Development of a polygenic risk score to predict recurrence and survival in patients with resected cholangiocarcinoma

Beatriz Val[1,2*], Ainhoa Lapitz[1,2,3*], Tania Pastor[1*], Laura Izquierdo-Sanchez[1,2], Colm J. O'Rourke[5], Alejandro Forner[2,6], Cedric Coulouarn[7], Cedric Menard[7], Giovanni Brandi[8], Simona Tavolari[8], Lewis Roberts[9], Ahmed Fowsiyo[9], Mohamed Ali[9], Rocio I. R. Macias[2,10], Vincenzo Cardinale[11], Jesper B. Andersen[5], Adelaida La Casta[4], Luis Bujanda[1,2,4], Maria J. Perugorria[1,2,4], Raúl Jimenez-Aguero[1,4], Jesus M. Banales[1,2,12,13$], Pedro M. Rodrigues[1,2,12$]

*Share co-authorship; $Share co-seniorship and correspondence

1Department of Liver and Gastrointestinal Diseases, Biogipuzkoa Health Research Institute - Donostia University Hospital -, University of the Basque Country (UPV/EHU), Donostia-San Sebastian, Spain; 2National Institute for the Study of Liver and Gastrointestinal Diseases (CIBERehd, "Instituto de Salud Carlos III"), Spain; 3Department of Biochemistry and Molecular Biology, Faculty of Science and Technology, University of the Basque Country (UPV/EHU), Leioa 48940, Spain; 4Departments of Medicine and Physiology, Faculty of Medicine and Nursing, University of the Basque Country UPV/EHU, Leioa, Spain; 5Biotech Research and Innovation Centre (BRIC), Department of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; 6Barcelona Clinic Liver Cancer (BCLC) group, Liver Unit, Hospital Clínic of Barcelona, Fundació Clínic per a la Recerca Biomédica (FCRB), IDIBAPS, University of Barcelona, Barcelona, Spain; 7Inserm, Univ Rennes, OSS (Oncogenesis Stress Signaling), UMR_S 1242, CHU Rennes, 35042 Rennes, France; 8Medical Oncology, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy; 9Department of Medicine, Mayo Clinic, Rochester, MN, USA; 10Tropical Medicine and Gastroenterology Department, Sohag Faculty of Medicine, Egypt; 11Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN, USA; 12Experimental Hepatology and Drug Targeting (HEVEPHARM) group, University of Salamanca, IBSAL, Salamanca, Spain; 13Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Rome, Italy; 14IKERBASQUE, Basque Foundation for Science, Bilbao, Spain; 15Department of Biochemistry and Genetics, School of Sciences, University of Navarra, Pamplona, Spain.

Contact information

pedromiguel.rodriguesvieira@bio-gipuzkoa.eus,
jesusmaria.banalesasurmendi@bio-gipuzkoa.eus

Introduction and Aims

- Cholangiocarcinoma (CCA) is a heterogeneous group of biliary tumors
- Poor prognosis, even among patients undergoing curative-intent surgery, with recurrence rates of 50–70% and five-year survival rates between 7% and 20%.
- Lack of reliable and clinically validated prognostic biomarkers

This study aimed to identify and validate **prognostic biomarkers** in **tumor tissue** using an integrative approach combining **mRNA expression**, measured by OpenArray RT-qPCR, and **clinical information**

Methods

- Identification of potential CCA tissue prognostic biomarkers through literature review and Cox PH overall survival (OS) analysis in five transcriptomic cohorts
- Collection of an international cohort including tumor tissue from 221 patients undergoing tumor resection
- Measurement of mRNA expression by high-throughput TaqMan OpenArray RT-qPCR
- Cox PH analyses on feature association with recurrence-free survival (RFS) and OS
- Machine learning-based models training, using nested cross-validation, and evaluation

