New immunological markers for the identification of lung cancer risk groups among workers in hazardous industries in Kuzbass region

V A Ustinov, M S Nekrash, A E Studennikov and A N Glushkov

Federal State Scientific Institute, Federal Research Centre Coal and Coal Chemistry, Siberian Branch of the Russian Academy of Sciences, Institute of Human Ecology, Leningradsky Pr., Kemerovo, 650065, Russia

E-mail: ustinovVA@ihe.sbras.ru

Abstract. Polycyclic aromatic hydrocarbons are widely known as risk factor in relation to lung cancer development. In this connection, we suggested to use idiotypic and anti-idiotypic antibodies IgG and IgA classes against polycyclic aromatic hydrocarbons in human blood serum as new biomarkers for lung cancer risk. The blood serums of 202 healthy men and 275 men with lung cancer were analyzed by ELISA based on idiotypic and anti-idiotypic antibodies IgG and IgA classes against polycyclic aromatic hydrocarbons. Obtained data were analyzed. It was suggested to use ratio of idiotypic and anti-idiotypic antibodies rather than distinguish level each of them separately. The neural networks for idiotypic and anti-idiotypic antibodies of IgG class (AUC = 0.95), idiotypic and anti-idiotypic antibodies of IgA class (AUC = 0.86), and idiotypic and anti-idiotypic antibodies of IgA and IgG classes (AUC = 0.93) were built as models for lung cancer predictions. Finally, the ELISA data of 52 Kuzbass region coal miners were identified as a group of lung cancer risk using obtained models. So, suggested markers antibodies in human blood serum were not only identified lung cancer patients also elicited group of lung cancer risk among healthy people.

1. Introduction

Polycyclic aromatic hydrocarbons (PAHs) are chemical carcinogens of the environment, which are significant factors in the occurrence of cancer in humans [1]. Getting into the human body they cause production of specific antibodies (Abs). It was assumed that Abs against environment chemical carcinogens modulate their biological properties in the human body and are able to influence the processes of initiation and promotion of malignant transformation of cells [1, 2]. Abs proteomics technology has the potential to become a fundamental technology in drug discovery for development of novel biomarkers and therapeutic targets.

Abs in human biological fluids are using for lung cancer (LC) immunodiagnostic. For example, the measurement of Abs against seven tumor-associated antigens (Ags) by immunoassay was in the early detection of LC [3]. It was evaluated Abs against nine tumorassociated Ags: p62, p16, Koc, p53, Cyclin B1, Cyclin E, Survivin, HCC1, and RalA by ELISA as serological markers in LC [4] or four different Ags were present in non-small cell LC cells in situ [5]. It was successfully identified oxysterol binding protein like 5 and calumenin as potential biomarkers related to metastasis in LC [6]. The accuracy of a panel of proteins and an autoAbs were validated in a population relevant to LC detection and suggested a benefit to combining clinical features with the biomarker results [7]. Other model for LC diagnosis
was built based on the blood biomarkers progastrin-releasing peptide, carcinoembryonic Ag, squamous cell carcinoma Ag, and cytokeratin 19 fragment [8].

Risk factors for the development of LC, such as tobacco smoking along [9-11], gender [12], tobacco smoking and gender [13], age, gender, and smoking status [14], immune alterations [15] have been summarized and integrated into comprehensive models of incidence. Previous studies of risk factor for PAHs status among LC patients have typically considered with air pollutants [16] or tobacco smoking and environmental risk factors at the same time [17].

However, many of these studies did not classify LC cases by Abs levels of idiotypic IgG class (IgG1), anti-idiotypic IgG class (IgG2), idiotypic IgA class (IgA1), and anti-idiotypic IgA class (IgA2) against PAHs in blood serum as LC markers. To further clarify the role of idiotypic and anti-idiotypic Abs against low molecular weight xenobiotics in carcinogenesis we considered them in a group of LC patients to compare with healthy people. Preliminarily we found higher levels Abs against PAHs in LC patients which correlated with benzo[a]pyrene immunized mouse [18]. This study concurred with previous studies that LC was diagnosed by levels of IgG1 and IgG2 class against PAHs in human blood serum [19].

2. Results

We collected blood serums of healthy men (n = 93) and men with LC (n = 191). The conjugate Bp-BSA and mouse idiotypic scFvs against PAHs were immobilized on 96 wells plates to determine of IgA1, IgA2, IgG1, and IgG2 against PAHs in human blood serum by horseradish peroxidase labeled anti-human immunoglobulin A and horseradish peroxidase labeled anti-human immunoglobulin G.

