CANCER RESEARCH AND TREATMENT (CRT)

Original Article

"Omic" Approach in Non-Smoker Female with Lung Squamous Cell Carcinoma Pinpoints to Germline Susceptibility and Personalized Medicine

Margherita Baldassarri, MD[1,2,†], Chiara Fallerini, PhD[1,†], Francesco Cetta, MD[3], Marco Ghisalberti, MD[4], Cristiana Bellan, MD[5], Simone Furini, PhD[6]; Ottavia Spiga, PhD[7], Sergio Crispino, MD[8], Giuseppe Gotti, MD[4], Francesca Ariani, PhD[1,2], Piero Paladini, MD[4], Alessandra Renieri, MD, PhD[1,2,*], Elisa Frullanti, PhD[1]

[1]Medical Genetics, University of Siena, Siena, Italy, [2]Genetica Medica, Azienda Ospedaliera Universitaria Senese, Siena, Italy, [3]IRCCS MultiMedica, Milan, Italy, [4]Thoracic Surgery Unit, Azienda Ospedaliera Universitaria Senese, Siena, Italy, [5] Department of Medical Biotechnology, Section of Pathology, University of Siena, Siena, Italy, [6] Department of Medical Biotechnology, University of Siena, Siena, Italy, [7] Department of Biotechnology, Chemistry and Pharmacy, University of Siena, Siena, Italy, [8] Medical Oncology Unit, Azienda Usl Toscana Sudest, Siena, Italy.

[†]Margherita Baldassarri and Chiara Fallerini contributed equally to the work.

Running Title: Integrative Omic approach for precision medicine

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi:10.4143/crt.2017.125

Correspondence : Alessandra Renieri, MD, PhD

Medical Genetics Unit, University of Siena, Azienda Ospedaliera Universitaria Senese, Viale Bracci, 2 - 53100 Siena, Italy

Tel: 39-0577-233303 Fax: 39-0577-233325 Email: alessandra.renieri@unisi.it

Abstract

Purpose

Lung cancer is strongly associated to tobacco smoking. However, global statistics estimate that in females the proportion of lung cancer cases that is unrelated to tobacco smoking reaches fifty percent, making questionable the etiology of the disease.

Materials and Methods

A never-smoker female with primary EGFR/KRAS/ALK-negative squamous cell carcinoma of the lung and their normal sibs were subjected to a novel integrative "omic" approach using a pedigree-based model for discovering genetic factors leading to cancer in the absence of well-known environmental trigger. A first-step whole-exome sequencing on tumor and normal tissue did not identify mutations in known driver genes. Building on the idea of a germline oligogenic origin of lung cancer, we performed whole-exome sequencing of DNA from patients' peripheral blood and their unaffected sibs. Finally, RNA-sequencing analysis in tumoral and matched non-tumoral tissues was carried out in order to investigate the clonal profile and the pathogenic role of the identified variants.

Results

Filtering for rare variants with CADD>25 and potentially damaging effect, we identified rare/private germline deleterious variants in 11 cancer-associated genes, none of which, except one, shared with the healthy sib, pinpointing to a "private" oligogenic germline signature. Noteworthy, among these, two mutated genes, namely *ACACA* and *DEPTOR*, turned to be potential targets for therapy because related to known drivers, such as *BRCA1* and *EGFR*.

Conclusions

In the era of precision medicine, this report emphasizes the importance of an "omic" approach to uncover oligogenic germline signature underlying cancer development and to identify suitable therapeutic targets as well.

Keywords: Exome, Massively-Parallel Sequencing, High-Throughput RNA Sequencing, Disease Susceptibility, Oligogenic Inheritance, Precision medicine

Introduction

Lung cancer (LC) is an important public health problem and the leading cause of cancer death worldwide [1]. The two major histological types of LC are non-small cell lung cancers (NSCLC) and small-cell lung cancers (SCLC). NSCLCs consist mostly of lung adenocarcinomas (ADCA) and lung squamous cell carcinomas (SQCC). These two NSCLC subtypes have both unique and common clinical and histopathological characteristics. For example, ADCAs are mainly observed in never smokers [2] while SQCC is the most frequent histological type among smoking patients.

The discovery of common, mutually exclusive, driver genes, epidermal growth factor receptor (*EGFR*), *KRAS* proto-oncogene GTPase (*KRAS*) and anaplastic lymphoma receptor tyrosine kinase (*ALK*), that represent crucial therapeutic targets, led to a remarkable improvement in NSCLC therapy. However, only about 50% of NSCLC acquire mutations in these genes. Identification of new cancer driver genes currently is a top research priority. Recent studies using next-generation sequencing (NGS) approach characterized NSCLC tissues in order to identify new less common candidate driver mutations [3, 4] highlighting the presence of great variability even within tumor samples belonging to the same histological type.

Up to now, application of NGS technology has been reported in many lung cancer samples leading to the discovery of a large set of new somatic mutations in key cancer genes. Nevertheless, no studies are present in literature that performed whole exome sequencing (WES) and RNA-sequencing (RNA-seq) from tumoral and matched non-tumoral lung tissues integrating these with WES of blood DNA from the same individual to identify possible tumor-driver variants which are already present in the germline background and that may predispose to lung cancer. Using WES analysis in a model of cases and controls deriving from the same kindred - discordant sib-pair design - we have recently showed that a set of

oligogenic germline mutations, that are detectable in the peripheral blood, may play a role in the development and growth of lung ADCA in never smokers [5]. Noteworthy, our previous results demonstrated that the oligogenic cancer predisposing signature is a “private signature”, varying from one patient to another.

Herein, we tested further the hypothesis of the oligogenic origin of lung cancer in a never-smoker female who developed a SQCC of the lung. We applied for the first time an integrative “omic” approach of WES and RNA-seq experiments in a discordant sib-pair design.

