UNIVERSITY OF MEDICINE AND PHARMACY
"CAROL DAVILA", BUCURESTI
DOCTORAL SCHOOL
MEDICINE

GENETIC EPIDEMIOLOGY OF LUNG AND COLORECTAL CANCER

PHD THESIS ABSTRACT

PhD supervisor:
PROF. UNIV. DR. STREINU-CERCEL ADRIAN

PhD student:
MATEI ADRIAN

2024

Contents

Introduction

In the 21st century, cancer is a significant social, public health and economic problem. Internationally, it is responsible for almost one in six (16.8%) and one in four (22.8%) deaths from non-communicable diseases, three out of ten premature deaths from non-communicable diseases are caused by cancer (30.3% in people aged 69 years and over), and in 177 out of 183 countries, cancer is one of the top three causes of death for this age group [1].

Cancer not only significantly reduces life expectancy, but also involves significant social and economic costs that differ according to the type, geographical area and gender of the patient [2].

Given the occupancy of each of the top 5 cancers worldwide in 2022 in terms of incidence and mortality according to the International Agency for Research on Cancer (Table 1) [1,3], we considered that addressing lung and colorectal cancer makes an important contribution to the scientific community at the international level.

Table 1. Top 5 cancer incidence and mortality worldwide in 2022 [1,3].

Top	**Cancer localization**	**Incidence**	**Top**	**Cancer localization**	**Mortality**
1	Trachea, bronchi and lungs	2480675	1	Trachea, bronchi and lungs	1817469
2	Sân	2296840	2	Colorectal	904019
3	Colorectal	1926425	3	Liver and intrahepatic bile ducts	758725
4	Prostate	1467854	4	Sân	666103
5	Stomach	968784	5	Stomach	660175

In terms of incidence and mortality by type of cancer in Romania in 2022, the data reported to the International Agency for Research on Cancer coincide for lung and colorectal cancer mortality, but colorectal cancer incidence is in first place (Table 2) [1,3].

Table 2. Top 5 cancer incidence and mortality in Romania in 2022 [1,3].

Top	**Cancer localization**	**Incidence**	**Top**	**Cancer localization**	**Mortality**
1	Colorectal	13541	1	Trachea, bronchi and lungs	10530
2	Sân	12685	2	Colorectal	7381

3	Trachea, bronchi and lungs	11716	3	Sân	3877
4	Prostate	10442	4	Liver and intrahepatic bile ducts	3495
5	Urinary bladder	5157	5	Pancreas	3209

I. General part

1. Lung cancer

1.1. Epidemiology of lung cancer

Lung cancer ranks first in cancer incidence worldwide, with 2.48 million new cases in 2022 and first among cancer deaths with 1.8 million deaths [1].

In 2022, according to the International Agency for Research on Cancer, 104,661 new cases of cancer were diagnosed and 56,216 cancer-related deaths were recorded in Romania. Romania has one of the highest mortality rates in Europe for most cancers [4].

According to a recent study, adenocarcinoma was the most prevalent lung cancer subtype worldwide in 2020, with higher incidence rates than squamous cell carcinoma in most male-dominated countries and in all 185 female-dominated countries [5].

The five-year survival rate for lung cancer is typically less than 20% in most countries [6]with little variation according to the development of society [7].

A study conducted in developed countries on survival differences by stage, histologic subtype and sex showed that treatment-related factors, healthcare systems and degree of comorbidity may play an important role [8].

1.2. Risk factors associated with lung cancer

Lung cancer is associated with alterations in tumor suppressor genes within the genome. The presence of an initiating event, such as a mutation, can alter the tumor suppressor gene by deleting or adding segments of DNA base pairs. In addition, expression of erbB1 and/or erbB2 genes contributes to the inhibition of apoptosis (programmed cell death). Alterations in the expression patterns of these genes lead to the transformation of target cells, especially lung cells, into cancer cells with uncontrolled proliferative capacities [9,10].

These cancers are marked by certain environmental and occupational risk factors such as gender, age, smoking, radon and air pollution [11-15] and, above all, genetic factors.

The main risk factors involved in lung cancer are:

a) Like

b) Age

c) Family history

d) Smoking

e) Radon

f) Air pollution

g) Genetic factors

The relevance of family history of cancer, particularly for people with early-onset lung cancer, has been emphasized by incorporating this information into various lung cancer risk assessment models [16,17].

In particular, smokers with a family history of lung cancer face a two-fold increased risk of developing the disease, and even non-smokers with a family history are at high risk [18].

Studies have identified high-effect genomic associations between rare genetic variants in BRCA2 and CHEK2 and an increased risk of squamous cell lung cancer, which have been described in a European population cohort [19].

One study reported 22 out of 21 genes (including ATM, CXCR2, CYP1A1, CYP2E1, ERCC1, ERCC2, FGFR4, SOD2, TERT and TP53) as having significant associations with lung cancer susceptibility [20].

Genome-wide association studies have identified 18 genetic loci associated with susceptibility to different histologic subtypes of lung cancer [21]. This research has highlighted RNASET2, SECISBP2L and NRG1 as potential candidate genes, in addition to the nicotinic cholinergic cholinergic receptor CHRNA2 and telomere-related genes OFBC1 and RTEL1, demonstrating the genetic complexity underlying the development of this malignancy [21].

Genetic factors implicated in lung cancer susceptibility include cell cycle-related genes such as the tumor suppressor genes p53 and p73, as well as the apoptosis genes FAS and FASL. In addition, biomarkers such as DNA adducts (a piece of DNA covalently linked to a chemical such as safrole, benzopyrenediol epoxide, acetaldehyde) can be used to assess carcinogenicity and lung cancer risk. In addition, epigenetic DNA modifications, including DNA methylation, histone deacetylation, and phosphorylation, have been linked to increased susceptibility to lung cancer by influencing gene expression [22].

1.3. Lung cancer screening

Lung cancer generally has a poor prognosis, as most patients are diagnosed at an advanced stage, resulting in a five-year survival rate of only 13%, with notable regional disparities across Europe. This unfavorable outcome is related to the late identification of the disease, with 50-70% of cases typically presenting at an advanced stage, as opposed to 15-25% detected early [23].

The implementation of screening programs to enable early detection of lung cancer has the potential to reduce mortality rates and improve patient outcomes, particularly among high-risk populations [24].

Although significant research has explored the potential of biomarkers as indicators of lung cancer, their integration into lung cancer screening study protocols remains limited to date [25-28].

1.4. GWAS in lung cancer

In the last ten years, genome-wide association studies in lung cancer have been able to identify more than 45 genomic loci [19,21,29-37]. A large proportion of these loci have been associated with different subgroups correlated with histologic subtypes and smoking status, indicating a complex biological process underlying the oncogenic process of lung cancer.

Biomarkers can be categorized into several groups: those used for diagnosis, disease staging, prognosis and guiding therapeutic decisions. In particular, there is a substantial body of research focused on the early detection of non-small cell lung cancer (NSCLC), while studies to assess and predict therapeutic response have also expanded rapidly [38-40].

Evaluation of circulating tumor cells (CTCs) is of great value in assessing both diagnostic markers and therapeutic efficacy [38].

