Received: 17 January 2023 | Revised: 1 March 2023 | Accepted: 3 March 2023

DOI: 10.1002/cncr.34797

ORIGINAL ARTICLE

Utility of *CYP2D6* copy number variants as prognostic biomarker in localized anal squamous cell carcinoma

Lucía Trilla-Fuertes PhD[1] | Angelo Gámez-Pozo PhD[1,2] | Miguel Nogué PhD, MD[3,4] | Isabel Busquier PhD, MD[3,5] | Fernando Arias PhD, MD[3,6] | Fernando López-Campos PhD, MD[3,7] | Ana Fernández-Montes PhD, MD[3,8] | Ana Ruiz PhD, MD[3,9] | Concepción Velázquez PhD, MD[3,10] | Celia Martín-Bravo MD[11] | Elisabeth Pérez-Ruiz PhD, MD[12] | Elena Asensio PhD, MD[3,13] | Xavier Hernández-Yagüe PhD, MD[3,14] | Aline Rodrigues PhD, MD[3,15] | Ismael Ghanem PhD, MD[3,16] | Rocío López-Vacas Tech[1] | Ahmed Hafez PhD[17] | Pedro Arias Tech[18] | Irene Dapía PhD[18] | Mario Solís BSc[19] | Antje Dittmann PhD[20] | Ricardo Ramos PhD[21] | Carlos Llorens PhD[17] | Joan Maurel PhD, MD[3,22,23] | Ángel Campos-Barros PhD[24] | Juan Ángel Fresno Vara PhD[1,2,25] | Jaime Feliu PhD[3,16,25,26]

Correspondence
Jaime Feliu, Hospital Universitario La Paz, Paseo de la Castellana 261, 28046, Madrid, Spain.
Email: jaime.feliu@uam.es

Funding information
IdiPAZ, Grant/Award Number: Jesús Antolín Garciarena Fellowship; European Proteomics Infrastructure Consortium, Grant/Award Number: 823839, Horizon 2020 Programme

Abstract

Background: Anal squamous cell carcinoma (ASCC) is an infrequent tumor whose treatment has not changed since the 1970s. The aim of this study is the identification of biomarkers allowing personalized treatments and improvement of therapeutic outcomes.

Methods: Forty-six paraffin tumor samples from ASCC patients were analyzed by whole-exome sequencing. Copy number variants (CNVs) were identified and their relation to disease-free survival (DFS) was studied and validated in an independent retrospective cohort of 101 ASCC patients from the Multidisciplinary Spanish Digestive Cancer Group (GEMCAD). GEMCAD cohort proteomics allowed assessing the biological features of these tumors.

Results: On the discovery cohort, the median age was 61 years old, 50% were males, stages I/II/III: 3 (7%)/16 (35%)/27 (58%), respectively, median DFS was 33 months, and overall survival was 45 months. Twenty-nine genes whose duplication was related to DFS were identified. The most representative was duplications of the CYP2D locus, including *CYP2D6*, *CYP2D7P*, and *CYP2D8P* genes. Patients with *CYP2D6* CNV had worse DFS at 5 years than those with two *CYP2D6* copies

(21% vs. 84%; $p < .0002$, hazard ratio [HR], 5.8; 95% confidence interval [CI], 2.7–24.9). In the GEMCAD validation cohort, patients with *CYP2D6* CNV also had worse DFS at 5 years (56% vs. 87%; $p = .02$, HR = 3.6; 95% CI, 1.1–5.7). Mitochondria and mitochondrial cell-cycle proteins were overexpressed in patients with *CYP2D6* CNV. **Conclusions:** Tumor *CYP2D6* CNV identified patients with a significantly worse DFS at 5 years among localized ASCC patients treated with 5-fluorouracil, mitomycin C, and radiotherapy. Proteomics pointed out mitochondria and mitochondrial cell-cycle genes as possible therapeutic targets for these high-risk patients.

Plain Language Summary

- Anal squamous cell carcinoma is an infrequent tumor whose treatment has not been changed since the 1970s.
- However, disease-free survival in late staged tumors is between 40% and 70%.
- The presence of an alteration in the number of copies of *CYP2D6* gene is a biomarker of worse disease-free survival.
- The analysis of the proteins in these high-risk patients pointed out mitochondria and mitochondrial cell-cycle genes as possible therapeutic targets.
- Therefore, the determination of the number of copies of *CYP2D6* allows the identification of anal squamous carcinoma patients with a high-risk of relapse that could be redirected to a clinical trial.
- Additionally, this study may be useful to suggest new treatment strategies to increase current therapy efficacy.

KEYWORDS
anal squamous cell carcinoma, cell cycle, copy number variants, *CYP2D6*, disease-free survival, mitochondrial metabolism, proteomics

INTRODUCTION

Anal squamous cell carcinoma (ASCC) is an infrequent tumor; it has been estimated that 9440 new cases will be diagnosed in the United States in 2022, representing approximately 2.8% of all gastrointestinal cancers.[1] The standard treatment is a combination of 5-fluorouracil (5-FU) with mitomycin C (MMC) or cisplatin, concomitantly with radiotherapy since the 1970s.[2,3] This treatment has demonstrated its efficacy in early-stage tumors. However, in T3-T4 or N1 tumors, disease-free survival (DFS) ranges are between 40% and 70%.[4,5] Therefore, a molecular characterization and the determination of potential therapeutic targets are still needed to improve the management of these patients.

With the bloom of high-throughput sequencing techniques, it is possible to study the multiple genetic alterations in clinical samples. Whole-exome sequencing (WES) is now incorporating into clinical practice,[6,7] and several works have associated genetic alterations with prognosis in cancer.[8–10] Mutations in *PIK3CA* have been associated with overall survival (OS) in recurrent ASCC patients after an abdominoperineal resection.[11] In a previous study, we identified genetic variants in five genes that were associated with DFS in ASCC.[12]

Our group has recently established the first molecular subtypes in ASCC,[13] using proteomics and genomics data. Two molecular subtypes were defined, one related to adhesion, T lymphocytes, mitochondria, and metabolism, among other biological processes, and the other with a higher frequency of mutations and increased expression of translation and ribosome proteins. However, this molecular classification did not have a prognostic implication or clinical application at the time.

A recently published study highlighted that point mutations in cancer driver genes contain little information about cancer patient prognosis, whereas copy number variants (CNVs) in these same driver genes had significant prognostic power.[9] Therefore, the aim of this study was the characterization of CNVs related to prognosis, and more specifically to DFS, in ASCC. In addition, high-throughput proteomics was used to molecularly characterize the impact of selected CNVs.

MATERIALS AND METHODS

Patient characteristics of discovery and validation cohorts

Formalin-fixed paraffin-embedded (FFPE) samples from patients diagnosed with localized ASCC were analyzed in this study.

