

## Germline mutations in young non-smoking women with lung adenocarcinoma

Donner, Iikki

2018-08

---

Donner, I, Katainen, R, Sipilä, L J, Aavikko, M, Pukkala, E & Aaltonen, L A 2018, ' Germline mutations in young non-smoking women with lung adenocarcinoma ' Lung Cancer , vol. 122 , pp. 76-82 . <https://doi.org/10.1016/j.lungcan.2018.05.027>

---

<http://hdl.handle.net/10138/239086>

<https://doi.org/10.1016/j.lungcan.2018.05.027>

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*



## Germline mutations in young non-smoking women with lung adenocarcinoma

Iikki Donner<sup>a</sup>, Riku Katainen<sup>a</sup>, Lauri J. Sipilä<sup>a</sup>, Mervi Aavikko<sup>a</sup>, Eero Pukkala<sup>b,c</sup>, Lauri A. Aaltonen<sup>a,\*</sup>

<sup>a</sup> Department of Medical and Clinical Genetics, Medicum, and Genome-Scale Biology Research Program, Research Programs Unit, Faculty of Medicine, University of Helsinki, Helsinki, Finland

<sup>b</sup> Finnish Cancer Registry, Institute for Statistical and Epidemiological Cancer Research, Helsinki, Finland

<sup>c</sup> Faculty of Social Sciences, University of Tampere, Tampere, Finland

### ARTICLE INFO

#### Keywords:

Lung adenocarcinoma  
Never smoker  
Women  
Exome sequencing  
Early onset  
Genetic predisposition to cancer

### ABSTRACT

**Objectives:** Although the primary cause of lung cancer is smoking, a considerable proportion of all lung cancers occur in never smokers. Gender influences the risk and characteristics of lung cancer and women are over-represented among never smokers with the disease. Young age at onset and lack of established environmental risk factors suggest genetic predisposition. In this study, we used population-based sampling of young patients to discover candidate predisposition variants for lung adenocarcinoma in never-smoking women.

**Materials and methods:** We employed archival normal tissue material from 21 never-smoker women who had been diagnosed with lung adenocarcinoma before the age of 45, and exome sequenced their germline DNA.

**Results and conclusion:** Potentially pathogenic variants were found in eight Cancer Gene Census germline genes: *BRCA1*, *BRCA2*, *ERCC4*, *EXT1*, *HNF1A*, *PTCH1*, *SMARCB1* and *TP53*. The variants in *TP53*, *BRCA1*, and *BRCA2* are likely to have contributed to the early onset lung cancer in the respective patients (3/21 or 14%). This supports the notion that lung adenocarcinoma can be a component of certain cancer predisposition syndromes. Fifteen genes displayed potentially pathogenic mutations in at least two patients: *ABCC10*, *ATP7B*, *CACNA1S*, *CFTR*, *CLIP4*, *COL6A1*, *COL6A6*, *GCN1*, *GJB6*, *RYR1*, *SCN7A*, *SEC24A*, *SP100*, *TTN* and *USH2A*. Four patients showed a mutation in *COL6A1*, three in *CLIP4* and two in the rest of the genes. Some of these candidate genes may explain a subset of female lung adenocarcinoma.

### 1. Introduction

Lung cancer is the most common cancer in the world and the leading cause of cancer death [1]. It is highly fatal, with an overall mortality to incidence ratio of 0.87. The primary risk factor for lung cancer is cigarette smoke. Other established non-inherited risk factors include ionizing radiation, asbestos, different kinds of occupational exposures, air pollution and pulmonary diseases such as tuberculosis and chronic obstructive pulmonary disease [2]. Lung cancer and pulmonary fibrosis are driven by similar biological pathways [3], and a recent study found an eightfold higher incidence ratio of lung cancer among sufferers of idiopathic pulmonary fibrosis (IPF) when compared to the general population [4]. A family history of lung cancer has been estimated to double the risk of developing the disease [5,6], and the heritability estimate based on the largest twin cohort to date (Nordic Twin Study on Cancer) is 0.38 [7]. A number of susceptibility loci and genes have been identified including 5p15 [8,9], 6p2 [10,11], 15q25.1

[10,11], *EGFR* [12–14], *HER2* [15], *BRCA* [16], *BAP1* [17], and *PARK2* [18]. Also, germline *TP53* and *DICER1* mutations causing Li-Fraumeni and *DICER1* syndromes, respectively, are linked to an increased risk of lung cancer [19,20].

The two main histological subtypes of lung cancer are small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). SCLC originates from neuroendocrine-cell precursors, while NSCLC originates from bronchial epithelial-cell precursors and can be further divided into three types: squamous cell-, adeno- and large cell carcinoma. All major histological subtypes of lung cancer are associated with smoking.

Epidemiological, pathological, and molecular evidence suggest that lung cancer in never smokers should be treated as a distinct disease entity [21]. Women are more likely to suffer from non-smoking-associated lung cancer than men, reflecting sex-based differences in exposure to risk factors and possibly also susceptibility [22]. A woman with an affected parent has a higher risk than a man of developing lung cancer, and the molecular profile and prognosis of the disease in

\* Corresponding author at: Biomedicum Helsinki, P.O. Box 63 (Haartmaninkatu 8), FI-00014, University of Helsinki, Finland.  
E-mail address: [lauri.aaltonen@helsinki.fi](mailto:lauri.aaltonen@helsinki.fi) (L.A. Aaltonen).

women differs from that in men [23]. A key driver of lung adenocarcinoma tumorigenesis especially in never-smoker women is somatic mutation of *EGFR* [24,25].

Here, we studied germline predisposition to lung adenocarcinoma (LUAD) in never-smoking women, first identifying the youngest patients within the population-based Finnish Cancer Registry (FCR) database. Young age at onset is a hallmark of hereditary susceptibility to many common cancers. Normal tissue archival samples from never-smoking female patients under 45 years of age were systematically collected as guided by the database. The FCR harbors data on all cancer cases diagnosed in Finland since 1953, totaling over one million cases.

## 2. Materials and methods

### 2.1. Ethics statement

This study was approved by the National Supervisory Authority for Welfare and Health (Valvira; 1423/06.01.03.01/2012), National Institute for Health and Welfare (THL; 151/5.05.00/2017), and the ethics committee of the Hospital District of Helsinki and Uusimaa (HUS; 408/13/03/03/09).

### 2.2. Patients and samples

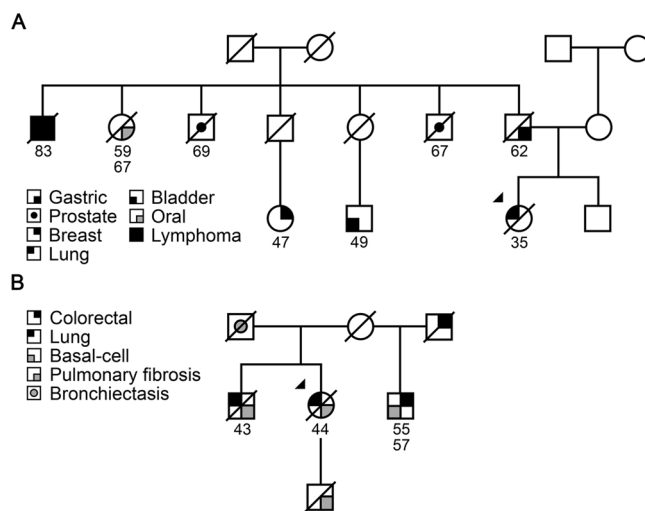
We included female never smokers listed in the FCR and diagnosed with LUAD between 1953 and 2013, and before the age of 45. Smoking histories were collected from patient records, and if there was no mention of smoking habits the patient was excluded from the study. Since the establishment of the FCR, altogether 105 women had been diagnosed with LUAD before the age of 45. Of these 49 were current or former smokers, 26 had no mention of smoking history or patient records were no longer available, and 30 were never smokers. Sufficient quality normal tissue formalin-fixed, paraffin-embedded (FFPE) samples were available for 21 of the 30 never-smoking patients. Median age at diagnosis was 39 for the exome sequenced samples (range 20–44).

