomalies associated with initiation, proliferation, and maintenance of cancerous neoplasms. Such a catalogue will allow not only to obtain multiple new data on molecular biology of cancer, but to improve the methods for cancer diagnostics, treatment, and prevention, to define new targets for developing anticancer agents. Systematic studies in this direction have already enabled identifying many new cancer genes and even whole classes of cancer genes. To date, the data have already been analysed, obtained using large "tumour-normal tissue" samples for more than 30 forms of cancer. The most significant progress is demonstrated by the Consortium mostly working for the USA institutions and universities, named "The Cancer Genome Atlas" [5] meaningfully abbreviated as TCGA – the same letters are used to designate four nucleotides, the components of DNA. Established in 2005, TCGA regularly publishes the results of its studies in the leading scientific journals. It turns to be impossible to tell about all the Consortium publications herein. The authors will present only the results from the last article devoted to squamous cell carcinoma of the head and neck. This heterogeneous group of cancers is the sixth by the incidence rate and amounts to ~ 5% of all cancers worldwide. 348 authors were involved in the studies. 279 "tumour-norm" pairs were analysed. Most tumours related to human papilloma viruses, had mutations in the helical domain of the PIK3CA oncogene. New anomalies were found, including loss of TRAF3, amplification of the E2F1 gene, involved in controlling the cell cycle. In smoking-related

DOI: 10.37200/IJPR/V24I5/PR2020200
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

tumours, inactivating mutations of the TP53 and CDKN2A genes were almost always observed, amplifications of segments of 3q26/28 and 11q13/22 chromosomes were seen. Oral tumours relatively favourable in terms of potential treatment and chances for recovery, contained activating mutations of the HRAS or PIK3CA genes in combination with inactivating mutations of the CASP8, NOTCH1 and TP53 genes. In cases of other subgroups of this cancer, inactivating mutations of the NSD1 gene were found, the product of which is linked to rearrangements of chromatin, inactivating mutations of the AJUBA_and FAT1 genes, controlling enzymes of the Wnt signalling pathway, mutations activating the NFE2L2 oxidative stress factor. Some cancer-related genes mutate rather often in multiple or, at least, in several types of cancer. Hence, there is little wonder that it has been these ones (TP53, in particular), which were first characterised. However, most cancer genes are found with medium frequency (2-20%) or less often. The main problems arise when identifying rare cancer genes. Thus, recent study of 183 pulmonary adenocarcinomas has shown that in 15% of patients no mutations are found in 10 classes of genes known for this disease, and in 38% of cases three or less mutations were identified [11]. Due to damaging the DNA reparation systems, mutagenesis is somewhat intensified in cancer cells. The frequency of these tumour-induced somatic mutations may vary by several orders of magnitude for different tumours. Hence, when identifying "medium" and especially "rare" genes, an essential problem arises: how genes and mutations related to cancer can be differentiated from the background, from multiple random mutations not associated with cancer? In two articles [10,13,14] the authors have studied the results, collected from different databases, of analysing exomes (coding regions of genes - exons with DNA sequences adjacent to them – non-coding intervals - introns) of 4742 "tumour-normal tissue" pairs, belonging to cancers of 21 various types. The number of samples for certain types varied from 35 to 892. 3078483 one-time replacements of nucleotides in tumours were found as compared with the normal tissue, 77270 one-time deletions or insertions of nucleotides, 29837 di-, tri- or oligonucleotide deletions or insertions. Most (2294935) of one-time replacements of nucleotides didn't change coding sequences. Out of the remaining one-time replacements, 540831 constituted the so-called missense-mutations_(causing replacements of amino acids in proteins), 207144 constituted synonymic (not causing replacements of amino acids) replacements of nucleotides, 46264 nonsense mutations_(causing long-term termination of protein synthesis), 33673 - mutations, damaging mRNA splicing (assembly of coding sequences of mRNA from the respective blocks). The data on the "depth" of sequencing and purity of tumours samples allow to assess more than 90% sensitivity of the analysis. Incidence rates of mutations per a unit of genome length for different types of cancer differed in more than 5 orders of magnitude (from 0.03 to 7000 per a million of DNA nucleotides), mutation spectra also greatly differed [6]. Mutations associated with cancer with certainty, were found in 224 different genes. For different types of cancer, the number of mutant genes greatly varied (from 1 to 58). For 7 types it was less than 10, and for two (breast and endometrial cancers) – more than 30. Only 22 genes were found definitely associated with more than three types of cancer. The analysis enabled identifying almost all previously known genes, related to carcinogenesis. 33 genes were also found, mutations in which were previously not associated with cancer. These genes relate to division of cells, apoptosis, genome stability, chromatin activity regulation, immune response, RNA transformations, and homeostasis of proteins. Among 81 genes more, there also should be genes associated with cancer. Based on the obtained results, the authors computed: how many "tumour-normal tissue" pairs should be analysed to find genes definitely mutating in cancer

DOI: 10.37200/IJPR/V24I5/PR2020200
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

depending on its type, incidence rate of mutations per a unit of genome length for a given type, incidence rate of mutations of this gene for a given type of cancer. Computations show that for 17 of 21 of the analysed cancer types, there are still not enough data to identify the genes mutating with the frequency not higher than 5% of the background, and for 7 types – even with not more than 10% frequency. It has also been identified that to catalogue cancer genes covering 90% of disease cases, it is required to analyse about 650 "tumour-norm" pairs, if the average frequency of mutations is ~0.5 per a million of DNA base pairs (as in neuroblastoma) [6, 20], or even about 5300 pairs, if the frequency is ~12.9 per a million (as in melanoma). All in all, to catalogue somatic mutations for genes, mutating with high (>20%) and medium (2-20%) frequency, for ~ 50 known types of cancer 2000 "tumour-norm" pairs should be analysed on average, i.e. about 100000 pairs in total.

