Journal of Pathology
J Pathol 2023
Published online 23 June 2023 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/path.6141

BRIEF REPORT

Defining the first *bona fide* cell model for SMARCA4-deficient, undifferentiated tumor

Alberto M Arenas[1,2,3], José Manuel Ruiz-Jiménez[2,4], Javier L López-Hidalgo[3,5], Juan Sanjuán-Hidalgo[1,2] and Pedro P Medina[1,2,3]*

[1] *Department of Biochemistry and Molecular Biology I, Faculty of Sciences, University of Granada, Granada, Spain*
[2] *GENYO, Centre for Genomics and Oncological Research: Pfizer/University of Granada/Andalusian Regional Government, Granada, Spain*
[3] *Health Research Institute of Granada (ibs.Granada), Granada, Spain*
[4] *FIBAO, Fundación Pública para la Investigación Biosanitaria de Andalucía Oriental Alejandro Otero, Granada, Spain*
[5] *Hospital Universitario Clínico San Cecilio (HUCSC), Granada, Spain*

Correspondence to: PP Medina, Avenida de la Ilustración 114, 18007, Granada, Spain. E-mail: pedromedina@ugr.es

Abstract

The World Health Organization's tumor classification guidelines are frequently updated and renewed as knowledge of cancer biology advances. For instance, in 2021, a novel lung tumor subtype named SMARCA4-deficient, undifferentiated tumor (SMARCA4-dUT, code 8044/3) was included. To date, there is no defined cell model for SMARCA4-dUT that could be used to help thoracic clinicians and researchers in the study of this newly defined tumor type. As this tumor type was recently described, it is feasible that some cell models formerly classified as lung adenocarcinoma (LUAD) could now be better classified as SMARCA4-dUT. Thus, in this work, we aimed to identify a *bona fide* cell model for the experimental study of SMARCA4-dUT. We compared the differential expression profiles of 36 LUAD-annotated cell lines and 38 cell lines defined as rhabdoid in repositories. These comparative results were integrated with the mutation and expression profiles of the SWI/SNF complex members, and they were surveyed for the presence of the SMARCA4-dUT markers SOX2, SALL4, and CD34, measured by RT-qPCR and western blotting. Finally, the cell line with the paradigmatic SMARCA4-dUT markers was engrafted into immunocompromised mice to assess the histological morphology of the formed tumors and compare them with those formed by a *bona fide* LUAD cancer cell line. NCI-H522, formerly classified as LUAD, displayed expression profiles nearer to rhabdoid tumors than LUAD tumors. Furthermore, NCI-H522 has most of the paradigmatic features of SMARCA4-dUT: hemizygous inactivating mutation of SMARCA4, severe SMARCA2 downregulation, and high-level expression of stem cell markers SOX2 and SALL4. In addition, the engrafted tumors of NCI-H522 did not display a typical differentiated glandular structure as other *bona fide* LUAD cell lines (A549) do but had rather a largely undifferentiated morphology, characteristic of SMARCA4-dUT. Thus, we propose the NCI-H522 as the first *bona fide* cell line model of SMARCA4-dUT.

Keywords: lung cancer; SMARCA4; biologic model; cell culture; adenocarcinoma of lung; undifferentiated carcinoma; SWI/SNF; SMARCA2; rhabdoid tumor; cell line

Received 10 January 2023; Revised 23 March 2023; Accepted 11 May 2023

No conflicts of interest were declared.

Introduction

A proper classification of the different subtypes of cancer is critical to better approach the prognosis and treatment of patients, as well as to define study models for preclinical research. Traditionally, tumor classification has been based on the histomorphological features that tumors show upon surgery. However, thanks to the advances in molecular genomics over the past decade, new cancer subtypes are emerging whose distinction from other subtypes is entirely 'omics'-based [1]. Thus, the World Health Organization (WHO) periodically updates its tumor classification guidelines, integrating the histomorphological and molecular distinctions of cancer subtypes.