For the discovery cohort, tumor samples were reviewed by an experienced pathologist, and all the samples contained at least 70% of tumor cells. This study was approved by the ethical committee of Hospital Universitario La Paz (PI-1926) and written informed consent was obtained for all the participants in the study. Twenty-seven patients from the VITAL clinical trial study (GEMCAD-09-02, NCT01285778), were treated with panitumumab, 5-FU, MMC, and radiotherapy.[14] The other 18 patients were included from the routine clinical practice at Hospital Universitario La Paz and Hospital Clinic and were treated with cisplatin–5-FU or MMC–5-FU and concomitant radiotherapy. The inclusion criteria were to have a histologically confirmed ASCC, to be 18 years old or older, to have an Eastern Cooperative Oncology Group performance status score (ECOG-PS) from 0 to 2, to have not received prior radiotherapy or chemotherapy for this malignancy, follow-up during >3 years or till relapse, and no presence of distant metastasis at diagnosis.

An independent validation cohort was recruited with the collaboration of the Spanish Multidisciplinary Group of Digestive Cancer (GEMCAD). Written consent was obtained for all the participants in the study, and the study was approved by the Hospital La Paz Ethical Committee (PI-3821). Inclusion criteria were presence of a tumor with a locally advanced-stage (TNM T2-4, NxM0), receiving a first line of treatment consisting of MMC–5-FU and radiotherapy concomitantly, follow-up longer than 3 years or till relapse, and FFPE sample available. Exclusion criteria were histological diagnosis different than ASCC, tumor <2 cm without adenopathies (T1N0M0), metastasis at diagnosis, receiving a first-line treatment different than MMC–5-FU and radiotherapy, or insufficient follow-up.

A third cohort consisting of seven patients with blood and primary tumor samples available was recruited to study germline and somatic CNVs. The inclusion criteria for this cohort were a diagnosis of ASCC and blood and primary tumor FFPE samples available.

Human papilloma virus characterization in the discovery cohort

Human papilloma virus (HPV) infection was determined by CLART HPV2 (Genomica), following manufacturers' instructions.

DNA isolation from FFPE samples

A 1 × 10 mm slide from each FFPE sample was deparaffinized, and DNA was isolated with GeneRead DNA FFPE Kit (Qiagen), following manufacturer's instructions. Once eluted, DNA was stored at −80°C until use.

Library preparation, exome capture, and Illumina sequencing

For the whole-exome sequencing (WES) experiments, purified DNA was quantified by Picogreen, and mean size was measured by gel electrophoresis. Genomic DNA was fragmented by mechanical means (Bioruptor) to an average size of approximately 200 bp. DNA tests were repaired, phosphorylated, A-tailed, and ligated to specific adaptors. Then, PCR-mediated labeling with Illumina-specific sequences and sample specific barcodes (Kapa DNA library generation kit) was performed.

Exome capture was done using the VCRome system (capture size 37 Mb, Roche Nimblegen) under a multiplexing of eight samples per capture per reaction. After the capture, libraries were cleaned, quantified and titrated by real-time PCR before sequencing. Sequencing was performed with coverage of 4.5 Gb per sample using an Illumina NextSeq NS500 (Illumina Inc). WES raw data files are available in Sequence Read Archive (SRA, https://www.ncbi.nlm.nih.gov/sra) under the name PRJNA 573670.

Bioinformatics processing of WES data

Preprocessing and mapping analyses

The workflow for characterization of SNPs and Indels consisted of the following steps: (1) FASTQC (http://www.bioinformmatics.babraham.ac.uk/projects/fastqc) was used for quality analysis; (2) Cutadapt[15] and PrinSeq[16] were used for adapter removal and reads preprocessing; and (3) BWA[17] in Samtools[18] and Picard Tools (http://picard.sourceforge.net) were used to map the preprocessed fastq files against the human genome h19 release and create bam files.

For the calling of CNVs, (1) the "Preprocess Intervals" command of GATK was used to create an interval list from the hg19 reference sequence, and (2) the above referred fastq files obtained from the SRA archive and the GATK commands "CollectReadCounts" and "CreateReadCountPanelOfNormals" were used to obtain a PON for CNVs. Next, the GATK commands "collectReadCounts," "denoiseReadCounts," "collectAllelicCounts," "modelSegments," and "callCopyRatioSegments" were used to process the bam files and generate the calling of CNVs.

For the annotation of effects, the variant effect predictor (VEP) of ENSEMBL was used to annotate the effects of the CNVs.[19]

Pipelines

The VariantSeq pipeline manager of the GPRO Suite[20] was used to execute all steps with the exception those of the CNV calling that were executed by the command line interface.

Identification of CNVs related to DFS

Duplications associated with DFS were determined using BRB Array Tools[21] by a Kaplan–Meier test. A p value under .001 was considered as statistically significant to correct the multi testing.

TaqMan quantitative polymerase chain reaction copy number assays

The reaction mixtures of 10 μL contained a total of 5 ng genomic DNA as a polymerase chain reaction (PCR) template, 1× TaqMan Gene Expression PCR Master Mix (5 μL), 1× TaqMan Copy Number Assay (Hs00010001_cn), and 1× TaqMan Copy Number Reference Assay (RNase P). Real-time PCR plates were run in sextuplicate in an QuantStudio 12K Flex Real-Time PCR System (Applied Biosystems, Thermo Fisher) under standard running conditions (95° for 10 min and 40 two-step cycles consisting of 95°C for 15 s and 60°C for 1 min) and analyzed with Copy Caller Software 2.1 (Applied Biosystems). A predicted number of copies <1.91 was considered as deletion, an interval between [1.92–2.40] was considered as two copies, and a predicted number of copies >2.41 was considered as duplication.

Analysis of germline and somatic CYP2D6 CNVs

With the aim of determining with certainty if the *CYP2D6* alterations had a germinal or a somatic origin, a study of paired samples of ASCC paraffin tumors and blood samples was performed.

DNA from blood samples was isolated using Gentra Puregene Blood core kit B (Qiagen, Germany) according to manufacturer's instructions. For this, 500 μL of whole blood were used. DNA concentration and purity were determined with the NanoDrop ND-1000 spectrophotometer (NanoDrop, USA). DNA was stored at −20°C after quantification.

Protein isolation

The GEMCAD ASCC validation cohort was also analyzed by mass-spectrometry proteomics. To isolate proteins, FFPE sections were deparaffinized in xylene and washed twice in absolute ethanol. Protein isolates were prepared in 2% sodium dodecyl sulfate (SDS) using a protocol based on heat-induced antigen retrieval. Protein quantification was done using MicroBCA Protein Assay Kit (Pierce-Thermo Scientific) and protein isolates were digested with trypsin (1:50). SDS was removed using Detergent Removal Spin Columns (Pierce). Finally, peptides were cleaned up using C18 stage-tips, re-solubilized in MS sample buffer, and spiked with indexed retention time peptides.

Liquid chromatography with tandem mass spectrometry proteomics experiments

For the data-independent acquisition (DIA) step, each sample was run individually. Mass spectrometry analysis was performed on an Orbitrap Fusion (Thermo Scientific) equipped with a Digital PicoView source (New Objective) and coupled to an M-Class UPLC (Waters), operated in trapping mode. Peptides were loaded onto a commercial MZ Symmetry C18 Trap Column (5 μm, 180 μm × 20 mm, Waters) followed by nanoEase MZ C18 HSS T3 Column (1.8 μm, 75 μm × 250 mm, Waters). The peptides were eluted at a flow rate of 300 nL/min with a gradient from 5% to 22% in 109 min. The DIA runs were acquired in Orbitrap-Orbitrap mode with isolation windows of 24 mass-to-charge ratio (m/z) covering a range from 385 to 1015 m/z (DIA) with 1 Da overlap. All relevant data were deposited in the ProteomeXchange Consortium via the PRIDE (http://www.ebi.ac.uk/pride) partner repository with the data set identifier PXD037816 and can be accessed through the reviewer account (username: reviewer_pxd037816@ebi.ac.uk; password: lQUp0ASW).

