

# **Gene-Set Meta-Analysis to Discover Molecular-Biological Pathways Associated to Lung Cancer**

**Dissertation**

zur Erlangung des humanwissenschaftlichen Doktorgrades in der Medizin  
der Georg-August-Universität Göttingen

vorgelegt von

*Mag. Albert Rosenberger*

aus Wien

Göttingen, 2017

## Betreuungsausschuss

Erstbetreuer/in: *Prof. Dr. Heike Bickeböller,*  
Institut für Genetische Epidemiologie,  
Universitätsmedizin Göttingen  
Georg-August Universität Göttingen

Weitere Betreuer/innen: *Prof. Dr. Dieter Kube*  
Klinik für Hämatologie und Onkologie,  
Universitätsmedizin Göttingen  
Georg-August Universität Göttingen

*Prof. Dr. Tim Friede*  
Institut für Medizinische Statistik  
Universitätsmedizin Göttingen  
Georg-August Universität Göttingen

Tag der mündlichen Prüfung: 7.6.2017

## Affidavit

Here I declare that my doctoral thesis entitled "*Gene-Set Meta-Analysis to Discover Molecular-Biological Pathways Associated to lung cancer*" has been written independently with no other sources and aids than quoted.

Albert Rosenberger

Göttingen, March 2017

## Acknowledgment

On the road to complete this thesis I have been supported by numerous people and tons of chocolate.

I wish to thank my supervisor Professor Heike Bickeböller for her constant support and encouragement. I wish to thank Andrew Entwistle who vamped up the manuscripts into proper English before being published. I wish to thank Melanie Sohns and Stefanie Friedrichs who supported me in accomplishing some of the calculations. I wish to thank the members of my thesis committee, Professor Friede and Professor Kube.

I wish to thank all members of the International Lung Cancer Consortium (ILCCO) for sharing their data, without which the whole project wouldn't have been able to be carried out. I would like to mention especially Christopher Amos and Rayjean Hung for the coordination of ILCCO and their teams for compiling harmonized data sets. I am also indebted to the co-authors for the joint work on my papers.

I wish to thank all participants in the studies included in this work, knowing that it is not something to be taken for granted, to fill questionnaires, undergo extra-examinations or to give blood. To those being diagnosed with lung cancer, I hope you survived and enjoy your life in health.

Finally, I wish to thank my wife and children for not asking why this all took so long until being completed.

## Abstract

During the last two decades the search for risk factors of many human diseases turned into a new direction, since it got feasible to gauge the human genome. Genome-wide association studies (GWAS) were carried out to identify genes or narrow genomic regions responsible for the susceptibility to health problems. Lung cancer is one of these conditions. Lung cancer is a major public health problem - worldwide. During the whole lifetime, one out of 14 men and one out of 17 women will develop an invasive lung or bronchial cancer.<sup>1</sup> Moreover, only one or two out of 5 patients survive the first 5 years after being diagnosed.<sup>2</sup>

The *International Lung Cancer Consortium* (ILCCO), overlapping with the *Transdisciplinary Research in Cancer of the Lung* (TRICL, another international group of cooperating lung cancer researchers) was established in 2004 with the aim of sharing comparable data from ongoing lung cancer case-control and cohort studies. The participating studies are from different geographical areas and ethnicities. On the basis of genomic data being shared within ILCCO, it was possible to identify and verify the existence of genomic risk loci for lung cancer in European population at chromosomes 5p15.33, 6p21-22 and 15q25.<sup>3-10</sup>

However, the applied genome-wide association studies suffer from several drawbacks. One thereof is the ignorance of the complexity of molecular-biological mechanisms. Gene-set analyses methods (GSA) were proposed as complementing approaches in the investigation of the genetic basis of diseases using GWAS results to overcome this disadvantage. These aim to discover a joint association of the markers belonging to the genes of considered biological pathways (denoted as gene-sets (GS)) with a disease of interest. Moreover, even if GSAs based on different but comparable studies successfully identify joint association for the same GS, one cannot simply consider this as replicated findings. The pattern of associations of the concerned markers needs to be taken into account.

I have proposed the quantitative approach META-GSA to combine results from GSAs, respectively  $p_{GS}$ -values of GSs, by incorporating *concordance* of single-marker *association patterns* between studies, relevant for the GS of interest.<sup>11</sup>

This new method has been applied to the data of ILCCO/TRICL. A pathway currently marked as specific to systemic lupus erythematosus was discovered as being significantly implicated in lung cancer. The gene region 6p21-22 in this pathway appears to be more extensively associated with lung cancer than previously assumed.<sup>3,6,8</sup> Given wide-stretched linkage disequilibrium in and around the area *APOM/BAG6/MSH5*<sup>12</sup>, there is currently not enough information or evidence to conclude whether the potential pleiotropy of lung cancer and systemic lupus erythematosus is spurious, biological, or mediated.<sup>13</sup> Further research on this pathway and gene region will be necessary.

## Zusammenfassung

In den letzten vier Jahrzehnten wandte sich die Suche nach Risikofaktoren vieler menschlicher Krankheiten dem menschlichen Genom zu, da es möglich wurde, dieses massenhaft und kostengünstig zu vermessen. Genomweite Assoziationsstudien (GWAS) wurden durchgeführt, um Gene oder enge genomische Regionen zu identifizieren, die für die Ermittlung der individuellen Krankheitsanfälligkeit, z.B. hinsichtlich Lungenkrebs (LK), von Bedeutung sind. Lungenkrebs ist weltweit ein großes Thema des Gesundheitswesens. Im Laufe ihres Lebens werden einer von 14 Männern und eine von 17 Frauen einen invasiven Lungen- oder Bronchialtumor entwickeln.<sup>1</sup> Außerdem überleben nur ein bis zwei von 5 Patienten die ersten 5 Jahre nach der Diagnose Lungenkrebs.<sup>2</sup>

Das *International Lung Cancer Consortium* (ILCCO), das mit der Gruppe *Transdisciplinary Research in Cancer of the Lung* (TRICL) überlappt, wurde 2004 mit dem Ziel gegründet, vergleichbare Daten von laufenden Studien zu Lungenkrebs-Erkrankung zusammen zu bringen. Die teilnehmenden Studien stammen aus verschiedenen geografischen Regionen und umfassen mehrere Ethnien. Auf der Basis aller genomischen Daten des ILCCO, war es möglich, die Existenz von genomischen Risikofaktoren für Lungenkrebs in europäisch stämmigen Populationen auf den Chromosomen 5p15.33, 6p21-22 und 15q25.3-10 zu identifizieren und zu verifizieren.<sup>3-10</sup>

Die angewandten statistischen Auswertemethoden sind jedoch nicht frei von Nachteilen. Einer davon ist das Ignorieren der Komplexität molekularbiologischer Mechanismen in der Genese einer Erkrankung. Gen-Set-Analyse-Methoden (GSA) wurden daher als ergänzende Ansätze zur Datenauswertung von GWAS vorgeschlagen. GSAs zielen darauf ab, eine gemeinsame Assoziation zwischen einer Zielerkrankung und den Markern, die zu den Genen der betrachteten biologischen Mechanismen gehören (im Weiteren als Gen-Sets (GS) bezeichnet), aufzudecken. Aber, auch wenn GSAs, die auf unterschiedlichen, aber vergleichbaren Studien basieren, erfolgreich eine gemeinsame Assoziation für dieselben GSs identifizieren konnten, kann das nicht als Replikation der Ergebnisse angesehen werden. Das Assoziationsmuster der betroffenen Marker muss berücksichtigt werden.

Ich habe den quantitativen Ansatz META-GSA entwickelt, um Ergebnisse von GSAs quantitativ valide zu kombinieren, indem die Konkordanz von Einzelmarker-Assoziationsmustern zwischen den Studien adäquat berücksichtigt wird.

Diese neue Methode wurde dann auf die Daten von ILCCO / TRICL angewandt. Gene, die der Krankheit *Systemischer Lupus Erythematoses* (SLE) zugeordnet werden, wurden auch als signifikant zu Lungenkrebs assoziiert erkannt. Daraus lässt sich folgern, dass die Genregion 6p21-22 mit Lungenkrebs stärker assoziiert ist als bisher angenommen.<sup>3,6,8</sup> Es besteht jedoch ein weit reichendes Kopplungsungleichgewicht dieser Region mit dem Bereich APOM/BAG6/MSH512, der einen bekannten Risikofaktor für Lungenkrebs darstellt. Aus den Daten kann nicht eindeutig geklärt werden, ob die potentielle *Pleiotropie* von Lungenkrebs und SLE scheinbar, biologisch oder vermittelt ist.<sup>13</sup> Weitere Untersuchungen dieser Genregion sind daher notwendig.

## Abbreviations

RR	relative risk
OR	odds ratio
95%-CI	95% confidence interval
ES	enrichment score
NES	normalized enrichment score
PDR	directed reversed p-value
GS	gene-set (of interest)
GS'	remaining genes in the genome
GSA	gene-set analysis
GWAS	genome-wide association studies
pooledGWAS-GSA	Alternative approach to META-GSA: The combining of studies is performed at the level of markers, followed by a single GSA performed on the pooled marker-specific associations.
MtG	annotation of markers to genes
GtP	annotation of genes to gene-sets/pathways
LC	Lung Cancer
SLE	Systemic Lupus Erythematosus
ILCCO	International Lung Cancer Consortium
TRICL	Transdisciplinary Research in Cancer of the Lung
DNA	deoxyribonucleic acid
AUC	area under the curve
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
ALIGATOR	Association List Go AnnoTatOR
GSEA	Gene-Set Enrichment Analysis

## Table of contents

1	Introduction.....	9
1.1	Lung cancer – and its risk factors .....	9
1.2	Genome-wide association studies .....	11
1.3	Drawbacks of genome-wide association studies.....	12
2	Gene-set analysis (GSA).....	14
2.1	Methods for the self-contained null-hypothesis .....	14
2.2	Methods for the competitive null-hypothesis .....	17
3	Summaries.....	19
3.1	META-GSA: A novel method to synthesise statistical evidence of several gene-set analyses.....	19
3.2	Gene-set analysis with respect to lung cancer: A meta-analysis.....	21
4	Discussion .....	22
A	References.....	25
B	Curriculum Vitae.....	32
	Publications .....	33



# 1 Introduction

## 1.1 Lung cancer – and its risk factors

Cancer is a major public health problem all over the globe. About 14.1 million new cancer cases, 8.2 million deaths by cancer and 32.6 million people living with cancer (within 5 years of diagnosis) were counted for 2012 worldwide.<sup>14</sup> In developed countries about 482,600 new male and 241,700 female cases per anno for cancer of the lung or bronchus (LC) have been estimated. This is in men the second largest, respectively in women third largest, cancer incidence.<sup>15</sup> In Germany 34,690 males and 18,810 women have been newly diagnosed with LC in 2013. 29,708 men and 15,140 women died from LC in the same year.<sup>2</sup> During the whole lifetime, one out of 14 men and one out of 17 women will develop an invasive lung or bronchial cancer (estimated for the United States of America, 2010 to 2012).<sup>1</sup> Furthermore, once diagnosed with lung cancer, one has to face a poor prognosis. According to the “Zentrum für Krebsregisterdaten”, 84% of men 79% women will die within the first 5 years after being diagnosed with LC.<sup>2</sup> Due to the combination of high incidence and poor prognosis, LC is the most common cause of cancer death worldwide. However, standardised incidence and prevalence of LC for men is decreasing since the mid of the 1980s, in Germany as in other developed countries. On average between 2003 and 2013 the incidence declined by -1.2% (percentage points) per anno, the prevalence by -1.5%. In contrast, there is a continuously increasing trend in incidence and prevalence for women, since the beginning of the data collection. On average between 2003 and 2013 the incidence raised by +3.1% and the prevalence by +2.7%, each year.<sup>2</sup>

Lifelong tobacco smoking remains the predominant cause of LC, even in former-smokers. It is known today that tobacco smoke is a mixture of more than 5,000 chemicals. It is toxic and carcinogenic.<sup>16</sup> At least 98 of these components are, once inhaled, hazardous for human health. The first relation between smoking and LC was drawn in 1939, considering a German case-series of 96 LC-patients.<sup>17</sup> The first scientific evidence was given by Doll and Hill in 1952, comparing 1.465 LC cases and 1.465 matched controls, but without calculating any excess risk estimates for smokers<sup>18</sup>. By summarizing the results of a total of 287 studies estimating the association between smoking and LC, all published between 1950 and 1995, an overall relative risk of RR=5.50 (95%-CI 5.07-5.96) was estimated for (ever-) smokers and an RR=8.43 (95%-CI: 7.63-9.31) for current smokers, each compared to never-smokers. The relative risk was stronger for squamous LC (current smoking RR=16.91, 95%-CI: 13.14-21.76) than adeno LC (RR=4.21, 95%-CI: 3.32-5.34).<sup>19</sup> LC only occurs in one out of 10 heavy smokers.<sup>19</sup>

The rate of male smokers in Germany has dropped since the 1990s (1992: 37% of men smoked; 2013: 29% of men smoked).<sup>20</sup> Even more, for German male smokers the 30-day prevalence of smoking more than 20 cigarettes a day fell from 27% to 2% between 1980 and 2013. Both changes act as the main explanation for the reduction of incidence and prevalence of LC. In contrast, about one out of 5 German women is smoking, without any change

of this rate in the last decades. (1992: 22% of women smoked; 2013: 20% of women smoked). There also exists a negative trend in the amount of smoked cigarettes a day in German women (as in men). For female smokers the 30-day prevalence of smoking more than 20 cigarettes per day fell from 16% to 1.4% between 1980 and 2013.