In 2021, the WHO revised its classification of lung cancers, defining a novel subtype called SMARCA4-deficient, undifferentiated tumors (SMARCA4-dUT, code 8044/3) [2]. *SMARCA4* (BRG1) is one of the two core catalytic subunits of the SWI/SNF chromatin remodeling complex, and it appears recurrently mutated in lung cancer [3,4]. SMARCA4-dUT tumors were previously classified as SMARCA4-deficient, non-small cell lung cancer (NSCLC). However, they now constitute a separate group outside of NSCLC with their own distinct histological and immunophenotypical features: *SMARCA4* mutational

inactivation; *SMARCA2* loss or severe downregulation; a rhabdoid-like or undifferentiated morphology, with layers of cells lacking any glandular architecture; and the expression of the pluripotency markers SOX2, SALL4, and/or CD34. These features allow the differentiation of SMARCA4-dUT from other SMARCA4-deficient lung adenocarcinomas (LUAD) and the morphologically similar malignant rhabdoid tumors (MRT) [2,5,6].

We recently characterized the mutational status and expression profile of 20 SWI/SNF members in 38 LUAD cell lines and tumor patient samples [7,8]. Because SMARCA4-dUT was just recently included in the 2021 WHO classification of lung tumors, there is a possibility that some of the cell lines previously classified as LUAD cell lines could now be better reinterpreted as SMARCA4-dUT. Therefore, in this study, we aimed to define the first *bona fide* SMARCA4-dUT cell line model.

Materials and methods

Expression data analysis from DepMap

We downloaded transcriptomic data and cell line information from the Cancer Dependency Map Project (DepMap) database (https://depmap.org/portal/download/). We selected 36 LUAD cell lines and a total of 38 cell lines annotated as 'rhabdoid' [MRT ($N = 14$); small cell carcinoma of the ovary, hypercalcemic type ($N = 2$); atypical teratoid rhabdoid tumor ($N = 5$); and rhabdomyosarcoma ($N = 17$)].

A hierarchical clustering analysis was performed and displayed using the pheatmap package in R version 4.2.1 (R Core Team 2022, Vienna, Austria). Before clustering, we filtered the genes with the highest variance ($N = 500$), setting a maximum threshold of correlation ($\rho < 0.8$), avoiding any redundancy, and normalized the expression values of every gene. The distance between samples was measured using the Euclidean distance method, and the Ward method was used for gene and cell line clustering.

Cell culture

The cell lines A427, A549, NCI-H23, NCI-H358, and NCI-H522 were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA). They were all cultured under standard conditions (37°C, 5% CO_2) in RPMI 1640 medium (Biowest, Nuaillé, France, #L0501), supplemented with 10% FBS (Thermo Fisher Scientific, Waltham, MA, USA, #10270-106), 1% penicillin–streptomycin (Merck, Darmstadt, Germany, #P0781), and 1% L-glutamine (VWR, Radnor, PA, USA, #X0550). Cells in culture were regularly tested for mycoplasma contamination.

Assessing expression of SMARCA4-dUT markers in cell lines

Total RNA and protein were extracted from dry pellets of the A427, A549, NCI-H23, NCI-H358, and NCI-H522 cell lines. RNA expression levels were measured by reverse transcription quantitative PCR (RT-qPCR). Total extracted RNA (2 µg) was treated with DNase I (Invitrogen, Waltham, MA, USA, #18068015) and then reverse-transcribed to complementary DNA using RevertAid RT kit (Thermo Fisher Scientific, #K1691). A SYBR Green quantitative PCR reaction was performed using KAPA SYBR® FAST (Merck, #SFUKB) and the QuantStudio™ 3 Real-Time PCR System (Thermo Fisher Scientific). Relative expression was calculated using the ΔΔCt method, with the ΔCt mean of all cell lines to normalize and *GAPDH* as the reference gene. All experiments were performed in triplicate. Statistical analysis was performed using GraphPad Prism 9.0's unpaired *t*-test comparing every cell line to the NCI-H522. The oligonucleotides used for each gene are listed in supplementary material, Table S1.