Evaluation of the expression of CEA, aldosterone, thymidylate synthase, PD-1/PD-L1 and ERCC1 has been shown to be essential in assessing therapeutic efficacy, aligning with the personalized treatment approach for non-small cell lung cancer [38].

Quantification of the expression of the long non-coding RNA SPRY4-IT1 has provided valuable insights into the mechanisms of lung cancer development and prognosis, making it a promising target for future research and investigation [40].

Among the biomarkers identified in different studies, folate receptors expressed on circulating tumor cells, which can be detected in peripheral blood samples, have demonstrated efficacy in distinguishing between benign lung diseases and non-small cell lung cancer (NSCLC). In addition, other biomarkers have shown promise in assessing

therapeutic efficacy, in supporting prognostic assessment and in guiding the selection of the most appropriate therapeutic approach [38].

2. Colorectal cancer

2.1. Epidemiology of colorectal cancer

According to global estimates, in 2022 there were more than 1.9 million new cases of colorectal cancer and an estimated 904,000 deaths attributable to the disease, accounting for about one-tenth of the total global cancer burden [1].

Colorectal cancer ranks 3rd in cancer incidence worldwide with 1.92 million new cases in 2022 and 2nd in cancer deaths with 903,859 deaths [1].

As an indicator of socioeconomic development, colorectal cancer incidence rates have steadily increased in developing countries [41,42], including countries in Eastern Europe, Southeast and South Central Asia and South America [43,44].

2.2. Risk factors associated with colorectal cancer

It is estimated that about 20% of colorectal cancer cases have a family history of the disease. In addition, certain genetic syndromes, such as hereditary non-polyposis colorectal cancer, account for about 3% of colorectal cancers, while conditions such as Gardner syndrome and familial adenomatous polyposis are strongly associated with colorectal cancer and account for about 1% of all cases. However, the predominant factor driving the majority of colorectal cancer cases appears to be environmental influences rather than inherited genetic alterations [45].

The development of colorectal cancer appears to be influenced by a complex set of factors, including dietary habits, lifestyle behaviors and various environmental exposures. Factors such as environmental and food-borne mutagens may contribute to the development of this malignancy [46]. In addition, gut microbiome disorders and subsequent chronic inflammation have been identified as additional risk factors for colorectal cancer, contributing to approximately 20% of all cases [47-49]. The dysregulation of the gut microbiota appears to promote a chronic inflammatory state that may precede and lead to the development of colorectal tumors [47,50].

Approximately 5% of all colorectal cancers are attributed to the most common cancer syndromes called familial adenomatous polyposis and Lynch syndrome [51,52].

This section will examine the various risk factors associated with colorectal cancer:

a) Age
b) Family history and genetics

c) Lifestyle factors

2.3. Colorectal cancer screening

Colonoscopic screening and surveillance has been shown to reduce the risk of colorectal cancer by allowing direct removal of precancerous lesions and early detection of colorectal cancer [53,54].

Studies have demonstrated that the SEPT9 DNA test has a reasonably high sensitivity and specificity for the detection of colorectal cancer [53].

Genetic testing can help determine whether people from certain families have an inherited cancer syndrome that increases their risk of colon cancer [55].

2.4. GWAS in colorectal cancer

Genome-wide association studies have not only identified common genetic variants associated with colorectal cancer susceptibility, but have also highlighted the importance of rare missense variants in disease risk [56].

Genome-wide association studies have further explored the interaction between genetic variation and environmental factors such as alcohol consumption [57].

II. Personal contributions

3. Working hypothesis and general objectives

Genome-wide association studies using large cohorts of Eastern European populations may be a first step towards identifying patients predisposed to oncologic diseases such as lung and colorectal cancer in Romania.

Establishing associations between the population genetic profile and the mechanisms underlying the genetic etiopathogenesis of colorectal and lung cancer is the objective of this thesis.

The association of an infectious disease with an oncologic disease adversely affects the quality of life, number of hospitalizations, complications and prognosis of patients' survival rate.

I mention that there are very few national genetic studies on colorectal and lung cancer at the time of publication of this thesis.

4. General research methodology

Study data come from the ROMCAN (*Genetic epidemiology of Cancer in Romania*) research grant. This project, which was coordinated by the University of Medicine and "Pharmacy Carol Davila", Faculty of Medicine in Bucharest, in collaboration with the

Institute "DeCode Genetics" from Iceland, was carried out between 2014 and 2017 . The project assessed risk factors for prostate, breast, colon, lung and rectal cancers.

The retrospective cross-sectional study included patients from two clinics in Bucharest (Theodor Burghele Urology Clinic and Sfânta Maria General Surgery Clinic) for various medical conditions.

Lung and colorectal cancer patients who were treated at the clinics participating in the project made up the case cohort. Patients without neoplasia constituted the control cohort.

Prior to enrollment in the study, patients gave written informed consent for the use of personal, clinical data and biological samples for genetic research. The study was approved by the Bioethics Commission of the Romanian College of Physicians and the study protocols were approved by the National Ethics Commission of the Romanian College of Physicians.

Face-to-face interviews were conducted by trained interviewers using standardized questionnaires to collect personal data (ethnicity, marital status, education, height and weight), lifestyle data (occupation, smoking, coffee, alcohol and tea consumption) and medical history (personal and family). The patients included in the study were of the following origin

5. Genetic epidemiology of lung cancer

5.1. Introduction

The specific objectives are to identify the markers with the best association results with lung cancer and to identify those common with those present in phenotype associations for respiratory tract infections.

5.2. Material and method

5.2.1. Study population

The study cohort consisted of 2165 hospitalized patients, as follows: 1092 cases of lung cancer confirmed on the basis of pathology reports from surgery such as biopsy or cytology results and using clinical criteria, and 1073 control cases constituted by including patients hospitalized for non-oncological medical pathologies.

Infections were defined using inpatient data based on International Classification of Diseases (ICD)-10 codes. Clinically relevant entities, such as lower respiratory tract infections in the context of chronic obstructive pulmonary disease (COPD), aspiration, obstruction and opportunistic infections caused by immunosuppression, were defined by reviewing all available ICD-10 codes.

5.2.2. Genotyping and data analysis

DNA was extracted using buccal swab samples and analyzed at deCODE Genetics (Reykjavik, Iceland). The genotyping platform used was Infinium OmniExpress-24 bead chips (Illumina). Bioinformatics data were analyzed using Plink! v1.07 [58,59]. 10% of available markers were filtered using the Hardy-Weinberg equilibrium significance threshold of 5×10^{-6} . SNPS with a minor allele frequency below 1% were excluded. Each chromosome was phased using SHAPEIT [59,60]. For imputation we used the 1,000 Genomes dataset [59,61], phase 3 October 3, 2014 used for patients genotyped in 2024 with IMPUTE2 software [59,62].

SNPTEST was used for the association test [63], using a single binary phenotype variable. All *p* values are two-tailed.

5.2.3. Selecting SNPs to replicate previous results

Based on the results of the NHGRI systematic literature review of the NHGRI catalog of variants previously associated with lung cancer in GWAS [64], completed on August 11, 2022, a search query was completed with "lung cancer" as keyword.

After filtering the observed variants, we obtained a final set of 49 unique variants used in our replication.