It is estimated that 10–29% of LC cases are attributable to factors other than smoking, representing between 16,000 and 24,000 LC deaths annually in the USA alone.<sup>2-4</sup> Exposure to radon, a naturally occurring radioactive gas, is regarded as the second frequent cause of LC. About 10% of LC cases are attributed to it.<sup>21,22</sup> Exposure to other environmental factors, like exhaust fumes of diesel engines, arsenic, asbestos, crystalline silica and some more chemical substances were found to be further risk factors for LC.<sup>21,23-25</sup> An increased risk of developing LC has also been observed in patients with other diseases, such as COPD, pneumonia, tuberculosis, or the autoimmune disorder *systemic lupus erythematosus* (SLE).<sup>26,27</sup> To my knowledge, the only identified factor lowering the risk for LC is the intake of soy food or soy products.<sup>28</sup>

Other lifestyle factors had been discussed to lower the risk of LC, these are e.g. physical activity, consumption of fruits, red meat, tea, beer and wine or supplements of vitamins or minerals.<sup>29-31</sup> With the exception of some particular subgroups, like non-smoking women drinking black tea every day, for none of these factors a convincing evidence of preventing lung cancer in general is given. Although relative risk estimates from meta-analyses of RR=0.9 or similar are reported, the observed associations between lung cancer and dietary factors or physical activity are hard to disentangle from cigarette smoking.<sup>32-36</sup>

Familial aggregation of lung cancer was identified as a further risk factor, already discovered before the turn of the century<sup>37-42</sup>, persistent even when corrected for smoking.<sup>5-8</sup> This indicates the existence of a genetic component which is relevant in the aetiology of LC. Today (27<sup>th</sup> February 2017) PubMed lists 222,009 scientific publications assigned to the key "Lung neoplasms/genetics", some dating back to the 1960s. In one of the early enlisted publications, "*a negro family is presented in which all four of eight siblings older than 50 have developed carcinoma of the lung. Attention is drawn to the possible genetic and environmental factors which are etiologically related to cancer of the human lung.*"<sup>43</sup> Although at this time no facilities for genotyping were available, it was concluded that "*the inter-relationship of smoking and genetic factors may explain the very high incidence of lung carcinoma in this family*". Investigations from Germany showed a 2.6-fold increased lung cancer risk in young patients (OR, 95% CI 1.6–6.0) if a first degree relative had been diagnosed with LC cancer<sup>44</sup> and a 5.6-fold increased risk (OR, 95% CI 0.7-46.9) if a parent or sibling was affected with LC at young age, too.<sup>45</sup> Even for non-smokers aged 40 to 59, an increase of the lung cancer risk up to 6-fold was seen in the presence of lung cancer in a first-degree relative.<sup>46</sup>

In 1990 the results of a segregation analyses performed on 337 families, each ascertained through a lung cancer patient indicated that early onset LC can be caused by a Mendelian codominant inheritance of a rare major autosomal gene. Segregation at this putative locus

could account for about 2/3 of the cumulative incidence of lung cancer in individuals up to age 50.<sup>47,48</sup> It was not until 14 years later that the first evidence for linkage of a lung cancer susceptibility locus was published. The putative locus was localized to a region on 6q23–25.<sup>49</sup>

This provides evidence that genetic factors contribute to the susceptibility of LC. During the last decade several genetic variants have been identified as associated to lung cancer or to a specific histological sub-type by genome-wide association studies (GWASs), candidate gene or pathway studies. Genomic risk loci in European population were identified at chromosomes e.g. 5p15.33, 6p21-22 and 15q25.<sup>3-10</sup> Most of these variants could only be detected by combining several GWASs meta-analytically within the International Lung Cancer Consortium (ILCCO) or the Transdisciplinary Research in Cancer of the Lung (TRICL), and hence increasing the sample size.

## 1.2 Genome-wide association studies

In 1909 Wilhelm Johannsen coined the word “*gene*” to label the Mendelian unit of heredity. He also introduced the terms “*genotype*” and “*phenotype*” to discriminate between an individual genetic traits and its physically or mental appearance. But only after in 1953, Watson and Crick reached their ground-breaking conclusion that the *deoxyribonucleic acid molecule* (DNA) exists in the form of a three-dimensional double helix, the basis to investigate the genomic contribution to human life was given. Since then, medical and biologic scientists turn their attention to the human DNA and its differences between individuals or populations. The so-called “genomic variations” in the human genome can be of different forms, including single nucleotide polymorphisms (SNPs) or substitutions, tandem repeats, insertions or deletions (indels) or copy number variations (CNVs) or other chromosomal rearrangements. Genetic variations can be of diverse sizes; from single nucleotides to several mega bases. Owing to their inherent features, variations of larger size like e.g. tandem repeats have first been used in linkage studies, aiming to locate the chromosomal regions harbouring the mutations or genes for monogenic or familial disorders or quantitative with high penetrance traits.<sup>50</sup> The drawback of such family-based approaches can mainly be attributed to their low statistical power, when several genomic variations, each with small effect sizes and/or low penetrance, contribute to the heritability of the trait of interest (e.g. a disease status). In such a situation family-based approaches would require impractically large sample sizes in order to detect genetic risk factors.<sup>51</sup> Most diseases are believed to have such a complex genetic architecture.<sup>52</sup>

Starting with the decoding of the human genome sequence in 2003 and the availability of high-throughput genotyping facilities, the focus of genetic epidemiologists shifted towards genetic variations commonly occurring in the entire population, rather than in selected families. They also started to investigate the whole sequence of the DNA, rather than pre-selected candidate genes. “*Association studies using common allelic variants are cheaper and simpler than the complete resequencing of candidate genes, and have been proposed as a powerful means of identifying the common variants that underlie complex traits. In their*

*simplest form, association studies compare the frequency of alleles or genotypes of a particular variant between disease cases and controls.*"<sup>52</sup> For association studies typically log-linear models were assumed, which are robust against the majority of all possible genetic models. There is also no need to make assumptions about the genomic location of the causal variants. Association studies make use of the principle of linkage disequilibrium (LD) at the population level. LD simply quantifies the probability of the alleles of two loci (e.g. disease and marker loci) to appear on the same gamete more often together than expected by random. One of the simplest measures of LD is:

$$D = p_{AB} - p_A \cdot p_B$$

where  $p_A$  is the frequency of allele A at the first locus,  $p_B$  is the frequency of allele B at the second locus, and  $p_{AB}$  is the frequency of the haplotype AB (the joint presence of A and B).<sup>53</sup>

Put simply, LD is the non-random association of alleles of loci. *"When evolutionary forces can be ignored, including marker and disease locus mutation, any decay in disequilibrium is due solely to recombination"*<sup>53</sup> and hence applies to adjacent markers. This assumption is necessary to relate an observed association between a marker locus and a trait of interest (e.g. a disease status) to a causal genomic disease locus in LD to the marker locus. However, a useful level of LD between two loci is unlikely to extend beyond an average distance of 3 kb, aside from some genomic regions with extended LD-patterns.<sup>54,55</sup>

A genome-wide association study (GWAS) is defined as an association approach that surveys most of the genome for causal genetic variants<sup>52</sup>. This is accomplished by genotyping the DNA of each study participant at nowadays at least 300,000 loci. Common SNPs, defined by a minor allele frequency >5%, are usually investigated. The density of the SNPs chosen as genomic markers is intended to be high enough to contain (or cover) almost all the genome by LD.<sup>56,57</sup> Association to the trait of interest is usually estimated and statistically tested marker by marker. *"The genome-wide association approach therefore represents an unbiased yet fairly comprehensive option that can be attempted even in the absence of convincing evidence regarding the function or location of the causal genes."*<sup>52</sup> With advances in genotyping technologies and the assembling of millions of SNPs to a reference sequence of the human genome, GWAS became affordable and popular for the investigation of genomic risk factors for common complex diseases.<sup>58,59</sup>

### **1.3 Drawbacks of genome-wide association studies**

GWASs provide the opportunity to identify single markers or narrow genomic regions which are associated to a disease using genotypes of thousands of SNPs throughout the whole genome. However, these analyses are not free of drawbacks. Two major aspects are a) missing heritability and b) low predictive ability.

Discussing the performance of published GWAS in 2010, Eichler et al. stated: *"Although recent genome-wide studies have provided valuable insights into the genetic basis of human disease, they have explained relatively little of the heritability of most complex traits, and the*

variants identified through these studies have small effect sizes.”<sup>60</sup> This drawback, noted as *missing heritability*, has been related to a variety of aspects in the design and analysis of GWAS. Issues like incomplete coverage; disregard of rare variants or CNVs; genes that map to regions of copy-number polymorphic (CNP) duplications, opposing effects of selection, population history, migration and mutation rates or population stratification have been deemed to cause missing heritability.<sup>56,60</sup> A further weaknesses is the often inadequate or completely neglected modelling of epigenetic effects, gene x environment interaction or gene x gene interaction. But exactly such interdependencies can be expected, taking into account the complexity of molecular-biological mechanisms. Yet, they require very large sample sizes to be discovered.<sup>61-63</sup>

Given the limited ability to identify susceptibility loci, it is not surprising that GWAS results are often not sufficient to distinguish between individuals with low and high genetic disease risk (noted as *low predictive ability*).<sup>64</sup> For breast cancer the “*Gail model for prediction*” achieves an AUC (area under the curve) of 58%, without regarding known genetic risk factors. After incorporating 10 relevant genetic variants identified by GWAS the AUC only increased to 61.8%. For prostate cancer an AUC of 86.2% for the prediction with the PSA (prostate specific antigen) alone was reported. After adding 33 genetic variants into the model the AUC increased only to 87.2%. For lung cancer a prediction model, mainly based on smoking history in a Chinese population, reached an AUC of 61.9%, after adding 5 relevant SNPs the AUC increased to only 63.9%.<sup>65</sup>

It was pointed out that GWAS are neither intended to explain all genetic variation nor to find appropriate prediction models, but to observe associations between single loci and complex traits.<sup>66</sup> However, the interplay of genes in the aetiology of the considered phenotype remains unconsidered in GWAS, but can be important regarding the biological nature of the trait.<sup>62</sup> The joint consideration of molecular-biologically meaningful sets of markers, respectively genes, has therefore been proposed as an additional approach to reveal genetic risk factors or pinpoint to involved molecular mechanisms. A further part of the “missing heritability” might be explained in this way. Simultaneously considering related markers can provide a boost of power and uncover genes that are relevant in the aetiology of a disease but with low effects. To jointly analyse of several markers, respectively genes, allow researchers to better explore the multifaceted genetic architectures of complex diseases.

## 2 Gene-set analysis (GSA)

Several methods for gene-set analysis (GSA) were proposed as complementing approaches to the investigation of the genetic basis of diseases using GWAS results or including hundreds of markers in single, very large statistical models.<sup>67-73</sup> All approaches have been developed to investigate pre-specified biological pathways, gene networks or gene families (further denoted as gene-sets (GS)). The annotation of markers to genes (MtG), respectively genes to gene-sets/pathways (GtP) can be obtained from public accessible data bases. For example dbSNP<sup>58,59</sup> or ENSEMBL<sup>74</sup> can be used for MtG assignment; KEGG<sup>75,76</sup>, GO<sup>77</sup> or DAVID<sup>78</sup> can be used for GtP assignment.<sup>68</sup>

The GSA approaches can be grouped according to the null-hypothesis being tested.<sup>70</sup> In the context of GWAS these are:

Self-contained null-hypothesis (Q1) — The genes in a gene-set are not associated with the disease phenotype;

Competitive null-hypothesis (Q2) — The genes in a gene-set show the same magnitude of associations with the disease phenotype compared with genes in the rest of the genome.

### 2.1 Methods for the self-contained null-hypothesis

Some methods for testing the self-contained null-hypothesis (Q1) need the original genotype data as input and build a common statistical model for all markers of a considered GS. They are time- and storage- intensive. Comprehensive comparisons of these methods outline the pros and cons, but without pointing to one approach as the overall best.<sup>68,69,71,79</sup> Two methods are exemplarily introduced here, to demonstrate the methodical challenges in the way to define such a common model:

The *Network-Based Kernel Machine Test*<sup>80</sup> makes use of a semi-parametric logistic regression model for the probability of being a case, including all markers belonging to genes of a considered GS. Genetic effects are modelled non-parametrically. Environmental effects are modelled parametrically. The fitted model is of the form:

$$\text{logit}(p(y_i = 1)) = \mathbf{x}_i^T \boldsymbol{\beta} + h(\mathbf{z}_i)$$

where  $y_i$  is the case-control indicator (control:  $y_i = 0$ , case:  $y_i = 1$ ) for  $i = 1, \dots, n$  individuals. The vector  $\boldsymbol{\beta}$  represents the intercept and regression coefficient terms related to the environmental covariates  $x_i$  for the  $i^{\text{th}}$  individual. The variable  $z_i$  denotes the genotype vector of selected marker for the  $i^{\text{th}}$  individual.

The nonparametric, unknown centred smooth function  $h$  describes how the risk of being affected by the disease depends on the observed genotypes and can take the form:

$$h(\mathbf{z}_i) = \sum_{j=1}^n \alpha_j K(\mathbf{z}_i, \mathbf{z}_j)$$

where  $\mathbf{K}$  (the so called kernel) can be understood as measuring the similarity between the individuals  $i$  and  $j$  based on their genotypes.

“... [F]or GWAS, the linear kernel  $K(z_i, z_j) = z_j^T z_i$ , was probably the most frequently applied kernel. Using this kernel in the logistic kernel machine test, is equivalent to using a logistic regression with a linearly defined random effect for all SNPs”.<sup>81</sup> This kernel is a special case of the more general defined  $d$ -th polynomial kernel:

$$K(z_i, z_j) = (z_j^T z_i + \rho)^d$$

where  $\rho$  and  $d$  are tuning parameters.<sup>82</sup> However, both kernels fail in case of gene-gene interaction within the GS. Alternative kernels have been defined and successfully applied in the context of GWAS<sup>80,83-85</sup>. For instance the identity-by-state (IBS) kernel rest upon the proportion of alleles shared between two individuals  $i$  and  $j$ . The IBS kernel is defined as:

$$K(z_i, z_j) = \sum_{l=1}^{n_m} \frac{2I(z_{il} = z_{jl}) + I(|z_{il} - z_{jl}| = 1)}{2n_m}$$

where  $I$  denotes an indicator function taking the values 0 or 1 and  $n_m$  is the number of markers belonging to GS. This kernel has been shown to be more robust in case of non-linearity of genotype effects than the linear kernel.<sup>85</sup>

For the upper mentioned *Network-Based Kernel Machine Test*<sup>80</sup> the matrix  $\mathbf{K}$  is built as product of the observed genotypes, the MtG assignment and a weighting of genes according their importance within the GS. The *network-based kernel* is defined as:

$$K(z_i, z_j) = \mathbf{Z} \mathbf{A} \mathbf{N} \mathbf{A}^T \mathbf{Z}^Z$$

where  $\mathbf{Z}$  is a matrix containing the genotype data per individual (coded in trinary fashion – 0,1 or 2; respectively in-between values for imputed genotypes; dimension: number of individuals times number of markers in GS) and  $\mathbf{A}$  is a matrix containing the MtG assignment (dimension: number of markers in GS times number of genes in GS). The network structure of a considered GS is converted to an undirected adjacency matrix  $\mathbf{N}$  (dimension: number of genes in the path - squared) with all diagonal elements equal 1, due to “self-interaction”. Any other element of  $\mathbf{N}$  represents the interaction of a pair of genes within the GS, where 1 represents activation and –1 represents inhibition. For the construction of  $\mathbf{N}$  the database KEGG can be used. Thus the *network-based kernel* incorporates external knowledge of the biological mechanisms within the GS of interest into the GSA. However, the matrices to build  $K(z_i, z_j)$  can get very bulky for large gene-sets which comprise hundreds of markers, in particular for a GWAS with a large sample size.

As alternative approach, Chen et al. proposed a “*gene-set ridge regression in association studies* (GRASS)” algorithm.<sup>86</sup> In order to reduce the amount of data finally used for modeling, the genetic variation of all SNPs belonging to a single gene is decomposed by applying principal component analysis. Thereby, orthogonal *eigenSNPs* per gene are generated, but

only those that explain 95% of the genetic variation of a gene will further be included in the statistical model. Hence the variable  $\mathbf{z}_i$  denotes here the vector of *eigenSNPs* of the  $i^{\text{th}}$  individual. In contrast to the method above, GRASS utilises a logistic fully-parametrical regression model:

$$\text{logit}(p(y_i = 1)) = \mathbf{x}_i^T \boldsymbol{\beta}_1 + \mathbf{z}_i^T \boldsymbol{\beta}_2$$

but the estimated parameters  $\boldsymbol{\beta}_2$  are constrained under the usually penalty function of a *ridge regression*:

$$\|\boldsymbol{\beta}_2\|_\gamma = \left( \sum_{j=1}^p |\beta_j|^\gamma \right)^{\frac{1}{\gamma}}$$

with  $0 < \gamma < 1$  and  $p$  the number of parameters  $\beta_j$ .

