Article

A Comprehensive Search of Non-Canonical Proteins in Non-Small Cell Lung Cancer and Their Impact on the Immune Response

Ehsan Irajizad [1,2], Johannes F. Fahrmann [1], James P. Long [2], Jody Vykoukal [1], Makoto Kobayashi [3], Michela Capello [1], Chuan-Yih Yu [1], Yining Cai [1], Fu Chung Hsiao [1], Nikul Patel [1], Soyoung Park [1], Qian Peng [1], Jennifer B. Dennison [1], Taketo Kato [4], Mei Chee Tai [5], Ayumu Taguchi [6,7], Humam Kadara [5], Ignacio I. Wistuba [5], Hiroyuki Katayama [1], Kim-Anh Do [2,*], Samir M. Hanash [1,*,†] and Edwin J. Ostrin [8,*,†]

1   Department of Clinical Cancer Prevention, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030, USA
2   Department of Biostatistics, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030, USA
3   Department of Basic Pathology, School of Medicine, Fukushima Medical University, Hikarigaoka, Fukushima 960-1247, Japan
4   Department of Thoracic Surgery, Nagoya University, Nagoya 464-8601, Japan
5   Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030, USA
6   Division of Molecular Diagnostics, Aichi Cancer Center, Nagoya 464-8601, Japan
7   Division of Advanced Cancer Diagnostics, Nagoya University Graduate School of Medicine, Nagoya 464-8601, Japan
8   Departments of General Internal Medicine, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030, USA
*   Correspondence: kimdo@mdanderson.org (K.-A.D.); shanash@mdanderson.org (S.M.H.); ejostrin@mdanderson.org (E.J.O.); Tel.: +1-713-745-5242 (S.M.H.)
†   These authors contributed equally to this work.

Abstract: There is substantial interest in mining neoantigens for cancer applications. Non-canonical proteins resulting from frameshift mutations have been identified as neoantigens in cancer. We investigated the landscape of non-canonical proteins in non-small cell lung cancer (NSCLC) and their induced immune response in the form of autoantibodies. A database of cryptoproteins was computationally constructed and comprised all alternate open reading frames (altORFs) and ORFs identified in pseudogenes, noncoding RNAs, and untranslated regions of mRNAs that did not align with known canonical proteins. Proteomic profiles of seventeen lung adenocarcinoma (LUAD) cell lines were searched to evaluate the occurrence of cryptoproteins. To assess the immunogenicity, immunoglobulin (Ig)-bound cryptoproteins in plasmas were profiled by mass spectrometry. The specimen set consisted of plasmas from 30 newly diagnosed NSCLC cases, pre-diagnostic plasmas from 51 NSCLC cases, and 102 control plasmas. An analysis of LUAD cell lines identified 420 cryptoproteins. Plasma Ig-bound analyses revealed 90 cryptoproteins uniquely found in cases and 14 cryptoproteins that had a fold-change >2 compared to controls. In pre-diagnostic samples, 17 Ig-bound cryptoproteins yielded an odds ratio ≥2. Eight Ig-bound cryptoproteins were elevated in both pre-diagnostic and newly diagnosed cases compared to controls. Cryptoproteins represent a class of neoantigens that induce an autoantibody response in NSCLC.

Keywords: noncanonical open reading frames; altORFs; neoantigens; autoantibodies; NSCLC

1. Introduction

There is an increasing appreciation of the role of humoral immunity in immune surveillance, with findings of tumor infiltrating B lymphocytes being documented in numerous cancers [1,2]. The B-cell response occurs early during tumor development, resulting in

the production of autoantibodies against tumor antigens [3–5]. Several strategies have been applied for the discovery of circulating autoantibodies in cancer, including the serological screening of cDNA expression libraries (SEREX) [6], recombinant arrays [7,8], and phage-display libraries [9]. Tumor cell lysate-derived protein arrays have been utilized to define autoantibody signatures [10,11]. A more global approach consists of utilizing mass spectrometry to identify circulating antigen–antibody complexes.

Recent studies have demonstrated that eukaryotic transcripts may encompass non-canonical alternate open reading frames (altORFs), resulting in proteins with altered subcellular localization signals or different biological activities [12]. Additionally, some transcripts contain short upstream open reading frames (uORFs) that have a well-described role in translational regulation [13]. A recent report indicated that 6.5% of MHC-bound peptides were derived from a non-canonical reading frame. These peptides originated from the frameshifted translation of protein-coding transcripts, and the resulting peptides were shown to be immunogenic to peripheral blood-derived mononuclear cells [14]. Another study incorporated transcriptomics, ribosomal profiling, and mass spectrometry to elucidate hundreds of shared and tumor-specific, non-canonical HLA-bound peptides [15]. A study of melanoma cell lines showed that the induction of IFNγ through depletion of tryptophan contributes to the immune recognition of melanoma cells through an aberrant peptidome [16]. However, to date, there have been no global searches for immunogenic non-canonical proteins in cancer.

In prior studies, we identified autoantibody signatures in lung cancer from samples collected at the time of diagnosis as well as samples collected one or more years preceding diagnosis [17–19]. These signatures consisted of proteins and peptides derived from the canonical ORFs. We hypothesized that, given widespread translational dysregulation in cancer [20,21], novel proteins derived from altORFs, pseudogenes, intronic regions, and other transcripts considered not to encode proteins represent a novel source of tumor antigens that can elicit an immune response resulting in autoantibodies. We designate these non-canonical proteins 'cryptoproteins'. We first constructed a database of novel cryptoproteins with no homology to the canonical human peptidome that could be identified with a low false discovery rate using mass spectrometry-based proteomic analysis. We then applied this approach to demonstrate the occurrence of cryptoproteins in lung adenocarcinoma cell lines. Next, we identified circulating cryptoprotein–antibody complexes elevated in plasma samples from newly diagnosed non-small cell lung cancer (NSCLC) patients as well as plasma samples collected prior to diagnosis compared to controls.