Protein levels were analyzed by western blotting. Total protein (50 µg) was loaded into an 8% polyacrylamide gel (AppliChem, Darmstadt, Germany, #A1577) and separated by molecular weight using SDS-PAGE. The NCI-H358 cell line (LUAD, SMARCA4-wildtype) was used as positive control for both SMARCA4 and SMARCA2. Proteins were transferred to a PVDF membrane, and then the membrane was blocked with 5% milk powder in PBS (VWR, #392-0442) + 1% Tween-20 (VWR, #0777). The membrane was incubated with primary antibodies against human SMARCA4 (Santa Cruz Biotechnology, Dallas, TX, USA #SC10768, 1:500 dillution), SMARCA2 (Cell Signaling Technology, Danvers, MA, USA, #11966S, 1:500), SOX2 (Cell Signaling Technology, #3579, 1:1000), SALL4 (Abcam, Cambridge, UK, #ab57577, 1:1000), and α-tubulin (Santa Cruz Biotechnology, #SC23948, 1:10000) as a loading control. The membrane was incubated with the secondary antibodies horseradish peroxidase (HRP)-Mouse IgG (Agilent Dako, Santa Clara, CA, USA, #P0447, 1:5000) or HRP-Rabbit IgG (Agilent Dako, #P0448, 1:2000), and the bands were revealed using SuperSignal™ substrate (Thermo Fisher Scientific, #A43841), and an ImageQuant™ LAS-4000 imaging system. Raw western blotting images can be found in supplementary material, Figures S1–S5. All experiments were carried out in triplicate.

Morphology of tumor xenografts in mice

Animal research procedures were performed following the European Directive 2010/63/EU. A549, A427, or NCI-H522 cells (1×10^6 cells/tumor) were injected subcutaneously [total volume of 200 µl of 1:1 RPMI:Matrigel (Corning, #354248)] into the rear flanks of 6- to 8-week-old male NOD Scid Gamma (NSG) mice. Seven weeks later, mice were sacrificed, and their tumors were extracted and fixed overnight at 4°C in 10% neutral buffered formalin.

Fixed tumors were processed, and 3-µm-thick sections were obtained and stained with Harris hematoxylin/eosin-Y (Merck) using standard procedures, and immunostaining for E-cadherin [E-cadherin (36), Roche Ventana (Basel, Switzerland), ready to use], cytokeratin-7 [CK7 (OV-TL 12/30), Agilent Dako, 1:300], and CD34

[CD34 (QBEnd-10), Agilent Dako, 1:25] was also performed by AtrysHealth SA (Barcelona, Spain). Representative images were then taken using an Olympus BX43 microscope (Olympus Life Science, Waltham, MA, USA), and then the morphology of tumors was assessed by a trained pathologist (JLL-H). A summary of the histopathological analysis is provided in supplementary material, Table S2.

Results

SMARCA4 and SMARCA2 loss are the main immunohistochemical markers of SMARCA4-dUT, although some SMARCA4-deficient NSCLC tumors can also present these same features [5]. In our previous study of the SWI/SNF expression and mutational status on a panel of 38 LUAD cell lines, we discovered that three cell lines harbored *SMARCA4* inactivating mutations, and they had a null expression of *SMARCA2*: A427, NCI-H23, and NCI-H522 [7].

Perret *et al* (2019) defined the immunohistochemical signature of SMARCA4-dUT as showing the co-occurring SMARCA4 and SMARCA2 expression loss, along with SOX2 overexpression, and they recommended SALL4 and CD34 as additional useful markers for challenging, extensively necrotic biopsies [9]. Thus, we first analyzed the expression of *SOX2*, *SALL4*, and *CD34* in the three *SMARCA4/2* negative cell lines (A427, NCI-H23, and NCI-H522) and a control *bona fide* LUAD cell line (A549). Our RT-qPCR analysis showed that, compared to the mean of all four cell lines, NCI-H522 has an overexpression of 139-, 16.8-, and 2.6-fold of *SOX2*, *SALL4*, and *CD34*, respectively, thus standing out as the candidate cell line with the highest expression of all three markers (Figure 1A–C). Later, western blot analysis confirmed that, in terms of protein levels, NCI-H522 is the only cell line that simultaneously lacks both SMARCA4/SMARCA2, greatly overexpresses SOX2, and additionally expresses SALL4 (Figure 1D).

Transcriptomic data further confirmed these results: interestingly, the NCI-H522 cell line was the top