The methods and criteria of modified Shapiro-Wilk were used to assess the normality distribution of the ELISA data besides the analysis of histograms. The analysis of distribution showed that the data belong to a unimodal distribution with a positive asymmetry. Ejection points were more than three standard deviations from the median removed by box-plots analysis. Then the data were analyzed by Z adjusted Mann-Whitney U-test.

The significantly difference in median levels between healthy men and males’ patients with LC was found only in the case of IgG1 (p < 0.001) and IgG2 in blood serum (p < 0.001) (Table 1). However, if use ratio of idiotypic and anti-idiotypic antibodies rather than distinguish level each of them separately, it was significant difference between healthy men and males’ patients with LC in all considered cases. So, we were suggested to measure and use together idiotypic and anti-idiotypic antibodies IgG and IgA classes as LC markers.

Table 1. ELISA data analysis by Z adjusted Mann-Whitney U-test test of healthy donors and LC patients based on median values of levels IgA1, IgA2, IgG1, and IgG2 classes against PAHs in human blood serum and the median values of ratio for the Abs2/Abs1. The medians levels of Abs with the lower and upper quartiles and P values are shown.

| Abs      | Healthy men, median value of Abs (P25:P75) | Men with LC, median value of Abs (P25:P75) | Z adjusted Mann-Whitney U-test for data analysis between groups of healthy people and LC patients, p values |
|----------|-------------------------------------------|-------------------------------------------|-----------------------------------------------------------------------------------------------|
| IgG1     | 0.35 (0.24:0.5)                            | 1.33 (0.7:2.66)                           | p < 0.001                                                                                     |
| IgG2     | 1.25 (0.92:2.14)                           | 2.67 (1.14:5.06)                          | p < 0.001                                                                                     |
| IgA1     | 0.8 (0.55:1.33)                            | 0.8 (0.44:1.47)                           | p > 0.05                                                                                      |
| IgA2     | 1.29 (0.82:2.06)                           | 1.02 (0.43:1.91)                          | p > 0.05                                                                                      |
| IgG2/IgG1| 3.65 (1.86:6.4)                            | 1.35 (0.24:3.8)                           | p = 0.001                                                                                     |
| IgA2/IgA1| 2.45 (0.92:4.67)                           | 1.17 (0.52:2.35)                          | p < 0.001                                                                                     |
| IgA2/IgG1| 3.38 (1.67:10.37)                          | 0.66 (0.23:1.75)                          | p < 0.001                                                                                     |
| IgG2/IgA1| 1.64 (1:3.19)                              | 2.89 (1.21:7.48)                          | p = 0.003                                                                                     |
The neural networks were done using levels of IgG1, IgG2, IgA1, IgA2 classes against PAHs, age, and smoking as predictors. The neural networks were used as a prognostic model for the diagnosis of LC. It was Multi-layer Perceptron classifier. Whole data base of ELISA data was subdivided on test and training groups. The test group was 25% of whole data base and did not included in training of final models. Received training base was used tuition for multi-layer perceptron. Finally, several models were built with different numbers of hidden layers and different amount of neurons in hidden layers. It was chosen the best model by recall score, precision score, and accuracy score, which were got using testing base. The validation fraction was 10% and batch size was 200 samples. The best model was hidden layer sizes = (150, 3), recall score = 0.85, precision score = 0.85, and accuracy score = 0.91. It was used for calculations. The best models (70% train sample size, 10% for validation, and 20% test sample size) calculated 0.95 area under the ROC curve for LC patients in case of using IgG1 and IgG2 (Figure 1A), 0.85 for IgA1 and IgA2 (Figure 1B), and for all together IgG1, IgG2, IgA1, and IgA2 (Figure 1C) with 90-93% positive predictive value for LC patients.
Figure 1. ROC curves for neural networks calculations by IgG1 and IgG2 (A); IgA1 and IgA2 (B); IgG1, IgG2, IgA1, and IgA2 (C) against PAHs in human blood serum based on ELISA data.