2. Results

The cryptoproteome of human lung adenocarcinoma cell lines

Using our customized pipeline and cryptoDB, we evaluated the cryptoproteome of 17 human lung adenocarcinoma cell lines previously profiled by mass spectrometry (Supplemental Table S1) [22,23]. A total of 420 cryptoproteins were identified with PSM $\geq$ 5 and that were quantifiable in two or more cell lines. The abundance of cryptoproteins was not associated with prevalent (*KRAS*, *EGFR*, *KEAP1*, *LKB1*, *TP53*) mutations commonly observed in lung adenocarcinomas (Figure 1A) [24]. Of the 420 quantified cryptoproteins, 183 (43.6%) were derived from protein-coding genes, 101 (24.0%) from retained introns, 46 (11.0%) from processed transcripts, 32 (7.6%) from antisense RNAs, 25 (6.0%) transcripts flagged as undergoing nonsense mediated decay, 14 (3.3%) from long intergenic noncoding RNA (lincRNA), 6 (1.4%) from processed pseudogenes, 5 (1.2%) from non-experimentally confirmed regions, two (<1%) from sense intronic, and one each (<1%) of macro-lincRNA, miscellaneous RNA, unspecified pseudogene, sense overlapping, transcribed processed pseudogene, and transcribed unprocessed pseudogene (Figure 1B).
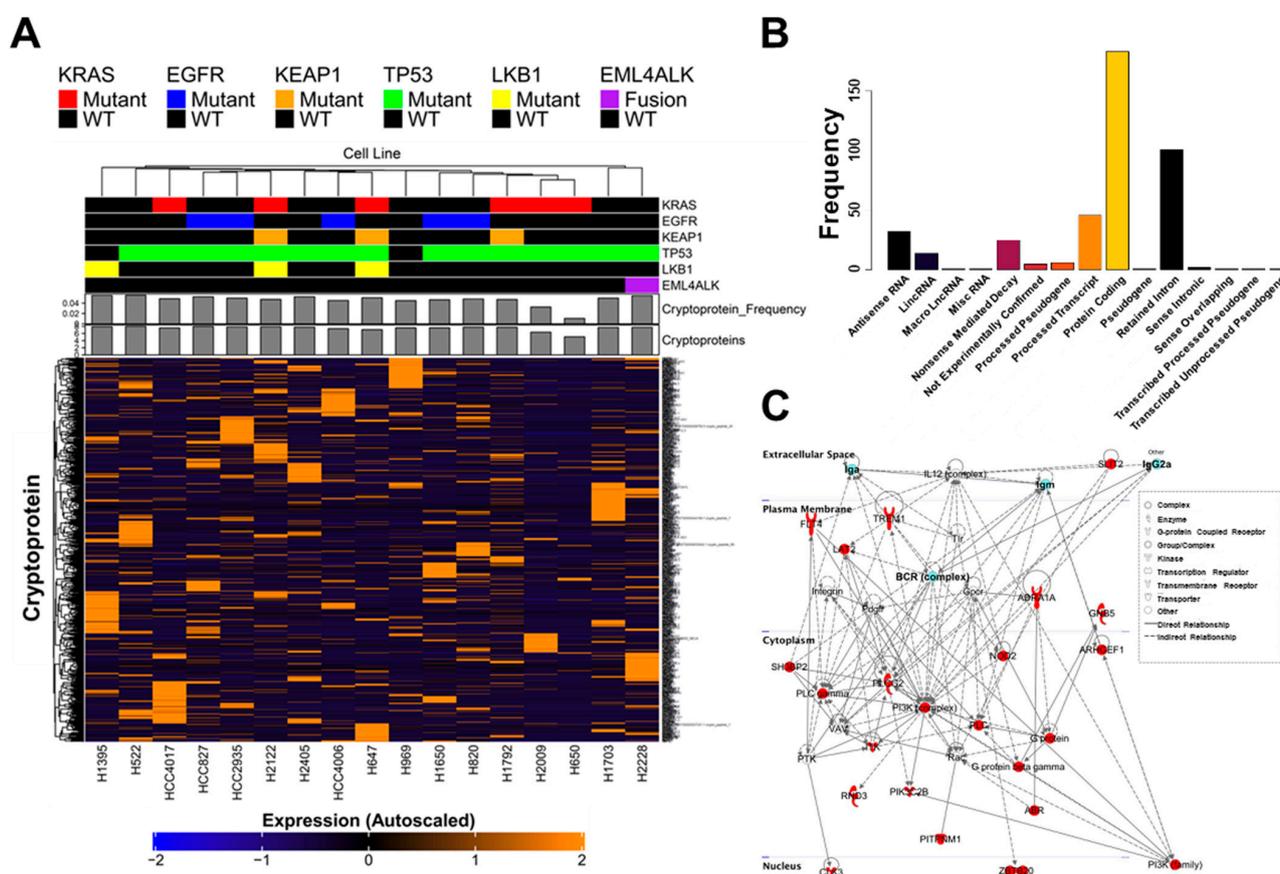
Figure 1. Cryptoproteins are widely detected in human lung adenocarcinoma cell lines: (**A**) A heatmap representation of cryptoprotein expression in lung adenocarcinoma cell lines with annotated driver mutation (KRAS, EGFR, KEAP1, TP53, LKB1, EML4/ALK) shows no association between mutational status and cryptoprotein production. (**B**) Cryptoproteins identified and quantified in lung adenocarcinoma cell lines. (**C**) Ingenuity pathway analyses of genes corresponding to the 420 cryptoproteins detected in lung adenocarcinoma cell lines revealed enrichment of inflammatory pathways with B-cell receptor and immunoglobulin families as central network nodes.

Ingenuity pathway analysis of the 420 canonical gene names corresponding to the identified cryptoproteins revealed immune-centric networks, with the B-cell receptor and immunoglobulin (Ig) genes family being represented as central nodes (Figure 1C), consistent with previous findings [14–16].