Materials and Methods

Samples and DNA / RNA extraction

Patient and sib signed the informed consent declaration at the Medical Genetics Unit of the Azienda Ospedaliera Universitaria Senese, Italy, for the use of their biological samples and clinical data for research purposes. The study protocol was approved by the Ethical Committee of the Azienda Ospedaliera Universitaria Senese. Information on histological diagnosis (by the Pathology Unit) was retrieved from the clinical records. Genomic DNA of patient and healthy sib was isolated from EDTA peripheral blood using QIAamp DNA Blood Kit (Qiagen®, Hilden, Germany) according to the manufacturer’s protocol. Formalin-fixed paraffin-embedded (FFPE) samples of patient’s tumoral and non tumoral lung tissues were obtained from the Pathology Unit of the recruiting hospital and analyzed through whole-exome sequencing and RNA sequencing. DNA was extracted from FFPE lung tumoral and non-tumoral tissue samples using MagCore® Genomic DNA FFPE One-Step Kit for MagCore® System (Diatech Pharmacogenetics srl, Jesi, Ancona, Italy). RNA was extracted from FFPE lung tumoral and non-tumoral tissue samples of our case using High Pure FFPE-

Tissue RNA Isolation Kit (Roche, Basel, Switzerland) following the manufacturer's instructions. Additional RNAs from 3 FFPE lung non-tumoral tissues from other lung cancer patients were isolated using the same kit and used as control tissues. RNA samples were processed to remove ribosomal RNA (rRNA) using Ribo-Zero rRNA Removal Kit for Human samples (Illumina, Grand Island, USA) following the manufacturer's instructions. RNA integrity was verified using the Agilent Eukaryote Total RNA Nano Kit (Agilent Technologies, Palo Alto, CA, USA) on Agilent2100 Bioanalyzer (Agilent Technologies). Both DNA and RNA were quantified by spectrophotometry (ND-2000c; NanoDrop Products, Wilmington, DE) and Qubit® Fluorometer with Qubit® dsDNA HS Assay and Qubit® RNA HS Assay Kits (LifeTechnologies, Carlsbad, CA, USA), respectively.

Whole Exome Sequencing and data analysis

Whole exome sequencing was performed using the Life Technologies Ion Proton sequencer (Life Technologies) on genomic DNA samples of cases and controls and tumor tissues as previously described [6]. Variant were then annotated against external datasets, including 1000 genomes (http://www.1000genomes.org/) and dbSNP. After removing variants with low coverage, in order to identify candidate susceptibility variants, we selected for rare variants with minor allele frequency-MAF ≤ 0.01) or not reported (**S1 Figure**). We then filtered excluding variants with clinical significance as "benign" or "likely benign" and present in an in-house database of variants. Additional filtering procedures were thus implemented for retrieving exonic rare variants with a potential detrimental impact on protein function, i.e., variants, i.e. insertions and deletions (indels) causing exonic frameshifts and missense variants predicted as deleterious applying the Combined Annotation Dependent Depletion (CADD>25) and the MetaSVM (Support Vector Machine) bioinformatics tools.

Genes affected by mutations were checked in literature in order to identify cancer-related genes. Validation of these variants was carried out using custom NGS panel for Ion PGM sequencer (Life Technologies).

RNA sequencing and data analysis

RNA sequencing was performed using Illumina HiSeq2500 platform (Illumina Inc., San Diego, CA, USA), in a 2x100bp paired-end (PE) configuration in High Output mode (V4 chemistry), with a total of at least 250 million reads per lane. After quality check, RNA (50 ng) was used to prepare sample libraries. Sequencing library construction included these steps: library construction using Illumina TruSeq RNA Sample Pre Kit (Illumina), library purification using Beckman AMPure XP beads (Beckman Coulter s.r.l., Milan, Italy), insert fragments test using Agilent High Sensitivity DNA Kit on Agilent 2100 Bioanalyzer (Agilent Technologies), and cBOT automatic cluster (TruSeq PE Cluster Kit v3-cBotHS). Post-library quality controls were performed using the Agilent RNA 6000 Nano kit (Agilent Technologies) on Agilent2100 Bioanalyzer (Agilent Technologies) and Qubit® Fluorometer (Life Technologies). Libraries were then loaded on HiSeq2500 sequencing platform (Illumina) and sequenced using 2x100bp pair-end High Output Mode (V4 chemistry) per lane. The reads generated on the HiSeq2500 were provided under FASTQ format.

Sequence reads in FASTQ format were processing using the Fastqc software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) for data quality check and removing excess adaptors to get high-quality and clean reads. The high-quality reads were aligned to the GRCh38/hg38 reference human genome (ftp://jgi-psf.org/pub/compgen/phytozome/v9.0/Ptrichocarpa/assembly/Ptrichocarpa_210.fa.gz) using the TopHat software (version 2.0.9) [7]. Transcript assembling and expression quantification were carried out

using Cufflinks (version 2) [8]. Gene expression was expressed as fragments per kilo-base transcript per million mapped reads (FPKM) value. This normalized value was used for visualization on a genome browser (http://genome.ucsc.edu/), as well as to compare read coverage between and throughout different genes. Statistical analysis was performed to compute the mean FPKM level with the associated P-value for lung normal tissues together with the mean FPKM level with the associated P-value for lung cancer tissues. Cuffdiff tool from Cufflinks was used to identify differentially expressed genes [9]. Gene Ontology (GO) and pathways analyses were performed using both Enrichr (http://amp.pharm.mssm.edu/Enrichr/) and Database for Annotation, Visualization and Integrated Discovery functional annotation tool (DAVID Bioinformatics Resources 6.7, https://david.ncifcrf.gov).

Results

Patient description

A still alive Caucasian female patient, who developed squamous cell carcinoma of the lung in the absence of smoking habit and/or asbestos exposure, was recruited. Her smoker father died at age 63 from lung cancer. She had no known other family history of cancers. At age of 77, after a “screening” chest and abdomen CT scan examination, a 12x7 mm nodule was found in the dorsal segment of the right superior lobe in association with a small “ground glass” area (9,5 mm in diameter) in the left superior lobe. The nodule showed a small volumetric increase after 3 months and the patient was scheduled for surgery at the Thoracic Surgery Unit of Azienda Ospedaliera Universitaria Senese, Italy. Intraoperative diagnosis was NSCLS. At gross pathologic examination, a 12x0.5 solid nodule with irregular margin was

detected. Histologic examination showed squamous cell carcinoma with a mainly perialveolar and perivascular growth pattern. Tumoral cells were P63+,CK5/6+,CK7-, TTF1-, Synaptophysin-, chromogranin-. Ki67 immunostaining was 20 per cent. There was moderate infiltration of the visceral pleura, but no lymphonodal infiltration. Two years before, patient had undergone to removal of a cutaneous nodule located in the left side of the neck that resulted to be a squamous cell carcinoma of the basal subtype. She subsequently underwent left cervical lymphadenectomy, that showed infiltration of 3 cervical lymphonodes. On the basis of histological analysis, site of tumors and type of lymphonodal involvement, it was stated that the patient had 2 different primary tumors, both belonging to the squamous subtype. The patient is alive and well 4 years after the first and 2 years after the second operation.