Infections are still common events in the natural course of cancer. During the course of the disease, lung cancer patients frequently develop an infection that can ultimately be fatal. Infections associated with lung cancer include lower respiratory tract infections in the context of concomitant chronic obstructive pulmonary disease (COPD), aspiration, obstruction and opportunistic infections due to immunosuppression [65].

A search query was performed with "infectious diseases of the respiratory tract" as keyword. A total of 63 markers were initially retrieved from the GWAS catalog based on the keyword search.

5.3. Results

A total of 24,295,558 markers were analyzed. Ten of the variants tested in the Romanian GWAS reached genome-wide statistical significance (*p-value* $< 5 \times 10^{-8}$), and 34 had *p-values* $< 10^{-6}$. Fig. 5.1 shows a Manhattan plot of the results. The strongest associations were observed in chromosomes 5,6,8,10,10,12 and 15.

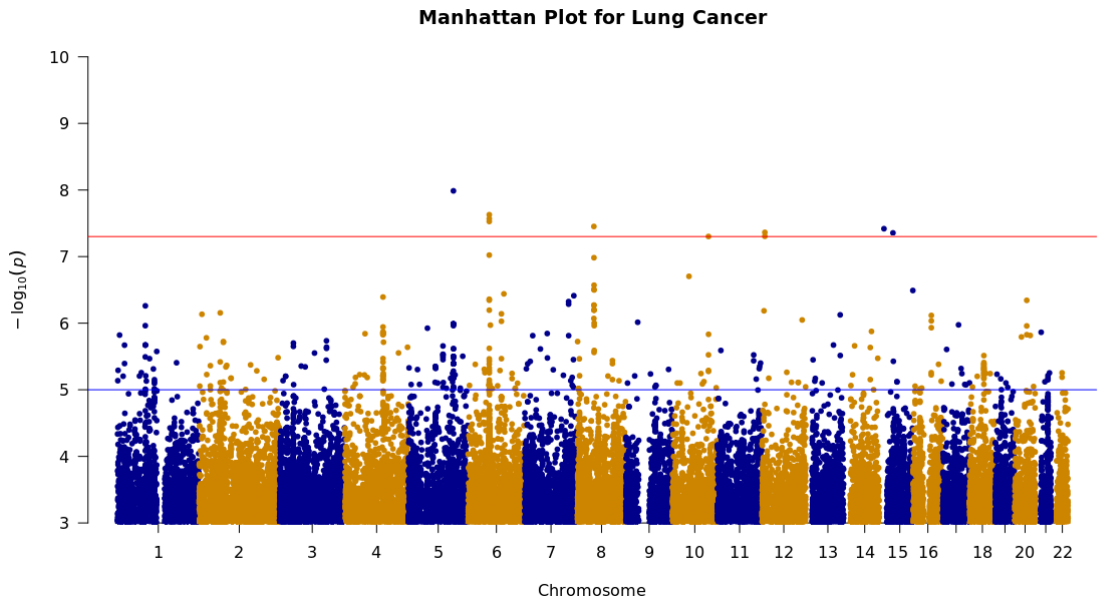

Fig. 5.1. Manhattan plot of -log (*p-value*) values for lung cancer genetic variants. The Y-axis shows the -log p-values10 and the X-axis shows the chromosomal position.

5.3.1. Genetic analysis of SNP RS2808630 - CRP gene variant

The association test revealed a *p-value* of 7×10^{-6} for SNP RS2808630 , which is found at position 159711078 on chromosome 1.

A localized variant after the CRP gene is represented by this marker.

5.3.2. Genetic analysis of SNP RS10849605 - RAD52 gene variant

The association test revealed a *p-value* of 6×10^{-7} for SNP RS10849605, which is found at position 1064438 on chromosome 12.

An intronic variant of the RAD52 gene is represented by this marker.

5.3.3. Genetic analysis of SNP RS8042374 - CHRNA3 gene variant

The association test showed a *p-value* of 8×10^{-12} for SNP RS8042374, which is found at position 78908032 on chromosome 15.

An intronic variant of the CHRNA3 gene is represented by this marker.

5.3.4. Genetic analysis of SNP RS8034191 - HYKK gene variant

The association test revealed a *p-value* of 3×10^{-18} for SNP RS8034191, which is found at position 78806023 on chromosome 15.

An intronic variant of the HYKK gene is represented by this marker.

5.3.5. Genetic analysis of SNP RS481519 - NEK10 gene variant

The association test revealed a *p-value* of 2×10^{-9} for SNP RS481519, which is found at position 27327214 of chromosome 3.

An intronic variant of the NEK10 gene is represented by this marker.

5.3.6. Genetic analysis of SNP RS55781567 - CHRNA5 gene variant

The association test revealed a *p-value* of 1×10^{-9} for SNP RS55781567, which is found at position 78857986 on chromosome 15.

A prime 5 UTR variant of the CHRNA5 gene is represented by this marker.

5.3.7. Genetic analysis of SNP RS748404 - TGM5 gene variant

The association test revealed a *p-value* of 1×10^{-6} for SNP RS748404, which is found at position 43559231 on chromosome 15.

A variant located upstream of the TGM5 gene is represented by this marker.

5.3.8. Genetic analysis of SNP RS402710 - CLPTM1L gene variant

The association test revealed a *p-value* of 4×10^{-6} for SNP RS402710, which is found at position 1320722 of chromosome 5.

An intronic variant of the CLPTM1L gene is represented by this marker.

5.3.9. Genetic analysis of SNP RS13181 - ERCC2 gene variant

The association test revealed a *p-value* of 9×10^{-7} for SNP RS13181, which is found at position 45854919 on chromosome 19.

An intronic variant of the ERCC2 gene is represented by this marker.

5.3.10. Genetic analysis of SNP RS1530057 - RBMS3 gene variant

The association test revealed a *p-value* of 3×10^{-6} for SNP RS1530057, which is found at position 29575463 on chromosome 3.

An intronic variant of the RBMS3 gene is represented by this marker.

We examined the expression position of each of the above genes in lung tissues by comparison with other available tissues, using data that can be found in the public database The Human Protein Atlas version 23.0 [66].

The expression of each gene in different lung tissue samples was analyzed using The Human Protein Atlas version 23.0 [66] and we analyzed the RNA expression of each gene in lung tissues and the results are shown in the table below:

SNPs	Gena	pTPM	nTPM consensus
rs2808630	CRP	0,63	0,4
rs10849605	RAD52	10,02	3,9
rs8042374	CHRNA3	0,08	0,2
rs803434191	HYKK	1,26	2
rs48481519	NEK10	0,69	2,2
rs5555781567	CHRNA5	0,21	0,5
rs74748404	TGM5	0,14	0

rs402710	CLPTM1L	91,99	48,2
rs1313181	ERCC2	9,76	5,6
rs1530057	RBMS3	10,05	12,2

5.4. Discuss

Analyzing the results on loci associated with lung cancer and respiratory tract infections, we have identified a number of important points that are involved in the biological mechanisms and pathways for both pathologies. One interesting marker is RS115602847, an intronic variant of the SPOCK1 gene. The SPOCK1 gene is known to be involved in the proliferation and migration of non-small cell lung cancer cells through Wnt/β-catenin signaling [67].