Circulating Ig-bound cryptoproteins in NSCLC

To expand upon our findings and determine the extent by which cryptoproteins were associated with a humoral response in the context of NSCLC, we performed the proteomic mass spectrometry-based analyses of Ig-bound cryptoproteins in plasmas from 30 newly diagnosed NSCLC cases (*n* = 3 patients per pool; 10 pools in total) and 102 controls (8–30 individuals per pool; six pools in total) (Supplemental Table S2) [17]. Our analysis yielded 162 Ig-bound cryptoproteins, of which 104 were quantified exclusively in NSCLC case samples (*n* = 90) or had fold-change >2 (*n* = 14) compared to controls (Supplemental Table S4, Figure 2A). Of the 104 cancer-associated, Ig-bound cryptoproteins, 46 (44.2%) were derived from protein-coding genes, 17 (16.3%) from retained introns, 13 (12.5%) from processed transcripts, 7 (6.7%) from antisense RNAs, 10 (9.6%) from lincRNA, 4 (3.8%) from processed pseudogenes, 3 (2.9%) from nonsense mediated decay, two (1.9%) from non-experimentally confirmed regions, and two (1.9%) from sense overlapping transcripts (Figure 2B). Limited overlap was found between Ig-bound, non-canonical cryptoproteins and respective Ig-bound canonical proteins in newly diagnosed cases, suggesting preferential immunogenicity directed against the non-canonical cyptoprotein (Figure 2C). Moreover,

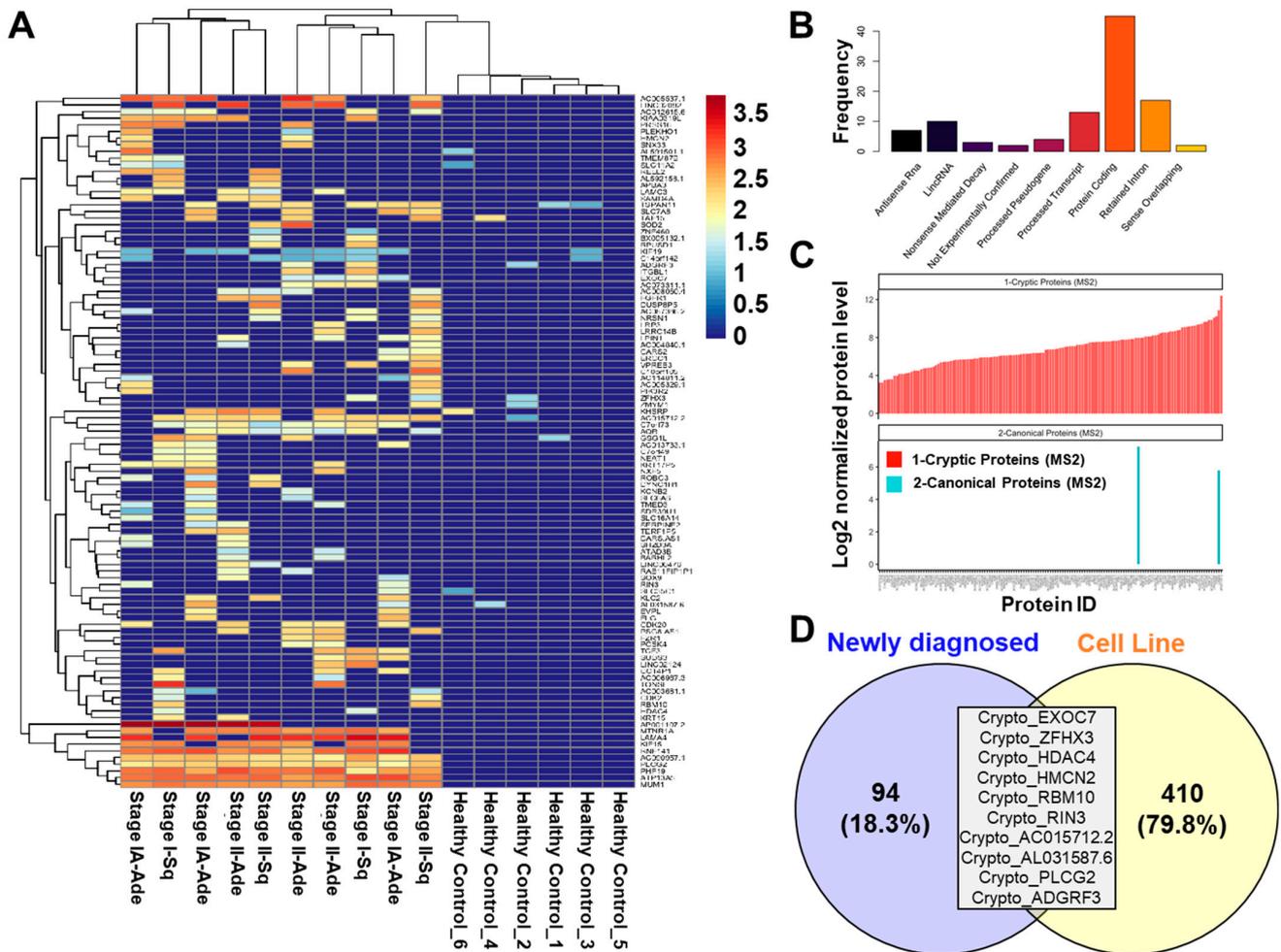10 out of 104 elevated cryptoproteins were also identified in adenocarcinoma cell lines (Figure 2D).

Figure 2. A unique Ig-bound cryptoprotein signature is detected in plasma from individuals newly diagnosed with NSCLC: (**A**) Heatmap depicting 104 Ig-bound cryptoproteins that are either uniquely identified or had a fold-change greater than 2 in cases compared to controls. Abbreviations: Sq, Squamous Cell Carcinoma; Adc, Adenocarcinoma. (**B**) Types of cryptoproteins that were elevated in the Ig-bound fraction of newly diagnosed lung cancer cases compared to controls. (**C**) Comparison of Ig-bound cryptoproteins and their canonical counterparts reveals near exclusivity of Ig reactivity against cryptoproteins. (**D**) Venn diagram illustrating overlap between cryptoproteins identified in lung adenocarcinoma cell lines with cryptoproteins that were elevated in the Ig-fraction of newly diagnosed lung cancer cases.