The second methods, as all others testing  $\boldsymbol{\beta}_2$  in the manner of Q1, may be invalid if e.g. the assumptions of the models are not fulfilled. This might be the case if variables (e.g. markers) present a multi-collinear structure (strong LD).<sup>87</sup> They can also lead to false positive findings, e.g. if not sufficiently adjusted population stratification causes spurious association.

Other methods to test Q1 are based on the weighted Fisher's inverse  $\chi^2$ -method to pool statistical evidence in combining several p-values (further denoted as SPP: simple p-value pooling):

$$M = -2 \sum_{m=1}^{n_m} w_m \ln(p_m)$$

with  $m$  an index for a marker,  $n_m$  the number of markers belonging to GS and  $w_m$  a weight.

Luo et al. proposed first to generate the correlation matrix  $R_m$  of the normal-quantiles of the  $p_m$ -values  $Z_m = \Phi^{-1}(1 - p_m)$  for all markers belonging to a gene  $g$ .<sup>88</sup> Then a linear combination in the form

$$T_g = \frac{e'Z_m}{\sqrt{e'R_m e}} \quad \text{with } e = (1, 1, \dots, 1)^T$$

is calculated and a gene-wide  $p_g$ -value, assuming  $T_g$  to be standard-normally distributed, is derived. Finally, all  $p_g$ -values of genes belonging to the GS are used to calculate  $M$  (instead of  $p_m$ ) and the corresponding GS-wide  $p_{GS}$ -value, setting  $w_g = 1$  (instead of  $w_m$ ) for all genes.

De la Cruz et al. proposed the method SLAT (Set Level Association Testing), which differs from the above mentioned approach by truncating  $p_m$  lower than a pre-specified threshold (=restriction of accounted markers) and assigning weights  $w_m$  according the number of markers in LD per gene.<sup>89</sup>



## 2.2 Methods for the competitive null-hypothesis

Alternative to such “one-model” approaches, methods testing the *competitive null-hypothesis* (Q2) have been developed. These usually compare single marker association results (e.g. p-values or odds ratios) of markers/genes belonging to GS with those of all remaining genes in the genome (remaining GS'). Thus, pathways enriched with moderate association signals at several markers may be discovered which would be missed by a single marker approach.

As before, comprehensive comparisons of these methods outline the pros and cons, but without pointing to one approach as the overall best.<sup>68,69,71,79</sup> Two methods are exemplarily introduced here, to demonstrate the methodical challenges in the way to define such a common model:

The first and most popular approach briefly described here is GSEA (Gene-Set Enrichment Analysis), proposed by Subramanian et al..<sup>90,91</sup> A given list of markers is first ranked by  $p_m$ -values. To evaluate the degree of “enrichment” the GSEA method calculates an Enrichment Score (ES) by walking down this ordered list. At the beginning, a cumulative sum C is set to zero and will be increased when a marker belongs to GS and decreased otherwise. The increment of increase and decrease is chosen in the way that C is equal to zero at the end of the list again. If a GS is enriched with markers of low  $p_m$ -values, then C will get very large quickly. If the GS is not associated to the trait of interest, C will follow a random walk around zero. The maximum departure of C from zero is taken as ES and can be interpreted as a weighted Kolmogorov-Smirnov statistic. Hence GSEA tests, whether the distribution of  $p_m$ -values follows a uniform distribution. GSEA normalizes the ES for each GS to account for the variation in set sizes, yielding a normalized enrichment score (NES). To avoid dependencies of markers in LD, one may choose the smallest  $p_m$ -value of markers belonging to the same gene.

The second approach briefly described here is ALIGATOR (Association List Go AnnoTatOR)<sup>92</sup>. It uses of all  $p_m$ -values per marker resulting from a previous GWAS, regardless whether a marker belongs to GS or GS'. The first step consists of counting the numbers of significant markers  $m_{sig,GS}$  belonging to a GS and  $m_{sig,GS'}$  belonging to the remaining GS'. These counts are then compared like in Fisher's exact test, however the null-distribution of the test statistic is generated by a Monte-Carlo-Permutation procedure. It is desirable to correct the achieved GS-specific p-values for the number of GSs being tested. Because the GSs are not generally independent, standard methods, such as the Bonferroni and Sidak corrections, are inappropriate; a bootstrap approach is applied instead. This method was used to accomplish GSA for this dissertation.

The input datasets for this method are small, because the thousands of genotypes of each individual are no more required; but the accomplishment is time-intensive. However, because of the massive use of permutation and boot-strapping almost no model assumption needs to be fulfilled.

The diversity of the upper mentioned methods illustrates that GSA itself is a generic term describing an analytical strategy rather than a single method. Approaches differ for example by the way to combine results of several markers to a single quantity for a gene, by the way to assign genes to pathways or to statistically contrast genes belonging to the GS of interest with all remaining genes (providing a  $p_{GS}$ -value for a GS).

The performance of some strategies has been compared<sup>25,31,32</sup>; e.g. by G. Fehring et al.<sup>33</sup> who compared GSA applied to two independent GWASs formed by several ILCCO/TRICL data sets. They found “*a highly plausible association for the acetylcholine receptor activity pathway*”, but concluded: “*Difficulty in replicating associations (between the independent GWAS datasets) hindered our comparison ...*”. They used a heuristic approach by declaring those pathways as highly plausibly associated to lung cancer which were highly ranked (low gene-set p-values) by at least two GSA methods in both GWAS data sets, without taking additional information into account.

From a scientific point of view such an ad-hoc approach is not satisfying, because significance and consistency need to be addressed in a systematic quantitative manner. GSA usually provides p-values for GSs without giving any kind of effect/association estimate. Hence, Fisher’s inverse  $\chi^2$  method (here also denoted as SPP) - a well-established method to pool significance by a meta-analysis - might be applied. However, even if a GS of interest is found significantly “enriched” within all of several independent GWAS data sets, it is not guaranteed that the underlying single-marker association results are consistent regarding the direction of the association (as risk or protection factor for the disease).

### 3 Summaries

#### 3.1 META-GSA: A novel method to synthesise statistical evidence of several gene-set analyses

I have proposed the quantitative approach META-GSA to combine results of GSAs, respectively  $p_{GS}$ -values of GSs, by incorporating *concordance* of single-marker *association patterns* between studies, relevant for the GS of interest.<sup>11</sup>

These  $p_{GS}$ -values usually result from one-sided statistical tests (e.g. the larger ES the lower  $p_{GS}$ ). Hence, low  $p_{GS}$ -values result from low single marker p-values  $p_m$  of genes in GS. Nevertheless, low  $p_{GS}$ -values can theoretically arise through *accumulated marker significance* in which the minor alleles of all markers are observed for example as protective factors in one study, while being seen at the same time as risk factors in another study. Thus, significance for GS can appear simultaneously in several studies without concordance of the patterns of observed associations of markers, respectively genes, contained in GS (briefly denoted as *association pattern*). Consequently, concordance of the *direction of the test* (of  $p_{GS}$ ) is not given *a priori*.

The main steps of META-GSA are first to determine the *concordance of association patterns*; second, to use these to derive a weight for each study; and third, to apply a weighted version of Fisher's inverse  $\chi^2$ -method<sup>93</sup> to summarize significance of GSAs in a single meta-analytical p-value. Thus *significance* of GSA-results and *concordance* of single-marker association are combined. META-GSA can be further considered as an approach testing *GS-significance* conditional to, or in the presence of *concordance of association patterns*.

META-GSA was intended to use as little input data as possible. This makes it feasible even if the access to original study data is restricted, e.g. if data owners are reluctant to share detailed information due to legal or ethical causes. Because META-GSA is designed to rely on p-values as quantity of significance, it was necessary to define a measure that reflects also the direction of the observed association for a marker (indicating a protective or risk factor). Thus, a *directed reversed p-value* (PDR) was defined as:

$$p'_{m,s} = d_{m,s} \cdot (1 - p_{m,s})$$

where  $m$  is an index for a marker,  $s$  is an index for a study and  $d_{m,s} \in \{-1, +1\}$  is the direction of the observed association.

Next, the correlation of all  $p'_{m,s}$ -values of all markers belonging to GS is calculated. Because it is not assumed that each study used the same GSA approach, a rank-correlation, which is independent of the scale of  $p'_{m,s}$ , is used. Once the correlation matrix is found, principal component analysis (PCA) is used to determine the load of each study on a common but unknown general *PDR-profile* of the GS. We assume that these loads can be represented by the

first principal component ( $PC1_s$ ). Study weights  $w_s$  are then calculated as the product of the *normalized load* and the *proportion of effective studies*:

$$w_s = \frac{PC1_s}{\sum_{j=1}^{n_s} PC1_j} \cdot \frac{EV1}{n_s}$$

where  $n_s$  denotes the number of studies in the meta-analysis and  $EV1$  denotes the first eigenvalue.  $EV1/\sum_{j=1}^{n_s} EV1_j$  is the fraction of explained variance, since  $\sum_{j=1}^{n_s} EV1_j = n_s$  for a correlation matrix,  $EV1/n_s$  can be considered as the *proportion of effective studies*.

Finally, the test statistic

$$M_{n_s} = -2 \sum_{s=1}^{n_s} w_s \ln(p_{GS,s})$$

is calculated and a corresponding  $p_{META-GSA,GS}$ -value is derived. Since the mathematical conditions of the weighted version of Fisher's inverse  $\chi^2$ -method to assume  $M_{n_s}$  as  $\chi^2$ -distributed<sup>93</sup> are not fulfilled, the application of a CPU-intensive permutation procedure is required.

Accordingly, the core test-statistic  $M_{n_s}$  of META-GSA depends on necessary aspects being considered: *GS-significance* by  $p_{GS,s}$ -values and *concordance of association patterns* by the *weights*  $w_s$ .

Furthermore, I have investigated the benefits and the effort of META-GSA in comparison with SPP, which is fast but does not address *concordance of association patterns* in any way. Both methods keep the *type 1 error* at the specified level. However, under  $H_0$  the results of META-GSA and SPP were found to be almost uncorrelated. False-positive gene sets found by META-GSA and SPP only partially overlap. We assume that this may result at least in part from including the concordance of association patterns, for the evaluation.

Moreover, META-GSA was found to be more powerful than SPP. The greater the number of studies combined, the larger the advantages in power became.

We also compared META-GSA to a pooledGWAS-GSA approach. For the latter, the combining of studies is switched to the level of markers, followed by a single GSA performed on the pooled marker-specific associations. In general, we found META-GSA to outperform pooledGWAS-GSA.

The method META-GSA is described in detailed in Rosenberger et al. *PLoS One* 2015<sup>11</sup>. All programs were implemented in SAS 9.3 (SAS Institute, NC, USA) and are provided as supplemental material.

### 3.2 Gene-set analysis with respect to lung cancer: A meta-analysis

I have further performed a meta-analysis of seven GSAs for lung cancer, applying the method META-GSA. Overall, information taken from 11,365 cases and 22,505 controls from within the TRICL/ILCCO consortia was used to investigate a total of 234 pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.<sup>75,76,94</sup>

In the original GWASs, a log-additive mode of inheritance was fitted for each marker, adjusting for age, sex, smoking status, study centre (if applicable), and the first three principal components to account for hidden genomic structure. There was agreement within TRICL/ILCCO on the model to be used for single-marker association analysis. The calculations had been performed by the study centres themselves and shared within the consortium. The results of marker-by-marker association tests were used as input information for the GSAs. The marker-to-gene annotation (MtG) for humans from the ENSEMBL database<sup>74</sup> was used. For some data set GSA results were already available. If necessary, GSA was performed using the program ALIGATOR.<sup>92</sup>

This revealed the systemic lupus erythematosus KEGG pathway *hsa05322*, driven by the gene region 6p21-22, but no other investigated KEGG pathway, as implicated in lung cancer ( $p=0.0306$ , corrected for multiple testing). This gene region is known to be associated with squamous cell lung carcinoma. The most important genes driving the significance of this pathway belong to the genomic areas *HIST1-H4L*, *-1BN*, *-2BN*, *-H2AK*, *-H4K* and *C2/C4A/C4B*. Within these areas, the markers most significantly associated with LC are rs13194781 (located within HIST12BN) and rs1270942 (located between *C2* and *C4A*).

The investigation of the association of KEGG pathways in lung cancer GWAS is reported in detail in Rosenberger et al. *PLoS One* 2017.<sup>95</sup>

## 4 Discussion

About 50,000 individuals are diagnosed with lung cancer in Germany every year. They face a poor prognosis, since only 21% of women 16% of men survive five years after being diagnosed. Most cases can be attributed to tobacco smoking, but there is also scientific evidence that genetic factors contribute to the susceptibility of Lung cancer. In the recent two decades genome-wide association studies (GWASs) have been carried out to identify genetic variants that are associated to lung cancer or to a specific histological subtype. Genomic risk loci in European population were identified at chromosomes e.g. 5p15.33, 6p21-22 and 15q25, after pooling data of several GWASs within *the International Lung Cancer Consortium* (ILCCO), overlapping with the *Transdisciplinary Research in Cancer of the Lung* (TRICL).<sup>3-10</sup> It has also been increasingly recognized that GSA can extend GWAS approaches by incorporating existing knowledge of biological processes, with the aim of identifying disease-related pathways. GSA has gained great popularity and several approaches have been proposed. Although the pros and cons have been discussed<sup>72,73</sup> and points to improve have been formulated<sup>70</sup>, it has only been mentioned that there is a need to replicate pathway association findings to avoid false positive results.<sup>70</sup> According to our knowledge, there is no formal method to combine the results of several GSAs. The basic criticism on applying simple p-pooling (SPP) is a lack of interpretability of results when single markers differ in their attributed role as risk or predictive factor. I have proposed the quantitative approach META-GSA to combine such results, respectively *GS-significance*, by incorporating *concordance* of single-marker *association patterns* between studies, relevant for the *GS* of interest. This method was then applied to the data shared within ILCCO/TRICL, to further investigate genetic risk factor for lung cancer.

META-GSA was found to outperform SPP and a pooledGWAS-GSA approach. One advantage of META-GSA is the fact that heterogeneity in the strength of association for single markers or genes, respectively, between studies does not necessarily cause lower power, as long as other genes belonging to *GS* compensate such deficiency. "Between-study heterogeneity ... can offer valuable insight for further clarification of gene-disease associations".<sup>96</sup>

Furthermore, META-GSA is applicable to any GSA method selected, even those using individual participants' genotype data, which may prove to be more suitable and more powerful than methods based on GWAS summary results (pooledGWAS-GSA).<sup>68,80</sup> Resting the GSA on common effect estimates can become critical in the case of strong study heterogeneity for few or many markers, since the existence of a common marker-specific association in such a situation is doubtful. For META-GSA, study heterogeneity results in low concordance of the patterns of study-specific association estimates and subsequently in low study weights, which simply reduces the power, however without violating such a critical assumption.

All the same, META-GSA has some critical points that need to be mentioned. First, it is necessary to estimate LD between neighboring markers to be able to calculate PDRs (the core

quantity of META-GSA) of differing markers across studies. It is known that LD-patterns in human subpopulations are different.<sup>97</sup> This can be problematic if one aims to combine results across different ethnicities or if the source population of a study is known to be admixed.