We identified the first and second degree relatives of these patients from the Population Information System, and checked their cancer histories from the FCR. Family history of cancer as well as relevant

**Table 1**

Age at diagnosis and family history of malignancy or lung disease of the lung adenocarcinoma patients.

Patient	Age at diagnosis	Family cancer/lung disease history
LuAd1	38	Multiple myeloma (maternal grandmother, 87); breast carcinoma NOS (maternal aunt, 33); bladder adenocarcinoma (maternal uncle, 69)
LuAd3	36	Prostate adenocarcinoma (father, 75; paternal grandfather, 59); lymphocytic leukemia NOS (paternal cousin, 5)
LuAd5	33	Colorectal adenocarcinoma and prostate adenocarcinoma (maternal grandfather, 78 and 86)
LuAd6	38	Hodgkin lymphoma NOS (brother, 26)
LuAd7	31	Brain carcinoma NOS (paternal grandmother, 77); prostate adenocarcinoma (paternal grandfather, 88; paternal uncle, 68; paternal uncle, 68); cervical squamous cell carcinoma (paternal cousin, 40); nasal rhabdomyosarcoma (paternal cousin, 0); ductal breast cancer (maternal cousin, 39)
LuAd8	20	Breast carcinoma NOS and uterine adenocarcinoma NOS (mother, 57 and 77); breast carcinoma NOS (maternal aunt, 52); prostate adenocarcinoma (maternal grandfather, 79)
LuAd10	29	Lung squamous cell carcinoma (father, 59)
LuAd11	35	Mucinous gastric adenocarcinoma (father, 62); prostate adenocarcinoma (paternal uncle, 67; paternal uncle, 69); oral squamous cell carcinoma (paternal aunt, 59 and 67); diffuse large B-cell lymphoma (paternal uncle, 83); ductal breast carcinoma (paternal cousin, 47)
LuAd13	29	Gastric adenocarcinoma (maternal grandmother, 67; father, 81); ductal breast carcinoma (mother, 70); cervical squamous cell carcinoma (daughter, 21); renal adenocarcinoma (paternal cousin, 66); diffuse large B-cell lymphoma (paternal cousin, 66)
LuAd14	21	Pancreatic carcinoma (maternal grandfather, 82); lymphocyte-rich classical Hodgkin lymphoma (father, 24); lung cancer NOS (paternal aunt, 51); urothelial carcinoma (paternal grandfather, 68); prostate adenocarcinoma (paternal grandfather, 68)
LuAd15	41	N/A
LuAd18	40	Breast carcinoma NOS (maternal aunt, 34)
LuAd19	39	Brain cancer NOS (paternal uncle, 47); Ewing's sarcoma (paternal cousin, 29)
LuAd20	41	Ductal breast carcinoma (paternal aunt, 61), colorectal adenocarcinoma (maternal grandmother, 87)
LuAd21	41	Ductal breast carcinoma (sister, 53), colorectal adenocarcinoma (maternal grandmother, 74), prostate adenocarcinoma (paternal grandfather, 73)
LuAd23	42	Urothelial carcinoma (maternal grandfather, 79)
LuAd24	44	Colorectal adenocarcinoma (half brother, 57); lung adenocarcinoma (brother, 43); usual interstitial pneumonia (son, XX)
LuAd26	42	Prostate adenocarcinoma (maternal half uncle, 62)
LuAd27	44	Lobular breast cancer (sister, 42); renal adenocarcinoma (father, 62); ductal breast cancer (paternal cousin, 48)
LuAd29	42	N/A
LuAd31	43	N/A



**Fig. 1.** A. Pedigree of LUAD11. B. Pedigree of LUAD24. Pedigrees have been slightly modified to prevent identification of individuals. Index cases are marked with arrows. Numbers under the individuals stand for the age at cancer diagnosis.

clinical information on the 21 patients that underwent exome analysis are listed in Table 1. For 12 of the patients we collected tumor tissue blocks in addition to the normal tissue blocks for loss of heterozygosity (LOH) analysis.

In addition, we collected normal tissue FFPE blocks from six cancer affected family members of LUAD11 (Fig. 1A), the parents of LUAD14, and the son of LUAD24 (Fig. 1B) who suffered from IPF.

### 2.3. Exome sequencing

We extracted genomic DNA from normal tissue archival tissue samples with the conventional phenol-chloroform method and prepared the DNA samples for exome sequencing with the KAPA Hyper Prep Kit (Kapa Biosystems Inc., Wilmington, MA) and SeqCap EZ System (Roche Nimblegen Inc., Madison, WI). The sources of the FFPE tissue and

**Table 2**  
Source of non-malignant DNA sample.

Patient	Source of DNA sample
LuAd1	Cervical lymph node
LuAd3	Endometrium
LuAd5	Lung
LuAd6	Pleura
LuAd7	Bronchus
LuAd8	Skin
LuAd10	Bronchus
LuAd11	Rib
LuAd13	Axillary lymph node
LuAd14	Lung <sup>a</sup>
LuAd15	Endometrium
LuAd18	Liver
LuAd19	Lung
LuAd20	Lung with lymph node
LuAd21	Posterior mediastinal lymph node
LuAd23	Duodenum
LuAd24	Lung
Son of LuAd24	Antrum ventriculi
LuAd26	Lung <sup>b</sup>
LuAd27	Gallbladder
LuAd29	Appendix
LuAd31	Lung

<sup>a</sup> variants validated by Sanger sequencing of DNA sample from breast.

<sup>b</sup> variants validated by Sanger sequencing of DNA sample from skin.

subsequent DNA samples are listed in Table 2. We used 300–1000 ng of DNA sheared to an average size of 200–300 bp with the Covaris S220 Focused-ultrasonicator (Covaris Inc., Woburn, MA). For target enrichment, we used the 96 Mb SeqCap EZ Exome + UTR kit (Roche Nimblegen Inc.). Paired-end sequencing was carried out with Illumina HiSeq 2000 or 4000 (Illumina, San Diego, CA) at Karolinska High Throughput Center.

DNA from 21 female patients and the IPF afflicted son of LUAD24 was exome sequenced successfully with an average coverage of targeted areas being 48 reads across all samples, and 91% of targeted areas being covered by more than six reads.

#### 2.4. Data analysis

The quality of raw sequencing data was examined using FastQC version 0.10.0 (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/>) and QualiMap version 2.1 (<http://qualimap.bioinfo.cipf.es/>). 3' ends of reads with high adapter similarity were removed with Trim Galore! version 0.3.07 ([http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)) and trimmed reads were mapped to 1000 genomes Phase 2 reference genome (GRCh37 / hg19) with Burrows-Wheeler Aligner (BWA)–MEM version 0.7.12 [26]. Overlapping read pairs were clipped using BamUtil version 1.0.13 (<http://genome.sph.umich.edu/wiki/BamUtil#Releases>) ClipOverlap. Duplicate reads were removed using Samtools version 1.0 (<http://www.htslib.org/>) rmdup on both paired-end and single-end reads. Aligned reads were locally realigned using the GATK version 3.5 [27] IndelRealigner and base scores were recalculated with GATK BaseRecalibrator. After realignment the final SNV and indel calls were made with the GATK HaplotypeCaller using a variant quality threshold of 1.0.