We further performed mass spectrometry-based, Ig-bound cryptoprotein analyses using 6 pools of pre-diagnostic samples consisting of 51 NSCLC subjects diagnosed within an average of six months following blood draw and compared findings with that of the 102 controls (Supplemental Table S3). A total of 101 Ig-bound cryptoproteins were quantified, 17 of which yielded statistically significant (1-sided $p < 0.05$) odds ratio ≥ 2 (Figure 3A,B, Supplemental Table S5). Of the 17 cancer-associated, Ig-bound cryptoproteins, 9 (52.9%) were derived from protein-coding genes, 3 (17.6%) from retained introns, 3 (17.6%) from processed transcripts, 1 (6.0%) from antisense RNAs, and 1 (6.0%) nonsense mediated decay (Figure 3C). Eight Ig-bound cyptoproteins consistently elevated the pre-diagnostic and newly diagnosed cases compared to controls (Supplemental Table S6).
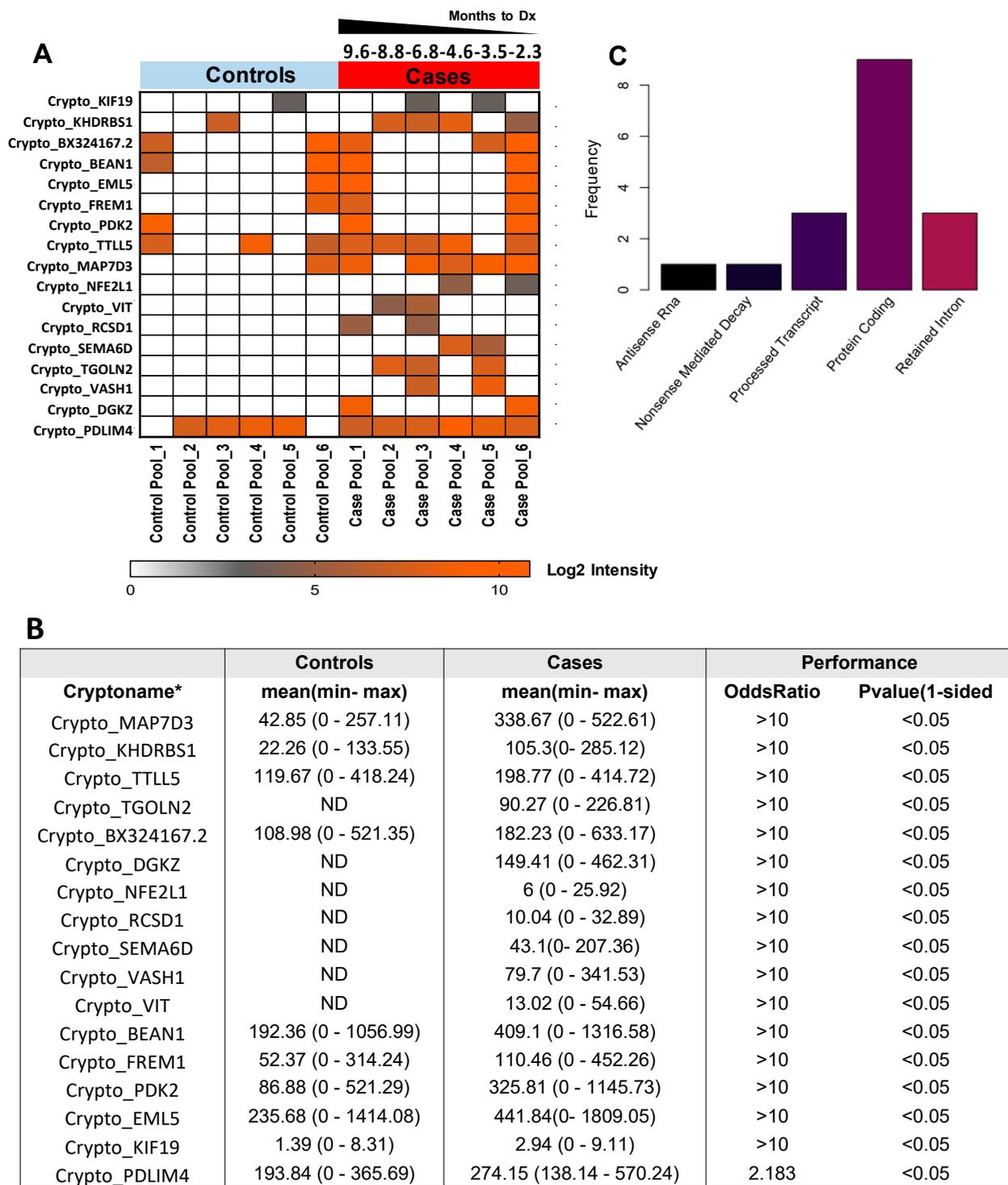
Figure 3. Ig-bound cryptoproteins are detectable in NSCLC pre-diagnostic plasma: (**A**) Heatmap showing 17 Ig-bound cryptoproteins that were elevated (OR > 2; 1-sided *p* value < 0.05) in plasmas collected within 1 year of a lung cancer diagnosis compared to controls. Abbreviations: Sq, Squamous Cell Carcinoma; Adc, Adenocarcinoma. (**B**) Tabular representation of the 17 Ig-bound cryptoproteins. ND = Not Detected. (**C**) Histogram detailing the cryptoprotein type for the 17 Ig-bound cryptoproteins that were elevated in cases diagnosed within 1 year of blood draw compared to controls.

3. Discussion

We constructed a cryptoprotein database of theoretical non-canonical proteins that we then applied to search untargeted proteomic datasets of proteomic profiling of lung adenocarcinoma cell lines and patient plasmas. This enabled the discovery of a previously unreported 'cryptoproteome' associated with NSCLC and provided evidence of a corresponding humoral response in the form of autoantibodies directed against cancer-associated cryptoproteins. These findings indicate translational potential in the form of candidate markers for early detection or immunotherapy targets.

In contrast to previous studies, which focused on aberrant peptides resulting from genomic alteration [14–16], we pursued an untargeted approach for broad non-canonical protein identification unrestricted by mutational status. The cryptoDB provides a resource of possible protein sequences non-homologous to previously described human proteins that can be detected in multiple samples using standard proteomic techniques. This suggests that increased attention may be paid to the protein-encoding potential of RNAs previously thought to be noncoding, and provides potential insights into where these RNAs may alter cellular physiology [25]. We demonstrate application of the database with statistically significant concordant findings across multiple samples, thereby demonstrating the validity of the approach.

Analysis of quantifiable cryptoproteins in newly diagnosed Ig-bound samples, pre-diagnosed Ig-bound samples, and cell lines revealed that quantified cryptoproteins were predominately derived from protein-coding transcripts (>40%), suggesting that the generation of these cryptoproteins is at the translation level. While much emphasis has been placed on proteins resulting from genetic aberrations in cancer, this finding provides compelling evidence of additional routes to the genesis of detectable neoantigens. For instance, this may partially explain the somewhat limited value of tumor mutational burden in predicting response to cancer immunotherapy [26]. An analysis of cell lines provides evidence for this association with translation errors. Several plausible mechanisms may be posited to account for the generation of a specific cryptoprotein, including a "slippery ribosome," altered nonsense mediated decay, relaxed translational fidelity, or aberrant transcription of ORF-containing pseudogenes [27–30]. An exploration of potential underlying mechanisms related to the generation of our identified cryptoproteins is warranted; however, such investigations are beyond the scope of the current study.