Secondly, the implemented way to combine marker-level estimates to a gene-level statistic is only one of several published alternatives.<sup>98-101</sup>

Moreover, to perform GWASs and GSAs one needs to annotate marker to genes (MtG) which are annotated to pathways (GtP) in a static way, referring to public databases. In contrast, genes act dynamically and may have for instance several transcripts which can be active or passive in a certain pathway. Thus, a static annotation maybe doesn't match well to a dynamic biological process. META-GSA is concerned to this drawback to the same extend as the GSA approaches aimed to be combined.

Next, META-GSA is applicable to any GSA method selected for the analysis of a single GWAS, irrespectively if the *self-contained null-hypothesis* (Q1) or the *competitive null-hypothesis* (Q2) was tested. Combining  $p_{GS,s}$ -values for Q1 and Q2 is calculative feasible but induces a lack of interpretability.

In addition, the use of a permutation procedure is time and CPU-intensive.

The biggest disadvantage is perhaps that p-values for each tested gene-set are the only results. META-GSA is designed to rely on p-values as quantity of significance. A p-value can be used to justify the existences of an association; however it is not solely determined by the strength of the observed effect, but also by factors like sample size, the used statistical model and the applied test procedure. Hence META-GSA is unable to estimate the fraction of the risk for an investigated disease that can be attributed to the identified driving genes or the whole gene-set. As with most GSA approaches META-GSA does not deliver any effect estimation.

All in all, I could demonstrate that META-GSA may be a powerful add-on tool in the research of the genetic architecture of complex traits or diseases. I have applied this new tool to discover an accumulation of genomic association with lung cancer in the KEGG pathway *hsa05322*, which comprises genes related to *systemic lupus erythematosus* (SLE). This suggests some cross-phenotype association with lung cancer and SLE.

Regarding the application, all  $p_{GS}$ -driving genes identified in the reported meta-analysis are located within or next to the major histocompatibility complex (MHC) on chromosome 6p21-22, albeit in two separate areas, about 3000 kb apart. The first area comprises the genes of histone cluster I: *HIST1-H4L*, *-1BN*, *-2BN*, *-H2AK*, *-H4K*. The second area comprises the genes *C2*, *C4A*, and *C4B*. Neither of both areas were before identified as associated to lung cancer. However, the identification of disease-relevant genes in the MHC region (6p21–6p22) and far beyond is complicated owing to the strong and extensive LD across both common and

rare haplotypes.<sup>12</sup> Hence any observed cross-phenotype association will probably tag plenty of genes.

In conclusion, the gene region 6p21-22 appears to be more extensively associated with lung cancer than previously assumed. Given wide-stretched linkage disequilibrium to the area *APOM/BAG6/MSH5*, there is currently simply not enough information or evidence to conclude whether the potential pleiotropy of lung cancer and SLE is spurious, biological, or mediated.



## A References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;**66**(1):7-30.
2. Barnes B, Kraywinkel K, Nowossadeck E, Schönfeld I, Starker A, Wienecke A, Wolf U. Bericht zum Krebsgeschehen in Deutschland 2016. Berlin: Zentrum für Krebsregisterdaten im Robert Koch-Institut, 2016.
3. Wang Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakidou A, Qureshi M, Dong Q, Gu X, Chen WV, Spitz MR, Eisen T, Amos CI, Houlston RS. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat. Genet* 2008;**40**(12):1407-1409.
4. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Chen C, Goodman G, Field JK, Liloglou T, Xinarianos G, Cassidy A, McLaughlin J, Liu G, Narod S, Krokan HE, Skorpen F, Elvestad MB, Hveem K, Vatten L, Linseisen J, Clavel-Chapelon F, Vineis P, Bueno-de-Mesquita HB, Lund E, Martinez C, Bingham S, Rasmuson T, Hainaut P, Riboli E, Ahrens W, Benhamou S, Lagiou P, Trichopoulos D, Holcatova I, Merletti F, Kjaerheim K, Agudo A, Macfarlane G, Talamini R, Simonato L, Lowry R, Conway DI, Znaor A, Healy C, Zelenika D, Boland A, Delepine M, Foglio M, Lechner D, Matsuda F, Blanche H, Gut I, Heath S, Lathrop M, Brennan P. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008;**452**(7187):633-637.
5. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, Dong Q, Zhang Q, Gu X, Vijayakrishnan J, Sullivan K, Matakidou A, Wang Y, Mills G, Doheny K, Tsai YY, Chen WV, Shete S, Spitz MR, Houlston RS. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* 2008;**40**(5):616-622.
6. Truong T, Hung RJ, Amos CI, Wu X, Bickeboller H, Rosenberger A, Sauter W, Illig T, Wichmann HE, Risch A, Dienemann H, Kaaks R, Yang P, Jiang R, Wiencke JK, Wrensch M, Hansen H, Kelsey KT, Matsuo K, Tajima K, Schwartz AG, Wenzlaff A, Seow A, Ying C, Staratschek-Jox A, Nurnberg P, Stoelben E, Wolf J, Lazarus P, Muscat JE, Gallagher CJ, Zienolddiny S, Haugen A, van der Heijden HF, Kiemeny LA, Isla D, Mayordomo JI, Rafnar T, Stefansson K, Zhang ZF, Chang SC, Kim JH, Hong YC, Duell EJ, Andrew AS, Lejbkowitz F, Rennert G, Muller H, Brenner H, Le Marchand L, Benhamou S, Bouchardy C, Teare MD, Xue X, McLaughlin J, Liu G, McKay JD, Brennan P, Spitz MR. Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium. *J Natl. Cancer Inst.* 2010;**102**(13):959-971.
7. Timofeeva MN, Hung RJ, Rafnar T, Christiani DC, Field JK, Bickeboller H, Risch A, McKay JD, Wang Y, Dai J, Gaborieau V, McLaughlin J, Brenner D, Narod SA, Caporaso NE, Albanes D, Thun M, Eisen T, Wichmann HE, Rosenberger A, Han Y, Chen W, Zhu D, Spitz M, Wu X, Pande M, Zhao Y, Zaridze D, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Krokan HE, Gabrielsen ME, Skorpen F, Vatten L, Njolstad I, Chen C, Goodman G, Lathrop M, Benhamou S, Vooder T, Valk K, Nelis M, Metspalu A, Raji O, Chen Y, Gosney J, Liloglou T, Muley T, Dienemann H, Thorleifsson G, Shen H, Stefansson K, Brennan P, Amos CI, Houlston R, Landi MT. Influence of common genetic variation on lung cancer risk: meta-analysis of 14 900 cases and 29 485 controls. *Hum Mol Genet* 2012;**21**(22):4980-95.
8. Brennan P, Hainaut P, Boffetta P. Genetics of lung-cancer susceptibility. *Lancet Oncol* 2011;**12**(4):399-408.
9. Wang Y, McKay JD, Rafnar T, Wang Z, Timofeeva MN, Broderick P, Zong X, Laplana M, Wei Y, Han Y, Lloyd A, Delahaye-Sourdeix M, Chubb D, Gaborieau V, Wheeler W, Chatterjee N, Thorleifsson G, Sulem P, Liu G, Kaaks R, Henrion M, Kinnersley B, Vallee M, LeCalvez-Kelm F,

- Stevens VL, Gapstur SM, Chen WV, Zaridze D, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Krokán HE, Gabrielsen ME, Skorpén F, Vatten L, Njolstad I, Chen C, Goodman G, Benhamou S, Vooder T, Valk K, Nelis M, Metspalu A, Lener M, Lubinski J, Johansson M, Vineis P, Agudo A, Clavel-Chapelon F, Bueno-de-Mesquita HB, Trichopoulos D, Khaw KT, Johansson M, Weiderpass E, Tjønneland A, Riboli E, Lathrop M, Scelo G, Albanes D, Caporaso NE, Ye Y, Gu J, Wu X, Spitz MR, Dienemann H, Rosenberger A, Su L, Matakidou A, Eisen T, Stefansson K, Risch A, Chanock SJ, Christiani DC, Hung RJ, Brennan P, Landi MT, Houlston RS, Amos CI. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. *Nat Genet* 2014;**46**(7):736-41.
10. Fehring G, Liu G, Pintilie M, Sykes J, Cheng D, Liu N, Chen Z, Seymour L, Der SD, Shepherd FA, Tsao MS, Hung RJ. Association of the 15q25 and 5p15 lung cancer susceptibility regions with gene expression in lung tumor tissue. *Cancer Epidemiol Biomarkers Prev* 2012;**21**(7):1097-104.
  11. Rosenberger A, Friedrichs S, Amos CI, Brennan P, Fehring G, Heinrich J, Hung RJ, Muley T, Müller-Nurasyid M, Risch A, Bickeboller H. META-GSA: Combining Findings from Gene-Set Analyses across Several Genome-Wide Association Studies. *PLoS One* 2015;**10**(10):e0140179.
  12. Ahmad T, Neville M, Marshall SE, Armuzzi A, Mulcahy-Hawes K, Crawshaw J, Sato H, Ling K-L, Barnardo M, Goldthorpe S, Walton R, Bunce M, Jewell DP, Welsh KI. Haplotype-specific linkage disequilibrium patterns define the genetic topography of the human MHC. *Human Molecular Genetics* 2003;**12**(6):647-656.
  13. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. *Nat Rev Genet* 2013;**14**(7):483-95.
  14. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin D, Forman D, Bray F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. <http://globocan.iarc.fr> Accessed 27. Feb. 2017.
  15. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin.* 2011;**61**(2):69-90.
  16. Talhout R, Schulz T, Florek E, van Benthem J, Wester P, Opperhuizen A. Hazardous compounds in tobacco smoke. *Int J Environ Res Public Health* 2011;**8**(2):613-28.
  17. Müller F. Tabakmißbrauch und Lungencarcinom *Z Krebsforsch* 1939;**49**:28.
  18. Doll R, Hill AB. A study of the aetiology of carcinoma of the lung. *Br Med J* 1952;**2**(4797):1271-86.
  19. Lee PN, Forey BA, Coombs KJ. Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. *BMC Cancer* 2012;**12**:385.
  20. Lampert T, Kuntz B. Tabak - Zahlen und Fakten zum Konsum. *Jahrbuch Sucht* 2015. Hamm, Germany: Deutsche Hauptstelle für Suchtfragen e.V., 2015.
  21. Sethi TK, El-Ghamry MN, Kloecker GH. Radon and lung cancer. *Clin Adv Hematol Oncol* 2012;**10**(3):157-64.
  22. *Health Effects of Exposure to Radon: BEIR VI*. Washington (DC), 1999.
  23. Hubaux R, Becker-Santos DD, Enfield KS, Lam S, Lam WL, Martinez VD. Arsenic, asbestos and radon: emerging players in lung tumorigenesis. *Environ Health* 2012;**11**:89.
  24. Bruske-Hohlfeld I, Mohner M, Pohlabein H, Ahrens W, Bolm-Audorff U, Kreienbrock L, Kreuzer M, Jahn I, Wichmann HE, Jockel KH. Occupational lung cancer risk for men in Germany: results from a pooled case-control study. *Am J Epidemiol.* 2000;**151**(4):384-395.

25. Celik I, Gallicchio L, Boyd K, Lam TK, Matanoski G, Tao X, Shiels M, Hammond E, Chen L, Robinson KA, Caulfield LE, Herman JG, Guallar E, Alberg AJ. Arsenic in drinking water and lung cancer: a systematic review. *Environ Res* 2008;**108**(1):48-55.
26. Ni J, Qiu LJ, Hu LF, Cen H, Zhang M, Wen PF, Wang XS, Pan HF, Ye DQ. Lung, liver, prostate, bladder malignancies risk in systemic lupus erythematosus: evidence from a meta-analysis. *Lupus* 2014;**23**(3):284-92.
27. Brenner DR, Boffetta P, Duell EJ, Bickeboller H, Rosenberger A, McCormack V, Muscat JE, Yang P, Wichmann HE, Brueske-Hohlfeld I, Schwartz AG, Cote ML, Tjonneland A, Friis S, Le Marchand L, Zhang ZF, Morgenstern H, Szeszenia-Dabrowska N, Lissowska J, Zaridze D, Rudnai P, Fabianova E, Foretova L, Janout V, Bencko V, Schejbalova M, Brennan P, Mates IN, Lazarus P, Field JK, Raji O, McLaughlin JR, Liu G, Wiencke J, Neri M, Ugolini D, Andrew AS, Lan Q, Hu W, Orlov I, Park BJ, Hung RJ. Previous lung diseases and lung cancer risk: a pooled analysis from the International Lung Cancer Consortium. *Am J Epidemiol* 2012;**176**(7):573-85.
28. Wu SH, Liu Z. Soy food consumption and lung cancer risk: a meta-analysis using a common measure across studies. *Nutr Cancer* 2013;**65**(5):625-32.
29. Slatore CG, Littman AJ, Au DH, Satia JA, White E. Long-term use of supplemental multivitamins, vitamin C, vitamin E, and folate does not reduce the risk of lung cancer. *Am J Respir Crit Care Med* 2008;**177**(5):524-30.
30. Kubik A, Zatloukal P, Tomasek L, Dolezal J, Syllabova L, Kara J, Kopecky P, Plesko I. A case-control study of lifestyle and lung cancer associations by histological types. *Neoplasma* 2008;**55**(3):192-9.
31. Gallicchio L, Boyd K, Matanoski G, Tao XG, Chen L, Lam TK, Shiels M, Hammond E, Robinson KA, Caulfield LE, Herman JG, Guallar E, Alberg AJ. Carotenoids and the risk of developing lung cancer: a systematic review. *Am J Clin Nutr* 2008;**88**(2):372-83.
32. Cortes-Jofre M, Rueda JR, Corsini-Munoz G, Fonseca-Cortes C, Caraballoso M, Bonfill Cosp X. Drugs for preventing lung cancer in healthy people. *Cochrane Database Syst Rev* 2012;**10**:CD002141.
33. Lung Cancer Prevention (PDQ(R)): Health Professional Version. *PDQ Cancer Information Summaries*. Bethesda (MD), 2016.
34. Vieira AR, Abar L, Vingeliene S, Chan DS, Aune D, Navarro-Rosenblatt D, Stevens C, Greenwood D, Norat T. Fruits, vegetables and lung cancer risk: a systematic review and meta-analysis. *Ann Oncol* 2016;**27**(1):81-96.
35. Sun Y, Li Z, Li J, Li Z, Han J. A Healthy Dietary Pattern Reduces Lung Cancer Risk: A Systematic Review and Meta-Analysis. *Nutrients* 2016;**8**(3):134.
36. Emaus A, Thune I. Physical activity and lung cancer prevention. *Recent Results Cancer Res* 2011;**186**:101-33.
37. Ooi WL, Elston RC, Chen VW, Bailey-Wilson JE, Rothschild H. Increased familial risk for lung cancer. *J.Natl.Cancer Inst.* 1986;**76**(2):217-222.
38. Osann KE. Lung cancer in women: the importance of smoking, family history of cancer, and medical history of respiratory disease. *Cancer Res.* 1991;**51**(18):4893-4897.
39. Radzikowska E, Rowinska-Zakrzewska E. [Familial aggregation of lung cancer in families of patients with lung cancer]. *Pneumonol.Alergol.Pol.* 1994;**62**(9-10):444-449.
40. Samet JM, Humble CG, Pathak DR. Personal and family history of respiratory disease and lung cancer risk. *Am Rev.Respir.Dis.* 1986;**134**(3):466-470.