“Omic” characterization

We firstly carried out whole-exome sequencing of gDNA from formalin-fixed paraffin-embedded (FFPE) tumoral and non-tumoral lung tissue. With the aim of identification of somatic driver mutations, we subtracted the germline background to 39,286 variants present in the tumoral lung tissue, leading to 2,543 exonic variants present only in the tumor. Among them, 1,381 were missense variants and 254 were truncating (**S2A Figure**). We then extracted the somatic mutational signature according to base substitutions, as described by Alexandrov et al., 2013 [10]. This analysis showed a predominance of C>T transitions in tumoral lung tissue, not corresponding to the specific cancer signature related to tobacco consumption that are characterized by predominance of C>A transversions (**S2B Figure**). Pathway analysis showed that mutations observed in squamous cell carcinoma tissue mostly affected genes involved in extracellular matrix organization (p-value= 0.0049), transmembrane transport of small molecules (p-value=0.0105) and collagen formation (p-

value=0.034) pathway, as previously reported [4] (**S2C Figure**). Overall, no mutations in the known driver genes, such as *EGFR, KRAS, AKT, ROS1*, nor in other putative driver genes described by Campbell et al. were present in our tumoral sample [3]. *EML4*-ALK fusion gene was also excluded. The failure of identification of somatic driver mutations rose the hypothesis that germline variants may have a role in etiology of the disease.

We then adopted a novel integrative "omic" approach using a combination of NGS techniques in a pedigree model that compares discordant sibs to test the hypothesis of oligogenic origin of lung cancer (**Figure 1**).

We carried out whole-exome sequencing of gDNA from peripheral blood of the SQCC patient and her unaffected sib as controls (**S1 Table**). Overall, we identified 37,727 germline variants in the patient and 37,355 germline variants in the healthy sib and 39,286 variants in the tumoral lung tissue.

To identify susceptibility variants, we compared case and control WES data and selected for rare variants (minor allele frequency-MAF ≤ 0.01) with a potentially damaging effect: truncating variants, i.e. insertions and deletions (indels) causing exonic frameshifts and nonsense mutation, and missense deleterious variants predicted with CADD>25 and filtering out those with MetaSVM as "tolerated" (**S1 Figure**). We obtained 111 potential deleterious variants of which 40 were shared by both siblings. Among the 111 potentially deleterious variants, 19 variants mapped in genes that have been found previously associated with cancer of the lung or other tissues (**S2 Table**). Validation of these variants led to the confirmation of all variants that were probably responsible for lung cancer susceptibility in our case. All the sequence variants identified were in the heterozygous state and were also present in the relative lung tumor tissue in heterozygous state as expected.

Finally, we performed RNA sequencing (RNA-seq) analysis of RNA extracted from both FFPE tumoral and non-tumoral lung tissues of our patient in order to detect differences in the expression level that could help the understanding of the functional role of germline variants. RNA-seq generated a mean of 75,134,860 reads per sample (**S3 Table**).

To assess the potential functional role of the 19 germline cancer-related variants identified by whole-exome sequencing in the patient and predicted as deleterious by bioinformatics tools, we combined results from WES experiments with the respective-gene expression profile from RNA-seq of non-tumoral and matched tumoral tissue. In 19 variants mapping in 19 genes, RNA sequencing data showed differential expression level (fold-change>± 1.5) for 11 genes, reinforcing the pathogenic role of the identified variants showing three different effects (**Table 1**). Three genes, namely *ACACA*, *CTSZ*, and *MYH9*, showed downregulation in lung cancer compared to normal tissue. Three genes, namely *DEPTOR*, *AVL9*, and *CARS*, showed downregulation of both normal and cancer tissue when we compared those tissues with a pool of 3 normal lung tissues. Lastly, five remaining genes (*PSCA*, *ENO3*, *PARP14*, *CTBP2*, *FOXM1*) showed an upregulation in lung cancer tissue. All the above mentioned variants were exclusive of the patient with the exception of variant in *FOXM1* gene that was present also in the healthy sib.

Among the 11 germline-mutated cancer-related genes, two genes, namely *ACACA* and *DEPTOR*, are known to be associated with the known driver genes, *BRCA1* and *EGFR*, respectively. In particular, we found a missense variant, that has been predicted as deleterious, c.C1948T, p.Arg650Trp (NM_198837.1, exon 16) in *ACACA* gene (OMIM* 200350) encoding Acetyl-CoA carboxylase alpha, the rate-limiting enzyme for the long-chain fatty acid synthesis [11]. The other mutation is a stopgain variant c.A631T, p.Arg211* in *DEPTOR* gene (OMIM *612974) encoding DEP domain-containing MTOR-interacting protein.

Interestingly, from RNA-seq data, we found a decreased expression level of *ACACA* in tumoral rather that in non-tumoral tissue. Sequence information was extracted from RNA-seq to confirm the presence of mutated mRNA and we found that 81,8% of reads carried the mutated allele pinpointing to a loss of heterozygosity (LOH) effect. The same analysis revealed a decreased expression level of *DEPTOR* both in normal and tumoral tissue at mRNA level pinpointing to LOH effect already present in the surrounding normal tissue.

Discussion

Among 111 potentially deleterious germline variants, 19 mapped in genes that have previously been associated with cancer of the lung or other tissues (S1 Table). Using the discordant-sib pair design, we confirmed the presence of an oligogenic signature present in the patient's genetic background [5]. In fact, only one out of 11 cancer-related variants was also found in the healthy sib. For all the 11 variants, RNA sequencing data reinforced the pathogenic role showing three different effects on the expression levels of harboring gene: three genes (*ACACA*, *CTSZ*, *MYH9*) showed a possible "second hit" in tumor suppressor genes responsible for gene inactivation; three genes (*AVL9*, *DEPTOR*, *CARS*) showed a possible transcript instability occurring in both tumoral and normal mutation-bearing tissues; five genes (*PSCA*, *ENO3*, *PARP14*, *CTBP2*, *FOXM1*) were upregulated in lung cancer tissue confirming previous published results and suggesting their role as oncogene. No mutations or gene expression alteration in genes belonging to the Hedgehog pathway, mainly involved in basal cell carcinoma development, were found in our case.

The oligogenic combination of cancer predisposing genes found in our patient supports the model of "private genetic epidemiology" for lung cancer susceptibility [5]. In

fact, we found that 11 cancer-related genes were mutated in the patient and only one in healthy sib (**Figure 2**), suggesting that the oligogenic control of lung cancer susceptibility is due to a peculiar signature of multiple germline mutations in cancer-related genes probably with a dosage effect. This personal signature may play a role in the development and growth of lung cancer and, therefore explain the non-hereditability of the condition, especially in non-smokers patients.