Interestingly, more than half (56%) of the Ig-bound cryptoproteins were exclusively quantified in NSCLC plasmas compared to controls. Moreover, the ability to detect Ig-bound cryptoproteins in pre-diagnostic plasmas provides the potential that these may be sensitive and specific markers of lung cancer risk and presence of disease. These could either serve to complement to existing markers [31–33] or may offer sufficient performance to stand alone as a new source of biomarkers for lung cancer early detection or risk assessment. Future studies, exploring the utility of autoantibodies directed against cancer-associated cryptoproteins alone or in combination with other biomarker types for risk prediction of lung cancer, are warranted.

On the other hand, there are some limitations to our studies. Detailed information regarding full smoking history including smoking duration were not available, thus limiting correlative analyses with smoking exposure. Similarly, the occurrence of cryptoproteins in cancer-associated exosomes or circulating tumor cells was not evaluated [34]. Our study focused on the occurrence of Ig-bound cryptoproteins in plasmas of adenocarcinoma and squamous cell carcinoma lung cancer cases. Whether there is a similar occurrence to be found in small cell lung cancer cases, as well as other NSCLC subtypes, such as large cell carcinoma, remains to be determined.

In conclusion, we establish cryptoproteins as a potential source of neoantigens in NSCLC. Autoantibodies against cancer-associated cryptoproteins are a promising source of biomarkers that may identify individuals at high risk of developing or harboring lung cancer. Future work will include validation in independent datasets, the collection of

additional samples, and the biological confirmation of autoantibody reactivity in plasmas of lung cancer patients.

4. Materials and Methods

Construction of a cryptoprotein proteomics pipeline and database

The fasta format transcript files from Genome Reference Consortium Human Build 38, release 27 (GRCh38.v27), were downloaded from GENCODE. In silico translation was completed for all transcripts including protein coding genes, pseudogenes, noncoding RNAs, including microRNAs (miRNAs) and long noncoding RNAs (lncRNAs), as well as variants including transcripts with retained introns. We selected all open reading frames (ORFs) greater than 50 codons, beginning with an AUG and ending with a canonical stop-codon (UAA, UGA, UAG), and eliminated the largest open reading frame from transcripts annotated as protein coding, which resulted in 1.1 million ORFs. These were then aligned against the human non-redundant (nr) protein database using the BLASTP algorithm. ORFs aligning with an E-score greater than 0.01, indicating a likely successful alignment, were discarded. This yielded 108,863 ORFs with no known homology to human proteins, which we termed the Cryptoprotein Database (cryptoDB, Figure 4). The fasta file of the cryptoDB is available on github (https://github.com/EhsanIrajizad/Cryptoprotein, uploaded April 2021, accessed 10 August 2022).
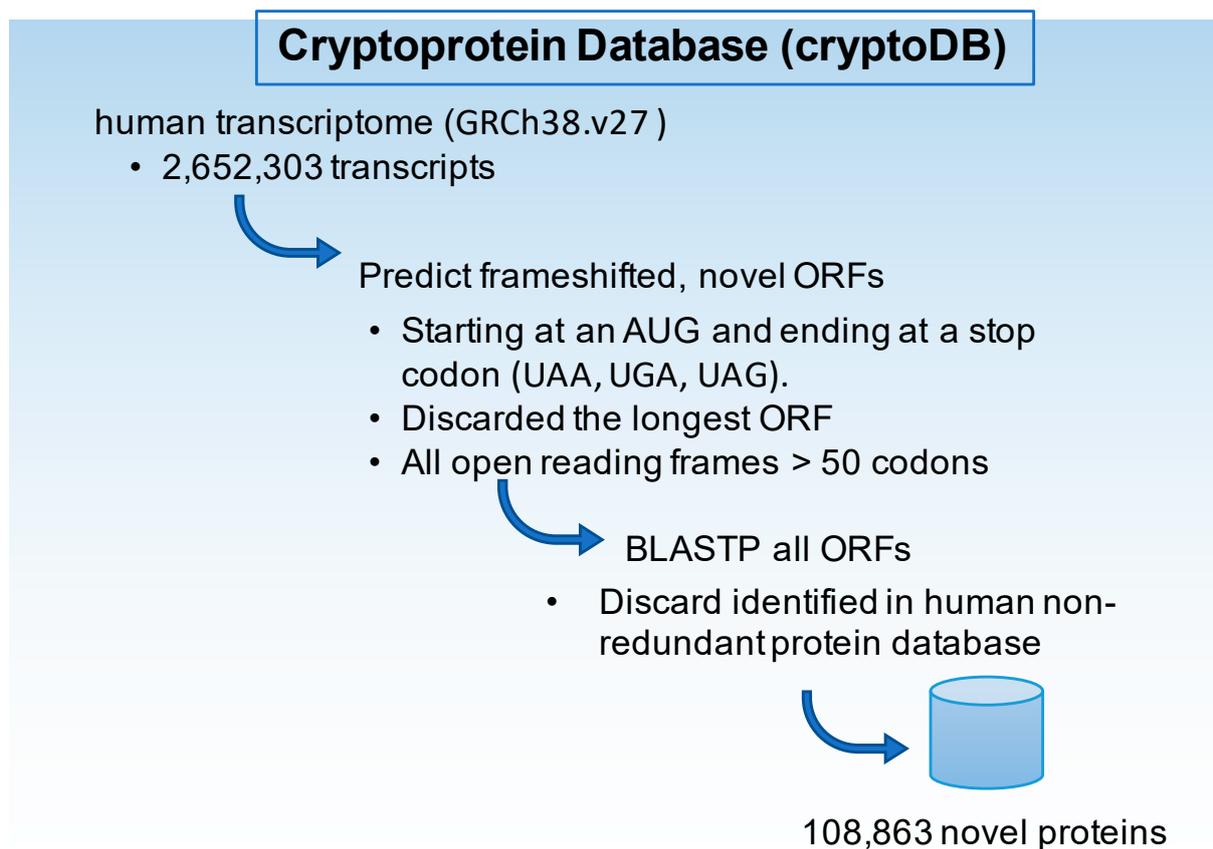

Figure 4. Construction of a cryptoprotein database (cryptoDB).