By citing gene expression databases, we observed that SPOCK1 expression at both protein and mRNA levels was also increased in human NSCLC tissues and cell lines. It is known that activation of the Wnt/β-catenin pathway is suppressed by SPOCK1 inactivation [67]. The Wnt/β-catenin pathway plays an important role in embryonic development and homeostasis of adult tissues [68]. Deregulation of Wnt/β-catenin signaling leads to various cancerous and non-cancerous diseases [69].

Similar studies show that SPARC/osteonectin, cwcv and kazal-like domains proteoglycan 1 (SPOCK1) is a novel transforming growth factor-β1 (TGF-β) target gene that modulates epithelial-mesenchymal transition of lung cancer cells. TGF-β was previously known to be an important inducer of epithelial-mesenchymal transition. We observed that the expression of SPOCK1 in lung cancer tumor tissues is significantly higher comparable to normal lung tissues [70]. SPOCK1 expression was also significantly higher in tumor tissues with metastases than in tumor tissues without metastases [71].

RS553613132 is an upstream variant of the SorCS1 gene, a gene known to be significantly correlated with prognosis in several types of cancer [72]. Recent studies indicate a strong association between SorCS1 and SKAM (stomach adenocarcinoma), TGCT (testicular germ cell tumors), THCA (thyroid carcinoma), UVM (uveal melanoma) and other cancers. The same study shows an increased mutation rate in SorCS1 [73].

Also, a comprehensive gene expression and distribution analysis during carcinogenesis in squamous cell carcinoma of the lung showed a strong association between SorCS1 and the studied phenotype (2.51×10).$^{-4}$

Previous results also suggest an important role for the CHRNA5 region as a genetic risk factor for airflow obstruction strongly correlated with pneumonia phenotype [74]. Identification of a candidate signal for this gene in the Romanian results may provide future insights into the correlation between lung cancer and pneumonia.

Although there are a number of genetically and phenotypically correlated data in Eastern European countries, the study provided valuable information on these two pathologies, which epidemiologic data indicate that more than half of people with lung cancer develop a respiratory tract infection.

The strong replication results validate the findings of the study and highlight genetic associations between these two phenotypes specific to the Romanian population.

Despite the fact that genome-wide association studies are a type of study that has been intensively used in the past for genetic prognostication, the novelty of the study approach is based on the comparison of three factors: the association of two correlated phenotypes with a large genetic database, the lack of similar information in Romanian, and the results of this study may prove useful for the development of genetic prognostic tests.

The study allows us to refine the association results and exclude associations due to differences in the definition of phenotypes between cohorts. Compared to previous studies, replication studies may use cohorts with different ethnic and pathological characteristics.

Differences in ethnic and genetic characteristics lead to different association results and, consequently, markers previously shown to be correlated with a risk variant may not show an association in a population of different ethnicity. The results presented indicate that at least nine genetic variants reach genome-wide statistical significance in the Romanian lung cancer cohort.

5.5. Conclusions

Identified markers that have previously shown an association with respiratory tract infections present in GWAS results may indicate a similar biological mechanism for both pathologies.

The study results validate previously known SNPs associated with lung cancer risk in the Romanian population.

6. Genetic epidemiology of colorectal cancer

6.1. Introduction

In the present study I attempt to identify genetic markers associated with colorectal cancer susceptibility using a genome-wide association study and to evaluate the identified genetic markers in the context of SARS-CoV-2 associated infections in the Romanian population.

6.2. Material and method

6.2.1. Study population

The study cohort consisted of 1645 hospitalized patients, as follows: 576 histopathologically confirmed colorectal cancer cases and 1069 control cases constituted by including patients hospitalized for non-oncologic medical pathologies.

6.2.2. Genotyping and data analysis

DNA was extracted from whole blood, analyzed at deCODE Genetics (Reykjavik, Iceland) and genotyped using Infinium OmniExpress-24 bead chips (Illumina). Genotyping data were filtered using Plink! v1.07 [58,59]. Approximately 10% of genotyped SNPs were removed using a Hardy-Weinberg equilibrium significance threshold of 5×10^{-6} and by excluding markers with a minor allele frequency less than 1%. Before imputation, each chromosome was phased in a single run using SHAPEIT [59,60]. October 3, 2014 phase 3 markers of 1,000 genomes [59,61] were imputed in 2024 individual chips using IMPUTE2 software [59,62]with a posterior probability of 0.9 as a threshold to determine genotypes. The genotype set was tested for population heterogeneity using principal component analysis in ADMIXTURE software [59,75], and the results were consistent with a homogeneous population.

The association test was calculated using SNPTEST [59,63], using a single binary variable as response; all reported *p* values are two-tailed.

6.2.3. Selecting SNPs to replicate previous results

On August 4, 2023, using the NHGRI catalog of published genome-wide association studies [59,64] as a starting point, a systematic literature review of colorectal cancer-associated variants from previous GWASs was completed.

After filtering the results we used a final set of 35 distinct variants for replication.

A secondary systematic literature review was performed using the NHGRI catalog of published genome-wide association studies [59,64] Searching for SNPs associated with COVID-19 symptom measurement as search keywords. For the number of 15 identified common markers, we found 5 markers in the Romanian GWAS results .

6.3. Results

The analysis was performed on a total of 24,295,558 markers. Among all the variants tested in the Romanian samples, two markers have genome-wide significance ($p < 5 \times 10^{-8}$). The Manhattan plot of the results is shown in Fig. 6.1.

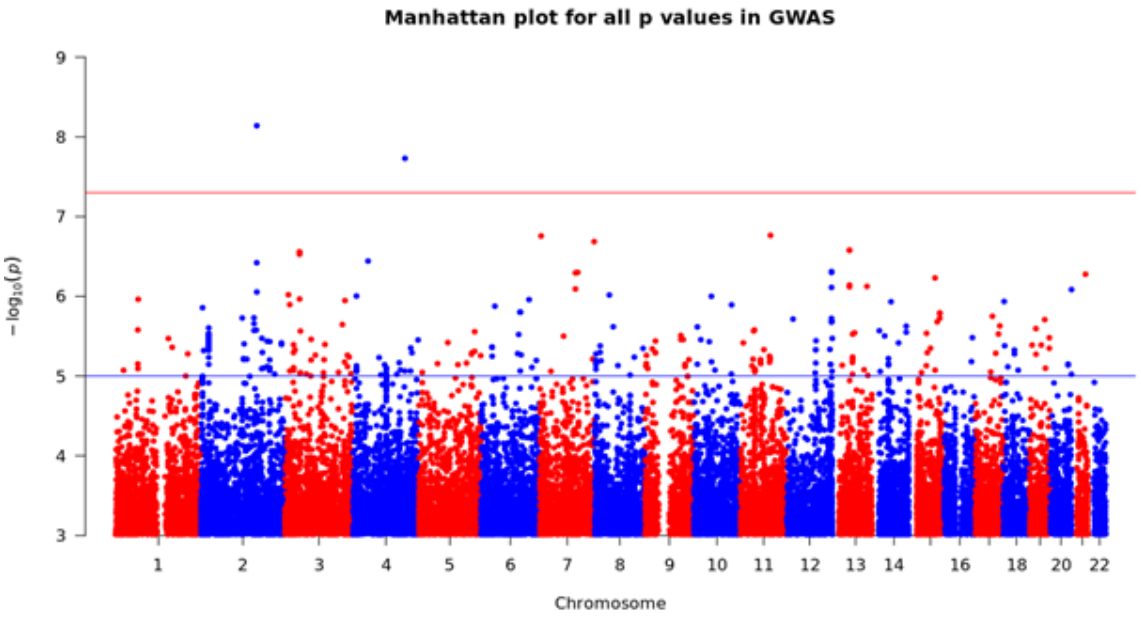

Fig. 6.1. Manhattan plot of -log (p-value) values for genetic variants of colorectal cancer. The Y-axis shows the -log p-values10 and the X-axis shows the chromosomal position.

