


RESEARCH ARTICLE

Open Access



Functional and genomic characterization of three novel cell lines derived from a metastatic gallbladder cancer tumor

Patricia García^{1†}, Carolina Bizama^{1†}, Lorena Rosa^{1,2†}, Jaime A. Espinoza³, Helga Weber⁴, Javier Cerda-Infante⁵, Marianela Sánchez⁶, Viviana P. Montecinos⁶, Justo Lorenzo-Bermejo⁷, Felix Boekstegers⁷, Marcela Dávila-López⁸, Francisca Alfaro¹, Claudia Leiva-Acevedo¹, Zasha Parra⁹, Diego Romero¹, Sumie Kato¹⁰, Pamela Leal⁴, Marcela Lagos¹¹ and Juan Carlos Roa^{12*} 

Abstract

Background: Gallbladder cancer (GBC) is the most common tumor of the biliary tract. The incidence of GBC shows a large geographic variability, being particularly frequent in Native American populations. In Chile, GBC represents the second cause of cancer-related death among women. We describe here the establishment of three novel cell lines derived from the ascitic fluid of a Chilean GBC patient, who presented 46% European, 36% Mapuche, 12% Aymara and 6% African ancestry.

Results: After immunocytochemical staining of the primary cell culture, we isolated and comprehensively characterized three independent clones (PUC-GBC1, PUC-GBC2 and PUC-GBC3) by short tandem repeat DNA profiling and RNA sequencing as well as karyotype, doubling time, chemosensitivity, in vitro migration capability and in vivo tumorigenicity assay. Primary culture cells showed high expression of CK7, CK19, CA 19-9, MUC1 and MUC16, and negative expression of mesothelial markers. The three isolated clones displayed an epithelial phenotype and an abnormal structure and number of chromosomes. RNA sequencing confirmed the increased expression of cytokeratin and mucin genes, and also of *TP53* and *ERBB2* with some differences among the three cell lines, and revealed a novel exonic mutation in *NF1*. The PUC-GBC3 clone was the most aggressive according to histopathological features and the tumorigenic capacity in NSG mice.

Conclusions: The first cell lines established from a Chilean GBC patient represent a new model for studying GBC in patients of Native American descent.

Keywords: Gallbladder cancer, Cancer cell lines, Ascites, Native American ancestry, Gene expression profile

Background

Gallbladder cancer (GBC; ICD-10 diagnosis code C23) accounts for 80–95% of biliary tract cancers and is the sixth most common gastrointestinal cancer worldwide [1, 2]. This aggressive disease is relative rare in most high-income countries, but GBC affects at least 219,420 persons every year worldwide. The geographic areas with the highest mortality rates include Chile, Bolivia, Korea, Nepal, Bangladesh, Japan, Peru, Czech Republic

*Correspondence: jcroa@med.puc.cl

[†]Patricia García, Carolina Bizama and Lorena Rosa contributed equally to this work

¹² Department of Pathology, Faculty of Medicine, Millennium Institute of Immunology and Immunotherapy, Pontificia Universidad Católica de Chile, Santiago, Chile

Full list of author information is available at the end of the article



and Slovakia [3]. In Chile, GBC is one of the most common neoplasms and represents the second cause of cancer-related death in women, the mortality rate in 2015 was 10.2 deaths per 100 000 women [4]. The regional distribution of GBC in Chile is correlated with the prevalence of its main risk factor, gallstone disease, together with the regional proportions of Mapuche Native American ancestry and other socioeconomic factors [5, 6].

One of the major characteristics of GBC is its late diagnosis and ineffective treatment [7]. The lack of specific symptoms in early stages and effective diagnostic biomarkers results in most patients being diagnosed with locally advanced or metastatic disease [8, 9]. Unfortunately, the 5-year survival rate is less than 10% for these patients [10].

To improve the clinical prognosis of GBC patients, more efficient methods for early diagnosis and more effective therapeutic approaches must be developed. Research on GBC is challenging not only because it is a rare tumor worldwide, but also due to the difficulty of accessing human tissue samples and the lack of experimental models. Obtaining fresh frozen tissue samples is extremely demanding because most of the early lesions are thin, flat, incidental and masked with gross acute inflammatory changes, and chemotherapy and palliative care is frequently offered to patients with advanced tumor after diagnosis confirmation by fine needle aspiration biopsy [11]. The prophylactic removal of gallbladders with stones hampers the study of the natural history of GBC, and the non-existence of animal models mimicking the progression of pre-neoplastic lesions to invasive cancer translates into the present basic and clinical knowledge relying on other types of gastrointestinal cancer.