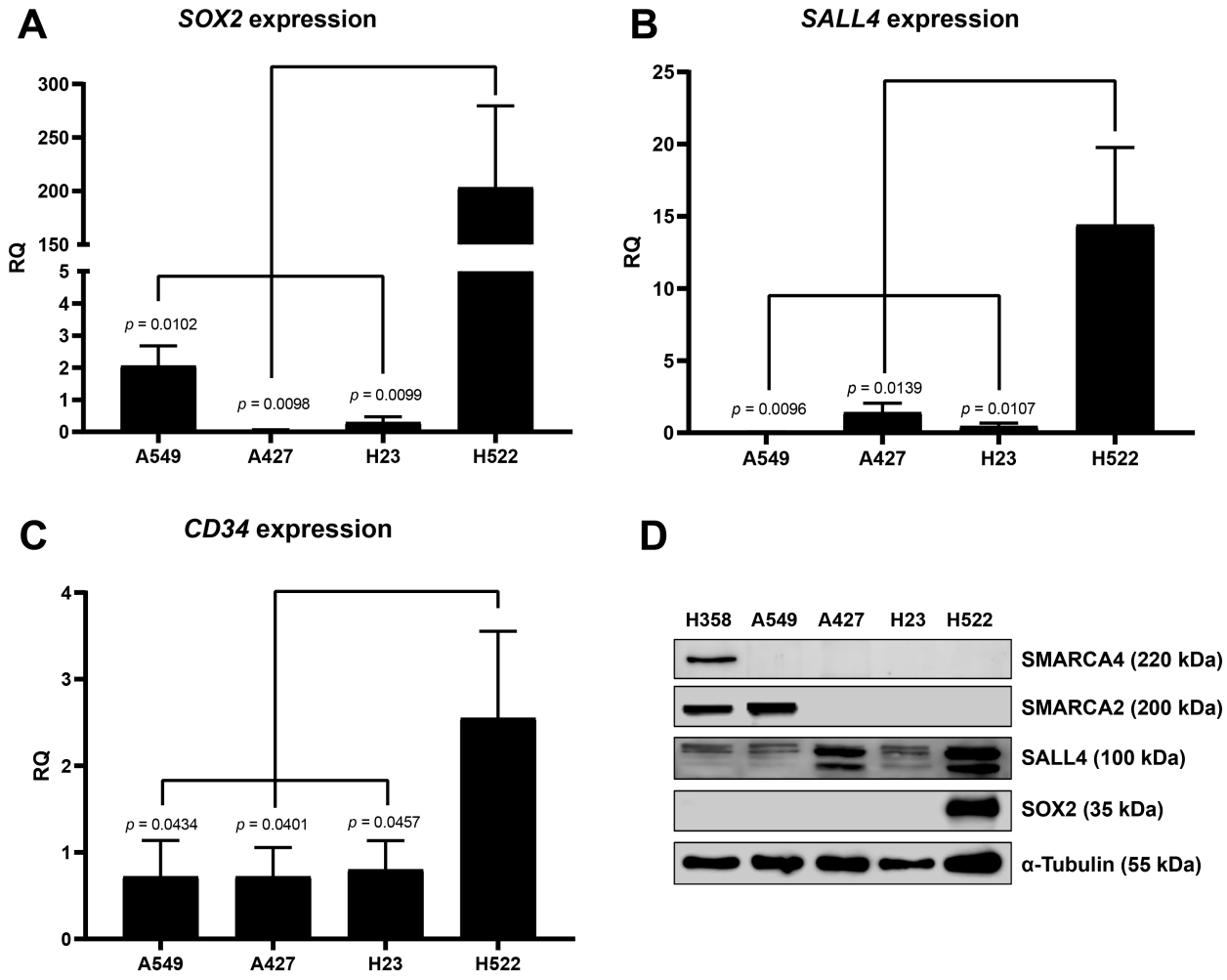

Figure 1. Expression levels of SMARCA4-dUT markers *SOX2* (A), *SALL4* (B), and *CD34* (C) in candidate cell lines (A427, NCI-H23, and NCI-H522) and a control, *bona fide* lung adenocarcinoma (LUAD) cell line (A549). Relative expression (RQ) values were normalized using *GAPDH* expression and the ΔCt mean of all cell lines. An unpaired *t*-test was performed to test the statistical significance of the data ($n = 3$, $p < 0.05$). (D) Protein levels assessed by western blotting of SMARCA4, SMARCA2, SALL4, and SOX2, with α-tubulin as a loading control and the LUAD cell line NCI-H358 as a SMARCA4/2-positive control.

SMARCA4-mutant LUAD cell line overexpressing *SOX2* and *SALL4* and the one with the lowest expression of *SMARCA2* (Figure 2A). In addition, because of the similar morphological features between SMARCA4-dUT and rhabdoid tumors [10], we analyzed the transcriptome expression data of cell lines annotated as 'LUAD' or 'rhabdoid' in the DepMap database. We clustered the 36 LUAD cell lines and 38 rhabdoid ones included in DepMap according to their expression profile, and we found that the NCI-H522 was the only LUAD-annotated cell line that clusters among the rhabdoid cell lines instead of the rest of LUAD cell lines (Figure 2B).