Focusing our analysis on genes that have already been reported in association with known druggable target genes -and likely having a major role in driving tumor development in our case-, we identified two possible suitable targets. Acetyl-CoA carboxylase alpha (ACACA) is a key enzyme in de novo lipogenesis catalyzing the carboxylation of acetyl-CoA to malonyl-CoA. ACACA is known to be one of the major BRCA1-interacting protein [12]. Cancer cells strongly need to de novo synthesize fatty acid for membrane and energy production [13] and interaction between ACACA and BRCA1 is likely to be important for BRCA1-mediated tumor suppressor activity [14]. BRCA1 indeed binds the phosphorylated and inactive form of ACACA and this interaction blocks ACACA activity by preventing its dephosphorylation [14]. The identified variant c.C1948T, p.Arg650Trp is outside the catalytic domain but may have a role in folding of domain conformation determining a different exposure for BRCA1 binding. Based on the evidence that in our tumoral tissue only the mutated transcript is present, we can assume that, even if protein level is decreased, the isoform loses negative control by BRCA1 and remains active. In this case, patient could potentially benefit from a therapy with ACACA-inhibitors such as miR-195 [15] or soraphen A [16] (**Figure 3A**).

The second druggable variant maps in *DEPTOR* gene, whose loss determines the activation of mTOR cascade and promotes cell growth and survival. It is known that

DEPTOR mainly acts as a tumor suppressor [17, 18]. A recent study in lung cancer demonstrated that activation of *EGFR* signaling resulted in downregulation of *DEPTOR* expression. In the same study, tumor progression mediated by *EGFR* ectopic expression was diminished by transfection with *DEPTOR* in lung adenocarcinoma cell lines and that treatment with Gefitinib, an *EGFR* tyrosine kinase inhibitor, stimulated *DEPTOR* expression leading to a significantly decreased cell proliferation, migration and invasion [19]. In line with these results, we observed an increased expression level of *DEPTOR* in normal and tumoral tissue, suggesting an enhanced predisposition of activation of mTOR signaling and thus a promotion of tumor progression (**Figure 3B**). These results suggested the possibility to use Gefitinib in our patient in order to restore *DEPTOR* level and consequently suppress cell growth and survival.

In summary, we suggest that both “primary” squamous cell carcinoma are triggered by the presence of an oligogenic germline signature consisting of at least 11 mutations, among which 2 lead to the activation of mTOR and *BRCA1*. Thus, this finding opens up the possibility of treating one of them with an already commercially available drug and the other with developing molecules.

This study has a number of limitations. First, alterations in non-coding regions and copy number variants would have been missed by our approach. Another limitation is that we focused only on cancer-associated genes, but we cannot exclude that non-cancer genes, in particular those involved in immune system or drug metabolism, could play a role in cancer development. Finally, additional functional studies will be required to translate our results into effective personalized therapy.

The present report provides evidence that an “omic” sequencing approach in cancer patient is feasible, and could impact on treatment decisions, namely in patients not suitable for

conventional therapy. In the present era of “precision medicine”, our study demonstrates the utility of applying this comprehensive sequencing protocol in order to effectively identify possible therapeutic targets enabling personalized therapies.

Conflicts of Interest

Conflict of interest relevant to this article was not reported.

Acknowledgments

This work was supported by grant from Regione Toscana – Istituto Toscano Tumori (ITT) and from ASSO (Associazione per lo Sviluppo della Scienza Oncologica) onlus. The authors thank Fondazione Umberto Veronesi for its support to EF through the Post-Doctoral Fellowships 2015. The funders had no role in the design or conduct of the study, in the collection, analysis or interpretation of the data, or in the preparation, review or approval of the manuscript.

References

1. Stewart BW & Wild CP. World Cancer Report. International Agency for Research on Cancer. 2014.
2. Samet JM, Avila-Tang E, Boffetta P, Hannan LM, Olivo-Marston S, Thun MJ, et al. Lung cancer in never smokers: clinical epidemiology and environmental risk factors. Clin Cancer Res. 2009;15:5626–45.
3. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat Genet. 2016;48:607-16.
4. Vanni I, Coco S, Bonfiglio S, Cittaro D, Genova C, Biello F, et al. Whole exome sequencing of independent lung adenocarcinoma, lung squamous cell carcinoma, and malignant peritoneal mesothelioma: A case report. Medicine (Baltimore). 2016;95:e5447.
5. Renieri A, Mencarelli MA, Cetta F, Baldassarri M, Mari F, Furini S, et al. Oligogenic germline mutations identified in early non-smokers lung adenocarcinoma patients. Lung Cancer. 2014;85:168-74.
6. Imperatore V, Mencarelli MA, Fallerini C, Bianciardi L, Ariani F, Furini S, et al. Potentially Treatable Disorder Diagnosed Post Mortem by Exome Analysis in a Boy with Respiratory Distress. Int J Mol Sci. 2016;17:306.
7. Trapnell C, Pachter L & Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25:1105-11.
8. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562-78.

9. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511-515.
10. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature. 2013;500:415-21.
11. Thong L. Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. Cell Mol Life Sci. 2005;62:1784-1803.
12. Magnard C, Bachelier R, Vincent A, Jaquinod M, Kieffer S, Lenoir GM, et al. BRCA1 interacts with acetyl-CoA carboxylase through its tandem of BRCT domains. Oncogene. 2002;21:6729-39.
13. Currie E, Schulze A, Zechner R, Walther TC, Farese RV Jr. Cellular fatty acid metabolism and cancer. Cell Metab. 2013;18:153-161.
14. Moreau K, Dizin E, Ray H, Luquain C, Lefai E, Foufelle F, et al. BRCA1 affects lipid synthesis through its interaction with acetyl-CoA carboxylase. J Biol Chem. 2006;281:3172-81.
15. Singh R, Yadav V, Kumar S, Saini N. MicroRNA-195 inhibits proliferation, invasion and metastasis in breast cancer cells by targeting FASN, HMGCR, ACACA and CYP27B1. Sci Rep. 2015;5:17454.
16. Cordonier EL, Jarecke SK, Hollinger FE, Zempleni J. Inhibition of acetyl-CoA carboxylases by soraphen A prevents lipid accumulation and adipocyte differentiation in 3T3-L1 cells. Eur J Pharmacol. 2016;780:202-8.
17. Peterson TR, Laplante M, Thoreen CC, Sancak Y, Kang SA, Kuehl WM et al. DEPTOR is an mTOR inhibitor frequently overexpressed in multiple myeloma cells and required for their survival. Cell. 2009;137:873-886