Spectronaut 15.4.210913 was used for directDIA analysis with default parameter settings. Spectra were searched against a canonical SwissProt database for human and common protein contaminants (NCBI taxonomy ID9606, release date July 9, 2019). Protein quantification was performed in Spectronaut using default settings and quantitative data were extracted using the BGS Factory Report (default) and used for follow-up analyses. To perform statistical modeling, fragment intensities were aggregated into precursor and peptide intensities. Log2 transformation and filtering according 75% of valid values were done using Perseus software.[22]

Systems biology analyses

Protein data obtained from mass-spectrometry experiments were analyzed by a systems biology approach based on probabilistic graphical models (PGMs) as previously reported.[23] Briefly, PGMs order the protein data in a network structure according to two criteria: first, calculating the spanning tree with the maximum likelihood, and then, performing a forward search adding edges that reduce Bayesian information criterion (BIC) but preserving the decomposability of the graph.[24] The obtained network has functional structure (i.e., proteins are grouped according biological functions). The overrepresented functions of each node of the network was established by gene ontology analyses using DAVID bioinformatics resources,[25] using "Homo sapiens" as background and GOTERM-FAT, KEGG, and Biocarta as categories. With the aim of comparing groups, functional node activities were calculated as the mean expression of these proteins related to the overrepresented function of each node.

Statistical analyses

Statistical analyses were done using GraphPad Prism v6. Survival analyses were performed using a Kaplan–Meier and long-rank test in BRB Array Tools.[21] DFS was defined as the time since the first dose of treatment to the first treatment failure or 5 years. OS was defined as the time since the first dose of treatment to death or 5 years. Hierarchical cluster was done in MeV software using correlation as associative method and average linkage as linkage method.[26] All *p* values were bilateral and considered as significant under 0.05, with the exception of the Kaplan–Meier test done in BRB Array Tool, where a *p* value <.001 was considered for multi-testing correction. Comparisons between groups were done using nonparametric tests such as Mann-Whitney test and Kruskal–Wallis test.

Resource availability

Data are available in a public, open access repository. Copy number variant data are available in SRA (https://www.ncbi.nlm.nih.gov/sra) under the name PRJNA573670. Proteomics data are available in ProteomeXchange Consortium via the PRIDE (http://www.ebi.ac.uk/pride) partner repository with the data set identifier PXD037816.

RESULTS

Discovery cohort

Forty-six patients diagnosed with localized ASCC were included in the discovery cohort. Twenty-seven patients were also included in the VITAL clinical trial (GEMCAD-09-02, NCT01285778). VITAL patients received panitumumab, 5-FU, and MMC, concomitantly with radiotherapy. Eighteen patients were recruited from the Hospital Universitario La Paz (Madrid, Spain) and the Hospital Clinic (Barcelona, Spain). Fourteen of 18 were treated with 5-FU and MMC or 5-FU and cisplatin, concomitantly with radiotherapy. Three patients were treated only with surgery, and one patient was treated only with radiotherapy. These four patients were excluded for the survival analyses. This cohort has been described in previous works.[12,13,27] A summary of the clinical characteristics of this cohort is shown in Table 1.

Whole-exome sequencing experiments

Discovery ASCC cohort was analyzed by WES. The mean coverage obtained in WES experiments was >42.6×, except for one sample with a coverage of 3.57× that was excluded for the subsequent analyses. The remaining samples displayed a mapping efficiency of 90%–98%, with the exception of a sample (75.4%). The human exome includes >195,000 exonic regions, of which 23,021 (11.21%) were not mapped in any sample.

TABLE 1 Clinical characteristics of the discovery and GEMCAD validation cohorts and *p* values comparing the distribution of each parameter in both cohorts.

	Discovery cohort (%)	Validation cohort (%)	*p*
No. of patients	46 (100)	88 (100)	
Age at diagnosis, years (median and range)	61 (41–86)	61 (32–84)	.78
Gender			.58
Male	23 (50)	39 (44)	
Female	23 (50)	48 (55)	
Unknown	0 (0)	1 (1)	
HPV			.021
Positive	34 (74)	15 (17)	
Negative	6 (13)	11 (13)	
Unknown	6 (13)	62 (70)	
HIV			.019
Positive	2 (4)	17 (19)	
Negative	44 (96)	71 (81)	
Lymph node status			1.00
N0	19 (42)	37 (42)	
N positive	24 (54)	49 (56)	
Unknown	3 (4)	2 (2)	
TNM stage AJCC 6th edition			.093
I	3 (7)	0 (0)	
II	16 (35)	36 (41)	
III	27 (58)	51 (58)	
IV	0 (0)	1 (1)	

Abbreviations: AJCC, American Joint Committee on Cancer; GEMCAD, Spanish Multidisciplinary Group of Digestive Cancer; HPV, human papillomavirus.

CNVs in CYP2D6 had prognostic value in ASCC patients

Twenty-nine genes in which having a duplication was significantly related with DFS in ASCC were determined by a Kaplan–Meier analysis (Table S1). Among them, the duplications located in the CYP2D locus, involving *CYP2D6*, *CYP2D7P*, and *CYP2D8P* genes were remarkable. When analyzing the length of the observed CYP2D duplications, the minimum genomic region always implicated *CYP2D6* (Sup Table 2). For this reason, *CYP2D6* CNV status was studied in depth in our cohort. Twenty-six (64%) patients presented two copies of *CYP2D6*, 11 (26%) had a duplication in *CYP2D6*, and four (10%) had a deletion.

Patients who presented a CNV (duplication or deletion) in *CYP2D6* had a worse DFS at 5 years than those patients with two *CYP2D6* copies ($p = .0035$; HR, 4.86; 95% CI, 1.78–18.44). The area under the receiver operating characteristic (ROC) curve was 0.84 with a p value of .0005 (Figure 1). A multivariate analysis showed that *CYP2D6* CNV status added prognostic information to stage, human papilloma virus infection, and gender (Table S3).

TABLE 2 *CYP2D6* copy number in paired ASCC blood and FFPE tumor samples from the same patient.

	CYP2D6 copy number in blood samples	*CYP2D6* copy number in tumor samples
Patient 1	4	19
Patient 2	3	3
Patient 3	2	3
Patient 4	3	6
Patient 5	2	5
Patient 6	3	5
Patient 7	2	3

Abbreviations: ASCC, anal squamous cell carcinoma; FFPE, formalin-fixed paraffin-embedded.