Finally, we assessed whether the main candidate NCI-H522 showed a typical LUAD morphology upon engraftment in mice, as expected from a LUAD cell line. Our results showed that the control A549 originated solid tumors with LUAD-characteristic acinar structures and tested positive for the epithelial and lung carcinoma markers E-cadherin and CK7 (Figure 3). Conversely, the NCI-H522 cell line tested negative for these markers, and it failed to generate such acinar structures, instead showing an undifferentiated morphology, with prominent and aberrant nuclei, high mitotic activity, and lots of intracytoplasmic inclusions, all of which are morphological characteristics from SMARCA4-dUT [2,5] (Figure 3). Regarding the A427 cell line, although it showed much poorer differentiation than A549, it was able to form some differentiated acinar structures with focal expression of E-cadherin (Figure 3), displaying some histological features from both the A549 and NCI-H522 cell lines.

Discussion

SMARCA4-dUT was addressed for the first time by Le Loarer *et al* [10], who described a cohort of patients suffering from what they named SMARCA4-deficient thoracic sarcoma (SMARCA4-DTS). The lung carcinomas of these patients bore some resemblance to MRT, such as their high undifferentiation state and the expression of the stem cell and tumorigenesis marker SOX2 [10]. Later, Yoshida *et al* [6] proposed the pluripotency markers SALL4 and CD34 as other differentially expressed markers to distinguish between SMARCA4-dUT, SMARCA4-deficient LUAD, and MRT [6].

NCI-H522 was first established in 1985 as a LUAD cell line by Banks-Schlegel, Gazdar and Harris [11], and it ranks as the 13th most widely used LUAD cell line in publications [7]. NCI-H522 was isolated from a male