A large proportion of the current knowledge on GBC biology comes from studies performed using GBC cell lines, which have played a critical role in identifying therapeutic targets and facilitating the rapid screening of new anticancer compounds. Most GBC cell lines derive from primary tumors and metastatic sites and have been isolated from Japanese [12–16], Korean [17] and Chinese [18, 19] patients. Cell lines from GBC patients with Native American ancestry are not available yet.

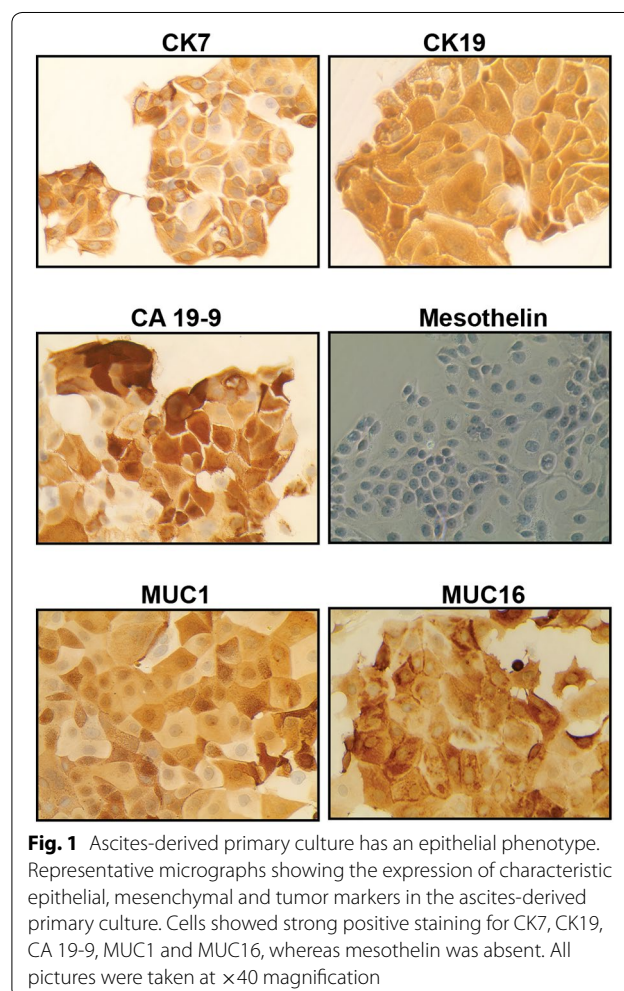
To our knowledge, this is the first report on the development and characterization of cell lines derived from a GBC patient of Native American descent. We describe the isolation of three cell lines from the ascites of a single patient, which conserved the genomic features and malignant behavior of the primary GBC tumor. Our findings suggest that the three established cell lines reflect the heterogeneity of GBC tumors and provide a useful model to investigate the mechanisms underlying late stage GBC progression and metastasis, gain new insight into

drug resistance mechanisms and test new therapeutic strategies.