41. Shaw GL, Falk RT, Pickle LW, Mason TJ, Buffler PA. Lung cancer risk associated with cancer in relatives. *J.Clin.Epidemiol.* 1991;**44**(4-5):429-437.
42. Tokuhata GK, Lilienfeld AM. Familial aggregation of lung cancer in humans. *J.Natl.Cancer Inst.* 1963;**30**:289-312.
43. Brisman R, Baker RR. Carcinoma of lung in four siblings. *Cancer* 1967;**20**(11):2048-53.
44. Kreuzer M, Kreienbrock L, Gerken M, Heinrich J, Bruske-Hohlfeld I, Muller KM, Wichmann HE. Risk factors for lung cancer in young adults. *Am J Epidemiol.* 1998;**147**(11):1028-1037.
45. Broman K, Pohlabein H, Jahn I, Ahrens W, Jockel KH. Aggregation of lung cancer in families: results from a population-based case-control study in Germany. *Am.J.Epidemiol.* 2000;**152**(6):497-505.
46. Schwartz AG, Yang P, Swanson GM. Familial risk of lung cancer among nonsmokers and their relatives. *Am J Epidemiol.* 1996;**144**(6):554-562.
47. Sellers TA, Bailey-Wilson JE, Elston RC, Wilson AF, Elston GZ, Ooi WL, Rothschild H. Evidence for mendelian inheritance in the pathogenesis of lung cancer. *J.Natl.Cancer Inst.* 1990;**82**(15):1272-1279.
48. Sellers TA, Chen PL, Potter JD, Bailey-Wilson JE, Rothschild H, Elston RC. Segregation analysis of smoking-associated malignancies: evidence for Mendelian inheritance. *Am J Med.Genet* 1994;**52**(3):308-314.
49. Bailey-Wilson JE, Amos CI, Pinney SM, Petersen GM, de Andrade M, Wiest JS, Fain P, Schwartz AG, You M, Franklin W, Klein C, Gazdar A, Rothschild H, Mandal D, Coons T, Slusser J, Lee J, Gaba C, Kupert E, Perez A, Zhou X, Zeng D, Liu Q, Zhang Q, Seminara D, Minna J, Anderson MW. A major lung cancer susceptibility locus maps to chromosome 6q23-25. *Am J Hum Genet* 2004;**75**(3):460-74.
50. Ku CS, Loy EY, Salim A, Pawitan Y, Chia KS. The discovery of human genetic variations and their use as disease markers: past, present and future. *J Hum Genet* 2010;**55**(7):403-15.
51. Risch N, Merikangas K. The future of genetic studies of complex human diseases. *Science* 1996;**273**(5281):1516-1517.
52. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;**6**(2):95-108.
53. Devlin B, Risch N. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* 1995;**29**(2):311-322.
54. Pritchard JK, Przeworski M. Linkage disequilibrium in humans: models and data. *Am J Hum Genet* 2001;**69**(1):1-14.
55. Price AL, Weale ME, Patterson N, Myers SR, Need AC, Shianna KV, Ge D, Rotter JI, Torres E, Taylor KD, Goldstein DB, Reich D. Long-range LD can confound genome scans in admixed populations. *Am J Hum Genet* 2008;**83**(1):132-5; author reply 135-9.
56. Ha NT, Freytag S, Bickeboeller H. Coverage and efficiency in current SNP chips. *Eur J Hum Genet* 2014;**22**(9):1124-30.
57. Edenberg HJ, Liu Y. Laboratory methods for high-throughput genotyping. *Cold Spring Harb Protoc* 2009;**2009**(11):pdb top62.
58. dbSNP [http://www.ncbi.nlm.nih.gov/projects/SNP/snp\\_summary.cgi](http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi)
59. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001;**29**(1):308-11.

60. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet* 2010;**11**(6):446-50.
61. Wang S, Zhao H. Sample size needed to detect gene-gene interactions using association designs. *Am J Epidemiol* 2003;**158**(9):899-914.
62. Maxwell CA, Moreno V, Sole X, Gomez L, Hernandez P, Urruticoechea A, Pujana MA. Genetic interactions: the missing links for a better understanding of cancer susceptibility, progression and treatment. *Mol Cancer* 2008;**7**:4.
63. Murcray CE, Lewinger JP, Conti DV, Thomas DC, Gauderman WJ. Sample size requirements to detect gene-environment interactions in genome-wide association studies. *Genet Epidemiol* 2011;**35**(3):201-10.
64. Witte JS. Genome-wide association studies and beyond. *Annu Rev Public Health* 2010;**31**:9-20  
4 p following 20.
65. Wang X, Oldani MJ, Zhao X, Huang X, Qian D. A review of cancer risk prediction models with genetic variants. *Cancer Inform* 2014;**13**(Suppl 2):19-28.
66. Visscher PM, Montgomery GW. Genome-wide association studies and human disease: from trickle to flood. *JAMA* 2009;**302**(18):2028-9.
67. Fehring G, Liu G, Briollais L, Brennan P, Amos CI, Spitz MR, Bickeboller H, Wichmann HE, Risch A, Hung RJ. Comparison of pathway analysis approaches using lung cancer GWAS data sets. *PLoS One* 2012;**7**(2):e31816.
68. Ramanan VK, Shen L, Moore JH, Saykin AJ. Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends Genet* 2012;**28**(7):323-32.
69. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol* 2012;**8**(2):e1002375.
70. Wang L, Jia P, Wolfinger RD, Chen X, Zhao Z. Gene set analysis of genome-wide association studies: methodological issues and perspectives. *Genomics* 2011;**98**(1):1-8.
71. Wang K, Li M, Hakonarson H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* 2010;**11**(12):843-54.
72. Fridley BL, Biernacka JM. Gene set analysis of SNP data: benefits, challenges, and future directions. *Eur J Hum Genet* 2011;**19**(8):837-843.
73. Ackermann M, Strimmer K. A general modular framework for gene set enrichment analysis. *BMC Bioinformatics* 2009;**10**.
74. Cunningham F, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, Gil L, Giron CG, Gordon L, Hourlier T, Hunt SE, Janacek SH, Johnson N, Juettemann T, Kahari AK, Keenan S, Martin FJ, Maurel T, McLaren W, Murphy DN, Nag R, Overduin B, Parker A, Patricio M, Perry E, Pignatelli M, Riat HS, Sheppard D, Taylor K, Thormann A, Vullo A, Wilder SP, Zadissa A, Aken BL, Birney E, Harrow J, Kinsella R, Muffato M, Ruffier M, Searle SM, Spudich G, Trevanion SJ, Yates A, Zerbino DR, Flicek P. Ensembl 2015. *Nucleic Acids Res* 2014.
75. KEGG Pathway Database. Kanehisa Laboratories, 1995-2017.
76. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res*. 2012;**40**(Database issue):D109-D114.
77. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE,

- Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000;**25**(1):25-29.
78. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;**4**(1):44-57.
79. Freytag S, Bickeboller H. Comparison of three summary statistics for ranking genes in genome-wide association studies. *Stat Med* 2014;**33**(11):1828-41.
80. Freytag S, Manitz J, Schlather M, Kneib T, Amos CI, Risch A, Chang-Claude J, Heinrich J, Bickeboller H. A network-based kernel machine test for the identification of risk pathways in genome-wide association studies. *Hum Hered* 2013;**76**(2):64-75.
81. Freytag S, Bickeboller H, Amos CI, Kneib T, Schlather M. A novel kernel for correcting size bias in the logistic kernel machine test with an application to rheumatoid arthritis. *Hum Hered* 2012;**74**(2):97-108.
82. Liu D, Lin X, Ghosh D. Semiparametric regression of multidimensional genetic pathway data: least-squares kernel machines and linear mixed models. *Biometrics* 2007;**63**(4):1079-88.
83. Schaid DJ. Genomic similarity and kernel methods I: advancements by building on mathematical and statistical foundations. *Hum Hered* 2010;**70**(2):109-31.
84. Schaid DJ. Genomic similarity and kernel methods II: methods for genomic information. *Hum Hered* 2010;**70**(2):132-40.
85. Wu MC, Kraft P, Epstein MP, Taylor DM, Chanock SJ, Hunter DJ, Lin X. Powerful SNP-set analysis for case-control genome-wide association studies. *Am J Hum Genet* 2010;**86**(6):929-42.
86. Chen LS, Hutter CM, Potter JD, Liu Y, Prentice RL, Peters U, Hsu L. Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. *Am J Hum Genet* 2010;**86**(6):860-71.
87. Harrell FE. *Regression modeling strategies with applications to linear models, logistic regression, and survival analysis*. Springer series in Statistics. New York Berlin Heidelberg: Springer Verlag, 2001.
88. Luo L, Peng G, Zhu Y, Dong H, Amos CI, Xiong M. Genome-wide gene and pathway analysis. *Eur J Hum Genet* 2010;**18**(9):1045-1053.
89. De la Cruz O, Wen X, Ke B, Song M, Nicolae DL. Gene, region and pathway level analyses in whole-genome studies. *Genet Epidemiol* 2010;**34**(3):222-31.
90. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc.Natl.Acad.Sci.U.S.A* 2005;**102**(43):15545-15550.
91. Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP. GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics* 2007;**23**(23):3251-3253.
92. Holmans P, Green EK, Pahwa JS, Ferreira MA, Purcell SM, Sklar P, Owen MJ, O'Donovan MC, Craddock N. Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. *Am J Hum Genet* 2009;**85**(1):13-24.
93. Makambi K. Weighted inverse chi-square method for correlated significance tests. *Journal of Applied Statistics* 2003;**30**(2):225-234.
94. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;**28**(1):27-30.

95. Rosenberger A, Friedrichs S, Amos C, Sohns M, Hung R, Fehringer G, McLaughlin J, Brennan P, Risch A, Brüske I, Caporaso N, Landi M, Christian iD, Wei Y, Bickböller H. Gene-set meta-analysis of lung cancer identifies pathway related to systemic lupus Erythematosus. *PLoS One* 2017.
96. Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS One* 2007;**2**(9):e841.
97. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science* 2002;**296**(5576):2225-9.
98. Zaykin DV, Zhivotovsky LA, Czika W, Shao S, Wolfinger RD. Combining p-values in large-scale genomics experiments. *Pharm Stat* 2007;**6**(3):217-26.
99. Dudbridge F, Koeleman BP. Rank truncated product of P-values, with application to genomewide association scans. *Genet Epidemiol*. 2003;**25**(4):360-366.
100. Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, Caporaso N, Kraft P, Chatterjee N. Pathway analysis by adaptive combination of P-values. *Genet Epidemiol* 2009;**33**(8):700-9.
101. Hsueh HM, Chen JJ, Kodell RL. Comparison of methods for estimating the number of true null hypotheses in multiplicity testing. *J Biopharm Stat* 2003;**13**(4):675-89.

## B Curriculum Vitae

Albert Rosenberger      born in Vienna,      16<sup>th</sup> July 1968  
Nationality:              Austrian

### Professional Career

since 11/2001	<b>University Medical Centre Göttingen</b> Department of Genetic Epidemiology chair: Prof. Dr. Heike Bickeböller position as <b>scientific employee</b>
08/1997 – 11/2001	<b>University Hospital Tübingen</b> <i>imi</i> – Department of Medical Information Processing chair: Prof. Dr. Hans-Konrad Selbmann position as <b>scientific employee</b>
01/1994 – 07/1997	<b>University of Vienna</b> Department of Social Medicine chair: Prof. Dr. Michael Kunze position as <b>research associate</b>
08/1996 – 03/1997	publishing house OSTAG / Vienna position as <b>technical co-worker</b> and <b>assistant of the management</b>
11/1996 – 03/1997	<b>AKH – General Hospital Vienna</b> Department of Cancer Research and Epidemiology chair: Prof. Dr. Christian Vutuc position as <b>freelancer and research associate</b>
11/1994 – 06/1995	<b>AKH – General Hospital Vienna</b> IMC - Department of Medical Computer Science <b>chair:</b> Prof. Dr. Michael Schemper position as <b>freelancer and research associate</b>
01/1990 – 09/1990	SIEMENS Austria Viennese factories Department of process planning and control position as <b>technical co-worker</b>

### Education

1995 - 1996	<b>University College Dublin:</b> 1 year academic courses in statistics and sociology
1990 - 1996	<b>University of Vienna:</b> regular studies of statistics and sociology <b>graduation: Mag. rer. soc. oec.</b>
1988 - 1989	<b>obligatory military service</b>
1983 - 1988	<b>TGM- Technical College in Vienna:</b> school of production engineering <b>graduation: Matura</b>
1974 - 1983	<b>elementary and secondary school in Vienna</b>



## Publications

### 2017

- Pathway-analysis of published genome-wide association studies of lung cancer: A potential role for the CYP4F3 locus  
Yin J, Liu H, Liu Z, Owzar K, Han Y, Su L, Wei Y, Hung RJ, Brhane Y, McLaughlin J, Brennan P, Bickeboeller H, **Rosenberger A**, Houlston RS, Caporaso N, Landi MT, Heinrich J, Risch A, Christiani DC, Amos CI, and Wei Q  
Mol Carcinog. 2017 Feb 2. doi: 10.1002/mc.22622. [Epub ahead of print]

### 2016

- A Novel Genetic Variant in Long Non-coding RNA Gene NEXN-AS is Associated with Risk of Lung Cancer  
Yuan H, Liu H, Liu Z, Owzar K, Han Y, Su L, Wei Y, Hung R, McLaughlin J, Brhane Y, Rafnar T, Thorleifsson G, Stefansson K, Brennan P, Bickeboeller H, **Rosenberger A**, Houlston R, Caporaso NE, Landi MT, Heinrich J, Risch A, Christiani DC, Amos CI  
Sci Rep. 2016 Oct 7;6:34234. doi: 10.1038/srep34234
- Genetic Risk Can Be Decreased: Quitting Smoking Decreases and Delays Lung Cancer for Smokers With High and Low CHRNA5 Risk Genotypes - A Meta-Analysis  
Chen L-S, Baker T, Hung RJ, Horton A, Culverhouse R, Hartz S, Saccone N, Cheng I, Deng B, Han Y, Hansen HM, Horsman J, Kim C, **Rosenberger A**, Aben KK, Andrew AS, Breslau N, Chang S-C, Saum K-U, Dienemanne H, Hatsukami DK, Johnson EO, Pande M, Wrensch MR, McLaughlin J, Skaug V, van der Heijden HF, Wampfler J, Wenzlaff A, Woll P, Zienolddiny S, Bickeböller H, Brenner H, Duell EJ, Haugen A, Brüske I, Kiemeny LA, Lazarus P, Le Marchand L, Liu G, Mayordomo J, Risch A, Schwartz AG, Teare D, Wu X, Wiencke JK, Yang P, Zhang Z-F, Spitz MR, I. Amos CI, Bierut LJ  
EBioMedicine. 2016 Sep;11:219-226. doi: 10.1016/j.ebiom.2016.08.012. Epub 2016 Aug 10
- Polymorphisms of the centrosomal gene (FGFR1OP) and lung cancer risk: a meta-analysis of 14,463 cases and 44,188 controls  
Kang X, Liu H, Onaitis MW, Liu Z, Owzar K, Su L, Han Y, Wei Y, Hung RJ, Brhane Y, McLaughlin J, Brennan P, Bickeböller H, **Rosenberger A**, Houlston RS, Caporaso N, Landi MT, Heinrich J, Risch A, Wu X, Ye Y, Christiani DC, Amos CI, Wei Q;  
Carcinogenesis. 2016 Mar;37(3):280-9
- Association of Patient Risk Factors and Frequency of Nevus-Associated Cutaneous Melanomas  
Haenssle H, Mograby N, Ngassa A, Buhl T, Emmert St, Schön M, **Rosenberger A**, Bertsch HP  
Dermatol. 2016 Mar;152(3):291-8  
zitiert in: *Nävusassoziierte Melanome: Risiken quantifiziert, um Früherkennung zu optimieren*; Dtsch Arztebl 2016; 113(25): A-1217 / B-1022 / C-1006