6.3.1. Genetic analysis of SNP RS10411210 - RHPN2 gene variant

The association test revealed a *p-value* of 5×10^{-9} for SNP RS10411210, which is found at position 33532300 on chromosome 19.

An intronic variant of the RHPN2 gene is represented by this marker.

6.3.2. Genetic analysis of SNP RS73376930 - GREM1 gene variant

The association test revealed a *p-value* of 1×10^{-11} for SNP RS73376930, which is found at position 33012502 on chromosome 15.

An intronic variant of the GREM1 gene is represented by this marker.

6.3.3. Genetic analysis of SNP RS12241008 - VTI1A gene variant

The association test revealed a *p-value* of 1×10^{-9} for SNP RS12241008, which is found at position 114280702 on chromosome 10.

An intronic variant of the VTI1A gene is represented by this marker.

6.3.4. Genetic analysis of SNP RS4939827 - SMAD7 gene variant

The association test revealed a *p-value* of 1×10^{-12} for SNP RS4939827, which is found at position 46453463 on chromosome 18.

An intronic variant of the SMAD7 gene is represented by this marker.

6.3.5. Genetic analysis of SNP RS10936599 - MYNN gene variant

The association test revealed a *p-value* of 3×10^{-8} for SNP RS10936599, which is found at position 169492101 on chromosome 3.

An intronic variant of the MYNN gene is represented by this marker.

We examined the expression position of each gene in tissues of the gastrointestinal tract by comparison with other available tissues, using data that can be found in the public database The Human Protein Atlas version 23.0 [66].

The expression of each gene in different tissue samples of the gastrointestinal tract was analyzed using The Human Protein Atlas version 23.0 [66] and we analyzed the RNA expression of each gene in colon and rectal tissues and the results are shown in the table below:

SNPs	Gena	pTPM	nTPM consensus colon	nTPM consensus rect
rs10411210	RHPN2	12,16	22,9	18,4
rs7333376930	GREM1	11,67	34,9	25,9
rs1224241008	VTI1A	4,3	6,2	7,2
rs493939827	SMAD7	20,45	13,7	13,1
rs10936599	MYNN	20,45	5,5	6,8

6.4. Discuss

In the aftermath of the COVID-19 pandemic, cancer screening programs were reduced or discontinued, creating difficulties in diagnosing cancers, including colorectal cancer [76]. Given this, a different approach involving genetic testing of high-risk individuals after the initial detection of mismatch repair mutations by risk assessment using risk calculators available for unaffected individuals is logically needed [77,78].

In most cancer cases, the impact of Severe Acute Respiratory Syndrome-Co coronavirus (SARS-CoV-2) infections is particularly marked. As a result of the immunosuppressive conditions associated with oncologic disease, most cancer patients are more vulnerable to SARS-CoV-2 infections, showing more severe signs and symptoms than the general population or most patients [79,80].

The effects of COVID-19 on cancer patients have been the subject of several published studies. A previous study demonstrated that because anticancer treatments such as chemotherapy and surgery induce a state of systemic immunosuppression, people with cancer are more susceptible to infection than people without cancer [81,82].

By comparing the Romanian GWAS results for colorectal cancer with previously published markers associated with COVID-19 symptom measures, we were able to identify

several loci that may be useful in the future for the assessment of treatment, prognosis and outcomes.

A particularly interesting finding is the TANK1 gene. Type I interferons, which possess essential antiviral characteristics such as the ability to counter SARS-CoV-2, have previously been identified to be generated largely via the signaling function of this gene [83]. Malignancy-associated infection is one of the most worrisome long-term consequences that could pose a serious threat to public health in the coming years. A number of key mechanisms involved in cancer initiation and progression are affected by SARS-CoV-2 infection [84].

RNA sequencing datasets of the two diseases have been used in previous studies to identify more than 200 common differentially expressed genes. Therefore, investigating the underlying cause of these two diseases is essential [85]. Susceptibility to SARS-CoV-2 infection may be increased in colorectal cancer patients overexpressing the TANK1 gene [86,87].

Also, in different populations, polymorphism of the TANK gene has also been associated with the effects of hepatitis B virus, this gene having been previously studied in relation to the prognosis of different infections and diseases [88].

The HIGD1A gene is another gene that plays a role in the specific suppression of basic mitochondrial genes during SARS-CoV-2 infection. In tissues of SARS-CoV-2 infected patients, there was increased expression of genes such as C2orf69 and HIGD1A, which regulate OXPHOS, as well as mitochondrial MT-CO3, MT-ATP6 and MT-ND6 transcripts. In addition, there was upregulation of genes controlling apoptosis and metabolism [89]. The HIGD1A gene has been evolutionarily conserved and is widely expressed in a variety of tissues. It is involved in the prevention of apoptosis and helps the survival of different cell types in hypoxic environments [90].

In addition, biomarkers that may aid in the prognosis of taste and olfactory impairment due to COVID-19 complications were identified by computational analysis. Previous studies have identified candidate genes that can help screen patients with COVID-19 and olfactory and gustatory impairment based on sophisticated machine learning techniques. CCDC13 was one of the most significant genes with possible interactions between the olfactory system and inflammatory responses caused by viral infection. Genetic variants for this gene were identified in a GWAS study that was performed using data from Romania [91].

It is important to understand that in particular epidemiological situations such as pandemics, reorganization and modification of screening strategies can be beneficial.

COVID-19 has had a significant impact on colorectal cancer screening as well as on how we will see it in the future.

6.5. Conclusions

Data obtained from 576 Romanian colorectal cancer patients revealed several SNPs (mainly located on chromosomes 2 and 3) that corresponded to previously published data (retrieved from GWAS Catalog and dbSNP databases).

We compared the results of GWAS performed on the Romanian population for colorectal cancer with markers known in the literature to be associated with COVID-19 symptom measures. This analysis led to the discovery of potential loci that could be used in future studies to assess future treatment, prognosis and outcomes.

7. Conclusions and personal contributions

7.1. Conclusions

During the PhD thesis two specialized articles were written: "A genome-wide association study in a romanian lung cancer cohort identifies multiple loci associated with susceptibility to respiratory tract infections" and "A genome-wide association study in a romanian colorectal cancer cohort identifies genetic markers associated with susceptibility to SARS-COV-2 infections" which were published in specialized journals.

Upon completion of my PhD thesis I set up a special database and biobank for lung and colorectal cancer research in the Romanian population.

7.2. Personal contributions

We have published two GWAS-type studies investigating lung and colorectal cancer in a Romanian population, as mentioned in the chapter "List of published scientific papers".