Results

Establishment of GBC cell lines from one primary cell culture

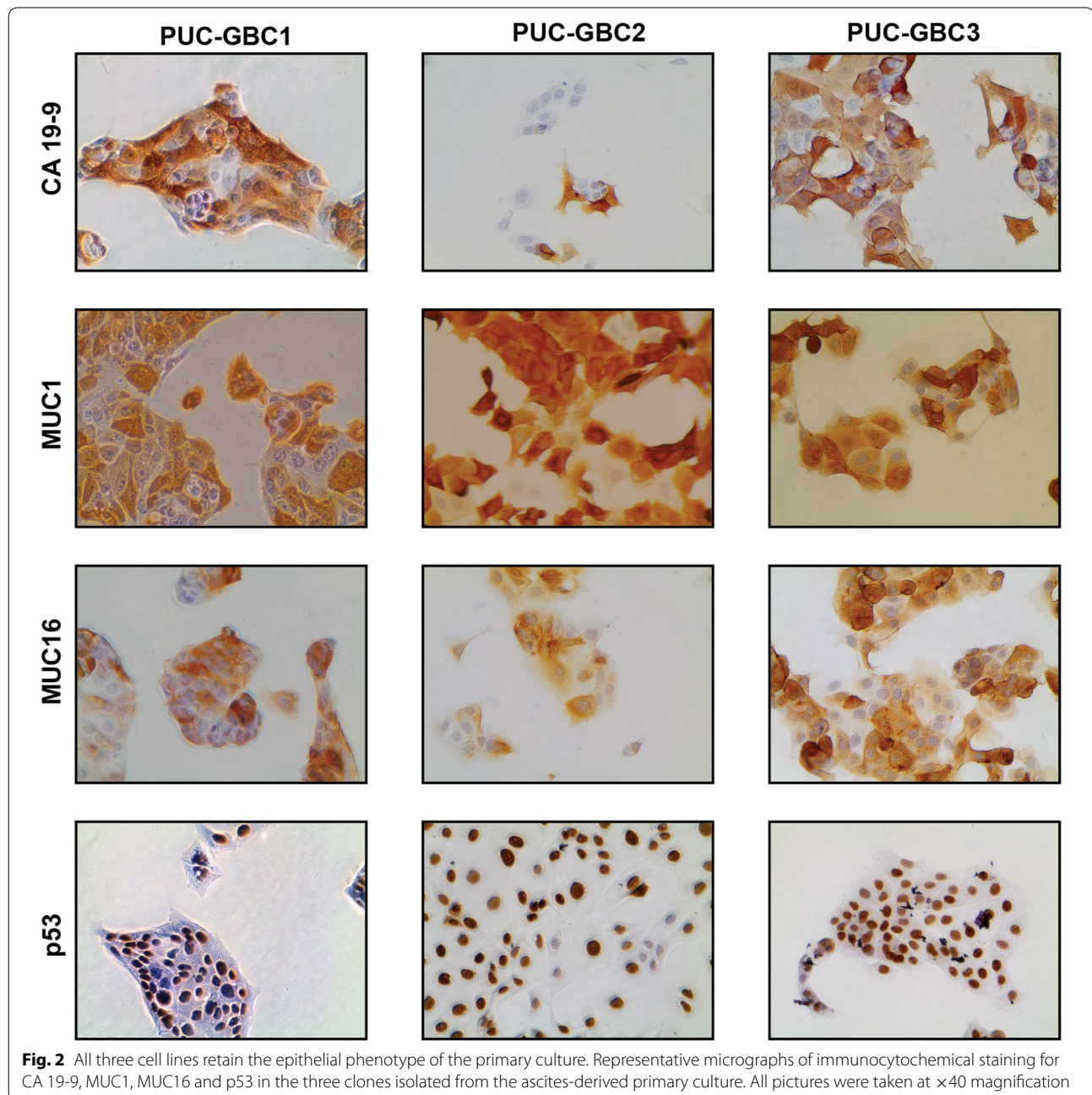
We successfully established a primary culture of tumor cells derived from the ascitic fluid of a 60-year-old man diagnosed with metastatic GBC, who showed the following percentages of genetic ancestry: 46% European, 36% Mapuche Native American, 12% Aymara Native American and 6% African. Moderate to strong expression (2+ to 3+) of the epithelial markers CK7 (cytokeratin 7) and CK19 (cytokeratin 19) was observed in 100% of the ascites-derived cells, mesothelin was not expressed. Almost 90% of the cells showed high expression levels (3+) of the tumor carbohydrate antigen 19-9 (CA 19-9) and moderate expression (2+) of the mucins MUC1 (mucin 1) and MUC16 (mucin 16) (Fig. 1). Taken



together, all these markers indicate an epithelial origin of the primary culture.

The primary culture consisted of epithelial cells with different morphologies, which grew as heterogeneous cell populations. In order to generate monoclonal cell lines, we used a limiting dilution procedure to isolate and spread individual cells. We established three cell lines as separate cultures, namely PUC-GBC1, PUC-GBC2 and PUC-GBC3. In consistency with the primary culture, the newly established cell lines showed high expression levels

of the epithelial markers CK7 and CK19 (3+, 2+) and no mesothelin expression (see Additional file 1). As depicted in Fig. 2, the three cell lines showed positive staining for CA 19-9, MUC1 and MUC16, with moderate to strong (2+ to 3+) intensity in more than 80% of tumor cells, with the exception of CA 19-9, which was only expressed in 20% of PUC-GBC2 cells. We also evaluated the protein expression of the tumor suppressor p53 since its stabilization/activation has been associated with mutations in the *TP53* (Tumor Protein P53) gene and poor prognosis



of GBC [20–25]. All three cell lines were p53 positive, with a strong brown nuclear staining in 100% of the tumor cells.