Annotation, comparative variant analysis, controlling and variant visualization were performed with BasePlayer (<https://doi.org/10.1101/126482>). Gene annotation was based on Ensembl GRCh37 release 84. Minimum variant quality score was set at 20, minimum coverage at six reads, and minimum allelic fraction at 30%. The data was filtered against the variant MAFs of gnomAD, available at the Genome Aggregation Database [28]. We required passed variants to have a MAF of < 0.001 in the gnomAD Finnish subset, which is made up of 12,897 individuals. Three patients (LUAD15, LUAD29 and LUAD31) were of Southeast Asian ancestry, however. For these patients gnomAD

populations East Asian (9435 individuals) and South Asian (15,391 individuals) were used, with a MAF threshold of < 0.001. Effects of missense variants were predicted with M-CAP [29]. Missense variants outside the range of M-CAP (mainly variants in other than the canonical transcript) as well as in-frame deletions and insertions were predicted by Provean [30]. Variants were required to get a deleterious prediction using the algorithms' default threshold values in order to pass filtering.

#### 2.5. Sanger sequencing and LOH analysis

The variants were verified by Sanger sequencing of DNA extracted from FFPE blocks with the GeneRead DNA FFPE Kit (Qiagen). Primers were designed with Primer3web version 4.0.0 [31] and capillary sequencing was performed at the Institute for Molecular Medicine Finland (FIMM) using the BigDye v.3.1 sequencing reaction and ABI3730xl DNA Analyzer electrophoresis (Applied Biosystems, Foster City, CA). Sequences were analyzed manually using FinchTV 1.5. LOH was determined by comparing peaks between normal and tumor (> 50% tumor tissue) samples when tumor tissue was available (12 cases). Each reaction was performed in triplicate to ensure consistency of the observations and exclude mono-allelic amplification. Normal tissue DNA from a total of eight relatives of LUAD11 and LUAD14 was also sequenced to see whether candidate variants in the patients segregated with malignancy.

### 3. Results and discussion

More than 100 cancer predisposition genes have been identified to date. A recent pan-cancer study that included more than 10,000 cases across 33 different cancers discovered plausible susceptibility variants in 8% of all patients [32]. However, the genetic architecture of lung cancer susceptibility is not well known. In the previously mentioned study [32], probable predisposing variants were found in more than 6% of LUAD. Another study, that analyzed the amount of rare germline truncating variants in cancer predisposition genes across 12 cancer types, found a similar rate (7%) for LUAD [33]. Here we performed a population based search through the FCR to identify young female patients diagnosed with the malignancy, as early age at onset is a hallmark of inherited cancer susceptibility. Family history was also looked into, but this was not a criteria for inclusion or exclusion.

We looked into genes listed in the Cancer Gene Census (CGC), available in the Catalogue of Somatic Mutations in Cancer (COSMIC) database [34], to see whether the patients could be suffering from known cancer predisposition syndromes. We observed eight potentially pathogenic heterozygous mutations (MAF < 0.001) in eight CGC germline genes (Table 3). The variants in *TP53*, *BRCA1*, and *BRCA2*, are in light of current knowledge likely to have contributed to the early-onset lung cancer in our patients. The *TP53*-nonsense mutation p.Gln52Ter was not found in any controls and has, to our knowledge, not previously been reported as a germline variant. Due to its truncating nature, it is likely to have an effect on the protein function. Constitutional mutations in *TP53* predispose to multiple different cancers in the carriers. The gene is mutated in approximately 70% of patients with Li-Fraumeni syndrome (LFS) and 40% of patients with Li-Fraumeni-Like syndrome (LFL) [35]. These conditions are characterized by predisposition to a wide spectrum of neoplasia, including soft-tissue sarcoma, osteosarcoma, leukemia, breast cancer, brain tumors, and adrenocortical carcinoma, often at early age at onset. Although lung cancer is not a classic LFS spectrum cancer, it has been linked to LFS [19,36]. Preliminary data from the LIFESCREEN randomized clinical trial suggest that LUAD is indeed part of the spectrum, being the most frequent tumor in the LFS cohort under study [37]. As a somatic mutation, the observed variant has been described 13 times in COSMIC. The patient carrying it was diagnosed at the age of 22. Both parents carried wild type alleles, indicating a de novo background. This finding further diminishes the possibility that the variant is a neutral change.

**Table 3**  
Potentially pathogenic variants found in Cancer Gene Census germline genes.

Gene	Variant location <sup>a</sup>	Base change <sup>b</sup>	Variant effect	rs-code	Population MAF	gnomAD MAF	Patient	LOH
<i>BRCA1</i>	chr17:41258546	c.139A > G	p.(Cys47Arg)	N/A	N/A	N/A	LuAd11	N/A
<i>BRCA2</i>	chr13:32944557	c.8350C > T	p.(Arg2784Trp)	rs80359075	5.80E-05*	1.22E-05	LuAd31	N/A
<i>ERCC4</i>	chr16:14031664	c.1853G > T	p.(Arg618Leu)	rs760922582	8.59E-04	1.29E-04	LuAd19	None
<i>EXT1</i>	chr8:119123269	c.17C > G	p.(Arg6Pro)	rs200815125	7.67E-05	4.38E-06	LuAd26	None
<i>HNFlA</i>	chr12:121434127	c.1018C > G	p.(Pro340Ala)	rs746055869	4.27E-04	4.33E-04	LuAd24	N/A
<i>PTCH1</i>	chr9:98231178	c.2105 G > C	p.(Pro702Arg)	rs368362152	5.30E-05*	1.44E-05	LuAd31	N/A
<i>SMARCB1</i>	chr22:24134046	c.197C > T	p.(Ser66Leu)	N/A	N/A	N/A	LuAd3	WT loss
<i>TP53</i>	chr17:7579533	c.154G > A	p.(Gln52*)	N/A	N/A	N/A	LuAd14	None

Abbreviations: MAF, minor allele frequency; LOH, loss of heterozygosity; WT, wild type; mut, mutant.

Population MAF refers to the gnomAD Finnish population, except for patients LuAd15, LuAd29 and LuAd31 who are of Southeast Asian ancestry. For these gnomAD populations East Asian (\*) and South Asian (\*\*) were used, depending on which one had the higher allele frequency.

<sup>a</sup> Genome assembly GRCh37.

<sup>b</sup> Position in longest translation (GRCh37 release 84).

The *BRCA1* variant p.Cys47Arg was found in a patient diagnosed with LUAD at the age of 35. The variant could not be found among the controls, nor has it been reported as a germline variant or to be somatically mutated. Residue 47 is a zinc coordinating residue located in the conserved RING motif of *BRCA1*. It is critical for the binding of *BARD1*. Functional studies have shown that another variant of this residue, Cys47Gly, is defective in homology-directed recombination (HDR) [38]. *BARD1* binding and HDR are critical for the tumor suppressive function of *BRCA1*. Furthermore, the Cys47Gly variant has been found to cause centrosome duplication [39]. *BRCA1* is a well-established breast and ovarian cancer susceptibility gene. Considering the rarity of the discovered variant and its critical location, it is plausible that it was involved in the carcinogenesis of LUAD in the patient in question. The variant was inherited from the father who was diagnosed with gastric cancer at the age of 62. A number of studies have found an increased risk of gastric cancer in *BRCA1* mutation carriers [40–42]. Of the other cancer afflicted family members of whom we were able to attain tissue material, the variant was also found in an uncle with lymphoma at the age of 83 as well as a female cousin with breast cancer at the age of 47 (Fig. 1A).

