An Extensive Examination of Non-Canonical Proteins in Non-Small Cell Lung Cancer and Their Effects on the Immune System

Deepti Pathak *

Department of Biotechnology, Guru Nanak Dev University, Amritsar, India

Corresponding Author*

Deepti Pathak
Department of Biotechnology, Guru Nanak Dev University, Amritsar, India
E-mail: pathakdeepti34@gmail.com

Received: 03-Jan-2023, Manuscript No. SCR-23-20635; **Editor assigned:** 06-Jan-2023, Pre QC No. SCR-23-20635 (PQ); **Reviewed**: 20-Jan-2023, QC No. SCR-23-20635 (Q); **Revised**: 22-Jan-2023, Manuscript No. SCR-23-20635 (R); **Published**: 26-Jan-2023, doi: 10.352482161-1076.23.13(1).430.

Abstract

Neoantigen mining is of great interest for cancer applications. Neoantigens in cancer have been found to be frameshift-mutated non-canonical proteins. We looked at the Non-Canonical Protein Landscape in Non-Small Cell Lung Cancer (NSCLC) and how autoantibodies were produced as an immune response. All alternate Open Reading Frames (altORFs) and ORFs discovered in pseudogenes, noncoding RNAs, and untranslated sections of mRNAs that did not align with known canonical proteins were compiled into a database of cryptoproteins. To determine the presence of cryptoproteins, the proteomic profiles of seventeen Lung Adenocarcinoma (LUAD) cell lines were examined. Plasma Immunoglobulin (Ig)-bound cryptoproteins were analyzed by mass spectrometry to determine the immunogenicity. The specimen collection included 102 control plasmas, pre-diagnostic plasmas from 51 NSCLC cases, and plasmas from 30 newly diagnosed NSCLC cases. 420 cryptoproteins were found after LUAD cell lines were examined. Analyses of plasma Ig-bound proteins discovered 14 cryptoproteins with a fold-change >2 compared to controls and 90 cryptoproteins that were only found in cases. 17 Ig-bound cryptoproteins in pre-diagnostic samples produced an odds ratio of 2. Both pre-diagnostic and recently diagnosed cases had higher levels of eight Ig-bound cryptoproteins when compared to controls. A class of neoantigens known as cryptoproteins causes an autoantibody response in NSCLC.

Keywords Noncanonical open reading frames • Neoantigens • Autoantibodies

Introduction

With the discovery of tumor-infiltrating B cells in several malignancies, the importance of humoral immunity in immune surveillance is being recognized more and more. Autoantibodies directed against tumor antigens are produced as a result of the B-cell response, which happens early in the tumor's growth. For the purpose of finding circulating autoantibodies in cancer, a variety of techniques have been used, such as serological screening of cDNA expression libraries, recombinant arrays, and phage-display libraries. Protein arrays produced from tumor cell lysates have been used to identify autoantibody signatures. Utilizing mass spectrometry to locate circulating antigen-antibody complexes is a more comprehensive strategy. Recent research has shown that eukaryotic transcripts can contain non-canonical alternate Open Reading Frames (altORFs), resulting in proteins with different biological functions or subcellular localization signals. A few transcripts also have brief upstream Open Reading Frames (uORFs), which play a crucial role in the control of translation. According to a recent study, 6.5% of MHC-bound peptides came from non-canonical reading frames. These peptides were produced by the frameshift translation of protein-coding transcripts, and peripheral blood-derived mononuclear cells demonstrated their immunogenicity to these peptides. Another work used transcriptomics, ribosome profiling, and mass spectrometry to identify hundreds of non-canonical HLA-bound peptides that were shared and particular to tumors. In earlier research, we discovered autoantibody signals in lung cancer in samples taken both at the time of diagnosis and samples taken one or more years beforehand. Proteins and peptides produced from canonical ORFs made up these signatures. According to our hypothesis, novel proteins derived from altORFs, pseudogenes, intronic regions, and other transcripts thought not to encode proteins represent a novel source of tumor antigens that can trigger an immune response and produce autoantibodies. Translational dysregulation is a common occurrence in cancer. These non-canonical proteins are referred to as "cryptoproteins" by us. In order to identify novel cryptoproteins with a low false discovery rate using mass spectrometry-based proteomic analysis, we first built a library of unique cryptoproteins that have no resemblance to the conventional human peptidome. We then used this method to show that lung cancer cell lines contain cryptoproteins. Next, we discovered higher levels of circulating cryptoprotein-antibody complexes in plasma samples from Non-Small Cell Lung Cancer (NSCLC) patients who had just received their diagnosis as well as plasma samples taken earlier in comparison to controls [1].