GEMCAD validation cohort

A total of 101 patients were recruited by GEMCAD. Of these 101 patients, 10 were excluded because they were not treated with chemoradiotherapy, two because they presented a T1 tumor, and another because tumor stage was unknown. Finally, 88 patients were included in the study. Clinical characteristics of the GEMCAD validation cohort are summarized in Table 1. There was a significant different distribution in HIV infection ($p = .019$) and in VPH infection ($p = .021$), although the latter may be due to the elevated number of missing data.

Validation of prognostic value of the *CYP2D6* CNVs

We studied the *CYP2D6* CNV status in 88 tumor samples using quantitative polymerase chain reaction (qPCR) CNVs TaqMan assays (Thermo Scientific). Twenty-four (27.3%) patients presented two copies of *CYP2D6*, 41 (46.6%) patients presented a *CYP2D6* duplication, and 20 (22.7%) patients had a deletion in *CYP2D6*. In three samples (3.4%), we did not obtain a qPCR result (Tables S4, S5, and S6).

The CYP2D6 CNV status showed prognostic value in the validation cohort. Patients with a *CYP2D6* duplication/deletion showed worse DFS at 5 years than those patients with two *CYP2D6* copies ($p = 0.021$; HR, 3.65; 95% CI, 1.15–5.78). The area under the ROC curve was 0.90 with a p value of <.0001 (Figure 2).

Analysis of germline and somatic *CYP2D6* CVNs

CYP2D6 CNVs were measured by qPCR in seven paired samples from blood and FFPE tumor from the same patient. In all the cases, with

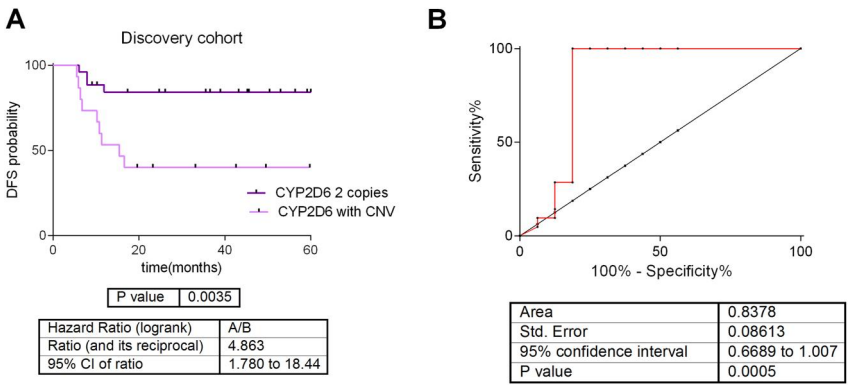

FIGURE 1 (A) DFS at 5 years according to *CYP2D6* CNV status in the discovery cohort. (B) ROC curve for *CYP2D6* prediction. CNV indicates copy number variant; DFS, disease-free survival; ROC, receiver operating characteristic.

the exception of one patient, tumor samples had higher number of copies than blood samples, suggesting that these *CYP2D6* CNVs generally occurs in the tumor (Table 2).

Proteomics characterization of the validation cohort

Eighty patients from the GEMCAD validation cohort have also been analyzed by DIA mass-spectrometry proteomics. A total of 3078 proteins were quantified and identified in these samples, and after applying a quality criterion of 75% of valid values, 2078 proteins were used for the analyses.

Relationship between *CYP2D6* CNVs and ASCC proteomics subtypes

In the GEMCAD validation cohort, each tumor was assigned to one of the two previously defined proteomics subtypes, as shown previously[13] (Figure S1). A network using all 2078 identified proteins was built using PGMs. The resulting network was formed by eight functional nodes with an overrepresented biological function (Figure S2A). Patients were classified into two groups with the same molecular characteristics as the original proteomics molecular subtypes, mainly characterized by a differential mitochondrial metabolism node activity (Figure S2B).

We studied the prognostic value of *CYP2D6* alterations in each proteomics subtype. Tumors analyzed by proteomics from 72 patients fulfilled clinical inclusion criteria for the *CYP2D6* study. We found that 31 (66%) V1 tumors and 20 (77%) V2 tumors presented a CNV in *CYP2D6*, whereas 15 (34%) V1 and six (23%) V2 tumors presented two *CYP2D6* copies. The presence of *CYP2D6* CNVs was related with a poor prognosis in both subtypes (Figure 3).

We found significant differences in the mitochondria functional node activity between V1 tumors with a CNV in *CYP2D6*, and therefore with a high-risk of relapse, and V1 tumors without CNVs in *CYP2D6* (Figure 4A, sup Figure 3). No differences in proteomics data according to the *CYP2D6* CNV status were found in V2 tumors.

This mitochondria node contained several genes related to mitochondrial metabolism but also genes located in the mitochondria

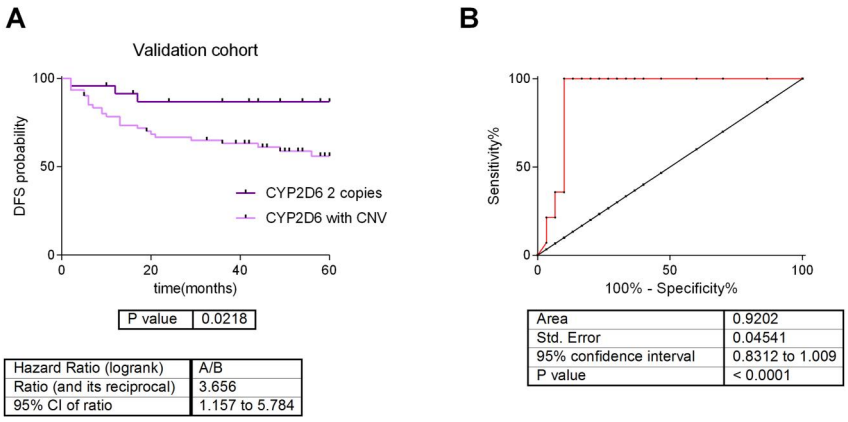

FIGURE 2 (A) DFS of GEMCAD validation cohort according *CYP2D6* status. (B) ROC curve for *CYP2D6* prediction. DFS indicates disease-free survival; GEMCAD, Spanish Multidisciplinary Group of Digestive Cancer; ROC, receiver operating characteristic.

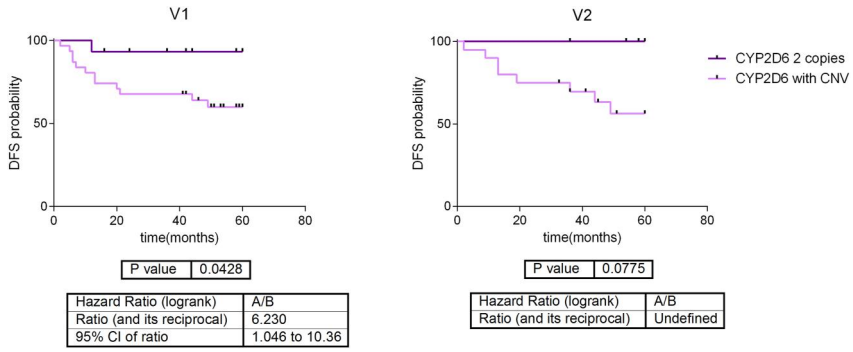

FIGURE 3 DFS at 5 years in GEMCAD validation cohort separated by ASCC proteomics molecular subtypes. ASCC indicates anal squamous cell carcinoma; DFS, disease-free survival; GEMCAD, Spanish Multidisciplinary Group of Digestive Cancer.