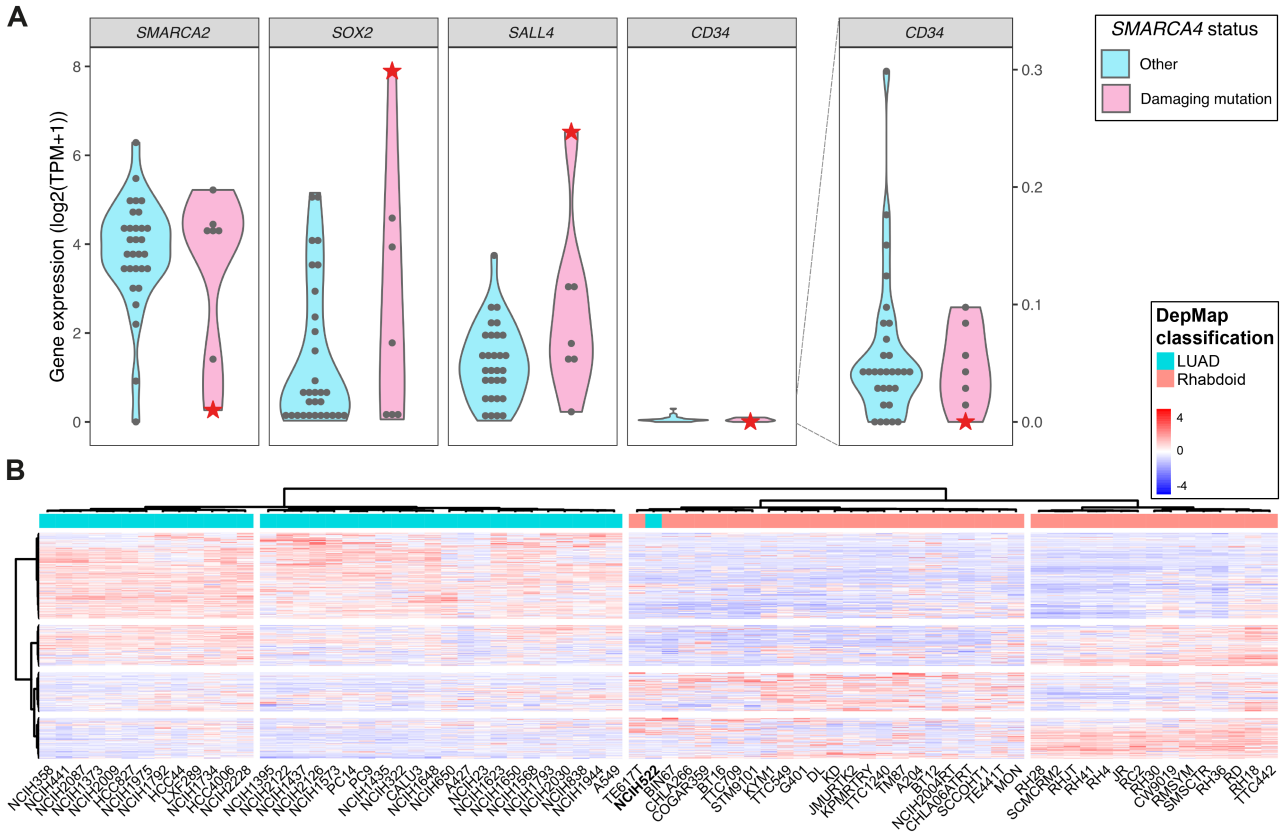

Figure 2. Expression analysis of 36 lung adenocarcinoma (LUAD) cell lines and hierarchical clustering of them and 38 rhabdoid-annotated cell lines from the DepMap portal. (A) Expression plots showing expression levels of *SMARCA2*, *SOX2*, *SALL4*, and *CD34* and the mutational status of *SMARCA4*. NCI-H522 is marked by a red star. Of note, the expression of *CD34* is very low in all LUAD cell lines, according to DepMap expression data. (B) Hierarchical clustering of the 36 LUAD- and 38 rhabdoid-annotated cell lines in the DepMap portal, according to the expression pattern of the top 500 genes with the highest variance. Heatmap units: Z-score normalization calculated from its log2(TPM + 1) expression values.