### 2015

- Inzidenz der Candidämie nach gastrointestinal-chirurgischen Eingriffen in deutschen Akutkrankenhäusern [Incidence of Candidaemia Following Abdominal Surgery in German Hospitals].  
Kuhns M, **Rosenberger A**, Bader O, Reichard U, Gross U, Weig M  
Zentralbl Chir 2015; 140(06): 617-623
- Identification of lung cancer histology-specific variants applying Bayesian framework variant prioritization approaches within the TRICL and ILCCO consortia  
Brenner D, Amos CI, Brhane Y, Timofeeva M, Caporaso N, Wang Y, Christiani D, Bickeböller H, Yang P, Albanes D, Stevens V, Gapstur S, McKay J, Boffetta P, Zaridze D, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Krokan HE, Skorpen F, Gabrielsen ME, Vatten L, Njølstad I, Chen C, Goodman G, Lathrop M, Vooder T, Välk K, Nelis M, Metspalu A, Broderick P, Eisen T, Wu X, Zhang D, Chen W, Spitz MR, Wei Y, Su L, Xie D, She J, Matsuo K, Matsuda F, Ito H, Risch A, Heinrich J, **Rosenberger A**, Muley Th, Dienemann H, Field J, Raji O, Chen Y, Gosney J, Li-loglou T, Davies M, Marcus M, McLaughlin J, Orlow I, Han Y, Li Y, Zong X, Johansson M; and EPIC Investigators, Liu G, Tworoger SS, Le Marchand L, Henderson BE, Wilkens LR, Dai J, Shen H, Houlston RS, Landi MT, Brennan P, Hung RJ  
Carcinogenesis. 2015 Nov;36(11):1314-26

- META-GSA: Combining findings from gene-set analyses across several genome-wide association studies  
**Rosenberger A**, Friedrichs St, Amos C.I., Brennan P, Fehring G, Heinrich J, Hung R.J., Muley Th., Müller-Nurasyid M, Risch A, Bickeböll H  
 PLoS One. 2015 Oct 26;10(10):e0140179
- Identification of shared and unique susceptibility pathways among cancers of the lung, breast, and prostate from genome-wide association studies and tissue-specific protein interactions.  
 Qian DC, Byun J, Han Y, Field JK, Hung RJ, Greene CS, Brhane Y, Fehring G, Landi MT, **Rosenberger A**, Bickeböll H, Malhotra J, Risch A, Heinrich J., Hunter DJ, Henderson BE, Haiman CA, Schumacher FR, Eeles RA, Seminara D, Easton DF, Amos CI  
 Hum Mol Genet. 2015 Dec 20;24(25):7406-20.]
- CHRNA5 Risk Variant Predicts Earlier Lung Cancer Diagnosis and Delayed Smoking Cessation - A Meta-analysis  
 Chen LS, Hung RJ, Baker T, Horton A, Culverhouse R, Saccone N, Cheng I, Deng B, Han Y, Hansen HM, Horsman J, Kim C, Lutz S, **Rosenberger A**, Aben K, Andrew A, Breslau N, Chang SC, Dieffenbach AK, Dienemann H, Frederiksen B, Han J, Hatsukami D, Johnson E, Pande M, Wrensch MR, McLaughlin J, Skaug V, van der Heijden HF, Wampfler J, Wenzlaff A, Woll P, Zienolddiny S, Bickeböll H, Brenner H, Duell EJ, Haugen A, Heinrich J, Hokanson JE, Hunter D, Kiemeny LA, Lazarus P, Le Marchand L, Liu G, Mayordomo J, Risch A, Schwartz AG, Teare D, Wu X, Wiencke JK, Yang P, Zhang ZF, Spitz MR, Kraft P, Amos CI, Bierut LJ  
 J Natl Cancer Inst. 2015 Apr 14;107(5). pii: djv100. doi: 10.1093/jnci/djv100. Print 2015 May
- Informed genome-wide association analysis with family history as a secondary phenotype identifies novel loci of lung cancer  
 Poirier JG, Brennan P, McKay JD, Spitz MR, Bickeböll H, Risch A, Liu G, Le Marchand L, Tworoger S, McLaughlin J, **Rosenberger A**, Heinrich J, Brüske I, Muley T, Henderson BE, Wilkens LR, Zong X, Li Y, Hao K, Timens W, Bossé Y, Sin DD, Obeidat M, Amos CI, Hung RJ  
 Genet Epidemiol. 2015 Mar;39(3):197-206. doi: 10.1002/gepi.21882. Epub 2015 Jan 19
- Assessment of melanoma histotypes and associated patient related factors: Basis for a predictive statistical model  
 Haenssle H, Hoffmann S, Buhl, T, Emmert St, Schön M; Bertsch HP, **Rosenberger A**  
 J Dtsch Dermatol Ges. 2015 Jan;13(1):37-45. doi: 10.1111/ddg.12561
- Melanoma thickness: The role of patients` characteristics, risk indicators and patterns of diagnosis  
 Haenssle H, Hoffmann S, Holzkamp R, Samhaber K, Lockmann A, Fliesser M, Emmert St, Schön M P, **Rosenberger A**, Buhl T  
 J Eur Acad Dermatol Venereol. 2015 Jan;29(1):102-8. doi: 10.1111/jdv.12471. Epub 2014 Mar 19

## 2014

- Permutationstest bei komplexen Teststrategien  
**Rosenberger A**  
 Proceedings der 18. Konferenz der SAS®-Anwender in Forschung und Entwicklung (KSFE 2014), Tim Friede, Reinhard Hilgers, Ralf Minkenber (Hrsg.), Shaker Verlag, Aachen 2014. ISBN 978-3-8440-3086-0
- Re: Time to Smoke First Morning Cigarette (TTFC) and Lung cancer in a Case-Control Study  
**Rosenberger A**  
 J Natl Cancer Inst. 2014 Nov 7;106(12). pii: dju349. doi: 10.1093/jnci/dju349. Print 2014 Dec  
<http://jnci.oxfordjournals.org/cgi/content/full/dju?ijkey=zatzRIVILGLYg&keytype=ref>
- Kernel score statistic for dependent data  
 Malzahn D, Friedrichs St, **Rosenberger A**, Bickeböll H  
 BMC Proc. 2014 Jun 17;8(Suppl 1):S41. doi: 10.1186/1753-6561-8-S1-S41. eCollection 2014
- Genetic factors of individual radiation sensitivity in lung cancer families  
 Hornhardt S, Rößler U, Sauter W, **Rosenberger A**, Illig T, Bickeböll H, Wichmann H-E, Gommelka M  
 DNA Repair (Amst). 2014 Apr;16:54-65. doi: 10.1016/j.dnarep.2014.02.001. Epub 2014 Mar 4

- Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer  
Wang Y, Lloyd A, Broderick P, Chubb D, Eisen T, Matakidou A, Amos CI, Han Y, Chen W, Spitz M, Caporaso N, Albanes D, Wang Z, Chatterjee N, Chanock S, Landi MT, Stevens V, Wheeler W, Rafnar T, Thorleifsson G, Sulem P, Stefansson K, Christiani D, Su L, Laplana M, Risch A, Kaaks R, Dienemann H, **Rosenberger A**, Hung R, Zong X, Liu G, McKay J, Timofeeva M, Delahaye M-Sourdeix, Gaborieau V, Vallee M, LeCalvez F-Kelm, Scelo G, Brennan P, Zaridze D, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates IN, Bencko V, Foretova L, Janout V, Skorpen F, Vatten L, Njølstad I, Chen C, Goodman G, Lathrop M, Benhamou S, Vooder T, Valk K, Nelis M, Metspalu A, Lener M, Johansson M, Vineis P, Agudo A, Clavel-Chapelon F, Bueno-de-Mesquita B, Trichopoulos D, Khaw KT, Johansson M, Weiderpass E, Tjønneland A, Riboli E, Gapstur S, Henrion M, Elvestad M, Lubinski J, Wei Y, Wu X, Krokan H, Kinnersley B, Ye Y, Gu J  
Nat Genet. 2014 Jul;46(7):736-41. doi: 10.1038/ng.3002. Epub 2014 Jun 1

## 2013

- Identification of lung cancer histology-specific variants applying Bayesian framework variant prioritization approaches within the TRICL and ILCCO consortia  
Brenner D.R., Amos C.I. et al. **Rosenberger A**, Hung RJ  
Carcinogenesis. 2015 Nov;36(11):1314-26. doi: 10.1093/carcin/bgv128. Epub 2015 Sep 10
- Mild overexpression of Mecp2 in mice causes a higher susceptibility toward seizures  
Bodda Ch, Tantra M, Mollajew R, Arunachalam JP, Can K, Mironov SL, **Rosenberger A**, Ehrenreich H, Mannan AU  
Am J Pathol. 2013 Jul;183(1):195-210. doi: 10.1016/j.ajpath.2013.03.019. Epub 2013 May 15
- Limited diagnostic value of Wells-score and D-dimer testing in hospitalized dermatologic patients with symptoms of deep vein thrombosis  
Haenssle H, Ayad N, Buhl T, **Rosenberger A**, Thoms KM, Schön MP, Zutt M  
Eur J Dermatol. 2013 Nov-Dec;23(6):830-6. doi: 10.1684/ejd.2013.2188
- Inzidenz der Candidämie nach gastrointestinal-chirurgischen Eingriffen in deutschen Akutkrankenhäusern (Incidence of Candidemia following abdominal surgery in German hospitals)  
Kuhns M\*, **Rosenberger A\***, Bader O, Reichard U, Groß U und Weig M (\* first authors)  
Zentralbl Chir. 2015 Dec;140(6):617-23. doi: 10.1055/s-0033-1350814. Epub 2013 Nov 15
- Influence of Common Genetic Variation on Lung Cancer Risk: Meta-Analysis of 14,900 Cases and 29,485 Controls  
Timofeeva M N, Hung R, Rafnar T, Christiani D, Field J K, Bickeboller H, Risch A, McKay J D, Wang Y, Dai J, Gaborieau V, McLaughlin J, Brenner D, Narod S, Caporaso N E, Albanes D, Thun M, Eisen T, Wichmann H-E, **Rosenberger A**, Han Y, Chen W, Zhu D, Spitz M, Wu X, Pande M, Zhao Y, Zaridze D, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Krokan H E, Elvestad M B, Skorpen F, Vatten L, Njølstad I, Chen C, Goodman G, Lathrop M, Benhamou S, Vooder T, Vålk K, Nelis M, Metspalu A, Raji O, Chen Y, Liloglou T, Muley Th, Dienemann H, Porleifsson G, Shen H, Stefansson K, Brennan P, Amos Ch I, Houlston R, Landi M T, for TRICL and the GAME-ON Consortium  
Hum Mol Genet. 2012 Nov 15;21(22):4980-95. doi: 10.1093/hmg/dds334. Epub 2012 Aug 16
- Hierarchical modeling identifies novel lung cancer susceptibility variants in inflammation pathways among 10,140 cases and 11,012 controls  
Brenner DR, Brennan P, Boffetta P, Amos CI, Spitz MR, Chen C, Goodman G, Heinrich J, Bickeboller H, **Rosenberger A**, Risch A, Muley T, McLaughlin JR, Benhamou S, Bouchardy C, Lewinger JP, Witte JS, Chen G, Bull S, Hung RJ  
Hum Genet. 2013 May;132(5):579-89. doi: 10.1007/s00439-013-1270-y. Epub 2013 Feb 1

## 2012

- PI3K inhibition enhances doxorubicin-induced apoptosis in sarcoma cells  
Marklein D, Graab U, Naumann I, Yan T, Nitzki F, **Rosenberger A**, Dittmann K, Wienands J, Wojnowski L, Fulda S, Hahn H  
PLoS One. 2012;7(12):e52898. doi: 10.1371/journal.pone.0052898. Epub 2012 Dec 31