18. Li H, Sun GY, Zhao Y, Thomas D, Greenson JK, Zalupski MM, et al. DEPTOR has growth suppression activity against pancreatic cancer cells. Oncotarget. 2014;5:12811-9.
19. Zhou X, Guo J, Ji Y, Pan G, Liu T, Zhu H, et al. Reciprocal Negative Regulation between EGFR and DEPTOR Plays an Important Role in the Progression of Lung Adenocarcinoma. Mol Cancer Res. 2016;14:448-57.
20. Chiu HC, Chang TY, Huang CT, Chao YS, Hsu JT. EGFR and myosin II inhibitors cooperate to suppress EGFR-T790M-mutant NSCLC cells. Mol Oncol. 2012;6:299-310.
21. Zavasnik-Bergant T, Turk B. Cysteine cathepsins in the immune response. Tissue Antigens. 2006;67:349-55.
22. Tang J, Li Y, Lyon K, Camps J, Dalton S, Ried T, et al. Cancer driver-passenger distinction via sporadic human and dog cancer comparison: a proof-of-principle study with colorectal cancer. Oncogene. 2014;33:814-22.
23. Linford A, Yoshimura S, Nunes Bastos R, Langemeyer L, Gerondopoulos A, Rigden DJ, et al. Rab14 and its exchange factor FAM116 link endocytic recycling and adherens junction stability in migrating cells. Dev Cell. 2012;22:952-66.
24. Hu RJ, Lee MP, Connors TD, Johnson LA, Burn TC, Su K, et al. A 2.5-Mb transcript map of a tumor-suppressing subchromosomal transferable fragment from 11p15.5, and isolation and sequence analysis of three novel genes. Genomics. 1997;46:9-17.
25. Kawaguchi T, Sho M, Tojo T, Yamato I, Nomi T, Hotta K, et al. Clinical significance of prostate stem cell antigen expression in non-small cell lung cancer. Jpn J Clin Oncol. 2010;40:319-26.
26. Li JL, Fei Q, Yu J, Zhang HY, Wang P, Zhu JD. Correlation between methylation profile of promoter cpg islands of seven metastasis-associated genes and their expression states in six cell lines of liver origin. Ai Zheng. 2004;23:985-91.

27. Barbarulo A, Iansante V, Chaidos A, Naresh K, Rahemtulla A, Franzoso G, et al. Poly(ADP-ribose) polymerase family member 14 (PARP14) is a novel effector of the JNK2-dependent pro-survival signal in multiple myeloma. Oncogene. 2013;32:4231-42.
28. Zhang C, Li S, Qiao B, Yang K, Liu R, Ma B, et al. CtBP2 overexpression is associated with tumorigenesis and poor clinical outcome of prostate cancer. Arch Med Sci. 2015;11:1318-23.
29. Wang Y, Che S, Cai G, He Y, Chen J, Xu W. Expression and prognostic significance of CTBP2 in human gliomas. Oncol Lett. 2016;12:2429-2434.
30. Kim IM, Ackerson T, Ramakrishna S, Tretiakova M, Wang IC, Kalin TV, et al. The Forkhead Box m1 Transcription Factor Stimulates the Proliferation of Tumor Cells during Development of Lung Cancer. Cancer Res. 2006;66:2153–2161.

Figure 1. Flowchart illustrating the study strategy and result.

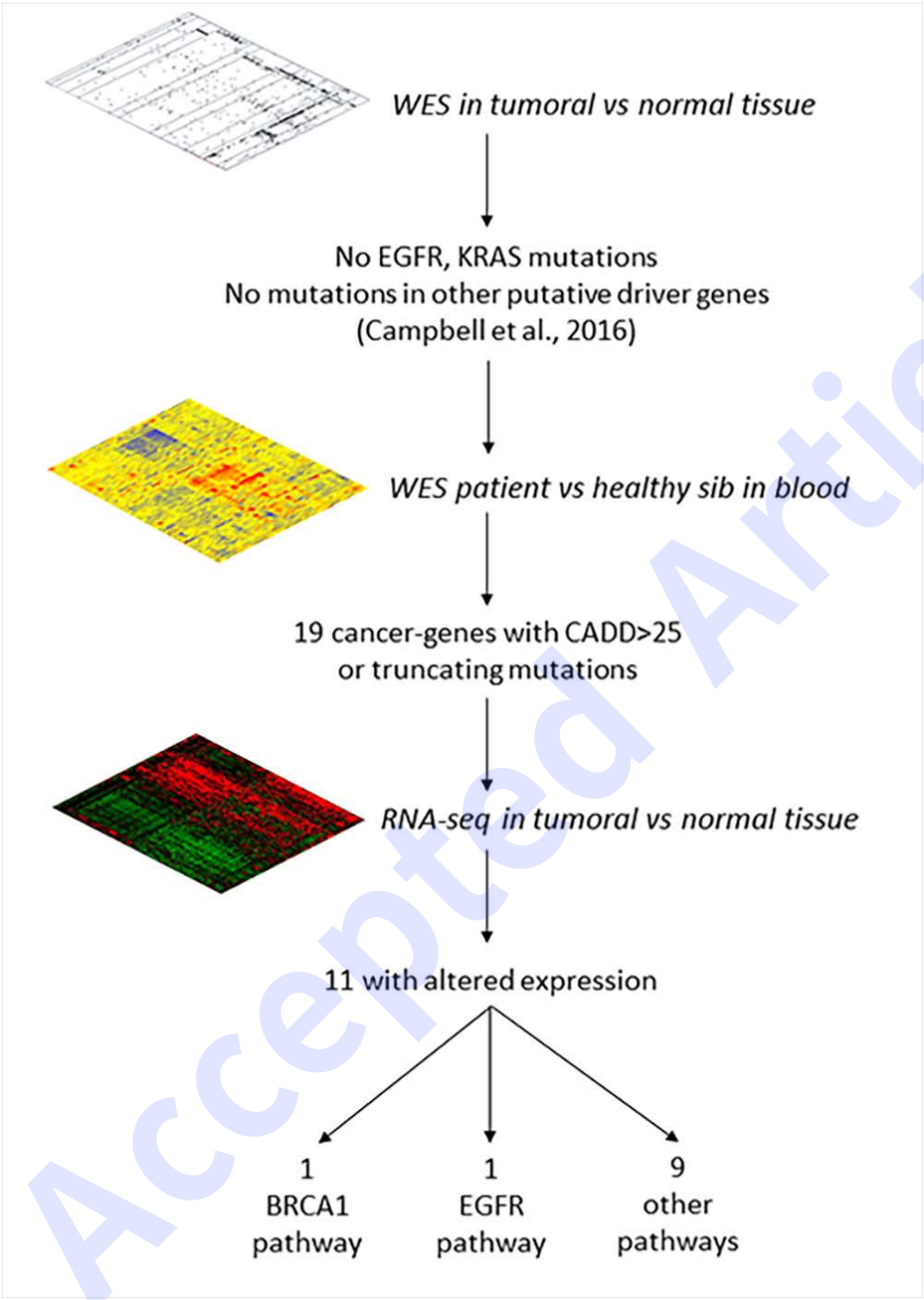


Figure 2. Pedigree of the two sib-pairs and representation of disruptive mutations in cancer-related genes found in the patient and in her unaffected sib.