Cell Lines

Seventeen lung adenocarcinoma cell lines (H2228, H1395, H2405, H522, H969, H1703, H1650, HCC827, HCC4006, H820, HCC2935, H2009, H650, H1795, H2122, H647, HCC4017, Supplemental Table S1) were acquired from American Type Culture Collection. Cell type was confirmed by short tandem repeat analysis. The cell lines were representative of different mutational backgrounds and were analyzed independently. Detailed information regarding mass spectrometry-based analysis of whole cell lysates is described elsewhere [35–39].

Lung cancer plasma collection

Blood samples were collected from two independent cohorts following Institutional Review Board approval and informed consent. One cohort consisted of plasma collected from individuals with newly diagnosed NSCLC at the University of Texas MD Anderson Cancer Center (MDA) (Supplemental Table S2). Another cohort consisted of plasma samples from individuals collected for the Beta-carotene and Retinol Efficacy Trial (CARET) cohort. CARET was a randomized, double-blind, placebo-controlled trial evaluating the cancer prevention efficacy and the safety of daily supplementation with beta-carotene and retinol palmitate in 18,314 individuals at high risk for lung cancer. Participants were enrolled at six US centers and were followed for cancer and mortality outcomes [40]. Six pools of healthy controls (n = 8–28 individuals per pool) were matched to six pools of pre-diagnostic NSCLC cases (n = 4–14 patients per pool) based on age, sex, and smoking history. All controls were followed-up for a minimum of four years to ensure that they were cancer-free (Supplemental Table S3). For the MDA cohort, controls from the CARET cohort were used to compare the distribution of Ig-bound cryptoproteins between cases and controls.

Mass spectrometry analysis of Ig-bound protein complexes

Mass spectrometry-based (MS) analysis of circulating Ig-bound protein complexes was performed as previously described [36,41]. Detailed has been presented in Supplementary Information.

Data Processing of Mass Spectrometry Data

Spectra from proteomic analyses of human lung cancer cell lines and immunoglobulin(Ig)-bound plasma proteins were reprocessed through a customized pipeline based on PeptideShaker [42]. To process only those spectra that did not align to a known UniProt sequence, spectra were first searched against the UniProt Database and spectra identified as UniProt peptides filtered out. Unaligned spectra were subsequently searched against the novel cryptoDB using PeptideShaker [42–46]. We compared the peptide-spectrum match score (PSM) for canonical peptides that matched to the UniProt database to the PSM for cryptopeptides. Peptides were considered a match if they had a false discovery rate (FDR) < 10%, consistent with prior approaches [14]. Furthermore, to reduce false positives, a selection of features in each experiment (cohort) was based on the identified cryptoprotein having a peptide-spectrum match (PSM) $\geq$ 5 and detection in 2 or more samples in each cohort. Cryptoprotein abundance was implied through summation of all aligned spectra.

Ingenuity Pathway Analyses

To identify potential pathway networks associated with cryptoproteins identified in lung adenocarcinoma cell lines, we used host, canonical gene names corresponding to each cryptoprotein, and performed Ingenuity Pathway Enrichment Analysis (IPA). Statistical significance of enriched pathways was determined by two-sided Fisher's Exact Test.

Statistical Analysis

Predictive performance of Ig-bound cryptoprotein complexes was assessed by odds ratios (ORs) using logistic regression for the newly diagnosed cohort and conditional logistic regression for the pre-diagnostic cohort. Analyses were carried out using the R software environment (version 3.6.1, The R Foundation, https://www.r-project.org, accessed 10 August 2022). p values are reported based on two-sided Wilcoxon rank sum test unless otherwise specified.

References

1. Tsou, P.; Katayama, H.; Ostrin, E.J.; Hanash, S.M. The Emerging Role of B Cells in Tumor Immunity. *Cancer Res.* **2016**, *76*, 5597–5601. [CrossRef] [PubMed]
2. Garaud, S.; Buisseret, L.; Solinas, C.; Gu-Trantien, C.; de Wind, A.; Van den Eynden, G.; Naveaux, C.; Lodewyckx, J.-N.; Boisson, A.; Duvillier, H. Tumor-infiltrating B cells signal functional humoral immune responses in breast cancer. *JCI Insight* **2019**, *5*, e129641. [CrossRef] [PubMed]
3. Hanash, S. Harnessing immunity for cancer marker discovery. *Nat. Biotechnol.* **2003**, *21*, 37–38. [CrossRef]
4. Le Naour, F. Identification of tumor antigens by using proteomics. *Methods Mol. Biol.* **2007**, *360*, 327–334. [PubMed]
5. Pereira-Faca, S.R.; Kuick, R.; Puravs, E.; Zhang, Q.; Krasnoselsky, A.L.; Phanstiel, D.; Qiu, J.; Misek, D.E.; Hinderer, R.; Tammemagi, M.; et al. Identification of 14-3-3 theta as an antigen that induces a humoral response in lung cancer. *Cancer Res.* **2007**, *67*, 12000–12006. [CrossRef] [PubMed]
6. Lu, H.; Goodell, V.; Disis, M.L. Humoral Immunity Directed against Tumor-Associated Antigens As Potential Biomarkers for the Early Diagnosis of Cancer. *J. Proteome Res.* **2008**, *7*, 1388–1394. [CrossRef]
7. Babel, I.; Barderas, R.; Díaz-Uriarte, R.; Martínez-Torrecuadrada, J.L.; Sanchez-Carbayo, M.; Casal, J.I. Identification of tumor-associated autoantigens for the diagnosis of colorectal cancer in serum using high density protein microarrays. *Mol. Cell. Proteom.* **2009**, *8*, 2382–2395. [CrossRef]
8. Massoner, P.; Lueking, A.; Goehler, H.; Höpfner, A.; Kowald, A.; Kugler, K.G.; Amersdorfer, P.; Horninger, W.; Bartsch, G.; Schulz-Knappe, P. Serum-autoantibodies for discovery of prostate cancer specific biomarkers. *Prostate* **2012**, *72*, 427–436. [CrossRef]
9. Babel, I.; Barderas, R.; Diaz-Uriarte, R.; Moreno, V.; Suarez, A.; Fernandez-Aceñero, M.J.; Salazar, R.; Capellá, G.; Casal, J.I. Identification of MST1/STK4 and SULF1 proteins as autoantibody targets for the diagnosis of colorectal cancer by using phage microarrays. *Mol. Cell. Proteom.* **2011**, *10*, M110.001784. [CrossRef]
10. Hong, S.-H.; Misek, D.E.; Wang, H.; Puravs, E.; Giordano, T.J.; Greenson, J.K.; Brenner, D.E.; Simeone, D.M.; Logsdon, C.D.; Hanash, S.M. An autoantibody-mediated immune response to calreticulin isoforms in pancreatic cancer. *Cancer Res.* **2004**, *64*, 5504–5510. [CrossRef]
11. Nam, M.J.; Kee, M.K.; Kuick, R.; Hanash, S.M. Identification of defensin α6 as a potential biomarker in colon adenocarcinoma. *J. Biol. Chem.* **2005**, *280*, 8260–8265. [CrossRef] [PubMed]
12. Landry, C.R.; Zhong, X.; Nielly-Thibault, L.; Roucou, X. Found in translation: Functions and evolution of a recently discovered alternative proteome. *Curr. Opin. Struct. Biol.* **2015**, *32*, 74–80. [CrossRef] [PubMed]
13. Ivanov, I.P.; Loughran, G.; Atkins, J.F. uORFs with unusual translational start codons autoregulate expression of eukaryotic ornithine decarboxylase homologs. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10079–10084. [CrossRef] [PubMed]
14. Laumont, C.M.; Daouda, T.; Laverdure, J.-P.; Bonneil, É.; Caron-Lizotte, O.; Hardy, M.-P.; Granados, D.P.; Durette, C.; Lemieux, S.; Thibault, P. Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat. Commun.* **2016**, *7*, 10238. [CrossRef]
15. Chong, C.; Müller, M.; Pak, H.; Harnett, D.; Huber, F.; Grun, D.; Leleu, M.; Auger, A.; Arnaud, M.; Stevenson, B.J. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat. Commun.* **2020**, *11*, 1293. [CrossRef]