By assessing the genetic profile associated with lung and colorectal cancer in a Romanian population, we have identified a set of markers with predictive value. The identified markers and their characteristics for lung cancer are summarized in chapter 5.3 and for colorectal cancer in chapter 6.3.

As a result of establishing the frequency of risk alleles for lung cancer in Chapters 5.3.1 to 5.3.10 and for colorectal cancer in Chapters 6.3.1 to 6.3.5, it will be possible to assess the hereditary risk in families at genetic risk.

We set up a special database and biobank for lung and colorectal cancer research in the Romanian population.

The data can form the basis for the development of national genetic screening programs for lung and colorectal cancer.

The analysis performed within the ROMCAN project led to a genome-wide assessment of a cohort of the Romanian population, representing one of the few GWAS studies conducted at national level and placing our country in the international research community using this type of studies.

Bibliography

1.Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229-63.

2.Chen S, Cao Z, Prettner K, Kuhn M, Yang J, Jiao L, et al. Estimates and Projections of the Global Economic Cost of 29 Cancers in 204 Countries and Territories From 2020 to 2050. JAMA Oncol. 2023 Apr 1;9(4):465-72.

3.Cancer (IARC) TIA for R on. Global Cancer Observatory [Internet] [cited 2024 Jun 11]. Available from: https://gco.iarc.fr/

4.Cancer Today [Internet] [cited 2024 Jun 11]. Available from: https://gco.iarc.who.int/today/

5.Zhang Y, Vaccarella S, Morgan E, Morgan E, Li M, Etxeberria J, Chokunonga E, et al. Global variations in lung cancer incidence by histologic subtype in 2020: a population-based study. Lancet Oncol. 2023 Nov 1;24(11):1206-18.

6.Allemani C, Matsuda T, Carlo VD, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. The Lancet. 2018 Mar 17;391(10125):1023-75.

7.Soerjomataram I, Cabasag C, Bardot A, Fidler-Benaoudia MM, Miranda-Filho A, Ferlay J, et al. Cancer survival in Africa, central and south America, and Asia (SURVCAN-3): a population-based benchmarking study in 32 countries. Lancet Oncol. 2023 Jan 1;24(1):22-32.

8.Araghi M, Fidler-Benaoudia M, Arnold M, Rutherford M, Bardot A, Ferlay J, et al. International differences in lung cancer survival by sex, histologic type and stage at diagnosis: an ICBP SURVMARK-2 Study. Thorax. 2022 Apr;77(4):378-90.

9.Lariscy JT, Hummer RA, Rogers RG. Lung cancer mortality among never-smokers in the United States: estimating smoking-attributable mortality with nationally representative data. Ann Epidemiol. 2020 May;45:5-11.

10.MacRosty CR, Rivera MP. Lung Cancer in Women: A Modern Epidemic. Clin Chest Med. 2020 Mar;41(1):53-65.

11.Ou JY, Fowler B, Ding Q, Kirchhoff AC, Pappas L, Boucher K, et al. A statewide investigation of geographic lung cancer incidence patterns and radon exposure in a low-smoking population. BMC Cancer. 2018 Jan 31;18(1):115.

12.Gaskin J, Coyle D, Whyte J, Krewksi D. Global Estimate of Lung Cancer Mortality Attributable to Residential Radon. Environ Health Perspect. 2018 May;126(5):057009.

13.Hamra GB, Guha N, Cohen A, Cohen A, Laden F, Raaschou-Nielsen O, Samet JM, et al. Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. Environ Health Perspect. 2014 Sep;122(9):906-11.

14.Gharibvand L, Shavlik D, Ghamsary M, Beeson WL, Soret S, Knutsen R, et al. The Association between Ambient Fine Particulate Air Pollution and Lung Cancer Incidence: Results from the AHSMOG-2 Study. Environ Health Perspect. 2017 Mar;125(3):378-84.

15.Sembiring YE, Effendi WI, Dillon JJ, Soebroto H, Winarno DJS, Puruhito P, et al. Lung Cancer: A Literature Review. J Respirasi. 2023 Sep 30;9(3):246-51.

16.Spitz MR, Etzel CJ, Dong Q, Amos CI, Wei Q, Wu X, et al. An Expanded Risk Prediction Model for Lung Cancer. Cancer Prev Res Phila Pa. 2008 Sep;1(4):250-4.

17.Cassidy A, Myles JP, van Tongeren M, Page RD, Liloglou T, Duffy SW, et al. The LLP risk model: an individual risk prediction model for lung cancer. Br J Cancer. 2008 Jan 29;98(2):270-6.

18.Matakidou A, Eisen T, Houlston RS. Systematic review of the relationship between family history and lung cancer risk. Br J Cancer. 2005 Oct 3;93(7):825-33.

19.Wang Y, McKay JD, Rafnar T, Rafnar T, Wang Z, Timofeeva MN, Broderick P, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat Genet. 2014 Jul;46(7):736-41.

20.Wang J, Liu Q, Yuan S, Xie W, Liu Y, Xiang Y, et al. Genetic predisposition to lung cancer: comprehensive literature integration, meta-analysis, and multiple evidence assessment of candidate-gene association studies. Sci Rep. 2017 Aug 21;7(1):8371.

21.McKay JD, Hung RJ, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nat Genet. 2017 Jul;49(7):1126-32.

22.Bade BC, Dela Cruz CS. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. Clin Chest Med. 2020 Mar;41(1):1-24.

23.De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D, et al. Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE--5---a population-based study. Lancet Oncol. 2014 Jan;15(1):23-34.

24.Urgent, Coordinated Global Action on Lung Cancer [World Economic Forum. [cited 2024 Jul 14]. Available from: https://www.weforum.org/publications/urgent-coordinated-global-action-on-lung-cancer/

25.Boutsikou E, Hardavella G, Fili E, Bakiri A, Gaitanakis S, Kote A, et al. The Role of Biomarkers in Lung Cancer Screening. Cancers. 2024 Jan;16(11):1980.

26.Carozzi FM, Bisanzi S, Carrozzi L, Falaschi F, Lopes Pegna A, Mascalchi M, et al. Multimodal lung cancer screening using the ITALUNG biomarker panel and low dose computed tomography. Results of the ITALUNG biomarker study. Int J Cancer. 2017 Jul 1;141(1):94-101.

27.Sullivan FM, Mair FS, Anderson W, Armory P, Briggs A, Chew C, et al. Earlier diagnosis of lung cancer in a randomized trial of an autoantibody blood test followed by imaging. Eur Respir J. 2021 Jan;57(1):2000670.

28.Larose TL, Meheus F, Brennan P, Johansson M, Robbins HA. Assessment of Biomarker Testing for Lung Cancer Screening Eligibility. JAMA Netw Open. 2020 Mar 2;3(3):e200409.

29.Wang Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakidou A, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet. 2008 Dec;40(12):1407-9.

30.McKay JD, Hung RJ, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet. 2008 Dec;40(12):1404-6.

31.Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet. 2009 Nov;85(5):679-91.

32.Wang Y, Broderick P, Matakidou A, Vijayakrishnan J, Eisen T, Houlston RS. Variation in TP63 is associated with lung adenocarcinoma in the UK population. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2011 Jul;20(7):1453-62.