Results

The levels of 11 and 25 genes are associated with RFS and OS, respectively

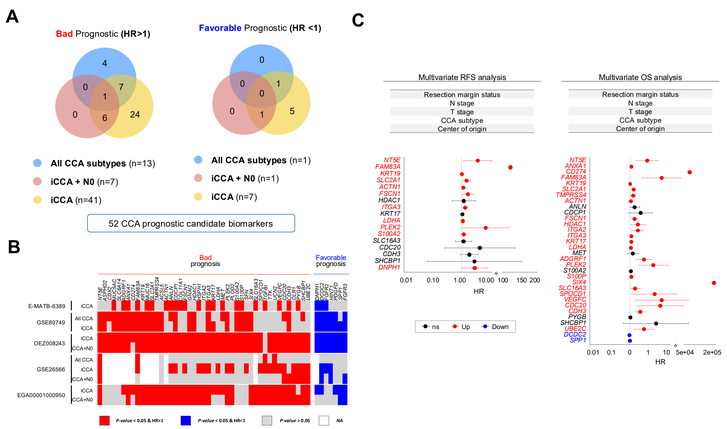


Figure 1. (**A**) Venn diagram illustrating the overlap of candidates of bad and favorable prognosis identified in three analysis groups: all CCA samples, iCCA samples, and iCCA with N0 lymph node status, resulting in a list of 52 final candidates. (**B**) Heatmap summarizing the 52 prognostic biomarkers across analysis groups. (**C**) Multivariate RFS (left) and OS (right) analysis of genes identified as significantly associated with RFS or OS, respectively, in univariate Cox PH analysis. Genes are classified according to the direction of their HR and statistical significance, as determined by the Wald test (HR > 1 & p < 0.05: up; HR < 1 & p < 0.05: down; p > 0.05: ns).

Predictive model:
LDHA, FSCN1, SLC2A1 +
Biological sex, age, CCA subtype, and surgery resection margin

Recurrence-free survival predictive model stratifies patients into 2-years recurrence risk groups

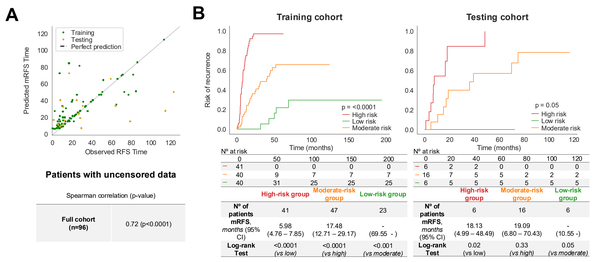


Figure 4. (A) Final RSF model model combining *LDHA*, *FSCN1*, *SLC2A1*, Biological sex, Age, CCA subtype, and surgery resection margin, predicted vs observed values for uncensored patients **(B)** KM curves for the RSF model 2-year recurrence risk groups prediction (high-, moderate-, and low-risk) in training (right) and testing (left) cohorts.

Overall Survival predictive model stratifies patients into 5-years OS risk groups

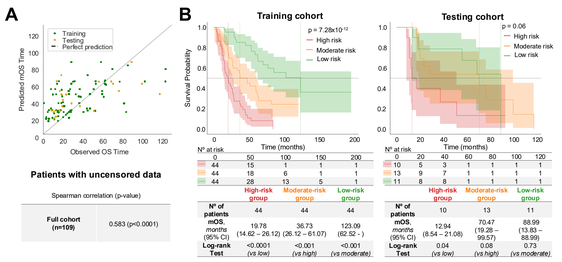
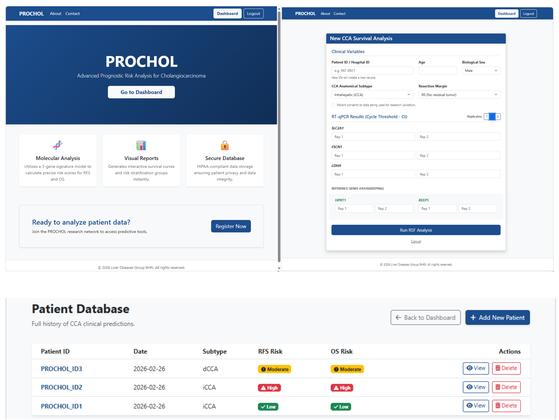