Finally, the group of 52 coal miners did not include in early calculation was analyzed to confirm LC predictions using levels of only IgG1 and IgG2 in blood serum as markers, age, and smoking as predictors by model of logistic regression which was statistical method that most closely parallels neural networks. Male patients with LC and healthy male donors used as control groups. The model distributed all the analyzed cases between lung cancer patients and healthy people with 80% sensitivity and 85% specificity. The value of average probability for lung cancer patients (value 0.85) was almost three times higher than for group of healthy males (value 0.29). It was interesting that the levels IgG1 in the group of coal miners was higher than in the groups of healthy men and LC patients. But the levels IgG2 was close to healthy male donors. The value of average LC probability in the group of coal miners by the model of logistic regression was in the middle of healthy donors and LC patients (value 0.59). It means that coal miners differed from healthy people and LC patients. So, it was probably belonged to group of cancer risk (Figure 2).

Than the all values risks of LC probabilities of 52 coal miners were placed on a scale from 0 till 1. Value zero was supposedly accepted for healthy people (absence of disease) and value one for LC patients (presence of disease). According to this analysis, the group of coal miners was divided into three subgroups: 16 coal miners with low risk of LC (value of probability 0-0.4), 13 coal miners with middle risk of LC (value of probability 0.4-0.6), and 23 coal miners with high risk of LC (value of probability 0.6-1) in the examined group. Also, it was seen from individual analysis data of each coal miner that LC risk in this population depended from age. For example, nonsmoking 62 years old coal miner had 0.95 probability of LC risk or nonsmoking 54 years old coal miner had 0.8 probability of LC risk compared to the nonsmoking 38 years old coal miner with 0.34 probability of LC risk or nonsmoking 43 years old coal miner with 0.41 probability of LC risk (data not shown). Definitely this study needs analyzes more blood serums samples of coal miners for further investigation.

![Figure 2](image_url)

Figure 2. Levels of the IgG1 and IgG2 against PAHs in blood serum of healthy people, LC patients, and coal miners. The scale on the ordinate is levels of Abs. Values are presented as the medians ± quartiles.

3. Discussion
In these prospective data, we observed simultaneously analyzes of the levels of IgG1, IgG2, IgA1, and IgA2 against PAHs in human blood serum in men with LC and healthy men. The present research was confirmed our previous data where the production of IgG1 and IgG2 against PAHs in nonimmunized mice was similar to those in healthy persons and the increasing levels the same Abs in immunized mice possessed similar to that in LC patients [18]. Also, the current study concurred with previous studies that LC was diagnosed by only levels of IgG1 and IgG2 against PAHs in human blood serum [19]. Finally built neural networks using levels of IgG1, IgG2, IgA1, IgA2, age, and smoking factors as predictors of LC risks shown high predictive value for LC.

The group of 52 coal miners (only men; the new group did not include in calculation above) were considered by a model that placed coal miners between LC patients and healthy people by the value of average LC probability. All the coal miners who participated in this study did not have LC. Therefore, we concluded that the coal miners belong to the so-called risk of LC group. It was confirmed by already published data that coal miners belonged to risk group of LC [20, 21]. That is why coal miners can have increased levels of IgG1 against PAHs in blood serum. We have broken down the coal miners’ data of values risk LC probabilities got from the logistic regression model test into three groups of coal miners with varying degrees of LC risk: low, medium, and high. It is interesting that the degree of risk (value of probability 0.8-0.9) of coal miners directly depended on the age of the individual. Definitely, this requires further work and analysis of large blood serum samples of coal miners.

So, the results of current manuscript could involve the formation of risk groups and preventive examination of people. The influence of individual factors including industrial and domestic factors on the levels and ratio of Abs1 and Abs2 would be the next subject of the research. Also, it would be interesting to figure out if Abs1 and Abs2 depends on the activity of chemical carcinogens metabolizing enzymes and levels of Abs against estradiol and progesterone in human blood serum (both men and women).

4. Conclusion

Thus, our proposed new biomarkers can be used to predict risk of LC, since the ratio of IgG1, IgG2, IgA1, and IgA2 against polycyclic aromatic hydrocarbons levels in human blood serum are differ in healthy people than in patients with LC. Despite the fact that the levels of IgA1 and IgA2 do not differ in healthy donors and patients with LC, they can nevertheless also be used to predict this disease.

Acknowledgments

Special thank Grebenshchikov I.S. for for assistance in ELISA and data preparation. Also we acknowledge Vavilov V.A. for help in computer processing, data preparation, and for a fruitful discussion of the results and Ivanov V.I. for data preparation. This study was supported by the State Basic Research Program of Russia, Siberian branch Russian Academy of Science (Goszadanie #0352-2019-0011).