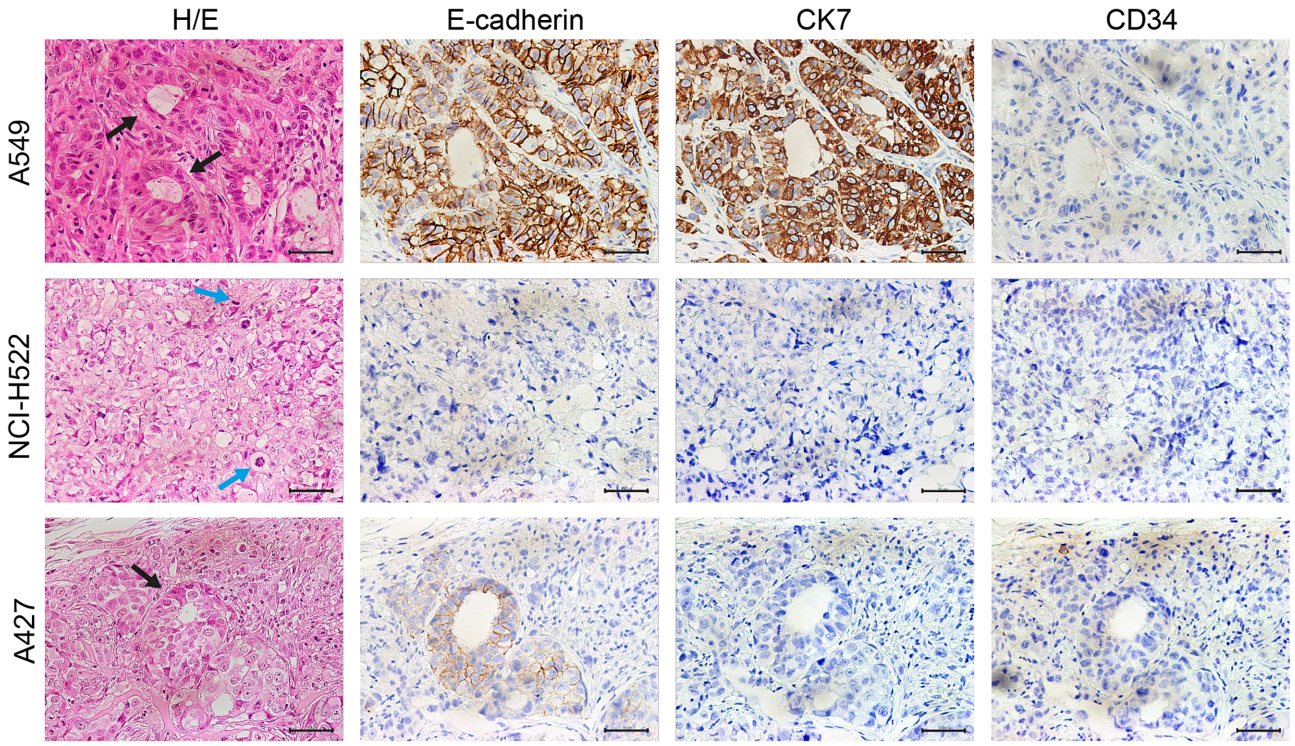

Figure 3. Histological images of A549, NCI-H522, and A427 cell line xenografts, stained with hematoxylin/eosin (H&E) and for E-cadherin, cytokeratin-7 (CK7), and CD34. A549 was able to form the acinar structures frequently found in lung adenocarcinomas (arrows in black). NCI-H522 failed to generate acinar structures or other differentiated structures displaying masses of undifferentiated cells presenting many intracytoplasmic intrusions, and a high mitotic activity (arrows in blue), typical of SMARCA4-dUT. A427 formed acinar structures, albeit to a much lesser extent than A549 (~5% of A427 cells presenting glandular differentiation versus 70–80% of A549, see supplementary material, Table S2). One of these structures is presented here. A549 tested strongly positive for both E-cadherin and CK7, whereas NCI-H522 tested negative for these markers of epithelial and adenocarcinoma tissues, respectively. A427 tested negative for CK7, and E-cadherin was also negative except in the acinar patches. None of the cell lines tested positive for the stem cell marker CD34. Scale bar, 50 µm.

58 years of age, with a smoking history of 60 packs/year [11]. These clinical features could fit perfectly in those observed in a typical SMARCA4-dUT cohort: predominantly male (73%), a mean age of 58 years, and a smoking history of 51 packs/year [12]. Regarding the SWI/SNF expression and mutational status in the NCI-H522 cell line, we discovered a hemizygous mutation in exon 5 of *SMARCA4* due to a deletion of two C nucleotides (c.805_806delCC), which creates a frameshift alteration (p. Pro270fs) and a truncated protein [4] that is responsible for its SMARCA4 null expression [13]. Later, we found that NCI-H522 had very low expression levels of *SMARCA2* [7], the other alternative SWI/SNF complex catalytic subunit. This silencing of *SMARCA2* is not due to a genomic mutation, as in the case of *SMARCA4*, because it harbors no mutations in its coding sequence [7].

According to our results, NCI-H522 (SMARCA4 negative, SMARCA2 negative, SOX2 high) has the paradigmatic SMARCA-dUT signature defined by Perret *et al* [9]. Interestingly, a possible explanation for the gain of such SOX2 expression in the NCI-H522 cell line may lie in the specific genetic amplification of *SOX4*, which we reported previously [14], as it is known that *SOX4* can enhance SOX2 expression [15]. In addition to the presence of the immunohistochemical markers of SMARCA4-dUT, NCI-H522 generated an undifferentiated morphology when engrafted *in vivo*, instead of the glandular differentiation features observed in LUAD [2,5]. Thus, our genetic expression profiling and biomarker and histology analyses postulate the NCI-H522 as a paradigmatic SMARCA4-dUT cell line, and we do not recommend using it as an archetypical LUAD cell line anymore. Considering all this, we propose the NCI-H522 as the first *bona fide* SMARCA4-dUT cell model for future studies about this novel and aggressive form of thoracic cancer.

Acknowledgements

PPM's laboratory is funded by the Spanish Ministry of Science and Innovation (PID2021-126111OB-I00), Junta de Andalucía (PIGE-0213-2020, PI-0203-2022), University of Granada (B-CTS-480-UGR20), and the Fundación Científica Asociación Española Contra el Cáncer (LAB-AECC-2018). The authors would also like to acknowledge the financial support of Mrs. Rosario Martín to promote lung cancer studies. AMA is supported by an FPU17/01258 fellowship. JMR-J is supported by a contract funded by Programa Operativo de Empleo Juvenil y de la Iniciativa de Empleo Juvenil (#04/2022-05). JS-H is supported by a grant from the Scientific Foundation of the Spanish Association Against Cancer in Granada (#PRDGR21428SANJ).