Materials and Methods

Construction of a cryptoprotein proteomics pipeline and database

The Genome Reference Consortium Human Build 38, release 27 (GRCh38.v27) fasta format transcript files were acquired from GENCODE. All transcripts, including protein-coding genes, pseudogenes, noncoding RNAs like microRNAs and long noncoding RNAs, as well as variations like transcripts with retained introns, underwent in silico translation. The longest open reading frame from transcripts labelled as protein coding was deleted, leaving us with 1.1 million ORFs. We then chose all Open Reading Frames (ORFs) longer than 50 codons, starting with an AUG and ending with a canonical stop-codon (UAA, UGA, UAG). The BLASTP programme was then used to align these with the human Non-Redundant (nr) protein database. The ORFs that aligned with an E-score of more than 0.01, which denotes a likely successful alignment, were eliminated.

Cell lines

The American Type Culture Collection provided seventeen lung adenocarcinoma cell lines (H2228, H1395, H2405, H522, H969, H1703, H1650, HCC827, HCC4006, H820, HCC2935, H2009, H650, H1795, H2122, H647, HCC4017, Supplemental. Short tandem repeat analysis was used to determine the cell type. The cell lines were examined individually and served as representations of various mutational backgrounds. The examination of entire cell lysates using mass spectrometry is explained in detail elsewhere.

Lung cancer plasma collection

After receiving informed consent and institutional review board approval, blood samples were taken from two separate cohorts. Plasma from people who had just been diagnosed with NSCLC was taken from one cohort at the University of Texas MD Anderson Cancer Center (MDA). Plasma samples from participants in the Beta-carotene and Retinol Efficacy Trial (CARET) cohort made up a different cohort. In the CARET study, 18,314 people at high risk for lung cancer were given daily

supplements of beta-carotene and retinol palmitate to see how well they prevented cancer and how safe they were. Six US centers accepted participants who were then monitored for cancer and mortality outcomes. Based on age, sex, and smoking history, six pools of pre-diagnostic NSCLC cases (n=4 patients per pool-14 patients per pool) were matched to six pools of healthy controls (n = 8people per pool-28 people per pool). For a minimum of four years, all controls were monitored to make sure they were cancer-free. The distribution of Ig-bound cryptoproteins between patients and controls was compared for the CARET cohort for the MDA cohort [2].

Data processing of mass spectrometry data

Reprocessing of the proteome spectrum of Immunoglobulin(Ig)-bound plasma proteins and human lung cancer cell lines was performed using a tailored PeptideShaker workflow. Spectra were first searched against the UniProt Database, and spectra recognized as UniProt peptides were filtered out, limiting processing to those spectra that did not align to a known UniProt sequence. The unique cryptoDB was later searched using PeptideShaker on unaligned spectra. We contrasted the PSM for cryptopeptides with the PSM for canonical peptides that matched to the UniProt database. In accordance with earlier methods, peptides were deemed a match if their False Discovery Rate (FDR) was less than 10%. A selection of features in each experiment (cohort) was made based on the detected cryptoprotein having a Peptide-Spectrum Match (PSM) 5 and detection in 2 or more samples in each cohort in order to reduce false positives. Summarizing all aligned spectra implied the abundance of the cryptoprotein [3].

Ingenuity pathway analyses

We used host, canonical gene names that corresponded to each cryptoprotein to conduct Ingenuity Pathway Enrichment Analysis (IPA) to find potential pathway networks linked to the cryptoproteins found in lung cancer cell lines. The two-sided Fisher's Exact Test was used to examine the statistical significance of the enriched pathways.

Statistical analysis

For the newly diagnosed group and the pre-diagnostic cohort, Odds Ratios (ORs) using logistic regression and conditional logistic regression were used to evaluate the predictive ability of Ig-bound cryptoprotein complexes. The R software environment was used to conduct the analyses. Unless otherwise stated, p values are presented using the two-sided Wilcoxon rank sum test.