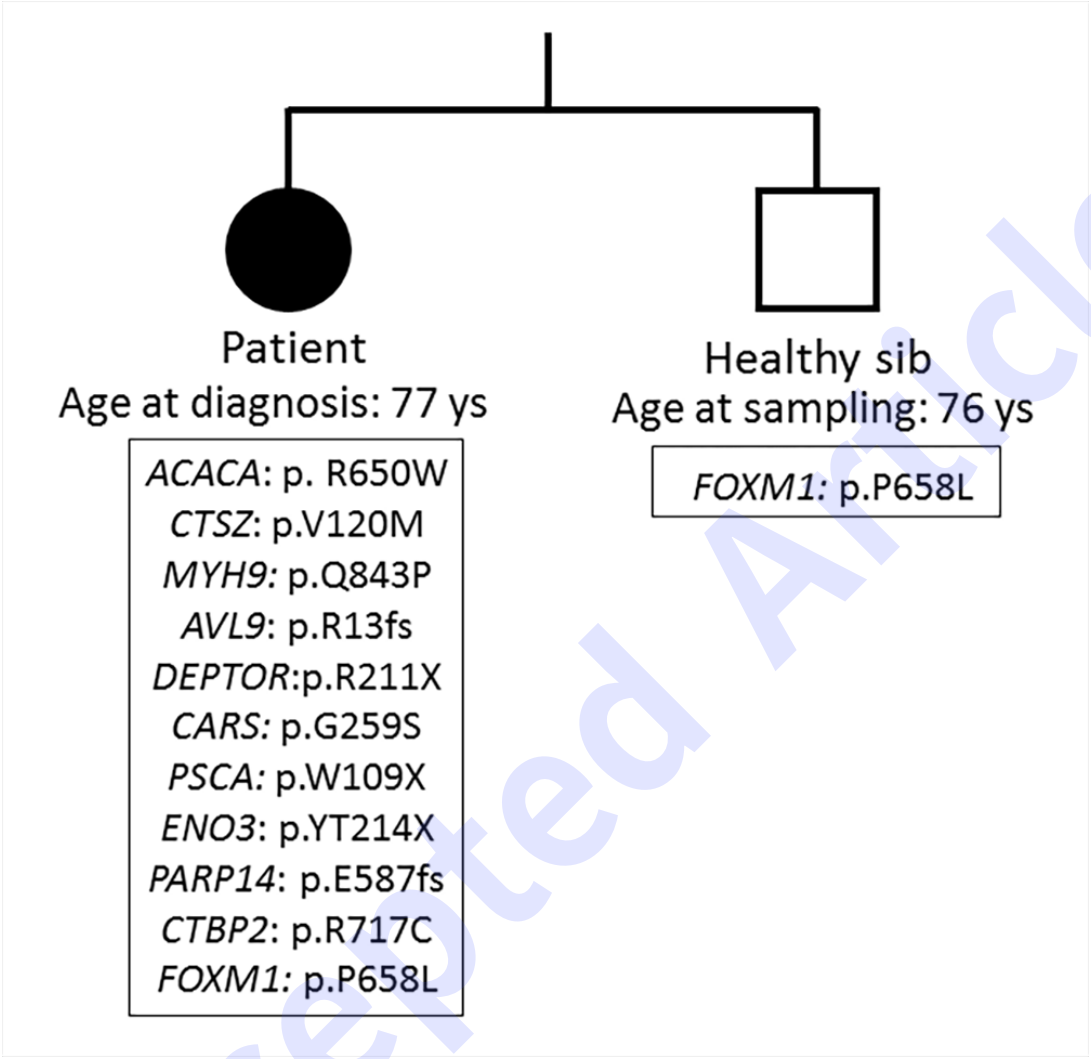


Figure 3. Likely personal germline driver genes in present case and possible therapeutic options. **A)** *ACACA* and **B)** DEPTOR pathway involvement.