Author contributions statement

AMA contributed to the conceptualization, investigation, validation, and writing of the original draft. JMR-J contributed to data analysis, data curation, and writing (review and editing). JLL-H contributed to histology analysis, and writing (review and editing). JS-H contributed to investigation, validation, and writing (review and editing). PPM contributed to conceptualization, research, project administration, funding acquisition, and writing of the original draft.

Data availability statement

The transcriptomic data and cell line information used in the analysis are available at the DepMap database (https://depmap.org/portal/download/), DepMap Public 22Q2 Dataset, 'sample_info.csv' and 'CCLE_expression.csv' files.

References

1. Bosman FT. Integrative molecular tumor classification: a pathologist's view. In *Encyclopedia of Cancer* (3rd edn). Academic Press: Cambridge, MA, 2019; 279–285.
2. Nicholson AG, Tsao MS, Beasley MB, *et al*. The 2021 WHO classification of lung tumors: impact of advances since 2015. *J Thorac Oncol* 2022; **17:** 362–387.
3. Medina PP, Sanchez-Cespedes M. Involvement of the chromatin-remodeling factor BRG1/SMARCA4 in human cancer. *Epigenetics* 2008; **3:** 64–68.
4. Medina PP, Romero OA, Kohno T, *et al*. Frequent BRG1/SMARCA4-inactivating mutations in human lung cancer cell lines. *Hum Mutat* 2008; **29:** 617–622.
5. Chatzopoulos K, Boland JM. Update on genetically defined lung neoplasms: NUT carcinoma and thoracic SMARCA4-deficient undifferentiated tumors. *Virchows Arch* 2021; **478:** 21–30.
6. Yoshida A, Kobayashi E, Kubo T, *et al*. Clinicopathological and molecular characterization of SMARCA4-deficient thoracic sarcomas with comparison to potentially related entities. *Mod Pathol* 2017; **30:** 797–809.
7. Peinado P, Andrades A, Cuadros M, *et al*. Comprehensive analysis of SWI/SNF inactivation in lung adenocarcinoma cell models. *Cancers (Basel)* 2020; **10:** 3712.
8. Peinado P, Andrades A, Cuadros M, *et al*. Multi-omic alterations of the SWI/SNF complex define a clinical subgroup in lung adenocarcinoma. *Clin Epigenetics* 2022; **14:** 42.
9. Perret R, Chalabreysse L, Watson S, *et al*. SMARCA4-deficient thoracic sarcomas: clinicopathologic study of 30 cases with an emphasis on their nosology and differential diagnoses. *Am J Surg Pathol* 2019; **43:** 455–465.
10. Le Loarer F, Watson S, Pierron G, *et al*. SMARCA4 inactivation defines a group of undifferentiated thoracic malignancies transcriptionally related to BAF-deficient sarcomas. *Nat Genet* 2015; **47:** 1200–1205.
11. Banks-Schlegel SP, Gazdar AF, Harris CC. Intermediate filament and cross-linked envelope expression in human lung tumor cell lines. *Cancer Res* 1985; **45:** 1187–1197.
12. Rekhtman N, Montecalvo J, Chang JC, *et al*. SMARCA4-deficient thoracic sarcomatoid tumors represent primarily smoking-related undifferentiated carcinomas rather than primary thoracic sarcomas. *J Thorac Oncol* 2020; **15:** 231–247.
13. Medina PP, Carretero J, Ballestar E, *et al*. Transcriptional targets of the chromatin-remodelling factor SMARCA4/BRG1 in lung cancer cells. *Hum Mol Genet* 2005; **14:** 973–982.
14. Medina PP, Castillo SD, Blanco S, *et al*. The SRY-HMG box gene, SOX4, is a target of gene amplification at chromosome 6p in lung cancer. *Hum Mol Genet* 2009; **18:** 1343–1352.
15. Ikushima H, Todo T, Ino Y, *et al*. Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through Sry-related HMG-box factors. *Cell Stem Cell* 2009; **5:** 504–514.

SUPPLEMENTARY MATERIAL ONLINE

Figures S1–S5. Western blotting raw images, including digitalization image from membrane and chemiluminescence signal from specific bands. Band images were obtained using an ImageQuant™ LAS-4000 imaging system and the reagents and antibodies detailed in the Materials and methods section.

Table S1. Oligonucleotide sequences used for qPCR

Table S2. Summary of histopathological analysis performed in tumor xenografts