A plausible susceptibility variant was also found in *BRCA2*: p.Arg2784Trp. The mouse equivalent of this residue, p.Arg2705, forms hydrogen bonds with the side chains of the residues p.Asn2702 (human p.Asn2781) and p.Ser2728 (human p.Ser2807), as well as the p.Asp41 side chain of *DSS1*, and thus holds together three loops [43]. The p.Arg2784Trp mutation in our patient may affect tertiary interactions between DNA-binding folds as well as the interaction between *BRCA2* and *DSS1*, which is required for the stability of *BRCA2* [44]. The variant also scored similar as known pathogenic *BRCA2* variants in an HDR assay [43]. The carrier was of Southeast Asian origin and diagnosed at the age of 43. Rare variants in *BRCA2* have been found to affect the risk of lung squamous cell carcinoma in a genome-wide association study [16]. In addition, a recent study on LUAD found a significant number of germline ultra-rare possibly pathogenic variants in *BRCA2* [45].

Altogether 15 genes harbored a rare (MAF < 0.001) potentially pathogenic variant in at least two patients within the series. The 33 variants are listed in Table 4. None of the recurrently mutated genes were among the CGC germline genes. The most frequently mutated candidate gene in the sample set was *COL6A1*. Three patients harbored one, and one patient harbored two heterozygous missense variants in the gene. The protein product *COL6A1* has three von Willebrand factor type A (VWA) domains, and all the variants were located in these domains. VWA domains are likely to mediate protein-protein interactions by coordination of divalent cations [46]. Exon skipping, nonsense and some missense mutations in *COL6A1* cause Bethlem myopathy 1 [47], and the gene has been found to be overexpressed in neoplastic lung tissues [48]. The second most frequently mutated gene with a variant in three patients was *CLIP4*. Little is known about the function of its

protein product.

*SCN7A* was heterozygously mutated in two patients, including LUAD24, who shared the c.1133C > T variant with her son who suffered from IPF. The family had a notable history of lung disease in general (Fig. 1B). Other potentially pathogenic variants shared by LUAD24 and her son are listed in Table 5. *SCN7A* is a sodium channel which is altered by sodium concentrations instead of membrane depolarization. The channels contribute to sodium absorption by cells when extracellular sodium levels are high, and channels in the circumventricular organs are also likely to be involved in the control of water and salt uptake behavior and thus body fluid homeostasis. *SCN7A* is expressed in specialized neurons, specialized ependymal and glial cells, non-myelinating Schwann cells, and lung alveolar type II (AT2) cells [49]. AT2 is the main cell type responsible for the secretion of surfactant proteins *SFTPC* and *SFTPA2*. Interestingly, mutations in the genes *SFTPC* and *SFTPA2* have been found to cause familial interstitial pneumonia [50], a hereditary form of IPF. The other variant, c.1003T > C, was found in LUAD21. We did not have tumor tissue from patient LUAD24, but Sanger sequencing of tumor-derived DNA from LUAD21 showed a decrease in the WT allele signal.

The remaining 13 shared genes (Table 4) were heterozygously mutated in two patients. *CFTR* is the most prominent candidate gene with regards to known function. Homozygous or compound heterozygous mutations in *CFTR* cause cystic fibrosis. A variant (c.1558G > T) affecting the same nucleotide as one of the two variants found in this study is listed as a cystic fibrosis causing mutation in the The Clinical and Functional Translation of CFTR (CFTR2) database (<http://cftr2.org>). The relationship between cystic fibrosis and lung cancer is unclear, but both frequent hypermethylation [51] and mutation [52] of *CFTR* have been observed in NSCLCs. A recent study found lower expression of *CFTR* in tumors and showed that the expression levels correlated with disease stage, lymph node metastasis, poor prognosis, and progression-free survival [53]. Knockdown of *CFTR* enhanced the malignant behavior of NSCLC, whereas its overexpression suppressed cancer progression *in vitro* and *in vivo*. The two patients in this study carrying *CFTR* variants were of Southeast Asian ancestry.

Somatic loss of heterozygosity was observed at 4/25 tested loci in 12 tumor samples. We analyzed the variants listed in Tables 3 and 4 in the respective patients in the cases in which tumor tissue material was available (LUAD3, LUAD5, LUAD10–11, LUAD13–15, LUAD18–21, and LUAD26). The results of the LOH analysis are listed along with the variants in Tables 3 and 4. A decrease of the wild type allele signal was observed for *SMARCB1* in LUAD3, and both *SP100* and *SCN7A* in LUAD21. *SMARCB1* is part of the SWF/SNF complex and a tumor suppressor gene, loss-of-function mutations in it cause rhabdoid predisposition syndrome, while non-truncating splice-site and missense mutations have been found to cause familial schwannomatosis. The variant here has to our knowledge not previously been found in the



**Table 4**  
Genes mutated more than once within the patient series.

Gene	Variant location <sup>a</sup>	Base change <sup>b</sup>	Variant effect	rs-code	Population MAF	Patient	LOH
COL6A1	chr21:47421948	c.2030G > A	p.(Arg677His)	rs373731596	N/A	LuAd20	None
	chr21:47423767	c.2927G > A	p.(Arg976His)	rs753485651	9.13E-05	LuAd20	None
	chr21:47404364	c.409C > G	p.(Leu137Val)	rs139648899	6.71E-04	LuAd8	N/A
	chr21:47404379	c.424G > C	p.(Val142Leu)	rs144282452	6.83E-04	LuAd21	None
CLIP4	chr21:47402620	c.170C > A	p.(Ala57Asp)	rs143502850	3.10E-04	LuAd10	None
	chr2:29356586	c.433C > T	p.(Arg145Trp)	rs373336528	1.94E-04	LuAd23	N/A
	chr2:29356586	c.433C > T	p.(Arg145Trp)	rs373336528	1.94E-04	LuAd13	Mut loss
	chr2:29375679	c.1150T > G	p.(Ser384Ala)	N/A	N/A	LuAd31	N/A
ABCC10	chr6:43415549	c.3833T > C	p.(Ile1278Thr)	rs200393514	7.88E-04	LuAd18	None
	chr6:43417692	c.4342C > T	p.(Arg1448Trp)	rs200628851	8.54E-04	LuAd11	None
ATP7B	chr13:52536004	c.1915G > A	p.(His639Tyr)	rs200728096	N/A	LuAd5	WT loss
CACNA1S	chr13:52513231	c.3649_3654delGTTCTG	p.(Val1217_Leu1218del)	rs781266802	2.24E-04	LuAd19	None
	chr1:201021734	c.3904G > A	p.(Arg1302Trp)	rs780390034	7.76E-05	LuAd20	None
CFTR	chr1:201046075	c.1800A > T	p.(Phe600Leu)	N/A	N/A	LuAd10	None
	chr7:117199683	c.1558G > A	p.(Val520Ile)	rs77646904	N/A	LuAd31	N/A
COL6A6	chr7:117232086	c.1865G > A	p.(Gly622Asp)	rs121908759	4.87E-04*	LuAd29	N/A
	chr3:130368133	c.5460_5461insC	p.(Lys1821Glnfs)	rs761569826	6.11E-05*	LuAd31	N/A
GCN1	chr3:130284164	c.989_990delAG	p.(Gln330Argfs)	rs767662986	N/A	LuAd26	None
	chr12:120580400	c.5740T > G	p.(Asn1914His)	rs746987106	8.97E-05	LuAd19	None
GJB6	chr12:120572361	c.7178C > T	p.(Arg2393His)	rs200245091	2.33E-04	LuAd1	N/A
	chr13:20796839	c.781T > A	p.(Ser261Cys)	rs143962007	4.49E-05	LuAd6	N/A
RYR1	chr13:20797241	c.379G > A	p.(Arg127Trp)	N/A	N/A	LuAd19	None
	chr19:39027398	c.11599C > T	p.(Arg3867Cys)	rs138593495	N/A	LuAd26	None
SCN7A	chr19:38966047	c.4250C > T	p.(Ala1417Val)	rs779275405	9.03E-05	LuAd14	None
	chr2:167313537	c.1133C > T	p.(Ser378Asn)	rs775112434	N/A	LuAd24	N/A
SEC24A	chr2:167318979	c.1003T > C	p.(Asn335Asp)	rs771820272	1.94E-04	LuAd21	WT loss
	chr5:134015331	c.1294A > G	p.(Arg432Cys)	rs757964453	N/A	LuAd25	N/A
SP100	chr5:134032900	c.2071A > T	p.(Leu691Phe)	rs199840978	7.76E-04	LuAd13	None
	chr2:231405625	c.2245A > T	p.(Thr749Ser)	rs550426045	6.83E-04**	LuAd29	N/A
TTN	chr2:231333785	c.1319C > T	p.(Thr440Ile)	rs749074736	8.99E-05	LuAd21	WT loss
	chr2:179425967	c.84892G > A	p.(Arg28298Trp)	rs779581886	4.50E-05	LuAd8	N/A
USH2A	chr2:179598232	c.15788A > T	p.(Ile5263Asn)	rs776892266	8.19E-04	LuAd6	N/A
	chr1:215960068	c.10331C > T	p.(Cys3444Tyr)	N/A	5.80E-05*	LuAd15	None
	chr1:216373413	c.3367A > G	p.(Tyr1123His)	rs794729204	5.05E-04	LuAd10	None