33.Hu Z, Wu C, Shi Y, Shi Y, Guo H, Zhao X, Yin Z, et al. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. Nat Genet. 2011 Jul 3;43(8):792-6.

34.Dong J, Hu Z, Wu C, Guo H, Zhou B, Lv J, et al. Association analyses identify multiple new lung cancer susceptibility loci and their interactions with smoking in the Chinese population. Nat Genet. 2012 Jul 1;44(8):895-9.

35.Li Y, Sheu CC, Ye Y, de Andrade M, Wang L, Wang L, Chang SC, et al. Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. Lancet Oncol. 2010 Apr;11(4):321-30.

36.Lan Q, Hsiung CA, Matsuo K, Hong YC, Hong YC, Seow A, Wang Z, et al. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. Nat Genet. 2012 Dec;44(12):1330-5.

37.Dai J, Lv J, Zhu M, Wang Y, Wang Y, Qin N, Ma H, et al. Identification of risk loci and a polygenic risk score for lung cancer: a large-scale prospective cohort study in Chinese populations. Lancet Respir Med. 2019 Oct;7(10):881-91.

38.Soares Costa AA, Veiga GL da, Alves B da CA, Gascón TM, Pereira EC, Azzalis LA, et al. Lung cancer biomarkers. A literature review. J Bras Patol E Med Lab. 2022 May 30;58:e4152022.

39.Marchetti A, Del Grammastro M, Felicioni L, Malatesta S, Filice G, Centi I, et al. Assessment of EGFR mutations in circulating tumor cell preparations from NSCLC patients by next generation sequencing: toward a real-time liquid biopsy for treatment. PloS One. 2014;9(8):e103883.

40.Sun M, Liu XH, Liu XH, Lu KH, Nie FQ, Xia R, Kong R, et al. EZH2-mediated epigenetic suppression of long noncoding RNA SPRY4-IT1 promotes NSCLC cell proliferation and metastasis by affecting the epithelial-mesenchymal transition. Cell Death Dis. 2014 Jun;5(6):e1298.

41.BW S, CP W. World Cancer Report 2014 [cited 2024 Jul 16]. Available from: https://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports/World-Cancer-Report-2014

42.Fidler MM, Soerjomataram I, Bray F. A global view on cancer incidence and national levels of the human development index. Int J Cancer. 2016;139(11):2436–46.

43.Arnold M, Abnet CC, Neale RE, Vignat J, Giovannucci EL, McGlynn KA, et al. Global Burden of 5 Major Types of Gastrointestinal Cancer. Gastroenterology. 2020 Jul 1;159(1):335-349.e15.

44.Arnold M, Sierra MS, Laversanne M, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality.Gut. 2017 Apr 1;66(4):683-91.

45.Rustgi AK. The genetics of hereditary colon cancer. Genes Dev. 2007 Oct 15;21(20):2525-38.

46.Yang K, Kurihara N, Fan K, Newmark H, Rigas B, Bancroft L, et al. Dietary induction of colonic tumors in a mouse model of sporadic colon cancer. Cancer Res. 2008 Oct 1;68(19):7803-10.

47.Terzić J, Grivennikov S, Karin E, Karin M. Inflammation and colon cancer. Gastroenterology. 2010 Jun;138(6):2101-2114.e5.

48.Chen LW, Egan L, Li ZW, Greten FR, Kagnoff MF, Karin M. The two faces of IKK and NF-kappaB inhibition: prevention of systemic inflammation but increased local injury following intestinal ischemia-reperfusion. Nat Med. 2003 May;9(5):575-81.

49.Hanada T, Kobayashi T, Chinen T, Saeki K, Takaki H, Koga K, et al. IFNgamma-dependent, spontaneous development of colorectal carcinomas in SOCS1-deficient mice. J Exp Med. 2006 Jun 12;203(6):1391-7.

50.Feagins LA, Souza RF, Spechler SJ. Carcinogenesis in IBD: potential targets for the prevention of colorectal cancer. Nat Rev Gastroenterol Hepatol. 2009 May;6(5):297-305.

51.Lotfollahzadeh S, Kashyap S, Tsoris A, Recio-Boiles A, Babiker HM. Rectal Cancer. Treasure Island (FL): StatPearls Publishing; 2024 [cited 2024 Jul 17]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK493202/

52.Stoffel EM, Murphy CC. Epidemiology and Mechanisms of the Increasing Incidence of Colon and Rectal Cancers in Young Adults. Gastroenterology. 2020 Jan;158(2):341-53.

53.Wu Z, Li Y, Zhang Y, Hu H, Wu T, Liu S, et al. Colorectal Cancer Screening Methods and Molecular Markers for Early Detection. Technol Cancer Res Treat. 2020 Dec 23;19:1533033820980426.

54.Doubeni CA, Corley DA, Quinn VP, Jensen CD, Jensen CD, Zauber AG, Goodman M, et al. Effectiveness of screening colonoscopy in reducing the risk of death

from right and left colon cancer: a large community-based study. Gut. 2018 Feb;67(2):291-8.

55.Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. Gastroenterology. 2010 Jun;138(6):2044-58.

56.Chang J, Tian J, Yang Y, Zhong R, Li J, Zhai K, et al. A Rare Missense Variant in TCF7L2 Associates with Colorectal Cancer Risk by Interacting with a GWAS-Identified Regulatory Variant in the MYC Enhancer. Cancer Res. 2018 Sep 4;78(17):5164-72.

57.Song N, Shin A, Shin A, Oh JH, Kim J. Effects of interactions between common genetic variants and alcohol consumption on colorectal cancer risk. Oncotarget. 2018 Jan 6;9(5):6391-401.

58.O'Connell J, Gurdasani D, Delaneau O, Pirastu N, Ulivi S, Cocca M, et al. A General Approach for Haplotype Phasing across the Full Spectrum of Relatedness. PLoS Genet. 2014 Apr 17;10(4):e1004234234.

59.Iordache PD, Mates D, Gunnarsson B, Eggertsson HP, Sulem P, Guðmundsson J, et al. Profile of common prostate cancer risk variants in an unscreened Romanian population. J Cell Mol Med. J Cell Mol Med. 2018 Mar;22(3):1574-82.

60.Delaneau O, Howie B, Cox AJ, Zagury JF, Marchini J. Haplotype Estimation Using Sequencing Reads. Am J Hum Genet. 2013 Oct 3;93(4):687-96.

61.1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015 Oct 1;526(7571):68-74.

62.Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009 Jun;5(6):e1000529.

63.Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet. 2010 Jul;11(7):499-511.

64.Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations.Nucleic Acids Res. 2014 Jan;42(Database issue):D1001-1006.

65.Akinosoglou KS, Karkoulias K, Marangos M. Infectious complications in patients with lung cancer. Eur Rev Med Med Pharmacol Sci. 2013 Jan;17(1):8-18.

66.Tissue expression of CRP - Summary - The Human Protein Atlas [Internet] [cited 2024 Jul 6]. Available from: https://www.proteinatlas.org/ENSG00000132693-CRP/tissue

67.Wang T, Liu X, Tian Q, Liang T, Chang P. Reduced SPOCK1 expression inhibits non-small cell lung cancer cell proliferation and migration through Wnt/β-catenin signaling. Eur Rev Med Med Pharmacol Sci. 2018 Feb;22(3):637-44.