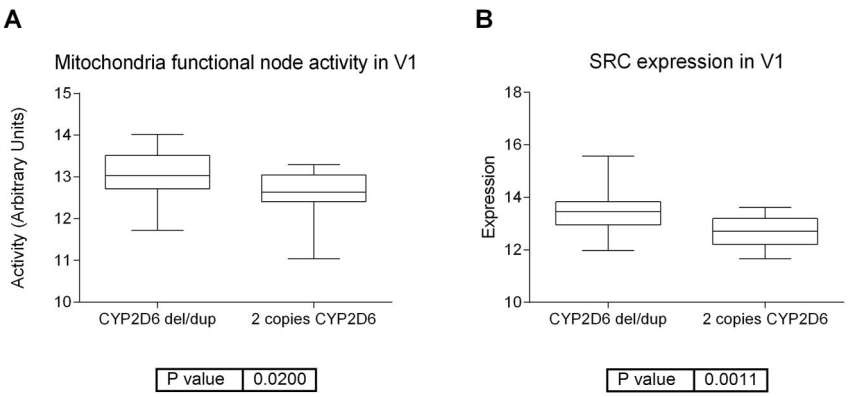

FIGURE 4 (A) Functional node activity of mitochondria node in V1 molecular group defined in the ASCC validation cohort according to *CYP2D6* status. (B) Protein expression of SRC in V1 molecular group according *CYP2D6* status. ASCC indicates anal squamous cell carcinoma; Del, deletion; dup, duplication.

and involved in cell cycle, as SRC, CDK1, or HSPA9. SRC had a differential expression according to the *CYP2D6* CNV status in V1, being higher in tumors with *CYP2D6* deletions and duplications ($p = .0011$) (Figure 4B).

DISCUSSION

ASCC is a rare tumor whose treatment has not changed since the 1970s, consisting of classical chemotherapy combined with radiotherapy. Despite this treatment, approximately 40%–70% of patients with ASCC tumors in advanced stages are predicted to suffer a relapse.[4,5] For these reasons, it is essential to molecularly characterize this disease to establish biomarkers and possible targeted therapies.

Studies about prognostic biomarkers in ASCC are scarce. T stage, N status, and gender have been associated with prognosis in ASCC.[28] HPV16 infection and several clinical factors as neutrophilia or anemia have also been reported as related to OS and/or DFS.[14,29,30] Regarding gene-based biomarkers, p53-positive tumors have a shorter DFS.[31,32] Moreover, p21, nuclear factor–κB, Ki67, and cyclin A levels have been associated with prognosis in ASCC.[33–35] Cacheux et al.[11] defined that mutations in *PIK3CA* were associated with worst OS in recurrent patients after an abdominoperineal resection. A recent study established that high levels of *Fusobacterium nucleatum* were associated with better OS and DFS in patients who suffered a progression and underwent abdominoperineal resection.[36] In a previous work, we identified that presenting a high or moderate impact of genetic variants in *BRCA2*, *ZNF750*, *FAM208B*, *ZNF599*, and *ZC3H13* genes was related to DFS in localized ASCC.[12] This is the first study in localized ASCC that associated CNVs with DFS.

The CYP2D locus contains three genes: *CYP2D6* and two pseudogenes, *CYP2D7P* and *CYP2D8P*. The three genes are formed by nine exons and share a high sequence similarity. Deletions and duplications of *CYP2D6* related to pharmacogenomics have been widely described. The described duplications included functional and nonfunctional alleles of *CYP2D6*.[37] *CYP2D6* is a gene of the cytochrome P450 family involved in phase 1 drug metabolism, including tamoxifen. *CYP2D6* variants can be classified into categories based on its metabolizer activity: decreased, normal, increased, or altered.[38] It has been described that the frequency in the general population of *CYP2D6* duplications is approximately 5%.[39] However, in our ASCC discovery cohort, *CYP2D6* duplications were present in 33% of the samples and in the GEMCAD validation cohort in 46.6% of patient's tumors. *CYP2D6* alterations have been previously associated with vulvar and anal cancer incidence, being hypothesized that incidence increases with an increasing number of *CYP2D6* inactivating alleles,[40] but they have been not associated with evolution of the disease until now. Nevertheless, *CYP2D6* alterations have been associated with several clinical factors in other neoplasia. CYP2D6*4 nonfunctional allele has been associated with a protective effect in breast cancer.[41] *CYP2D6* inactivating alleles were also related to a poor response to chemotherapy in head and neck squamous carcinoma,[42] but not with OS.[43] Moreover, a meta-analysis suggested that *CYP2D6* polymorphisms are related to the pathogenesis of various cancers.[44] Therefore, this is the first study associating *CYP2D6* CNVs with prognosis in cancer, specifically in ASCC.

Paired samples from blood and tumor from ASCC patients were used to confirm that *CYP2D6* CNVs are somatic alterations. Despite the fact that some of the patients presented germinal CNVs in *CYP2D6*, in all the cases the tumor sample presented a higher number of *CYP2D6* copies than the blood sample. This fact supports the relevance of *CYP2D6* somatic CNVs in ASCC. Additionally, the frequency of germline *CYP2D6* duplications determined in blood was higher than the frequency previously described in the general population,[39] suggesting a role of *CYP2D6* duplications in ASCC incidence as it has been previously suggested by other studies.[40]

A proteomics characterization of the GEMCAD validation cohort was also done with the aim of proposing new therapeutic targets in patients with high-risk of relapse determined by the *CYP2D6 CNV*

status. In a previous study, our group defined the first molecular classification in ASCC, consisting of two different proteomics subtypes.[13]

Additionally, V1 presented differences in mitochondria functional node activity according to the *CYP2D6 CNV* status. *CYP2D6* belongs to the CYP450 family, involved in the electron transfer chain.[45] This node also contained genes located in the mitochondria and related to the cell cycle. SRC presented a higher expression in V1 tumors with a CNV in *CYP2D6*. SRC is a tyrosine kinase of the PIK3 pathway involved in cellular proliferation. Src inhibitors have been tested in clinical trials for many solid tumors, including head and neck squamous cell carcinoma and colorectal cancer, with negative results.[46] A role of Src in the development of resistance to anti-EGFR therapies has been identified,[47,48] and preclinical models in colorectal cancer suggest that Src blockade can restore the sensitivity to EGFR inhibitors.[49] The higher activity of mitochondria metabolism also suggested the use of OXPHOS inhibitors, previously suggested as a promising therapy in cancer, as a possible therapy for patients with a CNV in CYP2D6. A relevant role of mitochondria in HPV-dependent cancers has been previously established.[50]

The study has some limitations. The main strength of the study is that *CYP2D6* prognostic value has been validated in an independent, multi-centric retrospective cohort, which supports the extrapolation of results, although a prospective validation of the prognostic value of *CYP2D6* CNVs would be interesting. Moreover, a wider study of paired blood and primary tumor samples would be helpful to exactly determine the role of germline and somatic *CYP2D6* CNVs and fully clarify the role of *CYP2D6* deletions. In addition, the limited number of samples did not allow determining in depth the relation with tumor stage. A larger cohort would be necessary for this purpose.