Abbreviations: MAF, minor allele frequency; LOH, loss of heterozygosity; WT, wild type; mut, mutant.

Population MAF refers to the gnomAD Finnish population, except for patients LuAd15, LuAd29 and LuAd31 who are of Southeast Asian ancestry. For these gnomAD populations East Asian (\*) and South Asian (\*\*\*) were used, depending on which one had the higher allele frequency.

<sup>a</sup> Genome assembly GRCh37.

<sup>b</sup> Position in longest translation (GRCh37 release 84).

germline, but twice as a somatic mutation in colorectal cancer, one being a de novo mutation in a lung metastasis [34]. Its role in the carcinogenic process in our patient remains elusive, however.

During the review process of this manuscript, a study on the mutational profile of 36 young LUAD patients was published [54]. Besides characterizing unique somatic mutational features of this patient subgroup, potential predisposing mutations were also identified. Such variants were found in *TP53* and *BRCA2* supporting some of the findings in our study, and showing the value of multiple independent efforts to elucidate the molecular basis of this devastating disease.

**Table 5**  
Variants shared between LuAd24 and IPF afflicted son.

Gene	Variant location <sup>a</sup>	Base change <sup>b</sup>	Variant effect	rs-code	gnomAD Finnish MAF
ACTR2	chr2:65495807	c.1139G > A	p.(Arg380Gln)	rs757353258	1.94E-4
SCN7A	chr2:167313537	c.1133C > T	p.(Ser378Asn)	rs775112434	N/A
RAET1L	chr6:150342163	c.508delG	p.(Glu170Lysfs)	rs745880461	8.97E-5
PLEC	chr8:144991271	c.13129C > T	p.(Ala4377Thr)	rs201688261	1.56E-4
ZFAND4	chr10:46159086	c.88G > A	p.(Leu30Phe)	rs147435712	6.99E-4
PDE6C	chr10:95422371	c.2339_2344delTTGATT	p.(Asp781_Phe782del)	rs765688777	1.35E-4
GPHN	chr14:66975273	c.28A > T	p.(Asn10Tyr)	rs121908539	N/A
HERC1	chr15:63926054	c.12686G > A	p.(Ser4229Leu)	rs199683306	3.11E-4
NKD1	chr16:50642262	c.250C > T	p.(Arg84Trp)	rs201400910	7.94E-5
FBXO39	chr17:6683762	c.576delG	p.(Asn193Metfs)	rs769388797	N/A

Abbreviations: MAF, minor allele frequency.

<sup>a</sup> Genome assembly GRCh37.

<sup>b</sup> Position in longest translation (GRCh37 release 84).

cancer often appear only at an advanced stage, and most cases are diagnosed when treatment options are limited. Identifying individuals with high genetic risk is important, since such knowledge provides opportunities for early diagnosis and improved management, including efforts towards cancer prevention in the at-risk relatives.

### Conflict of interest statement

None declared.

### Acknowledgements

This work was supported by the Academy of Finland, Finnish Center of Excellence Program 2012–2017 [grant number 250345].

We want to thank Heikki Metsola, Sini Nieminen, Alison Ollikainen, Marjo Rajalaakso, Inga-Lill Svedberg and Iina Vuoristo for their indispensable assistance.