IV. CONCLUSION

Overall, creation of the detailed catalogue of cancer genes is an important task, fulfilling which will allow to select the optimal cancer therapy for each patient: influence on certain signalling pathways or other processes, damaged in each specific case. Such a catalogue is also needed to choose the targets when developing anticancer agents, create new experimental models of animals and cell lines to study cancer, test new treatment means and methods.

ACKNOWLEDGEMENTS

The paper is written with co-financing of the MK-6388.2018.7 Grant of the President of the Russian Federation for the State Support of young Russian scientists.

REFERENCES

[1] Bamford S., Dawson E., Forbes S. et. al. The COSMIC (Catalogue of Somatic Mutations in Cancer) / *British Journal of Cancer*. - 2004 - 19 July (vol. 91, no.2). - Pp. 355-358.

[2] Boorjian S., Cowan J.E., Konety B.R., et. al. Cancer of the Prostate Strategic Urologic Research Endeavor Investigators. Bladder cancer incidence and risk factors in men with prostate cancer: results from Cancer of the Prostate Strategic Urologic Research Endeavor // *J Urol.* 2007. Vol.177, № 3. Pp. 883-887.

[3] Chen M., Liang J., Ji H. et al. Cdk8/19 Mediator Kinases Potentiate Induction of Transcription by Nfkappab // *Proc. Natl. Acad. Sci.* USA. 2017. V.114. N.38. Pp. 10208–10213.

[4] Chrouser K., Leibovich B. et al. Bladder cancer risk following primary and adjuvant external beam radiation for prostate cancer // *J. Urol.* 2006. Vol. 174, № 1. Pp. 107-110.

[5] Creighton C.J., Morgan M., Gunaratne P.H. et al. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma // *Nature.* - 2013. - V.499. - P.43-9.

[6] Duell E.J., Wiencke J.K., Cheng T.J. et. al. Polymorphisms in the DNA repair genes XRCC1 and ERCC2 and biomarkers of DNA damage in human blood mononuclear cells // *Carcinogenesis.* 2000. Vol. 21, № 5. P. 965-71.

[7] Garnett M.J., Edelman E.J., Heidorn S.J. et. al. Systematic identification of genomic markers of drug sensitivity in cancer cells.// Nature. - 2012. - 28 March (vol. 483, no. 7391). - Pp. 570-575.

[8] Hakimi A.A., Ostrovnaya I., Reva B. et. al. Adverse outcomes in clear cell renal cell carcinoma with mutations of 3p21 epigenetic regulators BAP1 and SETD2: a report by MSKCC and the KIRC TCGA research network // *Clin. Cancer Res.* - 2013. - V.19. - I.12. Pp. 3259-67.

[9] Liang J., Chen M., Hughes D. et al. Cdk8 Selectively Promotes the Growth of Colon Cancer Metastases in the Liver by Regulating Gene Expression of Timp3 and Matrix Metalloproteinases // *Cancer Res.* 2018. V.78. N.23. Pp. 6594-6606.

DOI: 10.37200/IJPR/V24I5/PR2020200
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020

[10] Li Ding & Michael C. Wendl. Differences that matter in cancer genomics // *Nat Biotechnol.* 2013. V.31. Pp. 892-893.
[11] Marcin Imielinski et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing // *Cell.* 2012. V.150. Pp. 1107-1120.
[12] McDermott M.S., Chumanevich A.A., Lim C.U. et.al. Inhibition of CDK8 mediator kinase suppresses oestrogen dependent transcription and the growth of oestrogen receptor positive breast cancer // *Oncotarget* 2017. V.8 (8) Pp.12558-12570.
[13] Michael S. Lawrence et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes// *Nature.* 2013. V. 499. Pp. 214-218.
[14] Michael S. Lawrence et al. Discovery and saturation analysis of cancer genes across 21 tumour types // *Nature.* 2014. V. 505. Pp. 495-501.
[15] Moullan N., Cox D.G., Angele S., Romestaing P., Gerard J.-P., Hall J. Polymorphisms in the DNA repair gene XRCC1, breast cancer risk and response to radiotherapy // *Cancer epidemiology, biomarkers & prevention.* 2003. Vol. 12. Pp. 1168.
[16] Park J.Y., Lee S.Y., Jeon H.-S. et. al. Polymorphism of the DNA repair gene XRCC1 and risk of primary lung cancer // *Cancer epidemiology, biomarkers & prevention.* 2002. Vol.11. Pp. 23-27.
[17] Porter D. C., Farmaki E., Altilia S. et al. Cyclin-dependent kinase 8 mediates chemotherapy-induced tumor-promoting paracrine activities // *Proc. Natl. Acad. Sci.* USA. 2012. V. 109. N. 34. Pp. 13799-13804.
[18] Roninson I. Seeking favors from nature // *Journal Cancer Biology & Therapy.* 2005. V.4. N.7. Pp. 794-799.
[19] Shen M., Hung R.J., Brennan P. Polymorphisms of the DNA repair genes XRCC1, XRCC3, XPD interaction with environmental exposures and bladder cancer risk in a case-control study in Northern Italy // *Cancer epidemiology, biomarkers & prevention.* 2003. Vol. 12. Pp. 1234-1240.
[20] van Haaften G., Dalgliesh G.L., Davies H. et. al. A.Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer./ *Nature Genetics.* - 2009.- May (vol. 41, no. 5). - Pp. 521-523.

DOI: 10.37200/IJPR/V24I5/PR2020200
Received: 12 Mar 2020 | Revised: 28 Mar 2020 | Accepted: 06 Apr 2020