5. References

[1] Glushkov A N 2002 Med. Hypotheses 59 501–3
[2] Glushkov A N 2014 Med. Hypotheses 83 166–71
[3] Dai L, Tsay J C, Li J, Yie T A, Munger J S, Pass H, Rom W N, Zhang Y, Tan E M and Zhang J Y 2016 Lung Cancer 99 172–9
[4] Li P, Shi J X, Xing M T, Dai L P, Li J T and Zhang J Y 2017 Tumour Biol. 39 1010428317711662
[5] Guergova-Kuras M, Kurucz I, Hempel W, Tardieu N, Kádas J, Malderez-Bloes C, Jullien A, Kieffer Y, Hincapie M, Guttman A, Csánky E, Dezso B, Karger B L and Takács L 2011 Mol Cell Proteomics 10 M111.010298
[6] Nagano K. 2018 Biol. Pharm. Bull. 41 663–9
[7] Mazzone P J, Wang X F, Han X, Choi H, Seeley M, Scherer R and Doseeva V 2018 Biomark Insights 13 1177271917751608
[8] Yang D, Zhang X, Powell C A, Ni J, Wang B, Zhang J, Zhang Y, Wang L, Xu Z, Zhang L, Wu G, Song Y, Tian W, Hu J-A, Zhang Y, Hu J, Hong Q, Song Y, Zhou J and Bai C 2018 Cancer 124 262–70

[9] Quaife S L, Vrinten C, Ruparel M, Janes S M, Beeken R J, Waller J and McEwen A 2018 BMC Cancer 18 497

[10] Park E R, Ostroff J S, Rakowski W, Gareen I F, Diefenbach M A, Feibelmann S and Rigotti N A 2009 Ann. Behav. Med. 37 268–79

[11] Carter-Harris L, Ceppa D P, Hanna N and Raw S M 2017 Health Expect. 20 59–68

[12] Pinto J A, Vallejos C S, Raez L E, Mas L A, Ruiz R, Torres-Roman J S, Morante Z, Araujo J M, Gómez H L, Aguilar A, Bretel D, Flores C J and Rolfo C 2018 ESMO Open 3 e000344

[13] Cheng T D, Darke A K, Redman M W, Zirpoli G R, Davis W, Payne Ondracek R, Bshara W, Omilian A R, Kratzke R, Reid M E, Molina J R, Kolesar J M, Chen Y, MacRae R M, Moon J, Mack P, Gandara D R, Kelly K, Santella R M, Albain K S and Ambrosone C B 2018 J. Natl. Cancer 110 734–42

[14] Miao J L, Cai J J, Qin X F and Liu R J 2018 BioMed Res. Int. 2018 8398156

[15] Kelsey K T and Wiencke J K 2018 Immunomethylomics: Ann. Am. Thorac. Soc. 15 S76–S80

[16] Lui K H, Bandowe B A, Tian L, Chan C S, Cao J J, Ning Z, Lee S C and Ho K F 2017 Chemosphere 169 660–8

[17] Li X, Wei J, Xu P, Yin X, Hu D, Zhang X, Liu L, Zhang K, Zhou C, Wang T, Zhang X, He M, Wu T, Yang M and Guo H 2015 Mol. Carcinog. 54 E103–E11

[18] Ustinov V A, Matveeva V A, Kostyanko M A and Glushkov A N 2013 Exp. Oncol. 35 207–10

[19] Grebenshchikov I S, Studennikov A E, Ivanov V I, Ivanova N V, Titov V A, Vergbickaya N E and Ustinov V A 2019 Oncotarget 10(49) 5070–81

[20] Taeger D, Pesch B, Kendzia B, Behrens T, Jöckel K H, Dahmann D, Siemiatycki J, Kromhout H, Vermeulen R, Peters S, Olsson A, Brüske I, Wichmann H E, Stücker I, Guida F, Tardón A, Merletti F, Mirabelli D, Richiard L, Pohlabeln H, Ahrens W, Landi M T, Caporaso N, Pesatori A C, Mukeriya A, Szeszenia-Dabrowska N, Lissowska J, Gustavsson P, Field J, Marcus M W, Fabianova E, Mannette J, Pearce N, Rudnai P, Bencko V, Janout V, Dumitru R S, Foretova L, Forastiere F, McLaughlin J, Demers P, Bueno-de-Mesquita B, Schüz J, Straif K and Brüning T 2015 Scand. J. Work Environ. Health 41 467–77

[21] Hosgood H D 3rd, Chapman R S, Wei H, He X, Tian L, Liu L Z, Lai H, Engel L S, Chen W, Rothman N and Lan Q 2012 Am. J. Ind. Med. 55 5–10