16. Bartok, O.; Pataskar, A.; Nagel, R.; Laos, M.; Goldfarb, E.; Hayoun, D.; Levy, R.; Körner, P.-R.; Kreuger, I.Z.; Champagne, J. Anti-tumour immunity induces aberrant peptide presentation in melanoma. *Nature* **2020**, *590*, 332–337. [CrossRef]

17. Yan, Y.; Sun, N.; Wang, H.; Kobayashi, M.; Ladd, J.J.; Long, J.P.; Lo, K.C.; Patel, J.; Sullivan, E.; Albert, T. Whole Genome–Derived Tiled Peptide Arrays Detect Prediagnostic Autoantibody Signatures in Non–Small-Cell Lung Cancer. *Cancer Res.* **2019**, *79*, 1549–1557. [CrossRef]

18. Hanash, S.M.; Ostrin, E.J.; Fahrmann, J.F. Blood based biomarkers beyond genomics for lung cancer screening. *Transl. Lung Cancer Res.* **2018**, *7*, 327. [CrossRef]

19. Qiu, J.; Choi, G.; Li, L.; Wang, H.; Pitteri, S.J.; Pereira-Faca, S.R.; Krasnoselsky, A.L.; Randolph, T.W.; Omenn, G.S.; Edelstein, C. Occurrence of autoantibodies to annexin I, 14-3-3 theta and LAMR1 in prediagnostic lung cancer sera. *J. Clin. Oncol.* **2008**, *26*, 5060. [CrossRef]

20. Pandolfi, P.P. Aberrant mRNA translation in cancer pathogenesis: An old concept revisited comes finally of age. *Oncogene* **2004**, *23*, 3134–3137. [CrossRef]

21. Robichaud, N.; Sonenberg, N.; Ruggero, D.; Schneider, R.J. Translational Control in Cancer. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032896. [CrossRef] [PubMed]

22. Tripathi, S.C.; Peters, H.L.; Taguchi, A.; Katayama, H.; Wang, H.; Momin, A.; Jolly, M.K.; Celiktas, M.; Rodriguez-Canales, J.; Liu, H. Immunoproteasome deficiency is a feature of non-small cell lung cancer with a mesenchymal phenotype and is associated with a poor outcome. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E1555–E1564. [CrossRef] [PubMed]

23. Schliekelman, M.J.; Taguchi, A.; Zhu, J.; Dai, X.; Rodriguez, J.; Celiktas, M.; Zhang, Q.; Chin, A.; Wong, C.-H.; Wang, H. Molecular portraits of epithelial, mesenchymal, and hybrid States in lung adenocarcinoma and their relevance to survival. *Cancer Res.* **2015**, *75*, 1789–1800. [CrossRef] [PubMed]

24. Shepherd, F.A.; Lacas, B.; Le Teuff, G.; Hainaut, P.; Jänne, P.A.; Pignon, J.-P.; Le Chevalier, T.; Seymour, L.; Douillard, J.-Y.; Graziano, S. Pooled analysis of the prognostic and predictive effects of TP53 comutation status combined with KRAS or EGFR mutation in early-stage resected non–small-cell lung cancer in four trials of adjuvant chemotherapy. *J. Clin. Oncol.* **2017**, *35*, 2018. [CrossRef]

25. Ameli Mojarad, M.; Ameli Mojarad, M.; Pourmahdian, A. Long Non-coding RNA snaR Promotes Proliferation in EGFR Wild Type Non-Small Cell Lung Cancer Cells. *Int. J. Mol. Cell. Med.* **2021**, *10*, 258–264.

26. Passaro, A.; Stenzinger, A.; Peters, S. Tumor Mutational Burden as a Pan-cancer Biomarker for Immunotherapy: The Limits and Potential for Convergence. *Cancer Cell* **2020**, *38*, 624–625. [CrossRef]

27. Brierley, I.; Jenner, A.J.; Inglis, S.C. Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.* **1992**, *227*, 463–479. [CrossRef]

28. Dinman, J.D. Slippery ribosomes prefer shapeshifting mRNAs. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 19225–19227. [CrossRef]

29. Lambert, J.-M.; Ashi, M.O.; Srour, N.; Delpy, L.; Saulière, J. Mechanisms and Regulation of Nonsense-Mediated mRNA Decay and Nonsense-Associated Altered Splicing in Lymphocytes. *Int. J. Mol. Sci.* **2020**, *21*, 1335. [CrossRef]