The multivariate analysis confirmed that *CYP2D6* had prognostic value independently of tumor size and nodal status. On the other hand, when analyzed independently, the low number of stage II patient make it difficult to extract conclusions, especially in the discovery cohort (where there are only two patients with stage II and *CYP2D6* CNV). A trend can be observed in the validation cohort as long as DFS at 5 years is 84% versus 60% in the non CNV versus the CNV group.

Regarding other genes involved in predicting ASCC outcome, we identified other genes with prognostic value regarding the presence of a CNV (Table S1). Most of these genes were included in one of two genomic regions at CHR22 (the *CYP2D6* region) and at CHR2. Regarding the CRH22 region, the minimum genomic region always implicated *CYP2D6*. Thus, the other CYP genes had prognostic and/or predictive value as well, but we focused on *CYP2D6* to be tested by qPCR in the validation cohort. Genes in CHR2 showed lower prognostic value. Combination of different genes' prognostic value as a signature is a common practice, being possible when the prognostic value is based on a quantitative variable (gene/protein expression) and the fact that importance of each gene in the combined prognostic value can be weighted by a factor. In our case, we used a quantitative variable (CNV yes/no), thus weight the prognostic value of each gene is not possible with standard methodologies. As long as tumor size showed no prognostic value by itself, it is unlikely it showed value in the multivariate analysis, as we showed in Table S3.

In conclusion, we established that the number of copies of *CYP2D6* is related to DFS in ASCC patients. Therefore, *CYP2D6* CNV status allows the identification of high-risk ASCC patients who could be redirected to a clinical trial. In addition, this study may be useful to explore new treatment strategies for these high-risk patients to increase current therapy efficacy, such combinations including Src or OXPHOS inhibitors.

AFFILIATIONS

[1]Molecular Oncology Lab, Hospital Universitario La Paz-IdiPAZ, Madrid, Spain

[2]Biomedica Molecular Medicine SL, Madrid, Spain

[3]Spanish Multidisciplinary Group of Digestive Cancer, Barcelona, Spain

[4]Medical Oncology Service, Hospital de Granollers, Barcelona, Spain

[5]Medical Oncology Service, Hospital Provincial de Castellón, Castellón, Spain

[6]Radiotherapy Oncology Service, Complejo Hospitalario de Navarra, Navarra, Spain

[7]Radiotherapy Oncology Service, Hospital Ramón y Cajal, Madrid, Spain

[8]Medical Oncology Service, Complejo Hospitalario Universitario de Ourense, Ourense, Spain

[9]Medical Oncology Service, Hospital Universitario Puerta de Hierro Majadahonda, IDIPHISA, Madrid, Spain

[10]Medical Oncology Service, Hospital Miguel Servet, Zaragoza, Spain

[11]Medical Oncology Service, Hospital Costa del Sol, Málaga, Spain

[12]Unidad de Gestión Clínica Intercentros de Oncología Médica, Hospitales Universitarios Regional y Virgen de la Victoria, Instituto de Investigación Biomédica de Málaga, Hospitales Universitarios Regional y Virgen de la Victoria, Málaga, Spain

[13]Medical Oncology Service, Hospital General Universitario de Elche, Elche, Spain

[14]Medical Oncology Service, Instituto Catalán de Oncología-Girona, Girona, Spain

[15]Medical Oncology Service, Hospital Universitario de Salamanca, Salamanca, Spain

[16]Medical Oncology Service, Hospital Universitario La Paz, Madrid, Spain

[17]Biotechvana SL, Parque Científico de Madrid, Madrid, Spain

[18]Pharmacogenetics Lab, Institute of Medical and Molecular Genetics-INGEMM, La Paz University Hospital-IdiPAZ, Autonomous University of Madrid, Madrid, Spain

[19]Bioinformatics Unit, Institute of Medical and Molecular Genetics-INGEMM, La Paz University Hospital-IdiPAZ, Madrid, Spain

[20]Functional Genomics Center Zurich, University of Zurich/ETH Zurich, Zurich, Switzerland

[21]Genomics Unit, Parque Científico de Madrid, Madrid, Spain

[22]Medical Oncology Department, Hospital Clínic of Barcelona, Translational Genomics and Targeted Therapeutics in Solid Tumors Group, IDIBAPS, University of Barcelona, Barcelona, Spain

[23]Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas, Madrid, Spain

[24]Institute of Medical and Molecular Genetics, IdiPAZ, Hospital Universitario La Paz/CIBERER Unit 753, ISCIII, Madrid, Spain

[25]Biomedical Research Networking Center on Oncology-CIBERONC, ISCIII, Madrid, Spain

[26]Cátedra UAM-Amgen, Universidad Autónoma de Madrid, Madrid, Spain

AUTHOR CONTRIBUTIONS

Ismael Ghanem: Conceptualization, design, interpretation, clinical data curation, sample recollection, and statistical analysis. **Joan Maurel:** Conceptualization, design, interpretation, clinical data curation, sample recollection, and statistical analysis. **Juan Ángel Fresno Vara:** Conceptualization, design, interpretation, probabilistic graphical models, and statistical analysis. **Jaime Feliu:** Conceptualization, design, and interpretation, clinical data curation, sample recollection, and study coordination. **Rocío López-Vacas:** DNA and protein extraction. **Pedro Arias:** DNA and protein extraction. **Irene Dapía:** DNA and protein extraction. **Miguel Nogué:** Clinical data curation and sample recollection. **Isabel Busquier:** Clinical data curation and sample recollection. **Fernando Arias:** Clinical data curation and sample recollection. **Fernando López-Campos:** Clinical data curation and sample recollection. **Ana Fernández-Montes:** Clinical data curation and sample recollection. **Ana Ruiz:** Clinical data curation and sample recollection. **Concepción Velázquez:** Clinical data curation and sample recollection. **Celia Martín-Bravo:** Clinical data curation and sample recollection. **Elisabeth Pérez-Ruiz:** Clinical data curation and sample recollection. **Elena Asensio:** Clinical data curation and sample recollection. **Xavier Hernández-Yagüe:** Clinical data curation and sample recollection. **Aline Rodrigues:** Clinical data curation and sample recollection. **Lucía Trilla-Fuertes:** Probabilistic graphical models, statistical analysis, and writing–original draft. **Angelo Gámez-Pozo:** Probabilistic graphical models and statistical analysis. **Antje Dittmann:** Mass spectrometry data. **Ricardo Ramos:** Sequencing analyses. **Ahmed Hafez:** Sequencing analyses. **Mario Solís:** Sequencing analyses. **Ángel Campos-Barros:** Sequencing analyses. **Carlos Llorens:** Sequencing analyses. **PG-A:** Statistical analyses. **JC:** Statistical analyses. **CC:** Statistical analyses. All the authors have directly participated in the preparation of this manuscript, have read and approved the final version, and declare no ethical conflicts of interest.