68.Pond KW, Doubrovinski K, Thorne CA. Wnt/β-catenin Signaling in Tissue Self-Organization. 2020 Aug 14;11(8):939. Genes. 2020 Aug 14;11(8):939.

69.Liu J, Xiao Q, Xiao J, Niu C, Li Y, Zhang X, et al. Wnt/β-catenin signaling: function, biological mechanisms, and therapeutic opportunities. Signal Transduct Target Ther. 2022 Jan 3;7(1):3.

70.Miao L, Wang Y, Xia H, Yao C, Cai H, Song Y. SPOCK1 is a novel transforming growth factor-β target gene that regulates lung cancer cell epithelial-mesenchymal transition.Biochem Biophys Res Commun. 2013 Nov 1;440(4):792-7.

71.Chen Q, Yao YT, Xu H, Chen YB, Gu M, Cai ZK, et al. SPOCK1 promotes tumor growth and metastasis in human prostate cancer. Drug Des Devel Ther. 2016;10:2311-21.

72.Callari M, Sola M, Magrin C, Rinaldi A, Bolis M, Paganetti P, et al. Cancer-specific association between Tau (MAPT) and cellular pathways, clinical outcome, and drug response. Sci Data. 2023 Sep 20;10(1):637.

73.Deng Y, Zhang X, Li D, Xu H. Pan-cancer analysis of the oncogenic effect of SorCS1 in human tumors and its correlation with LRP2 protein [Internet]. 2022 [cited 2024 Jun 11]. Available from: https://www.researchsquare.com/article/rs-2186632/v1

74.Ragland MF, Benway CJ, Lutz SM, Bowler RP, Hecker J, Hokanson JE, et al. Genetic Advances in Chronic Obstructive Pulmonary Disease. Insights from COPDGene. Am J Respir Crit Care Med. 2019 Sep 15;200(6):677-90.

75.Alexander DH, November J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009 Sep;19(9):1655-64.

76.Mazidimoradi A, Hadavandsiri F, Momenimovahed Z, Salehiniya H. Impact of the COVID-19 Pandemic on Colorectal Cancer Diagnosis and Treatment: a Systematic Review. J Gastrointest Cancer. 2023 Mar;54(1):171-87.

77.Chen S, Wang W, Lee S, Nafa K, Lee K, Lee J, Romans K, et al. Prediction of germline mutations and cancer risk in the Lynch syndrome. JAMA. 2006 Sep 27;296(12):1479-87.

78.Balmaña J, Stockwell DH, Steyerberg EW, Stoffel EM, Deffenbaugh AM, Reid JE, et al. Prediction of MLH1 and MSH2 mutations in Lynch syndrome. JAMA. 2006 Sep 27;296(12):1469-78.

79.Challen R, Brooks-Pollock E, Brooks-Pollock E, Read JM, Dyson L, Tsaneva-Atanasova K, Danon L. Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. BMJ. 2021 Mar 9;372:n579.

80.Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review. JAMA. 2020 Aug 25;324(8):782-93.

81.Al-Quteimat OM, Amer AM. The Impact of the COVID-19 Pandemic on Cancer Patients. Am J Clin Oncol. 2020 Jun;43(6):452-5.

82.Naimi A, Yashmi I, Jebeleh R, Imani Mofrad M, Azimian Abhar S, Jannesar Y, et al. Comorbidities and mortality rate in COVID-19 patients with hematological malignancies: A systematic review and meta-analysis. J Clin Lab Anal. 2022 Apr 6;36(5):e24387.

83.Ullah TR, Johansen MD, Balka KR, Ambrose RL, Gearing LJ, Roest J, et al. Pharmacological inhibition of TBK1/IKKε blunts immunopathology in a murine model of SARS-CoV-2 infection. Nat Commun. 2023 Sep 18;14(1):5666.

84.Jahankhani K, Ahangari F, Adcock IM, Mortaz E. Possible cancer-causing capacity of COVID-19: Is SARS-CoV-2 an oncogenic agent? Biochimie. 2023 Oct;213:130-8.

85.Song Y, Huang T, Pan H, Du A, Wu T, Lan J, et al. The influence of COVID-19 on colorectal cancer was investigated using bioinformatics and systems biology techniques. Front Med. 2023;10:1169562.

86.Howell R, Clarke MA, Reuschl AK, Reuschl AK, Chen T, Abbott-Imboden S, Singer M, et al. Executable network of SARS-CoV-2-host interaction predicts drug combination treatments. NPJ Digit Med. 2022 Feb 1;5(1):18.

87.Jahanafrooz Z, Chen Z, Bao J, Li H, Lipworth L, Guo X. An overview of human proteins and genes involved in SARS-CoV-2 infection. Gene. 2022 Jan 15;808:145963.

88.Real CI, Lu M, Liu J, Huang X, Trippler M, Hossbach M, et al. Hepatitis B virus genome replication triggers toll-like receptor 3-dependent interferon

responses in the absence of hepatitis B surface antigen. Sci Rep. 2016 Apr 28;6:24865.

89.Guarnieri JW, Dybas JM, Fazelinia H, Kim MS, Frere J, Zhang Y, et al. TARGETED DOWN REGULATION OF CORE MITOCHONDRIAL GENES DURING SARS-COV-2 INFECTION. BioRxiv Prepr Serv Biol. 2022 Feb 22;2022.02.19.481089.

90.Zhu JY, Chen M, Chen M, Mu WJ, Luo HY, Guo L. The functional role of Higd1a in mitochondrial homeostasis and in multiple disease processes. Genes Dis. 2022 Apr 22;10(5):1833-45.

91.Ren J, Zhang Y, Guo W, Feng K, Yuan Y, Huang T, et al. Identification of Genes Associated with the Impairment of Olfactory and Gustatory Functions in COVID-19 via Machine-Learning Methods. Life Basel Switz. 2023 Mar 15;13(3):798.

List of published scientific papers

1. **Matei A**, Iordache P D, Mates D, Rascu S, Radavoi D, Ursu R, et al. A genome-wide association study in a Romanian lung cancer cohort identifies multiple loci associated with susceptibility to respiratory tract infections. Romanian archives of microbiology and immunology, Volume 82, Issue 3, 187-195, 2023. (Chapter 5, pages 34-84)

Indexed in EBSCOhost Academic Search, ProQuest, Index Copernicus, CABBELS Scholarly Analytics, CABI, Google Scholar and CAS databases.

Impact factor: 0.242

Publication link: https://roami.ro/index.php/volume-82-issue-3-july-september-2023/

2. **Matei A**, Iordache P D, Mates D, Rascu S, Radavoi D, Ursu R, et al. A genome-wide association study in a Romanian colorectal cancer cohort identifies genetic markers associated with susceptibility to SARS-COV-2 infections. Romanian archives of microbiology and immunology, Volume 82, Issue 4, 251-257, 2023.

Indexed in EBSCOhost Academic Search, ProQuest, Index Copernicus, CABBELS Scholarly Analytics, CABI, Google Scholar and CAS databases.

Impact factor: 0.242

Publication link: https://roami.ro/index.php/volume-82-issue-4-october-november-2023/