### References

- [1] J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman, F. Bray, GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11, Available from: International Agency for Research on Cancer, Lyon, France, 2013, (Accessed 01 October 2017) <http://globocan.iarc.fr>.
- [2] J. Malhotra, M. Malvezzi, E. Negri, C. La Vecchia, P. Boffetta, Risk factors for lung cancer worldwide, *Eur. Respir. J.* 48 (2016) 889–902.
- [3] I. Giapanou, K.A.M. Arendt, G.T. Stathopoulos, Lung carcinogenesis and fibrosis taken together: just coincidence? *Curr. Opin. Pulm. Med.* 23 (2017) 290–297.
- [4] W.I. Choi, S.H. Park, B.J. Park, C.W. Lee, Interstitial lung disease and lung cancer development: a 5-year nationwide population-based study, *Cancer Res. Treat.* 40 (2018) 374–381.
- [5] S. Jonsson, U. Thorsteinsdottir, D.F. Gudbjartsson, H.H. Jonsson, K. Kristjansson, S. Arnason, V. Gudnason, H.J. Isaksson, J. Hallgrímsson, J.R. Gulcher, L.T. Amundadottir, A. Kong, K. Stefansson, Familial risk of lung carcinoma in the Icelandic population, *JAMA* 292 (2004) 2977–2983.
- [6] M.L. Coté, M. Liu, S. Bonassi, M. Neri, A.G. Schwartz, D.C. Christiani, M.R. Spitz, J.E. Muscat, G. Rennert, K.K. Aben, A.S. Andrew, V. Bencko, H. Bickeböller, P. Boffetta, P. Brennan, H. Brenner, E.J. Duell, E. Fabianova, J.K. Field, L. Foretova, S. Friis, C.C. Harris, I. Holcatova, Y.C. Hong, D. Isla, V. Janout, L.A. Kiemeny, C. Kiyohara, Q. Lan, P. Lazarus, J. Lissowska, L. Le Marchand, D. Mates, K. Matsuo, J.I. Mayordomo, J.R. McLaughlin, H. Morgenstern, H. Müeller, I. Orlov, B.J. Park, M. Pinchev, O.Y. Raji, H.S. Rennert, P. Rudnai, A. Seow, I. Stucker, N. Szeszenia-Dabrowska, M. Dawn Teare, A. Tjønneland, D. Ugolini, H.F. van der Heijden, E. Wichmann, J.K. Wiencke, P.J. Woll, P. Yang, D. Zaridze, Z.F. Zhang, C.J. Etzel, R.J. Hung, Increased risk of lung cancer in individuals with a family history of the disease: a pooled analysis from the International Lung Cancer Consortium, *Eur. J. Cancer* 48 (2012) 1957–1968.
- [7] J. Hjelmborg, T. Korhonen, K. Holst, A. Skytthe, E. Pukkala, J. Kutschke, J.R. Harris, L.A. Mucci, K. Christensen, C. Czene, H.O. Adami, T. Scheike, J. Kaprio, Nordic Twin Study of Cancer (NorTwinCan) collaboration. Lung cancer, genetic predisposition and smoking: the Nordic Twin Study of Cancer, *Thorax* 72 (2017) 1021–1027.
- [8] J.D. McKay, R.J. Hung, V. Gaborieau, P. Boffetta, A. Chabrier, G. Byrnes, D. Zaridze, A. Mukeria, N. Szeszenia-Dabrowska, J. Lissowska, P. Rudnai, E. Fabianova, D. Mates, V. Bencko, L. Foretova, V. Janout, J. McLaughlin, F. Shepherd, A. Montpetit, S. Narod, H.E. Krokan, F. Skorpen, M.B. Elvestad, L. Vatten, I. Njølstad, T. Axelsson, C. Chen, G. Goodman, M. Barnett, M.M. Loomis, J. Lubinski, J. Matyjasik, M. Lener, D. Oszutowska, J. Field, T. Liloglou, G. Xinarianos, A. Cassidy, EPIC Study, P. Vineis, F. Clavel-Chapelon, D. Palli, R. Tumino, V. Krogh, S. Panico, C.A. González, J. Ramón Quirós, C. Martínez, C. Navarro, E. Ardanaz, N. Larrañaga, K.T. Kham, T. Key, H.B. Bueno-de-Mesquita, P.H. Peeters, A. Trichopoulou, J. Linseisen, H. Boeing, G. Hallmans, K. Overvad, A. Tjønneland, M. Kumle, E. Riboli, D. Zelenika, A. Boland, M. Delepine, M. Foglio, D. Lechner, F. Matsuda, H. Blanche, I. Gut, S. Heath, M. Lathrop, P. Brennan, Lung cancer susceptibility locus at 5p15.33, *Nat. Genet.* 40 (2008) 1404–1406.
- [9] Y. Wang, P. Broderick, E. Webb, X. Wu, J. Vijaykrishnan, A. Matakidou, M. Qureshi, Q. Dong, X. Gu, W.V. Chen, M.R. Spitz, T. Eisen, C.I. Amos, R.S. Houlston, Common 5p15.33 and 6p21.33 variants influence lung cancer risk, *Nat. Genet.* 40 (2008) 1407–1409.
- [10] C.I. Amos, X. Wu, P. Broderick, I.P. Gorlov, J. Gu, T. Eisen, Q. Dong, Q. Zhang, X. Gu, J. Vijaykrishnan, K. Sullivan, A. Matakidou, Y. Wang, G. Mills, K. Doheny, Y.Y. Tsai, W.V. Chen, S. Shete, M.R. Spitz, R.S. Houlston, Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q.1, *Nat. Genet.* 40 (2008) 616–622.
- [11] R.J. Hung, J.D. McKay, V. Gaborieau, P. Boffetta, M. Hashibe, D. Zaridze, A. Mukeria, N. Szeszenia-Dabrowska, J. Lissowska, P. Rudnai, E. Fabianova, D. Mates, V. Bencko, L. Foretova, V. Janout, C. Chen, G. Goodman, J.K. Field, T. Liloglou, G. Xinarianos, A. Cassidy, J. McLaughlin, G. Liu, S. Narod, H.E. Krokan, F. Skorpen, M.B. Elvestad, K. Hveem, L. Vatten, J. Linseisen, F. Clavel-Chapelon, P. Vineis, H.B. Bueno-de-Mesquita, E. Lund, C. Martínez, S. Bingham, T. Rasmussen, P. Hainaut, E. Riboli, W. Ahrens, S. Benhamou, P. Lagiou, D. Trichopoulos, I. Holcátová, F. Merletti, K. Kjaerheim, A. Agudo, G. Macfarlane, R. Talamini, L. Simonato, R. Lowry, D.I. Conway, A. Znaor, C. Healy, D. Zelenika, A. Boland, M. Delepine, M. Foglio, D. Lechner, F. Matsuda, H. Blanche, I. Gut, S. Heath, M. Lathrop, P. Brennan, A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q, *Nature* 452 (2008) 633–637.
- [12] D.W. Bell, I. Gore, R.A. Okimoto, N. Godin-Heymann, R. Sordella, R. Mulloy, S.V. Sharma, B.W. Brannigan, G. Mohapatra, J. Settleman, D.A. Haber, Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR, *Nat. Genet.* 37 (2005) 1315–1316.
- [13] A. Gazdar, L. Robinson, D. Oliver, C. Xing, W.D. Travis, J. Soh, S. Toyooka, L. Watumull, Y. Xie, K. Kernstine, J.H. Schiller, Hereditary lung cancer syndrome targets never smokers with germline EGFR gene T790M mutations, *J. Thorac. Oncol.* 9 (2014) 456–463.