ACKNOWLEDGMENTS

We would like to thank the pathology departments of the Clinica Girona, Palamòs Hospital, Calella Hospital, and Figueras Hospital for their help. This research was funded by EPIC-XS (project 823839), the Horizon 2020 Programme of the European Union, and a Jesús Antolín Garciarena Fellowship from IdiPAZ.

CONFLICT OF INTEREST STATEMENT

Ricardo Ramos-Ruiz reports independent contracting fees from Fundación Parque Científico de Madrid. Antje Dittmann reports grant funding from the Horizon 2020 Framework Programme. Jaime Feliu reports consulting fees from Amgen, Eisai, F. Hozmann-La Roche, Ipsen Biopharm Limited, Merck, Novartis, Organon & Co, and Sirtex Medical Inc. Angelo Gámez-Pozo reports stock with Biomedica Molecular Medicine SL and holding patent EP19382982.7 (In Vitro Method for the Prognosis of Anal Squamous Cell Carcinoma). Xavier Hernández-Yagüe reports consulting fees for Merck and Sanofi-Aventis US LLC. Fernando López-Campos reports consulting fees from Hospital Ramón y Cajal. Ana Ruiz reports independent contractor fees from the Abbott Fund, AstraZeneca, and Nestle HealthCare Nutrition Inc; consulting fees from Pierre Fabre Pharmaceuticals and Sanofi US Services; expert witness fees from Amgen and Servier Azaires Medicales; and travel fees from Merck. Aline Rodrigues reports independent contractor fees from Aalborg Universitetshospital. Lucía Trilla-Fuertes reports holding patent EP19382982 (In Vitro Method for the Prognosis of Anal Squamous Cell Carcinoma). Juan Ángel Fresno Vara stock with Biomedica Molecular Medicine SL. The other authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data are available in a public, open access repository. Copy number variant data are available in SRA (https://www.ncbi.nlm.nih.gov/sra) under the name PRJNA573670. Proteomics data are available in ProteomeXchange Consortium via the PRIDE (http://www.ebi.ac.uk/pride) partner repository with the data set identifier PXD037816 and can be accessed through the reviewer account with username reviewer_pxd037816@ebi.ac.uk and password lQUp0ASW.

ORCID

Lucía Trilla-Fuertes https://orcid.org/0000-0002-4452-3474
Fernando López-Campos https://orcid.org/0000-0002-4077-0507
Ana Fernández-Montes https://orcid.org/0000-0003-4630-5830
Elisabeth Pérez-Ruiz https://orcid.org/0000-0001-7204-0800
Ricardo Ramos https://orcid.org/0000-0002-6331-9786
Jaime Feliu https://orcid.org/0000-0003-4867-1420

REFERENCES

1. Siegel RL, Miller KD, Fuchs HE, Jemal A, Cancer statistics, 2022, *CA Cancer J Clin*. 2022;72(1):7-33. doi:10.3322/caac.21708
2. Benson AB, Venook AP, Al-Hawary MM, et al. Anal carcinoma, version 2.2018, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2018;16(7):852-871.
3. Glynne-Jones R, Nilsson P, Aschele C, et al. Anal cancer: ESMO-ESSO-ESTRO clinical practice guidelines for diagnosis, treatment and follow-up. *Eur J Surg Oncol*. 2014;40(10):1165-1176. doi:10.1016/j.ejso.2014.07.030
4. Gunderson LL, Winter KA, Ajani JA, et al. Long-term update of US GI intergroup RTOG 98-11 phase III trial for anal carcinoma: survival, relapse, and colostomy failure with concurrent chemoradiation involving fluorouracil/mitomycin versus fluorouracil/cisplatin. *J Clin Oncol*. 2012;30(35):4344-4351. doi:10.1200/jco.2012.43.8085
5. Ajani JA, Winter KA, Gunderson LL, et al. Prognostic factors derived from a prospective database dictate clinical biology of anal cancer: the intergroup trial (RTOG 98-11). *Cancer*. 2010;116(17):4007-4013. doi:10.1002/cncr.25188
6. Tetreault M, Bareke E, Nadaf J, Alirezaie N, Majewski J. Whole-exome sequencing as a diagnostic tool: current challenges and future opportunities. *Expert Rev Mol Diagn*. 2015;15(6):749-760. doi:10.1586/14737159.2015.1039516
7. Rabbani B, Tekin M, Mahdieh N. The promise of whole-exome sequencing in medical genetics. *J Hum Genet*. 2014;59(1):5-15. doi:10.1038/jhg.2013.114
8. Hieronymus H, Murali R, Tin A, et al. Tumor copy number alteration burden is a pan-cancer prognostic factor associated with recurrence and death. *Elife*. 2018;7. doi:10.7554/elife.37294