- Informed conditioning on clinical covariates increases power in case-control association studies  
Zaitlen N, Lindström S, Pasaniuc B, Cornelis M, Genovese G, Pollack S, Barton A, Bowden DW, Eyre S, Freedman BI, Friedman DJ, Field JK, Groop L, Haugen A, Henderson B, Hicks P, Hocking LJ, Kolonel L, Landi MT, Langefeld CD, Le Marchand L, Meister M, Morgan AW, Raji OJ, Risch A, **Rosenberger A**, Scherf D, Steer S, Walshaw M, Waters KM, Wilson AG, Wordsworth P, Zienolddiny S, Tchetgen Tchetgen E, Haiman Ch, Hunter DJ, Plenge RM, Worthington J, Christiani D, Schaumberg DA, Chasman DI, Altshuler D, Voight B, Kraft P, Patterson N, Price AL  
*PLoS Genet.* 2012;8(11):e1003032. doi: 10.1371/journal.pgen.1003032. Epub 2012 Nov 8
- Previous Lung Diseases and Lung Cancer Risk: A Pooled Analysis from the International Lung Cancer Consortium  
Brenner DR E, Boffetta P, Duell EJ, Bickebölller H, **Rosenberger A**, McCormack V, Muscat JE, Yang P, Wichmann H-E, Brueske-Hohlfeld I, Schwartz AG, Cote ML, Tjønneland A, Friis S, LeMarchand L, Zhang Z-F, Morgenstern H, Szeszenia-Dabrowska N, Lissowska J, Zaridze D, Rudnai P, Fabianova E, Foretova L, Janout V, Bencko V, Schejbalova M, Brennan P, Mates IN, Lazarus P, Field JK, Raji O, McLaughlin JR, Liu G, Wiencke J, Neri M, Ugolini D, Andrew AS, Lan Q, Orlow I, Park BJ, Hung RJ  
*Am J Epidemiol.* 2012 Oct 1;176(7):573-85. Epub 2012 Sep 17
- Increased genetic vulnerability to smoking at CHRNA5 in early-onset smokers  
Hartz SM, Short S, Saccone NL, Culverhouse R, Chen L, Schwantes-An T-H, Coon H, Han Y, Stephens SH, Sun J, Chen X, Ducci F, Dueker N, Franceschini N, Frank J, Geller F, Guðbjartsson D, Hansel NN, Jiang C, Keskitalo-Vuokko K, Liu J, -Pekka Lyytikäinen L, Michel M, **Rosenberger A**, Scheet P, R. Shaffer J, Teumer A, R. Thompson J, M. Vink J, Vogelzangs N, Wenzlaff A, Wheeler W, Xiangjun X, Yang B-Z, Aggen SH, Balmforth AJ, Baumeister SE, Beaty T, Bennett S, Bergen A, Boyd H, Broms U, Campbell H, Chatterjee N, Chen J, Cheng Y-C, Cichon S, Couper D, Cucca F, Dick D, Foroud T, Furberg H, Giegling I, Gu F, Hall AS, Hällfors J, Han S, Hartmann AM, Hayward C, Heikkilä K, Hottenga JJ, Jousilahti P, Kaakinen M, J. Kittner S, Konte B, Korhonen T, Landi MT, Laatikainen T, Leppert M, Levy SM, Mathias R, McNeil DW, Medland SE, Montgomery G, Muley T, Murray T, Nauck M, North K, Paré PD, Peltonen L, Pergadia M, Polasek O, Ramos E, Ripatti S, Risch A, Ruczinski I, Rudan I, Salomaa V, Schlessinger D, Styrkársdóttir U, Terracciano A, Uda M, Willemsen G, Abecasis G, Barnes K, Bickebölller H, Boerwinkel E, Boomsma DI, Caporaso N, Duan J, Edenberg HJ, Francks C, Gejman PV, Gelernter J, Jörge Grabe H, Hops H, Jarvelin M-R, Jorma V, Kähönen M, Kendler KS, Lehtimäki T, Levinson DF, Marazita ML, Marchini J, Melbye M, Mitchell BD, Murray JC, Nöthen MM, Penninx BW, Raitakari O, Ritschel M, Rujescu D, Samani NJ, Sanders AR, Schwartz A, Shete S, Shi J, Spitz M, Stefansson K, Swan G, Thorgeirsson T, Veijola J, Völzke H, Wei Q, Wichmann H-E, Amos C, Breslau N, Cannon D, Ehringer M, Gruzca R, Hatsukami D, Heath A, Johnson EO, Kaprio J, Madden P, Martin NG, Stevens VL, Stitzel JA, Weiss RB, Kraft P, Bierut LJ  
*Arch Gen Psychiatry.* 2012 Aug;69(8):854-60. doi: 10.1001/archgenpsychiatry.2012.124
- Controlled-rate freezer cryopreservation of highly concentrated peripheral blood mononuclear cells results in higher cell yields and superior autologous T-cell stimulation for dendritic cell-based immunotherapy  
Buhl T, Legler TJ, **Rosenberger A**, Schardt A, Schön MP, Haenssle HA  
*Cancer Immunol Immunother.* 2012 Nov;61(11):2021-31. doi: 10.1007/s00262-012-1262-0. Epub 2012 Apr 22
- Heritability of radiation induced DNA damage response in lung cancer families  
**Rosenberger A**, Rössler U, Hornhardt S, Sauter W, Bickebölller H, Wichmann H-E, Gomolka M  
*Genes (Basel).* 2012 Mar 29;3(2):248-60. doi: 10.3390/genes3020248. Special issue: Radiation-Related Cancer 25 Years After Chernobyl
- Asthma and Lung Cancer Risk: A Systematic Investigation by the International Lung Cancer Consortium  
**Rosenberger A** R, Bickebölller H, McCormack V, Brenner DR, J. Duell EJ., Tjønneland A, Friis S, Muscat JE, Yang P, Wichmann H-E, Heinrich J, Szeszenia-Dabrowska N, Lissowska J, Zaridze D, Rudnai P, Fabianova E, Janout V, Bencko V, Brennan P, Mates D, Schwartz AG, Cote ML, Zhang Z-F, Morgenstern H, Oh SS, Field JK, Raji O, McLaughlin JR, Wiencke J, LeMarchand L, Neri M, Bonassi St, Andrew AS, Lan Q, Hu W, Orlow I, Park BJ, Boffetta P, Hung RJ  
*Carcinogenesis.* 2012 Mar;33(3):587-97. doi: 10.1093/carcin/bgr307. Epub 2011 Dec 22

- Calcitriol inhibits Hedgehog signaling and induces vitamin D receptor-signaling and differentiation in the Patched mouse model of embryonal rhabdomyosarcoma  
Uhmann A, Niemann H, Lammering B, Henkel C, Heß I, Nitzki F, **Rosenberger A**, Dullin C, Schraepler A, Reifenberger J, Schweyer S, Pietsch T, Strutz F, Schulz-Schaeffer W, Hahn H  
Sarcoma. 2012;2012:357040. doi: 10.1155/2012/357040. Epub 2012 Feb 21
- Integrating static and dynamic features of melanoma: The DynaMel algorithm  
Buhl T, Hansen-Hagge Ch, Korpas B, Kaune KM, **Rosenberger A**, Schön MP, Emmert St, Haenssle HA  
J Am Acad Dermatol. 2012 Jan;66(1):27-36. doi: 10.1016/j.jaad.2010.09.731. Epub 2011 Jun 11
- Validation of three BRCA1/2 mutation-carrier probability models MYRIAD, BRCAPRO and BOADICEA in a population-based series of 183 German families  
Schneegans SM\*, **Rosenberger A\***, Engel U, Sander M, Emons G, Shoukier M  
\* both first authors  
Fam Cancer. 2012 Jun;11(2):181-8. doi: 10.1007/s10689-011-9498-y
- A novel method for testing the veridicality of dental colour assessments  
Knösel M, Reus M, **Rosenberger A**, Ziebolz D  
Eur J Orthod. 2012 Feb;34(1):19-24. doi: 10.1093/ejo/cjq142. Epub 2011 Feb 2
- **mentioned in the acknowledgements**  
Novel Roles of Caenorhabditis elegans Heterochromatin Protein HP and Linker Histone in the Regulation of Innate Immune Gene Expression  
Studencka M, Konzer A, Moneron G, Wenzel D, Opitz L, Salinas-Riester G, Schulenburg H, Palladino F, Krüger M, Hell St, Wisniewski J.R., Schulze E and Jedrusik-Bode M  
Mol Cell Biol. 2012 Jan;32(2):251-65. doi: 10.1128/MCB.05229-11. Epub 2011 Nov 14

## 2011

- Interaction of human multidrug and toxin extrusion 1 (MATE1) transporter with antineoplastic agents  
Grottker J, **Rosenberger A**, Burckhardt G, Hagos Y  
Drug Metabol Drug Interact. 2011;26(4):181-9. doi: 10.1515/DMDI.2011.024. Epub 2011 Nov 7
- Antitumoral Effects of Calcitriol in Basal Cell Carcinomas Involve Inhibition of Hedgehog Signaling and Induction of Vitamin D Receptor Signaling and Differentiation  
Uhmann A, Niemann H, Lammering B, Henkel C, Heß I, Nitzki F, **Rosenberger A**, Dullin C, Schraepler A, Reifenberger J, Schweyer S, Pietsch T, Strutz F, Schulz-Schaeffer W, Hahn H  
Mol Cancer Ther. 2011 Nov;10(11):2179-88. doi: 10.1158/1535-7163.MCT-11-0422. Epub 2011 Aug 30
- Thrombophilia in patients with chronic venous leg ulcers-a study on patients with or without post-thrombotic syndrome  
Zutt M, Krüger U, **Rosenberger A**, Schön M.P., Neumann Ch., von Ahsen N., Kretschmer L  
J Eur Acad Dermatol Venereol. 2011 Dec;25(12):1432-9. doi: 10.1111/j.1468-3083.2011.04001.x. Epub 2011 Mar 11
- Aspirin and NSAID use and risk of lung cancer: International Lung Cancer Consortium pooled analysis  
McCormack V, Boffetta P, Brenner DR, Bickeböller H, **Rosenberger A**, Muscat JE, Lazarus P, Tjønneland A, Friis S, Christiani DC, Chun E, LeMarchand L, Rennert G, Rennert HS, Andrew AS, Orlow I, Hung RJ, Duell EJ  
Cancer Causes Control. 2011 Dec;22(12):1709-20. doi: 10.1007/s10552-011-9847-z. Epub 2011 Oct 11
- Increased Lipoprotein (a) concentrations in patients with chronic venous leg ulcers: a study on patients with or without postthrombotic syndrome  
Zutt M, Krüger U, von Ahsen N, **Rosenberger A**, Schön MP, Neumann C, Kretschmer L  
Wound Repair Regen. 2011 Mar-Apr;19(2):168-72. doi: 10.1111/j.1524-475X.2010.00657.x

- Durability of bleaching results achieved with 15% carbamide peroxide and 38% hydrogen peroxide in vitro  
Knösel M, Reus M, **Rosenberger A**, Attin T, Ziebolz D  
Eur J Esthet Dent. 2011 Autumn;6(3):342-56

INAUGURAL – DISSERTATION of Reus M. unter dem Titel „Vergleich der Farbstabilität milder und hochkonzentrierter Bleichmittel - Eine In-vitro-Untersuchung“  
(2010, Georg-August-Universität zu Göttingen)

- A Phenotype-Based Genetic Association Study Reveals the Contribution of Neuregulin1 Gene Variants to Age of Onset and Positive Symptom Severity in Schizophrenia  
Papiol S, Begemann M, **Rosenberger A**, Friedrichs H, Ribbe K, Grube S, Schwab MH, Jahn H, Gunkel S, Benseler F, Nave KA, Ehrenreich H  
Am J Med Genet B Neuropsychiatr Genet. 2011 Apr;156B(3):340-5. doi: 10.1002/ajmg.b.31168. Epub 2011 Jan 13
- Validation of a fully automated COMET assay: 1.75 million single cells measured over a year period  
**Rosenberger A**, Rössler U, Hornhardt S, Sauter W, Bickeböller H, Wichmann H-E, Gomolka M  
DNA Repair (Amst). 2011 Mar 7;10(3):322-37. doi: 10.1016/j.dnarep.2010.12.003. Epub 2011 Jan 14

## 2010

- GPC rs2352028 variant and risk of lung cancer in never smokers  
Landi MT, Chatterjee N, Caporaso NE, Rotunno M, Albanes D, Thun M, Wheeler W, **Rosenberger A**, Bickeböller H, Risch A, Wang Y, Gaborieau V, Thorgeirsson Th, Gudjartsson D, Sulem P, Spitz M, Wichmann HE, Rafnar Th, Stefansson K, Houlston RS, Brennan P  
Lancet Oncol. 2010 Aug;11(8):714-6; author reply 716. doi: 10.1016/S1470-2045(10)70096-6
- Activation of the hedgehog pathway confers a poor prognosis in embryonal and fusion gene-negative alveolar rhabdomyosarcoma  
Zibat A, Missiaglia E, **Rosenberger A**, Pritchard-Jones K, Shipley J, Hahn H, Fulda S  
Oncogene. 2010 Dec 2;29(48):6323-30. doi: 10.1038/onc.2010.368. Epub 2010 Sep 6
- Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium  
Truong Th, Hung RJ, Amos Chl, Wu X, Bickeböller H, **Rosenberger A**, Sauter W, Illig Th, Wichmann H-E, Risch A, Dienemann H, Kaaks R, Yang P, Jiang R, Wiencke JK, Wrensch M, Hansen H, Kelsey KT, Matsuo K, Tajima K, Schwartz AG, Wenzlaff A, Seow A, Ying C, Staratschek-Jox A, Nürnberg P, Stoelben E, Wolf J, Lazarus Ph, Muscat JE, Gallagher C, Zienolddiny Sh, Haugen A, van der Heijden HFM, Kiemeny LA, Isla D, Mayordomo JI, Rafnar Th, Stefansson K, Zhang Z-F, Chang Sh-Ch, Kim JH, Hong Y-Ch, Duell EJ, Andrew AS, Lejbkovicz F, Rennert G, Müller H, Brenner H, Marchand LL, Benhamou S, Bouchardy Ch, Teare MD, Xue X, McLaughlin J, Liu G, McKay JD, Brennan P, Spitz M  
J Natl Cancer Inst. 2010 Jul 7;102(13):959-71. doi: 10.1093/jnci/djq178. Epub 2010 Jun 14
- Inverse relationship between cerebrovascular lesions and severity of Lewy body pathology in patients with Lewy body diseases  
Ghebremedhin E, Rosenberger A, Rüb U, Vuksic M, Berhe T, Bickeböller H, de Vos, RAI, Thal DR, Deller Th  
J Neuropathol Exp Neurol. 2010 May;69(5):442-8. doi: 10.1097/NEN.0b013e3181d88e63
- Tumor stroma-derived Wnta induces differentiation of basal cell carcinoma of Ptch mutant mice via CaMKII  
Nitzki F, Zibat A, König S, Carstens P-O, Frommhold A, Uhmman A, Wijgerde M, **Rosenberger A**, Brembeck F, Klingler St, Reifenberger J, Pukrop T, Aberger F, Schulz-Schaeffer W, Hahn H  
Cancer Res. 2010 Apr 1;70(7):2739-48. doi: 10.1158/0008-5472.CAN-09-3743. Epub 2010 Mar 16
- Selection of patients for long-term digital dermoscopy by assessment of melanoma risk factors  
Haenssle HA, Korpas B, Hansen-HaggeCh, Buhl T, Kjell M, Kaune KM, **Rosenberger A**, Krueger U, Schön MP, Emmert St  
Arch Dermatol. 2010 Mar;146(3):257-64. doi: 10.1001/archdermatol.2009.370

- Seven-point checklist for dermoscopy: Performance during years of prospective surveillance of patients at increased melanoma risk  
Haenssle HA, Korpas B, Hansen-HaggeCh, Buhl T, Kaune KM, **Rosenberger A**, Krueger U, Schön MP, Emmert St  
J Am Acad Dermatol. 2010 May;62(5):785-93. doi: 10.1016/j.jaad.2009.08.049. Epub 2010 Mar 11
- Genetic polymorphism of MPO, GSTT, GSTM, GSTP, EPHX and NQO as risk factors of early onset lung cancer  
Timofeeva M, Kropp S, Sauter W, Beckmann L, **Rosenberger A**, Illig Th, Jäger B, Mittelstrass K, Diene-mann H, LUCY-Consortium, Bartsch H, Bickeböller H, Chang-Claude J, Risch A, H-E Wichmann H-E  
Int J Cancer. 2010 Oct 1;127(7):1547-61. doi: 10.1002/ijc.25175