- [14] H.A. Yu, M.E. Arcila, M. Harlan Fleischut, Z. Stadler, M. Ladanyi, M.F. Berger, M. Robson, G.J. Riely, Germline EGFR T790M mutation found in multiple members of a familial cohort, *J. Thorac. Oncol.* 9 (2014) 554–558.
- [15] H. Yamamoto, K. Higasa, M. Sakaguchi, K. Shien, J. Soh, K. Ichimura, M. Furukawa, S. Hashida, K. Tsukuda, N. Takigawa, K. Matsuo, K. Kiura, S. Miyoshi, F. Matsuda, S. Toyooka, Novel germline mutation in the transmembrane domain of HER2 in familial lung adenocarcinomas, *J. Natl. Cancer Inst.* 106 (2014) djt338.
- [16] Y. Wang, J.D. McKay, T. Rafnar, Z. Wang, M.N. Timofeeva, P. Broderick, X. Zong, M. Laplana, Y. Wei, Y. Han, A. Lloyd, M. Delahaye-Sourdeix, D. Chubb, V. Gaborieau, W. Wheeler, N. Chatterjee, G. Thorleifsson, P. Sulem, G. Liu, R. Kaaks, M. Henrion, B. Kinnerley, M. Vallée, F. LeCalvez-Kelm, V.L. Stevens, S.M. Gapstur, W.V. Chen, D. Zaridze, N. Szeszenia-Dabrowska, J. Lissowska, P. Rudnai, E. Fabianova, D. Mates, V. Bencko, L. Foretova, V. Janout, H.E. Krokan, M.E. Gabrielsen, F. Skorpen, L. Vatten, I. Njølstad, C. Chen, G. Goodman, S. Benhamou, T. Voeder, K. Välik, M. Nelis, A. Metspalu, M. Lener, J. Lubinski, M. Johansson, P. Vineis, A. Agudo, F. Clavel-Chapelon, H.B. Bueno-de-Mesquita, D. Trichopoulos, K.T. Khaw, M. Johansson, E. Weiderpass, A. Tjønneland, E. Riboli, M. Lathrop, G. Scelo, D. Albanes, N.E. Caporaso, Y. Ye, J. Gu, X. Wu, M.R. Spitz, H. Dienemann, A. Rosenberger, L. Su, A. Matakidou, T. Eisen, K. Stefansson, A. Risch, S.J. Chanock, D.C. Christiani, R.J. Hung, P. Brennan, M.T. Landi, R.S. Houlston, C.I. Amos, Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer, *Nat. Genet.* 46 (2014) 736–741.
- [17] M.H. Abdel-Rahman, R. Pilarski, C.M. Cebulla, J.B. Messingill, B.N. Christopher, G. Boru, P. Hovland, F.H. Davidorf, Germline BAP1 mutation predisposes to uveal melanoma, lung adenocarcinoma, meningioma, and other cancers, *J. Med. Genet.* 48 (2011) 856–859.
- [18] D. Xiong, Y. Wang, E. Kupert, C. Simpson, S.M. Pinney, C.R. Gaba, D. Mandal, A.G. Schwartz, P. Yang, M. de Andrade, C. Pikielny, J. Byun, Y. Li, D. Stambolian, M.R. Spitz, Y. Liu, C.I. Amos, J.E. Bailey-Wilson, M. Anderson, M. You, A recurrent mutation in PARK2 is associated with familial lung cancer, *Am. J. Hum. Genet.* 96 (2015) 301–308.
- [19] K.J. Park, H.J. Choi, S.P. Suh, C.S. Ki, J.W. Kim, Germline TP53 mutation and clinical characteristics of Korean patients with Li-fraumeni syndrome, *Ann. Lab. Med.* 36 (2016) 463–468.
- [20] L. de Kock, I. Bah, Y. Wu, M. Xie, J.R. Priest, W.D. Foulkes, Germline and somatic DICER1 mutations in a well-differentiated fetal adenocarcinoma of the lung, *J. Thorac. Oncol.* 11 (2016) e31–33.
- [21] K.C. Toh, F. Gao, W.T. Lim, S.S. Leong, K.W. Fong, S.P. Yap, A.A. Hsu, P. Eng, H.N. Koong, A. Thirugnanam, E.H. Tan, Never-smokers with lung cancer: epidemiologic evidence of a distinct disease entity, *J. Clin. Oncol.* 24 (2006) 2245–2251.
- [22] H.A. Wakelee, E.T. Chang, S.L. Gomez, T.H. Keegan, D. Feskanich, C.A. Clarke, L. Holmberg, L.C. Yong, L.N. Kolonel, M.K. Gould, D.W. West, Lung cancer incidence in never smokers, *J. Clin. Oncol.* 25 (2007) 472–478.
- [23] J. Remon, E. Molina-Montes, M. Majem, P. Lianes, D. Isla, P. Garrido, E. Felip, N. Viñolas, J. de Castro, A. Artal, M.J. Sánchez, Lung cancer in women: an overview with special focus on Spanish women, *Clin. Transl. Oncol.* 16 (2014) 517–528.
- [24] E. Esteban, M. Majem, M. Martínez Aguilero, N. Martínez Banaclocha, M. Dómine, L. Gómez Aldaravi, O. Juan, R. Cajal, M.C. Gonzalez Arenas, M. Provencio, Prevalence of EGFR mutations in newly diagnosed locally advanced or metastatic non-small cell lung cancer Spanish patients and its association with histological subtypes and clinical features: the Spanish REASON study, *Cancer Epidemiol.* 39 (2015) 291–297.
- [25] Y. Matsumura, Y. Owada, T. Inoue, Y. Watanabe, T. Yamaura, M. Fukuhara, S. Muto, N. Okabe, T. Hasegawa, M. Hoshino, J. Osugi, M. Higuchi, H. Suzuki, Epidermal growth factor receptor mutation status is strongly associated with smoking status in patients undergoing surgical resection for lung adenocarcinoma, *Interact. Cardiovasc. Thorac. Surg.* 25 (2017) 690–695.
- [26] Li Heng, Durbin Richard, Fast and accurate short read alignment with Burrows–Wheeler transform, *Bioinformatics* 25 (2009) 1754–1760.
- [27] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytzky, G. Garimella, D. Altshuler, S. Gabriel, M. Daly, M.A. DePristo, The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, *Genome Res.* 20 (2010) 1297–1303.
- [28] M. Lek, K.J. Karczewski, E.V. Minikel, K.E. Samocha, E. Banks, T. Fennell, A.H. O'Donnell-Luria, J.S. Ware, A.J. Hill, B.B. Cummings, T. Tukiainen, D.P. Birnbaum, J.A. Kosmicki, L.E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D.N. Cooper, N. DeFlaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M.I. Kurki, A.L. Moonshine, P. Natarajan, L. Ruzoco, G.M. Peloso, R. Poplin, M.A. Rivas, V. Ruano-Rubio, S.A. Rose, D.M. Ruderfer, K. Shakir, P.D. Stenson, C. Stevens, B.P. Thomas, G. Tiao, M.T. Tusie-Luna, B. Weisburd, H.H. Won, D. Yu,