9. Smith JC, Sheltzer JM. Systematic identification of mutations and copy number alterations associated with cancer patient prognosis. *Elife*. 2018;7. doi:10.7554/elife.39217
10. Ono A, Isaka M, Serizawa M, et al. Genetic alterations of driver genes as independent prognostic factors for disease-free survival in patients with resected non-small cell lung cancer. *Lung Cancer*. 2019;128:152-157. doi:10.1016/j.lungcan.2018.12.005
11. Cacheux W, Rouleau E, Briaux A, et al. Mutational analysis of anal cancers demonstrates frequent PIK3CA mutations associated with poor outcome after salvage abdominoperineal resection. *Br J Cancer*. 2016;114(12):1387-1394. doi:10.1038/bjc.2016.144
12. Trilla-Fuertes L, Ghanem I, Maurel J, et al. Comprehensive characterization of the mutational landscape in localized anal squamous cell carcinoma. *Transl Oncol*. 2020;13(7):100778. doi:10.1016/j.tranon.2020.100778
13. Trilla-Fuertes L, Ghanem I, Gámez-Pozo A, et al. Genetic profile and functional proteomics of anal squamous cell carcinoma: proposal for a molecular classification. *Mol Cell Proteomics*. 2020;19(4):690-700. doi:10.1074/mcp.ra120.001954
14. Feliu J, Garcia-Carbonero R, Capdevila J, et al. VITAL phase 2 study: Upfront 5-fluorouracil, mitomycin-C, panitumumab and radiotherapy treatment in nonmetastatic squamous cell carcinomas of the anal canal (GEMCAD 09-02). *Cancer Med*. 2020;9(3):1008-1016. doi:10.1002/cam4.2722
15. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J*. 2011;17(1):10-12. doi:10.14806/ej.17.1.200
16. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27(6):863-864. doi:10.1093/bioinformatics/btr026
17. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760. doi:10.1093/bioinformatics/btp324
18. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
19. Cunningham, F, Amode MR, Barrell D, et al. Ensembl 2015. *Nucleic Acids Res*. 2015;43(Database issue):D662-D669.
20. Futami R, Muñoz-Pomer A, Viu JM, et al. GPRO The professional tool for annotation, management and functional analysis of omic databases. *Biotechvana Bioinforma*. 2011:SOFT3.
21. Simon R. Roadmap for developing and validating therapeutically relevant genomic classifiers. *J Clin Oncol*. 2005;23(29):7332-7341. doi:10.1200/jco.2005.02.8712
22. Tyanova S, Temu T, Sinitcyn P, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods*. 2016;13(9):731-740. doi:10.1038/nmeth.3901
23. Gámez-Pozo A Berges-Soria J, Arevalillo JM, et al. Combined label-free quantitative proteomics and microRNA expression analysis of breast cancer unravel molecular differences with clinical implications. *Cancer Res*, 2015;75(11):2243-2253.
24. Lauritzen S. *Graphical Models*. Oxford University Press; 1996.
25. Huang dW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57. doi:10.1038/nprot.2008.211
26. Saeed AI, Sharov V, White J, et al. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*. 2003;34(2):374-378. doi:10.2144/03342mt01
27. Trilla-Fuertes L, Gámez-Pozo A, Maurel J, et al. Description of the genetic variants identified in a cohort of patients diagnosed with localized anal squamous cell carcinoma and treated with panitumumab. *Sci Rep*. 2021;11(1):7402. doi:10.1038/s41598-021-86966-w
28. Theophanous S, Samuel R, Lilley J, et al. Prognostic factors for patients with anal cancer treated with conformal radiotherapy-a systematic review. *BMC Cancer*. 2022;22(1):607. doi:10.1186/s12885-022-09729-4
29. Schernberg A, Huguet F, Moureau-Zabotto L, et al. External validation of leukocytosis and neutrophilia as a prognostic marker in anal carcinoma treated with definitive chemoradiation. *Radiother Oncol*. 2017;124(1):110-117. doi:10.1016/j.radonc.2017.06.009
30. Zhu X, Jamshed S, Zou J, et al. Molecular and immunophenotypic characterization of anal squamous cell carcinoma reveals distinct clinicopathologic groups associated with HPV and TP53 mutation status. *Mod Pathol*. 2021;34(5):1017-1030. doi:10.1038/s41379-020-00729-y
31. Allal AS, Waelchli L, Bründler MA. Prognostic value of apoptosis-regulating protein expression in anal squamous cell carcinoma. *Clin Cancer Res*. 2003;9:6489-6496.
32. Wong CS, Tsao MS, Sharma V, Chapman WB, Pintilie M, Cummings BJ. Prognostic role of p53 protein expression in epidermoid carcinoma of the anal canal. *Int J Radiat Oncol Biol Phys*. 1999;45(2):309-314. doi:10.1016/s0360-3016(99)00188-1
33. Holm R, Skovlund E, Skomedal H, Flørenes VA, Tanum G. Reduced expression of p21WAF1 is an indicator of malignant behaviour in anal carcinomas. *Histopathology*. 2001;39(1):43-49. doi:10.1046/j.1365-2559.2001.01161.x
34. Ajani JA, Wang X, Izzo JG, et al. Molecular biomarkers correlate with disease-free survival in patients with anal canal carcinoma treated with chemoradiation. *Dig Dis Sci*. 2010;55(4):1098-1105. doi:10.1007/s10620-009-0812-6
35. Nilsson PJ, Lenander C, Rubio C, Auer G, Ljungqvist O, Glimelius B. Prognostic significance of Cyclin A in epidermoid anal cancer. *Oncol Rep*. 2006;16:443-449. doi:10.3892/or.16.3.443
36. Hilmi M, Neuzillet C, Lefèvre JH, et al. Prognostic value of Fusobacterium nucleatum after abdominoperineal resection for anal squamous cell carcinoma. *Cancers (Basel)*. 2022;14(7):1606. doi:10.3390/cancers14071606
37. Nofziger C, Turner AJ, Sangkuhl K, et al. PharmVar GeneFocus: CYP2D6. *Clin Pharmacol Ther*. 2020;107(1):154-170. doi:10.1002/cpt.1643
38. Ingelman-Sundberg M. Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity. *Pharmacogenomics J*. 2005;5(1):6-13. doi:10.1038/sj.tpj.6500285
39. Menoyo A, del Rio E, Baiget M. Characterization of variant alleles of cytochrome CYP2D6 in a Spanish population. *Cell Biochem Funct*. 2006;24(5):381-385. doi:10.1002/cbf.1258
40. Chen C, Cook LS, Li XY, et al. CYP2D6 genotype and the incidence of anal and vulvar cancer. *Cancer Epidemiol Biomarkers Prev*. 1999;8:317-321.
41. Fernández-Santander A, del Saz Sánchez M, Tejerina Gómez A, Bandrés Moya F. CYP2D6*4 allele and breast cancer risk: is there any association? *Clin Transl Oncol*. 2012;14(2):157-159. doi:10.1007/s12094-012-0776-4
42. Shukla P, Gupta D, Pant MC, Parmar D. CYP 2D6 polymorphism: a predictor of susceptibility and response to chemoradiotherapy in head and neck cancer. *J Cancer Res Therapeut*. 2012;8(1):40-45. doi:10.4103/0973-1482.95172
43. Azad AK, Bairati I, Samson E, et al. Validation of genetic sequence variants as prognostic factors in early-stage head and neck squamous cell cancer survival. *Clin Cancer Res*. 2012;18(1):196-206. doi:10.1158/1078-0432.ccr-11-1759
44. Zhou LP, Luan H, Dong XH, Jin GJ, Man DL, Shang H. Genetic variants of CYP2D6 gene and cancer risk: a HuGE systematic review and meta-analysis. *Asian Pac J Cancer Prev*. 2012;13(7):3165-3172. doi:10.7314/apjcp.2012.13.7.3165
45. Hannemann F, Bichet A, Ewen KM, Bernhardt R. Cytochrome P450 systems--biological variations of electron transport chains. *Biochim Biophys Acta*. 2007;1770(3):330-344. doi:10.1016/j.bbagen.2006.07.017
46. Wheeler DL, Iida M, Dunn EF. The role of Src in solid tumors. *Oncology*. 2009;14(7):667-678. doi:10.1634/theoncologist.2009-0009

47. Shah AN, Gallick GE. Src, chemoresistance and epithelial to mesenchymal transition: are they related? *Anti Cancer Drugs*. 2007;18(4):371-375. doi:10.1097/cad.0b013e32801265d7
48. Avizienyte E, Brunton VG, Fincham VJ, Frame MC. The SRC-induced mesenchymal state in late-stage colon cancer cells. *Cells Tissues Organs*. 2005;179(1-2):73-80. doi:10.1159/000084511
49. Lu Y, Li X, Liang K, et al. Epidermal growth factor receptor (EGFR) ubiquitination as a mechanism of acquired resistance escaping treatment by the anti-EGFR monoclonal antibody cetuximab. *Cancer Res*. 2007;67(17):8240-8247. doi:10.1158/0008-5472.can-07-0589
50. Cruz-Gregorio A, Aranda-Rivera AK, Pedraza-Chaverri J. Human papillomavirus-related cancers and mitochondria. *Virus Res*. 2020;286:198016. doi:10.1016/j.virusres.2020.198016

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Trilla-Fuertes L, Gámez-Pozo A, Nogué M, et al. Utility of *CYP2D6* copy number variants as prognostic biomarker in localized anal squamous cell carcinoma. *Cancer*. 2023;129(16):2581-2592. doi:10.1002/cncr.34797