## 2009

- Integration of a priori gene set information into genome-wide association studies  
Sohns M, **Rosenberger A**, Bickeböller H  
BMC Proc. 2009 Dec 15;3 Suppl 7:S95. doi: 10.1186/1753-6561-3-S7-S95
- Effect of DNA repair host factors on temozolomide or dacarbazine melanoma treatment in Caucasians  
Boeckmann L, Schirmer M, **Rosenberger A**, Struever D, Thoms KM, Gutzmer R, Has C, Kunz M, Kuschal C, Laspe P, Schoen MP, Brockmoeller J, Emmert S  
Pharmacogenet Genomics. 2009 Oct;19(10):760-9. doi: 10.1097/FPC.0b013e3283307cd9
- A genome-wide association study of lung cancer identifies a region of chromosome p associated with risk for adenocarcinoma  
Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, Mirabello L, Jacobs K, Wheeler W, Yeager M, Bergen WA, Li Q, Consonni D, Pesatori AC, Wacholder S, Thun M, Diver R, Oken M, Wang Z, Burdett L, Doheny FK, Pugh WE, Laurie C, Brennan P, Hung R, Gaborieau V, McKay DJ, Lathrop M, McLaughlin J, Wang Y, Tsao M-S, Spitz MR, YWang u, Krokan H, Vatten L, Skorpen F, Arnesen E, Ben-hamou S, Bouchardy Ch, Metsapalu A, Vooder T, Nelis M, Vålk K, Field JK, Chen C, Goodman G, Sulem P, Thorleifsson G, Rafnar T, Eisen T, Sauter W, **Rosenberger A**, Bickeböller H, Risch A, Chang-Claude J, Wichmann H-E, Stefansson K, Houlston R, Amos ChI, Fraumeni FJ Jr., Savage ShA, Bertazzi AP, Tucker MA, Chanock St, Caporaso NE  
Am J Hum Genet. 2009 Nov;85(5):679-91. doi: 10.1016/j.ajhg.2009.09.012. Epub 2009 Oct 15
- Intracellular delivery of MHC class I-binding epitopes: DCs loaded and matured with cationic peptide/poly(I:C) complexes efficiently activate T cells  
Haenssle HA, Riedl P, Buhl T, Scharadt A, **Rosenberger A**, Schön MP, Schirmbeck R  
Exp Dermatol. 2010 Jan;19(1):19-28. doi: 10.1111/j.1600-0625.2009.00954.x. Epub 2009 Sep 16
- Associations between the oxytocin receptor gene (OXTR) and affect, loneliness and intelligence in normal subjects  
Lucht MJ, Barnow S, Sonnenfeld Ch, **Rosenberge A**, Grabe HJ, Schroeder W, Voelzke H, Freyberger HJ, Herrmann FH, Kroemer h Rosskopf D  
Prog Neuropsychopharmacol Biol Psychiatry. 2009 Aug 1;33(5):860-6. doi: 10.1016/j.pnpbp.2009.04.004. Epub 2009 Apr 17
- CYP polymorphisms as risk factors for early onset lung cancer: gender specific differences  
Timofeeva M, Kropp S, Sauter W, Beckmann L, **Rosenberger A**, Illig T, Jaeger B, Mittelstrass K, Diene-mann H, LUCY-Cornsortium, Bartsch H, Bickeböller H, Chang-Claude J, Risch A, Wichmann H-E  
Carcinogenesis. 2009 Jul;30(7):1161-9. doi: 10.1093/carcin/bgp102. Epub 2009 May 4
- TLR Gene Polymorphisms and Liver Disease Manifestations in Chronic Hepatitis C  
Askar E, Bregadze R, Mertens J, Schweyer S, **Rosenberger A**, Ramadori G, Mihm S  
J Med Virol. 2009 Jul;81(7):1204-11. doi: 10.1002/jmv.21491
- A multiplex real-time PCR method for detection of GSTM and GSTT  
Timofeeva M, Jäger B, **Rosenberger A**, Sauter W, Wichmann H-E, KORA Study Group, Bickeböller H, Risch A  
Clin Biochem. 2009 Apr;42(6):500-9. doi: 10.1016/j.clinbiochem.2008.12.011. Epub 2008 Dec 30

- Antitumor effects of a combined -aza-'deoxycytidine and valproic acid treatment on rhabdomyosarcoma and medulloblastoma in Ptch mutant mice  
Ecke, I, Petry F, **Rosenberger A**, Tauber S, Monkemeyer S, Hess I, Dullin Ch, Kimmina S, Pirngruber J, Johnsen SA, Uhmann A, Nitzki F, Wojnowski L, Schulz-Schaeffer W, Witt O, Hahn H  
Cancer Research. 2009 Feb 1;69(3):887-95. doi: 10.1158/0008-5472.CAN-08-0946. Epub 2009 Jan 20

## 2008

- Impact of Endothelin- LysAsn polymorphism for coronary artery disease and endorgan damage in hypertensives  
Popov AF, Schulz EG, Hinz J, Schmitto JD, Seipelt R, Koziolk MJ, **Rosenberger A**, Schoendube FA, Müller GA  
Coronary Artery Disease. 2008 Nov;19(7):429-34. doi: 10.1097/MCA.0b013e32830936e5
- Matrix metalloproteinase (MMP) is associated with early-onset lung cancer  
Sauter W, **Rosenberger A**, Beckmann L, Kropp S, Mittelstrass K, Timofeeva M, Wölke G, Steinwachs A, Scheiner D, Meese E, Sybrecht G, Kronenberg F, Dienemann H; LUCY-Consortium, Chang-Claude J, Illig T, Wichmann HE, Bickeböller H, Risch A  
Cancer Epidemiol Biomarkers Prev. 2008 May;17(5):1127-35. doi: 10.1158/1055-9965.EPI-07-2840
- Early onset lung cancer, cigarette smoking and the SNP of the murine double minute- (MDM) gene  
Mittelstrass K, Sauter W, **Rosenberger A**, Illig T, Timofeeva M, Klopp N, Dienemann H, Meese E, Sybrecht G, Woelke G, Cebulla M, Degen M, Morr H, Drings P, Groeschel A, grosse Kreymborg C, Haeussinger K, Hoeffken G, Schmidt C, Jilge B, Schmidt W, Ko Y, Tauscher D, Chang-Claude J, Wichmann H-E, Bickeboeller H, Risch A  
BMC Cancer. 2008 Apr 23;8:113. doi: 10.1186/1471-2407-8-113
- Detection of ATM Gene Mutations in Young Lung Cancer Patients: A Population-based Control Study  
Schneider J, Illig Th, **Rosenberger A**, Bickeböller H, Wichmann H.-E  
Arch Med Research. 2008 Feb;39(2):226-31. doi: 10.1016/j.arcmed.2007.08.004. Epub 2007 Oct 31
- Do genetic factors protect for early onset lung cancer? - A case-control study before the age of years  
**Rosenberger A**, Illig T, Korb K, Klopp N, Zietemann V, Wölke G, Meese E, Sybrecht G, Kronenberg F, Cebulla M, Degen M, Drings P, Gröschel A, Konietzko N, Kreymborg KG, Häussinger K, Höffken G, Jilge B, Ko YD, Morr H, Schmidt C, Schmidt EW, Täuscher D, Bickeböller H, Wichmann H-E  
BMC Cancer. 2008 Feb 25;8:60. doi: 10.1186/1471-2407-8-60
- CD ligation during dendritic cell maturation reduces cell death and prevents interleukin--induced regression to macrophage-like monocytes  
Haenssle H, Buhl T, Knudsen S, Krueger U, **Rosenberger A**, Reich K, Neumann C  
Exp Dermatol. 2008 Mar;17(3):177-87. Epub 2007 Dec 13
- Cyclopamine treatment of full-blown Hh/Ptch-associated RMS results in interference with Hh/Ptch signalling, but not in inhibition of tumor growth  
Ecke I, **Rosenberger A**, Obenauer S, Dullin Ch, Aberger F, Kimmina S, Stecca B, Ruiz i Altaba A, Schweyer St, Hahn H  
Molecular Carcinogenesis. 2008 May;47(5):361-72

## 2007

- Spontaneous elimination of hepatitis C virus infection: A retrospective study on demographic, clinical, and serological correlates  
Wietzke-Braun P, Mänhardt LB, **Rosenberger A**, Uy A, Ramadori G, Mihm S  
World J Gastroenterol. 2007 Aug 21;13(31):4224-9
- Sex-dependent genetic markers of CYP1A2 expression and activity in human liver microsomes  
Schirmer M, **Rosenberger A**, Klein K, Kulle B, Toliat MR, Nürnberg P, Zanger UM, Wojnowski L  
Pharmacogenomics. 2007 May;8(5):443-53
- Meta-analysis of whole-genome linkage scans for Parkinson's disease  
**Rosenberger A**, Sharma M, Müller-Mysok B, Gasser T, Bickeböller H  
BMC-Genetics. 2007 Jul 2;8:44



## 2006

- Significant association of a MV independent polymorphism in the 5' UTR of the PRNP gene with sporadic Creutzfeldt-Jakob disease in a large German case-control study  
Vollmert C, Windl O, Xiang W, **Rosenberger A**, Zerr I, Wichmann H-E, Bickeböllner H, Illig T, the KORA group, Kretzschmar H A  
Journal of Medical Genetics 2006 Oct;43(10):e53
- A functional polymorphism within plasminogen activator urokinase (PLAU) is associated with Alzheimer's disease  
Riemenschneider M, Konta L, Friedrich P, Schwarz S, Taddei K, Neff F, Padovani A, Kolsch H, Laws SM, Klopp N, Bickeboller H, Wagenpfeil S, Mueller JC, **Rosenberger A**, Diehl-Schmid J, Archetti S, Lautenschlager N, Borroni B, Muller U, Illig T, Heun R, Egensperger R, Schlegel J, Forstl H, Martins RN, Kurz A  
Hum Mol Genet. 2006 Aug 15;15(16):2446-56. Epub 2006 Jul 6
- Results from an observational trial: digital epiluminescence microscopy follow-up of atypical nevi increases the sensitivity and the chance of success of conventional dermoscopy in detecting melanoma  
Haenssle HA, Krueger U, Vente C, Thoms KM, Bertsch HP, Zutt M, **Rosenberger A**, Neumann C, Emmert S  
J Invest Dermatol. 2006 May;126(5):980-5

The working group of Dr. Hänßle had been awarded with the **Förderpreis der Niedersächsischen Krebsgesellschaft**

- Relationship of apolipoprotein E and age at onset to Parkinson disease neuropathology  
Ghebremedhin E, Del Tredici K, Vuksic M, Rub U, Thal DR, Burbach GJ, **Rosenberger A**, Bickeboller H, Deller T, de Vos RA, Jansen Steur EN, Braak H  
J Neuropathol Exp Neurol. 2006 Feb;65(2):116-23

## 2005

- Detection of the CCR-Delta HIV resistance gene in Bronze Age skeletons  
Hummel S, Schmidt D, Kremeyer B, Herrmann B, Oppermann M. (**Rosenberger A** acknowledged for detailed statistical analysis)  
Genes Immun. 2005 Jun;6(4):371-4
- Surrogate phenotype definition for alcohol use disorders: a genome-wide search for linkage and association  
**Rosenberger A**, Janicke N, Köhler K, Korb K, Kulle B, Bickeböllner H  
BMC Genet. 2005 Dec 30;6 Suppl 1:S55
- Dissection of heterogeneous phenotypes for quantitative trait mapping  
Bickeböllner H, Bailey JN, Papanicolaou GJ, **Rosenberger A**, Viel KR  
Genet Epidemiol. 2005;29 Suppl 1:S41-7
- NAD(P)H oxidase and multidrug resistance protein genetic polymorphisms are associated with doxorubicin-induced cardiotoxicity  
Wojnowski L, Kulle B, Schirmer M, Schluter G, Schmidt A, **Rosenberger A**, Vonhof S, Bickeböllner H, Toliat MR, Suk EK, Tzvetkov M, Krüger A, Seifert S, Kloess M, Hahn H, Löffler M, Nürnberg P, Pfreundschuh M, Trümper L, Brockmüller J, Hasenfuss G  
Circulation. 2005 Dec 13;112(24):3754-62. Epub 2005 Dec 5
- Effects of socio-demographic variables on health-related quality of life determined by the quality of life index - German version  
Prause W, Saletu B, Tribl GG, Rieder A, **Rosenberger A**, Bolitschek J, Holzinger B, Kapfhammer G, Katschnig H, Kunze M, Popovic R, Graetzhofner E, Zeitlhofer J  
Hum Psychopharmacol. 2005 Jul;20(5):359-65
- (**mentioned in the acknowledgment**)  
Detection of the CCR-Delta HIV resistance gene in Bronze Age skeletons  
Hummel S, Schmidt D, Kremeyer B, Herrmann B, Oppermann M  
Genes Immun. 2005 Jun;6(4):371-4

## 2004

- Results of a surveillance programme for patients at high risk of malignant melanoma using digital and conventional dermoscopy  
Haenssle HA, Vente C, Bertsch HP, Rupprecht R, Abuzahra F, Junghans V, Ellinghaus B, Emmert S, Hallermann C, **Rosenberger A**, Neumann C  
Eur J Cancer Prev 2004 Apr;13(2):133-8

## 2003

- Nonparametric longitudinal allele-sharing model  
Kulle B, Köhler K, **Rosenberger A**, Loesgen S, Bickeböller H  
BMC Genet 2003 Dec 31;4 Suppl 1:S85
- Modeling and dissection of longitudinal blood pressure and hypertension phenotypes in genetic epidemiological studies  
Bickeböller H, Barrett JH, Jacobs KB, **Rosenberger A**  
Genet Epidemiol. 2003;25 Suppl 1:S72-7
- Retrospective study of malignant melanoma patients treated with mistletoe extracts  
Stumpf C, **Rosenberger A**, Rieger S, Tröger W, Schietzel M, Stein GM  
Forsch Komplementarmed Klass Naturheilkd. 2003 Oct;10(5):248-55

## 2002

- Sleeping habits in the Austrian population  
Tribl GG, Schmeiser-Rieder A, **Rosenberger A**, Saletu B, Bolitschek J, Kapfhammer G, Katschnig H, Holzinger B, Popovic R, Kunze M, Zeitlhofer J  
Sleep Med. 2002 Jan;3(1):21-8
- Methods in evidence-based medicine consensus findings -a case study in trauma surgery  
**Rosenberger A**, Selbmann HK  
Z Arztl Fortbild Qualitatssich 2003 Jul;97(4-5):311-9
- Childhood diseases, infectious diseases, and fever as potential risk factors for cancer?  
Hoffmann C, **Rosenberger A**, Tröger W, Bühring M  
Forsch Komplementarmed Klass Naturheilkd. 2002 Dec;9(6):324-30
- Validation of questionnaires from several medical fields regarding the constitution of patients  
Hoffmann C, **Rosenberger A**, Tröger W, Stange R, Bühring M  
Forsch Komplementarmed Klass Naturheilkd. 2002 Feb;9(1):37-44

## 2001

- N-CPAP rejecters--a specific group of noncompliant patients with obstructive sleep apnea syndrome  
Bolitschek J, Rieder A, Broinger G, **Rosenberger A**, Kunze M, Aigner K  
Wien Klin Wochenschr. 2001 Apr 17;113(7-8):245-8

## 2000

- Sleep and quality of life in the Austrian population  
Zeitlhofer J, Schmeiser-Rieder A, Tribl G, **Rosenberger A**, Bolitschek J, Kapfhammer G, Saletu B, Katschnig H, Holzinger B, Popovic R, Kunze M  
Acta Neurol Scand. 2000 Oct;102(4):249-57
- Mistletoe extracts in the therapy of malignant, hematological and lymphatic diseases -a monocentric, retrospective analysis over years  
Stumpf C, **Rosenberger A**, Rieger S, Tröger W, Schietzel M  
Forsch Komplementarmed Klass Naturheilkd. 2000 Jun;7(3):139-46

## 1999

- Development of lymphocyte subsets in tumor patients after subcutaneous administration of mistletoe extracts  
Büssing A, **Rosenberger A**, Stumpf C, Schietzel M  
Forsch Komplementarmed. 1999 Aug;6(4):196-204-

## 1998

- Self-reported prevalence of venous diseases in the general population of Austria--results of the SERMO (Self-Reported Morbidity) Study  
Schmeiser-Rieder A, Kunze U, Mitsche N, **Rosenberger A**, Kunze M  
Acta Med Austriaca. 1998, 25(2):65-68
- Technical report: public health monitor: SERMO Study--significance for immunization programs  
Schmeiser-Reider A, Kunze U, Mitsche N, **Rosenberger A**, Kunze M  
Wien Med Wochenschr. 1998;148(8-9):211-4
- Impact of nasal continuous positive airway pressure treatment on quality of life in patients with obstructive sleep apnoea  
Bolitschek J, Schmeiser-Rieder A, Schobersberger R, **Rosenberger A**, Kunze M, Aigner K  
Eur Respir J. 1998 Apr;11(4):890-4