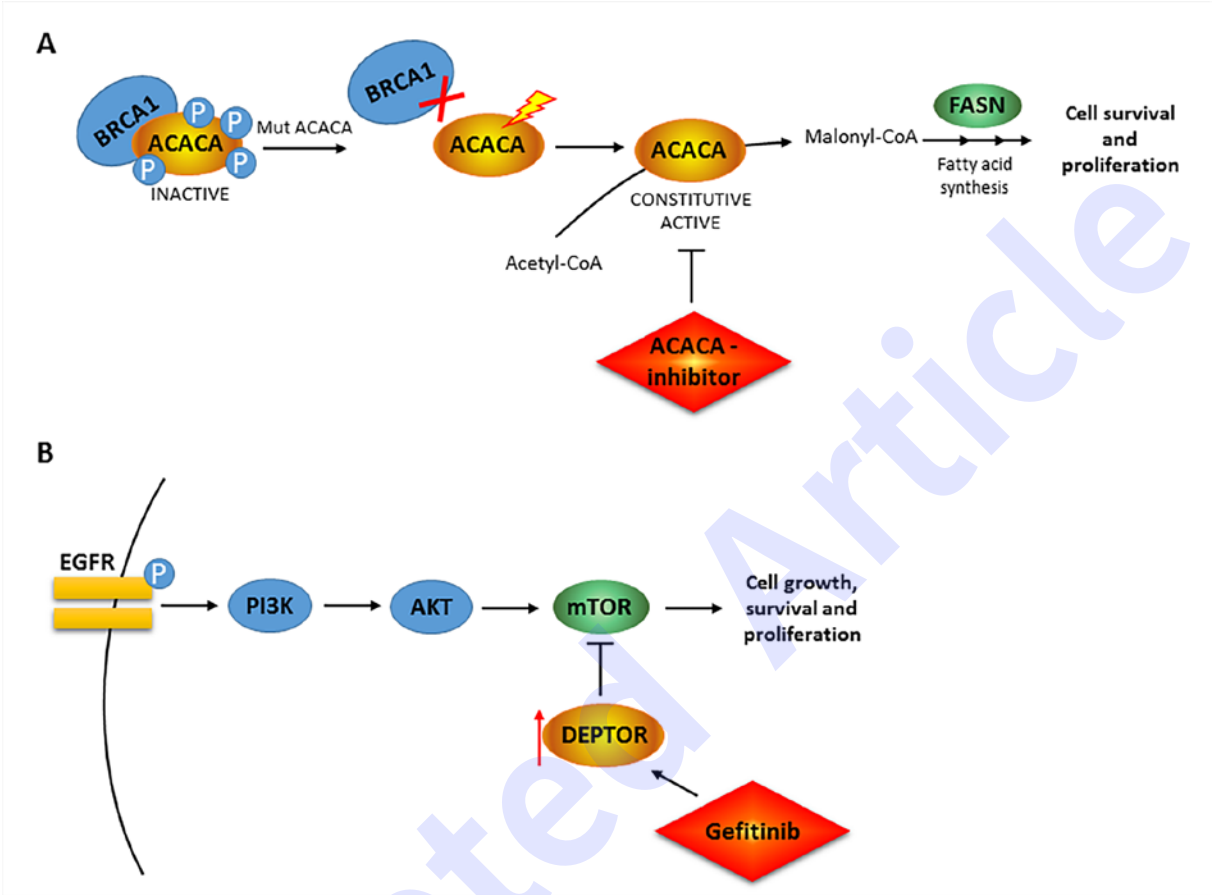


Table 1. Expression levels of 11 cancer-related germline variants (Fold-change>±1.5).

Group	Gene symbol	Amino acid change (CADD value)	Fold-change T/N	Gene function related to cancer development
GROUP 1: Under-expressed in tumoral tissue (Fold-change≤0.66)	*ACACA*	p. R650W (33)	-4.35	Acetyl-CoA carboxylase alpha has a role in de novo lipogenesis catalyzing the carboxylation of acetyl-CoA to malonyl-CoA. The enzyme is under long term control at the transcriptional and translational levels. It is known to specifically interact with *BRCA1* [12].
	MYH9	p.Q843P (31)	-1.64	Myosin heavy chain 9 is a myosin IIA heavy chain that interacts with β-actin to maintain cytoskeleton integrity. It is a *EGFR*-associating protein in NSCLC cell lines overexpressing *EGFR* [20].
	CTSZ	p.V120M (26)	-1.59	Cathepsin Z is a lysosomal cysteine proteinase. *CTSZ* is mostly expressed in immune cells, such as macrophages and monocytes, and involved in antigen processing and maturation of the major histocompatibility complex class II molecules [21]. This gene is expressed ubiquitously in cancer cell lines and primary tumors.
GROUP 2: Under-expressed in both normal and tumoral tissue (Fold-change≤0.66)	*AVL9*	p.R13fs (NA)	-3,13*	*AVL9* cell migration associated is among the 73 cancer driver candidate genes previously identified via dog-human comparison, deleted in colorectal cancer [22] and classified as putative tumor suppressors. *AVL9*-depletion inhibits the migration of A549 human adenocarcinomic alveolar basal epithelial cells [23].
	CARS	p.G259S (33)[a]	-3.03*	Cysteinyl-tRNA synthetase is one of several located near the imprinted gene domain on chromosome 11p15.5, an important tumor-suppressor gene region [24].
	DEPTOR	p.R211X (46)	-1.96*	DEP-domain containing MTOR interacting protein has growth suppression activity against pancreatic cancer cells, and its expression is gradually lost during pancreatic tumorigenesis [18]. Tumor progression mediated by *EGFR* ectopic expression was diminuished by transfection with *DEPTOR* in lung adenocarcinoma. Gefitinib, a specific *EGFR* inhibitor used for lung cancer therapy, stimulated *DEPTOR* accumulation [19].

GROUP 3: Over-espressed in tumoral tissue (Fold-change≥+ 1.5)	*PSCA*	p.W109X (NA)[a]	391.96	Prostate stem cell antigen encodes a glycosylphosphatidylinositol-anchored cell membrane glycoprotein that is upregulated in a large proportion of prostate cancers and other cancers. It is highly expressed in non-small cell lung cancer and may be functionally important for the disease: small interfering RNA-mediated knockdown of PSCA resulted in the inhibition of LC growth [25].
	ENO3	p.YT214X (35)	7.04	Enolase 3 encodes one of the three enolase isoenzymes. It was found upregulated in liver cancer mouse model [26].
	PARP14	p.E587fs (NA)	4.93	Poly(ADP-ribose) polymerase family member 14 is an anti-apoptotic protein that may regulate aerobic glycolysis and promote survival of cancer cells [27]. Its increased expression has been reported in a variety of tumor types.
	CTBP2	p.R717C (34)	4.34	C-terminal binding protein 2 is upregulated in a number of tumors, such as hepatocellular carcinoma, prostate cancer and gliomas [28,29].
	FOXM1	p.P658L (27) [a]	3.28	Forkhead box M1 is a transcriptional activator involved in cell proliferation. Consistent with an important role of Foxm1 in cell cycle progression, increased expression of Foxm1 was found in many human tumors. Increased Foxm1 expression in human lung adenocarcinomas and squamous cell carcinomas was associated with increased proliferation of tumor cells [30].

NA, not applicable

[a] All listed variants were private except those in *CARS* (MAF=0.002), PSCA (MAF=0.0076) and *FOXM1* (MAF=0.0044) genes.

*****Fold-change calculated by comparing expression levels of normal and tumor tissues with those of a pool of normal lung tissues.

S1 Table. Clinical characteristics and sequencing data of case and healthy sibling.

	Case	Sib
Sex	F	M
Age at diagnosis, years	77	NA
Age at sampling, years	NA	76
Smoker status	Never	Never
Histologic type	SQCC	NA
Follow-up status	Alive	Alive
Age at follow-up	78	77 [a)]
Exome sequence data, Mbp	4,06	5,47
Number of Reads, Mil	23,5	31,6