30. Rozov, A.; Demeshkina, N.; Westhof, E.; Yusupov, M.; Yusupova, G. Structural insights into the translational infidelity mechanism. *Nat. Commun.* **2015**, *6*, 7251. [CrossRef]

31. Fahrmann, J.F.; Marsh, T.; Irajizad, E.; Patel, N.; Murage, E.; Vykoukal, J.; Dennison, J.B.; Do, K.-A.; Ostrin, E.; Spitz, M.R.; et al. Blood-Based Biomarker Panel for Personalized Lung Cancer Risk Assessment. *J. Clin. Oncol.* **2020**, *40*, 876–883. [CrossRef] [PubMed]

32. Ostrin, E.J.; Bantis, L.E.; Wilson, D.O.; Patel, N.; Wang, R.; Kundnani, D.; Adams-Haduch, J.; Dennison, J.B.; Fahrmann, J.F.; Chiu, H.T.; et al. Contribution of a Blood-Based Protein Biomarker Panel to the Classification of Indeterminate Pulmonary Nodules. *J. Thorac. Oncol.* **2021**, *16*, 228–236. [CrossRef] [PubMed]

33. Fahrmann, J.F.; Kim, K.; DeFelice, B.C.; Taylor, S.L.; Gandara, D.R.; Yoneda, K.Y.; Cooke, D.T.; Fiehn, O.; Kelly, K.; Miyamoto, S. Investigation of metabolomic blood biomarkers for detection of adenocarcinoma lung cancer. *Cancer Epidemiol. Biomark. Prev.* **2015**, *24*, 1716–1723. [CrossRef] [PubMed]

34. Rossi, E.; Aieta, M.; Tartarone, A.; Pezzuto, A.; Facchinetti, A.; Santini, D.; Ulivi, P.; Ludovini, V.; Possidente, L.; Fiduccia, P.; et al. A fully automated assay to detect the expression of pan-cytokeratins and of EML4-ALK fusion protein in circulating tumour cells (CTCs) predicts outcome of non-small cell lung cancer (NSCLC) patients. *Transl. Lung Cancer Res.* **2021**, *10*, 80–92. [CrossRef] [PubMed]

35. Wang, H.; Hanash, S. Intact-protein analysis system for discovery of serum-based disease biomarkers. In *Serum/Plasma Proteomics*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 69–85.

36. Kobayashi, M.; Katayama, H.; Irajizad, E.; Vykoukal, J.V.; Fahrmann, J.F.; Kundnani, D.L.; Yu, C.-Y.; Cai, Y.; Hsiao, F.C.; Yang, W.-L. Proteome Profiling Uncovers an Autoimmune Response Signature That Reflects Ovarian Cancer Pathogenesis. *Cancers* **2020**, *12*, 485. [CrossRef]

37. Fahrmann, J.F.; Tanaka, I.; Irajizad, E.; Mao, X.; Dennison, J.B.; Murage, E.; Casabar, J.; Mayo, J.; Peng, Q.; Celiktas, M.; et al. Mutational Activation of the NRF2 Pathway Upregulates Kynureninase Resulting in Tumor Immunosuppression and Poor Outcome in Lung Adenocarcinoma. *Cancers* **2022**, *14*, 2543. [CrossRef] [PubMed]

38. Tanaka, I.; Dayde, D.; Tai, M.C.; Mori, H.; Solis, L.M.; Tripathi, S.C.; Fahrmann, J.F.; Unver, N.; Parhy, G.; Jain, R.; et al. SRGN-Triggered Aggressive and Immunosuppressive Phenotype in a Subset of TTF-1-Negative Lung Adenocarcinomas. *J. Natl. Cancer Inst.* **2022**, *114*, 290–301. [CrossRef]

39. Taguchi, A.; Politi, K.; Pitteri, S.J.; Lockwood, W.W.; Faca, V.M.; Kelly-Spratt, K.; Wong, C.H.; Zhang, Q.; Chin, A.; Park, K.S.; et al. Lung cancer signatures in plasma based on proteome profiling of mouse tumor models. *Cancer Cell* **2011**, *20*, 289–299. [CrossRef]

40. Goodman, G.E.; Thornquist, M.D.; Balmes, J.; Cullen, M.R.; Meyskens, F.L., Jr.; Omenn, G.S.; Valanis, B.; Williams, J.H., Jr. The Beta-Carotene and Retinol Efficacy Trial: Incidence of lung cancer and cardiovascular disease mortality during 6-year follow-up after stopping β-carotene and retinol supplements. *J. Natl. Cancer Inst.* **2004**, *96*, 1743–1750. [CrossRef]

41. Capello, M.; Vykoukal, J.V.; Katayama, H.; Bantis, L.E.; Wang, H.; Kundnani, D.L.; Aguilar-Bonavides, C.; Aguilar, M.; Tripathi, S.C.; Dhillon, D.S. Exosomes harbor B cell targets in pancreatic adenocarcinoma and exert decoy function against complement-mediated cytotoxicity. *Nat. Commun.* **2019**, *10*, 254. [CrossRef]

42. Vaudel, M.; Burkhart, J.M.; Zahedi, R.P.; Oveland, E.; Berven, F.S.; Sickmann, A.; Martens, L.; Barsnes, H. PeptideShaker enables reanalysis of MS-derived proteomics data sets. *Nat. Biotechnol.* **2015**, *33*, 22–24. [CrossRef] [PubMed]

43. Vaudel, M.; Burkhart, J.M.; Sickmann, A.; Martens, L.; Zahedi, R.P. Peptide identification quality control. *Proteomics* **2011**, *11*, 2105–2114. [CrossRef] [PubMed]

44. Reiter, L.; Claassen, M.; Schrimpf, S.P.; Jovanovic, M.; Schmidt, A.; Buhmann, J.M.; Hengartner, M.O.; Aebersold, R. Protein identification false discovery rates for very large proteomics data sets generated by tandem mass spectrometry. *Mol. Cell. Proteom.* **2009**, *8*, 2405–2417. [CrossRef] [PubMed]

45. Li, H.; Joh, Y.S.; Kim, H.; Paek, E.; Lee, S.-W.; Hwang, K.-B. Evaluating the effect of database inflation in proteogenomic search on sensitive and reliable peptide identification. *BMC Genom.* **2016**, *17*, 151–162. [CrossRef] [PubMed]

46. Nesvizhskii, A.I. A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. *J. Proteom.* **2010**, *73*, 2092–2123. [CrossRef] [PubMed]