Discussion

In order to search untargeted proteome datasets of proteomic profiling of lung cancer cell lines and patient plasmas, we created a cryptoprotein database of hypothetical non-canonical proteins. As a result, a hitherto unreported "cryptoproteome" associated with NSCLC was discovered, and autoantibodies directed against cancer-related cryptoproteins were found to be a corresponding humoral response. These findings suggest translational potential in the form of potential early detection or immunotherapy targets in the form of candidate markers. We pursued an untargeted strategy for wide non-canonical protein identification unrestrained by mutational status, in contrast to other research that concentrated on aberrant peptides emerging from chromosomal change [4]. The cryptoDB offers a database of potential protein sequences that are not similar to previously identified human proteins and that can be found in several samples by applying conventional proteomic methods. In addition to providing possible insights into how these RNAs might impact cellular physiology, this suggests that more attention may be paid to the protein-encoding potential of RNAs that were previously assumed to be noncoding. We use the database to demonstrate the approach's validity with results that are statistically significant and concordant across different samples [5]. Quantifiable cryptoproteins were analysed in newly diagnosed Ig-bound samples, pre-diagnosed Ig-bound samples, and cell lines. It was discovered that the majority of quantified cryptoproteins (>40%) were derived from protein-coding transcripts, indicating that these cryptoproteins are produced at the translational level. While proteins emerging from genetic abnormalities in cancer have received a lot of attention, our discovery offers solid proof that there are other ways that detectable neoantigens might develop. The somewhat limited utility of tumour mutational burden in predicting response to cancer immunotherapy, for example, may be partially explained by this.

This link with translation mistakes is supported by a cell line analysis. A "slippery ribosome," altered nonsense-mediated decay, relaxed translational fidelity, or aberrant transcription of ORF-containing pseudogenes are a few probable explanations for how a particular cryptoprotein is produced. It is necessary to investigate potential underlying mechanisms involved in the production of the cryptoproteins we have identified, although such research is outside the purview of the current study.

It's interesting to note that, when compared to controls, more than half (56%) of the Ig-bound cryptoproteins were only detected in NSCLC plasmas. Additionally, there is a chance that Ig-bound cryptoproteins in pre-diagnostic plasmas will be sensitive and precise indicators of lung cancer risk and disease presence. These might either be used in addition to current indicators or potentially perform well enough to be used alone as a new source of biomarkers for lung cancer early detection or risk assessment. Future research is necessary to determine the effectiveness of autoantibodies against cancer-associated cryptoproteins alone or in conjunction with other types of biomarkers for the risk prediction of lung cancer [6].

On the other hand, our research do have certain limitations. Limited correlation studies with smoking exposure resulted from the lack of detailed information covering the whole smoking history, including smoking duration. The presence of cryptoproteins in exosomes linked to cancer or in circulating tumour cells was also not examined. Our investigation concentrated on the presence of Ig-bound cryptoproteins in the plasma of lung cancer patients with adenocarcinoma and squamous cell carcinoma. It is unknown if the same phenomenon can be observed in cases of small cell lung cancer or other NSCLC subtypes, such as big cell carcinoma [7].

Finally, we identify cryptoproteins as a possible neoantigen source in NSCLC. A promising source of biomarkers that may help identify people who are highly susceptible to developing or harbouring lung cancer is autoantibodies against cancer-associated cryptoproteins. Future work will involve confirming the results using different datasets, gathering more samples, and biologically verifying autoantibody reactivity in plasma from lung cancer patients.

References

1. Tsou, P., et al. "The Emerging Role of B Cells in Tumor ImmunityB Cells, Cancer Immunity, and Autoantibodies." *Cancer Res.* 76.19 (2016): 5597-5601.
2. Garaud, S., et al. "Tumor-Infiltrating B Cells Signal Functional Humoral Immune Responses in Breast Cancer." *JCI Insight* 4.18 (2019): e129641
3. Hanash, S. "Harnessing Immunity for Cancer Marker Discovery." *Nat Biotechnol.* 21.1 (2003): 37-38.
4. Naour, F.L. "Identification of Tumor Antigens by Using Proteomics." *Target Discov Valid Rev Protoc.* Humana Press, 2007. 327-334.
5. Pereira-Faca, S.R., et al. "Identification of 14-3-30as an Antigen that Induces a Humoral Response in Lung Cancer." *Cancer Res.* 67.24 (2007): 12000-12006.
6. Lu, H., et al. "Humoral Immunity Directed against Tumor-Associated Antigens as Potential Biomarkers for the Early Diagnosis Of Cancer." *J Proteome Res* 7.4 (2008): 1388-1394.
7. Babel, I., et al. "Identification of Tumor-Associated Autoantigens for the Diagnosis of Colorectal Cancer in Serum using High Density Protein Microarrays." *Mol Cell Proteom.* 8.10 (2009): 2382-2395.

Cite this article: Pathak, D. An Extensive Examination of Non-Canonical Proteins in Non-Small Cell Lung Cancer and Their Effects on the Immune System. Surg Curr Res. 2022, 12 (11), 001-002