NA, not applicable
[a)] cancer-free

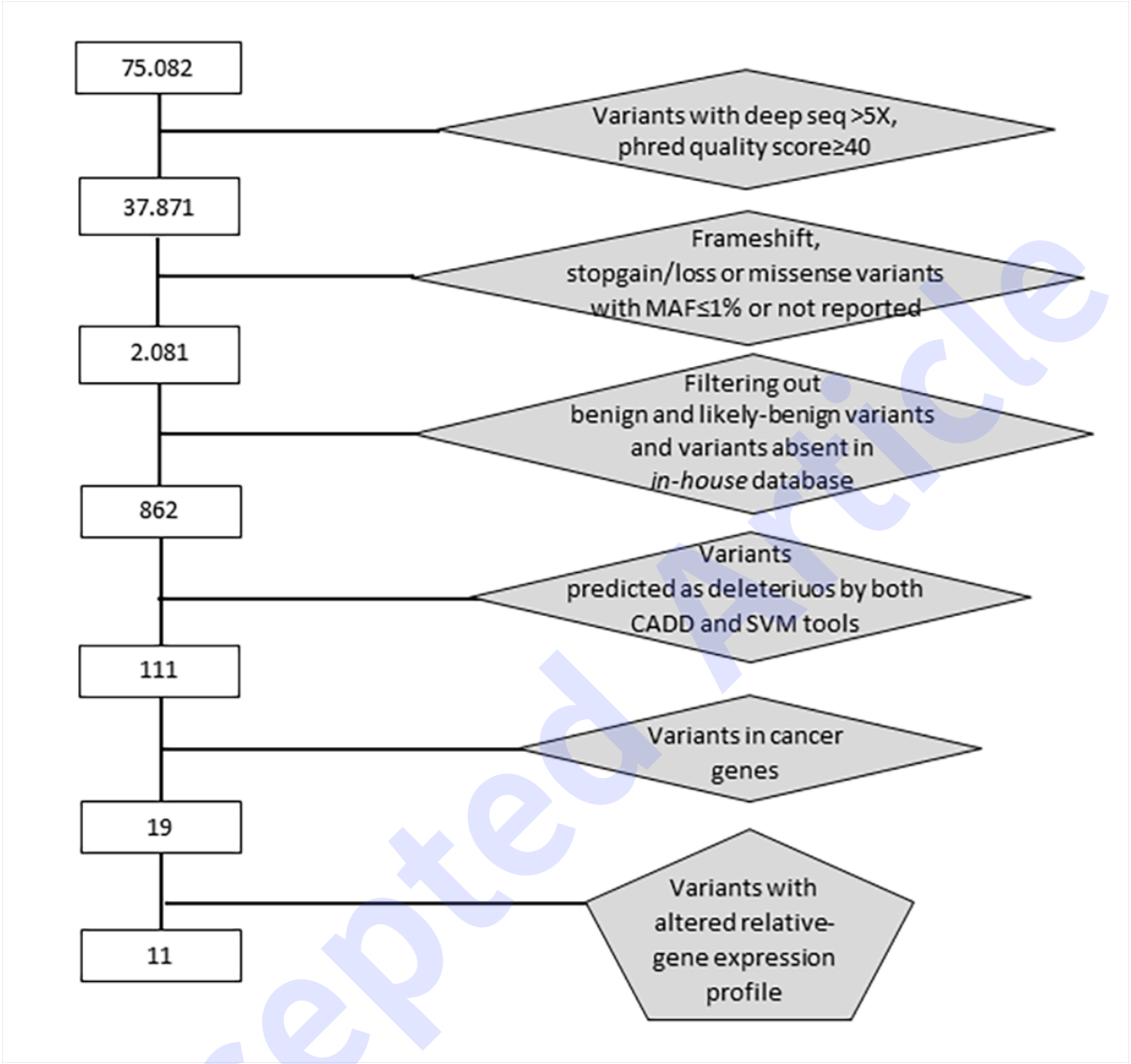
S2 Table. Dull genetic data of 19 cancer-related germline rare variants identified by whole-exome sequencing.

Gene	Chr	Start	End	Ref	Alt	Mutation	Accession number	Exon	Nucleotide change	Amino acid change
ACACA	chr17	35605011	35605011	G	A	nonsynonymous	NM_198837	exon16	c.C1948T	p.R650W
ATP2C1	chr3	130712791	130712791	G	A	nonsynonymous	NM_001199182	exon20	c.G1891A	p.G631S
AVL9	chr7	32535358	32535359	CG	-	frameshift deletion	NM_015060	exon1	c.37_38del	p.R13fs
CARS	chr11	3050233	3050233	C	T	nonsynonymous	NM_001751	exon8	c.G775A	p.G259S
CRYGD	chr2	208988920	208988920	G	C	Stopgain	NM_006891	exon2	c.C168G	p.Y56X
CTBP2	chr10	126686569	126686569	G	A	nonsynonymous	NM_022802	exon4	c.C2149T	p.R717C
CTSZ	chr20	57576649	57576649	C	T	nonsynonymous	NM_001336	exon3	c.G358A	p.V120M
DEPTOR	chr8	121019052	121019052	A	T	Stopgain	NM_001283012	exon5	c.A631T	p.R211X
ENO3	chr17	4858805	4858805	C	G	Stopgain	NM_001193503	exon6	c.C642G	p.Y214X
FOXM1	chr12	2968078	2968078	G	A	nonsynonymous	NM_001243088	exon8	c.C1973T	p.P658L
GPRC6A	chr6	117128308	117128308	C	T	nonsynonymous	NM_001286354	exon3	c.G560A	p.R187Q
LRP2	chr2	170042245	170042245	T	C	nonsynonymous	NM_004525	exon50	c.A9613G	p.N3205D
MYH9	chr22	36697683	36697683	T	G	nonsynonymous	NM_002473	exon21	c.A2528C	p.Q843P
PARP14	chr3	122419160	122419160	-	AAGCAGAA	frameshift insertion	NM_017554	exon6	c.1759_1760insAAGCAGAA	p.E587fs
PHF12	chr17	27254016	27254016	C	T	nonsynonymous	NM_001033561	exon3	c.G314A	p.R105H
PRMT2	chr21	48078714	48078714	G	A	nonsynonymous	NM_001242866	exon7	c.G712A	p.G238R
PSCA	chr8	143763531	143763531	G	A	Stopgain	NM_005672	exon3	c.G326A	p.W109X
SDF2L1	chr22	21998280	21998280	G	A	nonsynonymous	NM_022044	exon3	c.G482A	p.R161H
TPPP3	chr16	67424196	67424196	G	A	nonsynonymous	NM_015964	exon4	c.C412T	p.R138C

S3 Table. RNA-seq data summary

Tissue type	Yield (Mbases)	Total reads	% of ≥ Q30 Bases
Non-tumoral	7,040	69,701,750	91.06
Tumoral	8,137	80,567,970	91.42

S1 Figure. Flowchart illustrating WES filtering process and variants selection.



S2 Figure. Mutation spectra of SQCC cancer. **A)** Prevalence of somatic mutations. **B)** Frequency of the eight possible base pair substitutions/deletion. **C)** Results of pathway analysis of mutated genes in lung tumoral tissue. Pathways sorted from smallest to largest p-value (indicated within the bars).