Figure 5. (A) OS-RSF model combining *LDHA*, *FSCN1*, *SLC2A1*, Biological sex, Age, CCA subtype, and surgery resection margin, predicted mOS vs observed OS values for uncensored patients (B) KM curves for the OS-RSF model 5-year recurrence risk groups prediction (high-, moderate-, and low-risk) in training (right) and testing (left) cohorts.

PROgnostic CHOLangiocarcinoma (PROCHOL) web based calculator



RSF-selected candidate prognostic genes are upregulated in CCA tumors, specifically in malignant cholangiocytes and immune cell populations

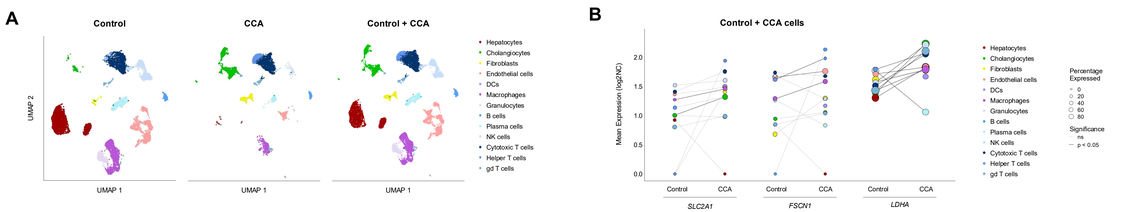


Figure 6. (A) UMAP plot of single-cell transcriptomic data from CCA samples (GSE151530) and control samples (GSE115469), with cells colored according to their assigned cell type clusters. (B) Dotplot showing the expression and proportion of expressing cells for the final candidate prognostic biomarkers. (D) Paired dotplots displaying the mean expression and proportion of expressing cells for each biomarker across cell type clusters in control and CCA samples. Statistical significance for each gene–cell type pair was assessed using a two-sided Student's t-test or Wilcoxon rank-sum test, depending on Shapiro–Wilk normality test.

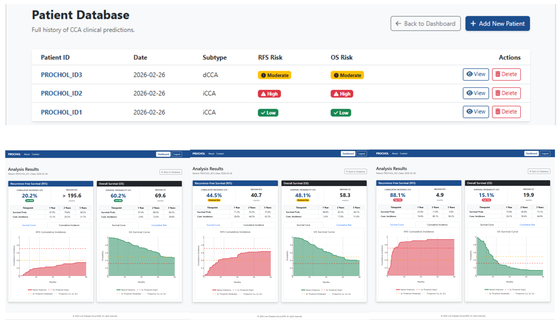


Figure 6. Representation of PROCHOL web-based calculator, including open page and dedicated data entry interface for clinical variables —age, biological sex, CCA subtype, and resection margin— alongside RT-qPCR cycle threshold values for the three-gene signature (SLC2A1, FSCN1, LDHA) and HK references (HPRT1, REEP5) (top) and patient-level visual reports (bottom).

Conclusions

- **496 candidate prognostic genes** were identified from the literature and **52** were validated by transcriptomic analysis in 5 human CCA cohorts
- **11 biomarkers** showed **independent** prognostic value for RFS in the international multicenter cohort
- The best predictive model for RFS and OS included **3 genes** (*LDHA, FSCN1*, *SLC2A1*) and **4 clinical variables** (age, biological sex, CCA subtype, surgery resection margin)
- Biomarkers were expressed in CCA cells and in tumor microenvironment

References

Banales JM, *et al.* Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol. 2020.

Macias RIR, *et al.* Diagnostic and prognostic biomarkers in cholangiocarcinoma. Liver Int. 2019.