- D.M. Altshuler, D. Ardissono, M. Boehnke, J. Danesh, S. Donnelly, R. Elosua, J.C. Florez, S.B. Gabriel, G. Getz, S.J. Glatt, C.M. Hultman, S. Kathiresan, M. Laakso, S. McCarroll, M.I. McCarthy, D. McGovern, R. McPherson, B.M. Neale, A. Palotie, S.M. Purcell, D. Saleheen, J.M. Scharf, P. Sklar, P.F. Sullivan, J. Tuomilehto, M.T. Tsuang, H.C. Watkins, J.G. Wilson, M.J. Daly, D.G. MacArthur, Exome aggregation consortium. Analysis of protein-coding genetic variation in 60,706 humans, *Nature* 536 (2016) 285–291.
- [29] K. Jagadeesh, A. Wenger, M. Berger, H. Guturu, P. Stenson, D. Cooper, J. Bernstein, G. Bejerano, M-CAP eliminates a majority of variants with uncertain significance in clinical exomes at high sensitivity, *Nat. Genet.* 48 (2016) 1581–1586.
- [30] Y. Choi, G.E. Sims, S. Murphy, J.R. Miller, A.P. Chan, Predicting the functional effect of amino acid substitutions and indels, *PLoS One* 7 (2012) e46688.
- [31] A. Untergasser, I. Cutcutache, T. Koressaar, J. Ye, B.C. Faircloth, M. Remm, S.G. Rozen, Primer3 - new capabilities and interfaces, *Nucleic Acids Res.* 40 (2012) e115.
- [32] K.L. Huang, R.J. Mashl, Y. Wu, D.I. Ritter, J. Wang, C. Oh, M. Paczkowska, S. Reynolds, M.A. Wyczalkowski, N. Oak, A.D. Scott, M. Krassowski, A.D. Cherniack, K.E. Houlahan, R. Jayasinghe, L.B. Wang, D.C. Zhou, D. Liu, S. Cao, Y.W. Kim, A. Koire, J.F. McMichael, V. Huchtagowder, T.B. Kim, A. Hahn, C. Wang, M.D. McLellan, F. Al-Mulla, K.J. Johnson, Cancer Genome Atlas Research Network, O. Lichtarge, P.C. Boutros, B. Raphael, A.J. Lazar, W. Zhang, M.C. Wendl, R. Govindan, S. Jain, D. Wheeler, S. Kulkarni, J.F. Dipersio, J. Reimand, F. Meric-Bernstam, K. Chen, I. Shmulevich, S.E. Plon, F. Chen, L. Ding, Pathogenic germline variants in 10,389 adult cancers, *Cell* 173 (2018) 355–370.
- [33] C. Lu, M. Xie, M.C. Wendl, J. Wang, M.D. McLellan, M.D. Leiserson, K.L. Huang, M.A. Wyczalkowski, R. Jayasinghe, T. Banerjee, J. Ning, P. Tripathi, Q. Zhang, B. Niu, K. Ye, H.K. Schmidt, R.S. Fulton, J.F. McMichael, P. Batra, C. Kandath, M. Bharadwaj, D.C. Koboldt, C.A. Miller, K.L. Kanchi, J.M. Eldred, D.E. Larson, J.S. Welch, M. You, B.A. Ozenberger, R. Govindan, M.J. Walter, M.J. Ellis, E.R. Mardis, T.A. Graubert, J.F. Dipersio, T.J. Ley, R.K. Wilson, P.J. Goodfellow, B.J. Raphael, F. Chen, K.J. Johnson, J.D. Parvin, L. Ding, Patterns and functional implications of rare germline variants across 12 cancer types, *Nat. Commun.* 6 (2015) 10086.
- [34] S.A. Forbes, D. Beare, H. Boutselakis, S. Bamford, N. Bindal, J. Tate, C.G. Cole, S. Ward, E. Dawson, L. Ponting, R. Stefancsik, B. Harsha, C.Y. Kok, M. Jia, H. Jubbs, Z. Sondka, S. Thompson, T. De, P.J. Campbell, COSMIC: somatic cancer genetics at high-resolution, *Nucleic Acids Res.* 45 (2017) D777–D783.
- [35] L.L. Bachinski, S.-E. Olufemi, X. Zhou, C.-C. Wu, L. Yip, S. Shete, G. Lozano, C.I. Amos, L.C. Strong, R. Krahe, Genetic mapping of a third Li-Fraumeni syndrome predisposition locus to human chromosome 1q23, *Cancer Res.* 65 (2005) 427–431.
- [36] G. Bougeard, M. Renaux-Petel, J.M. Flaman, C. Charbonnier, P. Fermey, M. Belotti, M. Gauthier-Villars, D. Stoppa-Lyonnet, E. Consolino, L. Brugières, O. Caron, P.R. Benusiglio, B. Bressac-de Paillerets, V. Bonadona, C. Bonaiti-Pellié, J. Tinat, S. Baert-Desurmont, T. Frebourg, Revisiting Li-Fraumeni syndrome from TP53 mutation carriers, *J. Clin. Oncol.* 33 (2015) 2345–2352.
- [37] O. Caron, T. Frebourg, P.R. Benusiglio, S. Foulon, L. Brugières, Lung adenocarcinoma as part of the li-fraumeni syndrome spectrum: preliminary data of the LIFSCREEN randomized clinical trial, *JAMA Oncol.* 3 (2017) 1736–1737.
- [38] D.J. Ransburgh, N. Chiba, C. Ishioka, A.E. Toland, J.D. Parvin, Identification of breast tumor mutations in BRCA1 that abolish its function in homologous DNA recombination, *Cancer Res.* 70 (2010) 988–995.
- [39] Z. Kais, N. Chiba, C. Ishioka, J.D. Parvin, Functional differences among BRCA1 missense mutations in the control of centrosome duplication, *Oncogene* 31 (2012) 799–804.
- [40] M.S. Brose, T.R. Rebbeck, K.A. Calzone, J.E. Stopfer, K.L. Nathanson, B.L. Weber, Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program, *J. Natl. Cancer Inst.* 94 (2002) 1365–1372.
- [41] B. Friedenson, BRCA1 and BRCA2 pathways and the risk of cancers other than breast or ovarian, *MedGenMed* 7 (2005) 60.
- [42] M. Gallardo, A. Silva, L. Rubio, C. Alvarez, C. Torrealba, M. Salinas, T. Tapia, P. Faundez, L. Palma, M.E. Riccio, H. Paredes, M. Rodriguez, A. Cruz, C. Rousseau, M.C. King, M. Camus, M. Alvarez, P. Carvallo, Incidence of BRCA1 and BRCA2 mutations in 54 Chilean families with breast/ovarian cancer, genotype-phenotype correlations, *Breast Cancer Res. Treat.* 95 (2006) 81–87.
- [43] L. Guidugli, V.S. Pankratz, N. Singh, J. Thompson, C.A. Erding, C. Engel, R. Schmutzler, S. Domchek, K. Nathanson, P. Radice, C. Singer, P.N. Tonin, N.M. Lindor, D.E. Goldgar, F.J. Couch, A classification model for BRCA2 DNA binding domain missense variants based on homology-directed repair activity, *Cancer Res.* 73 (2013) 265–275.
- [44] J. Li, C. Zou, Y. Bai, D.E. Wazer, V. Band, Q. Gao, DSS1 is required for the stability of BRCA2, *Oncogene* 25 (2006) 1186–1194.
- [45] E.M. Parry, D.L. Gable, S.E. Stanley, S.E. Khalil, V. Antonescu, L. Florea, M. Armanios, Germline mutations in DNA repair genes in lung adenocarcinoma, *J. Thorac. Oncol.* 12 (2017) 1673–1678.
- [46] C.A. Whittaker, R.O. Hynes, Distribution and evolution of von Willebrand/Integrin a domains: widely dispersed domains with roles in cell adhesion and elsewhere, *Mol. Biol. Cell* 13 (2002) 3369–3387.
- [47] G.J. Jöbsis, H. Keizers, J.P. Vreijling, M. de Visser, M.C. Speer, R.A. Wolterman, F. Baas, P.A. Bolhuis, Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures, *Nat. Genet.* 14 (1996) 113–115.
- [48] L. Voiles, D.E. Lewis, L. Han, I.P. Lupov, T.L. Lin, M.J. Robertson, I. Petrache, H.C. Chang, Overexpression of type VI collagen in neoplastic lung tissues, *Oncol. Rep.* 32 (2014) 1897–1904.
- [49] E. Watanabe, T.Y. Hiyama, R. Kodama, M. Noda, NaX sodium channel is expressed in non-myelinating Schwann cells and alveolar type II cells in mice, *Neurosci. Lett.* 330 (2002) 109–113.
- [50] J.A. Kropski, W.E. Lawson, L.R. Young, T.S. Blackwell, Genetic studies provide clues on the pathogenesis of idiopathic pulmonary fibrosis, *Dis. Models Mech.* 6 (2013) 9–17.
- [51] J.W. Son, Y.J. Kim, H.M. Cho, S.Y. Lee, S.M. Lee, J.K. Kang, J.U. Lee, Y.M. Lee, S.J. Kwon, E. Choi, M.J. Na, J.Y. Park, D.S. Kim, Promoter hypermethylation of the CFTR gene and clinical/pathological features associated with non-small cell lung cancer, *Respirology* 16 (2011) 1203–1209.
- [52] R. Govindan, L. Ding, M. Griffith, J. Subramanian, N.D. Dees, K.L. Kanchi, C.A. Maher, R. Fulton, L. Fulton, J. Wallis, K. Chen, J. Walker, S. McDonald, R. Bose, D. Ornitz, D. Xiong, M. You, D.J. Dooling, M. Watson, E.R. Mardis, R.K. Wilson, Genomic landscape of non-small cell lung cancer in smokers and never-smokers, *Cell* 150 (2012) 1121–1134.
- [53] J. Li, J.T. Zhang, X. Jiang, X. Shi, J. Shen, F. Feng, J. Chen, G. Liu, P. He, J. Jiang, L.L. Tsang, Y. Wang, R. Rosell, L. Jiang, J. He, H.C. Chan, The cystic fibrosis transmembrane conductance regulator as a biomarker in non-small cell lung cancer, *Int. J. Oncol.* 46 (2015) 2107–2115.
- [54] W. Luo, P. Tian, Y. Wang, H. Xu, L. Chen, C. Tang, Y. Shu, S. Zhang, Z. Wang, J. Zhang, L. Zhang, L. Jiang, L. Liu, G. Che, C. Guo, H. Zhang, J. Wang, W. Li, Characteristics of genomic alterations of lung adenocarcinoma in young never-smokers, *Int. J. Cancer* (2018), <http://dx.doi.org/10.1002/ijc.31542> Accepted for publication.