Short tandem repeat (STR) profiling, chromosome analyses and RNA sequencing

The STR DNA profile analysis revealed very minor discrepancies among the three cell lines: only PUC-GBC3 lost one allele at D7S820 (Table 1).

The modal chromosome number was in the hyperdiploid range (>50) for the three cell lines. The karyotype for PUC-GBC1 was: XY,-X, +3, +4, +5, +5, -7, +10, +11, +12, +12, +12, +13, add(13)(q33), +14, +15, +16, +16, -17,+ 19, -21, -21, -22, -22. For PUC-GBC2: XY, +1, +2, +3, +4, +5, +9, +10, +10, +13, +14, +14, add(q31), +19, -20, -22 [10]. And for PUC-GBC3: XY, +2, +4, + 4, der(5), -6, -8, -9, -11, i(12)(q10), add(13)(q34), +14, +15 ps+, +16, -18, -19, -20, -21, -21, -21, -22 (Fig. 3).

The analysis of RNA sequencing data confirmed the increased expression of cytokeratin (*KRT7* and *KRT19*) and mucin (*MUC1* and *MUC16*) genes, as well as *TP53* with some differences in expression among the cell lines. For example, *TP53* was weakly expressed in PUC-GBC3. The principal component analysis (PCA) of the three cell lines and the 20 genes with the largest variability in gene expression (highest coefficient of variation) revealed that the expression profiles of PUC-GBC2 and PUC-GBC3 were more similar to each other than to PUC-GBC1 (Fig. 4). Based on the genes with the largest expression variability, PUC-GBC2 showed an enrichment of genes coding for proteins that form the extracellular matrix, such as *COL3A1*, *POSTN* and *HAPLN1*. On the other hand, genes involved in the regulation of cancer stem cell properties (*HNF1A*, *OLFM4*, *PCK1* and *REG4*) and

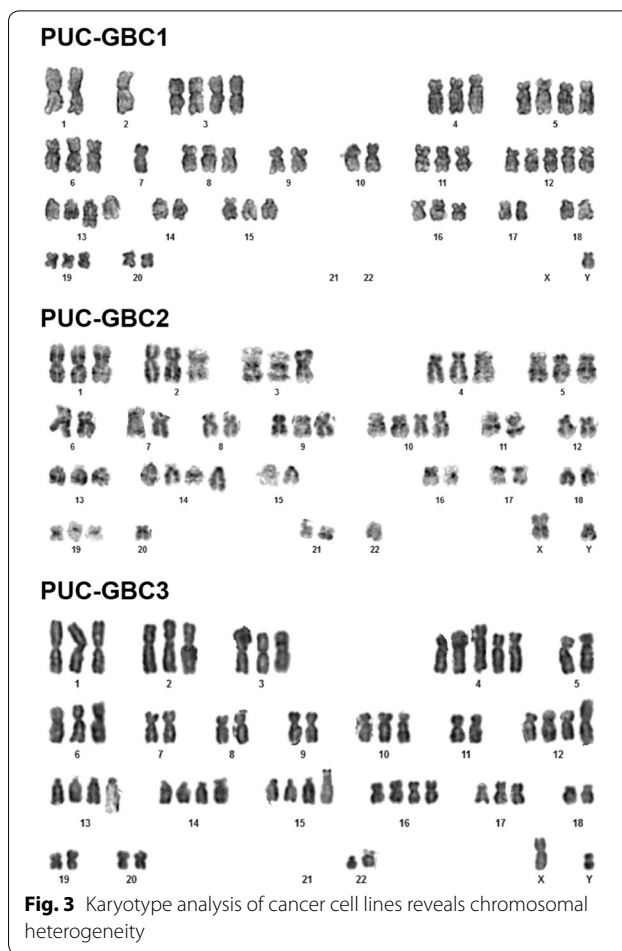


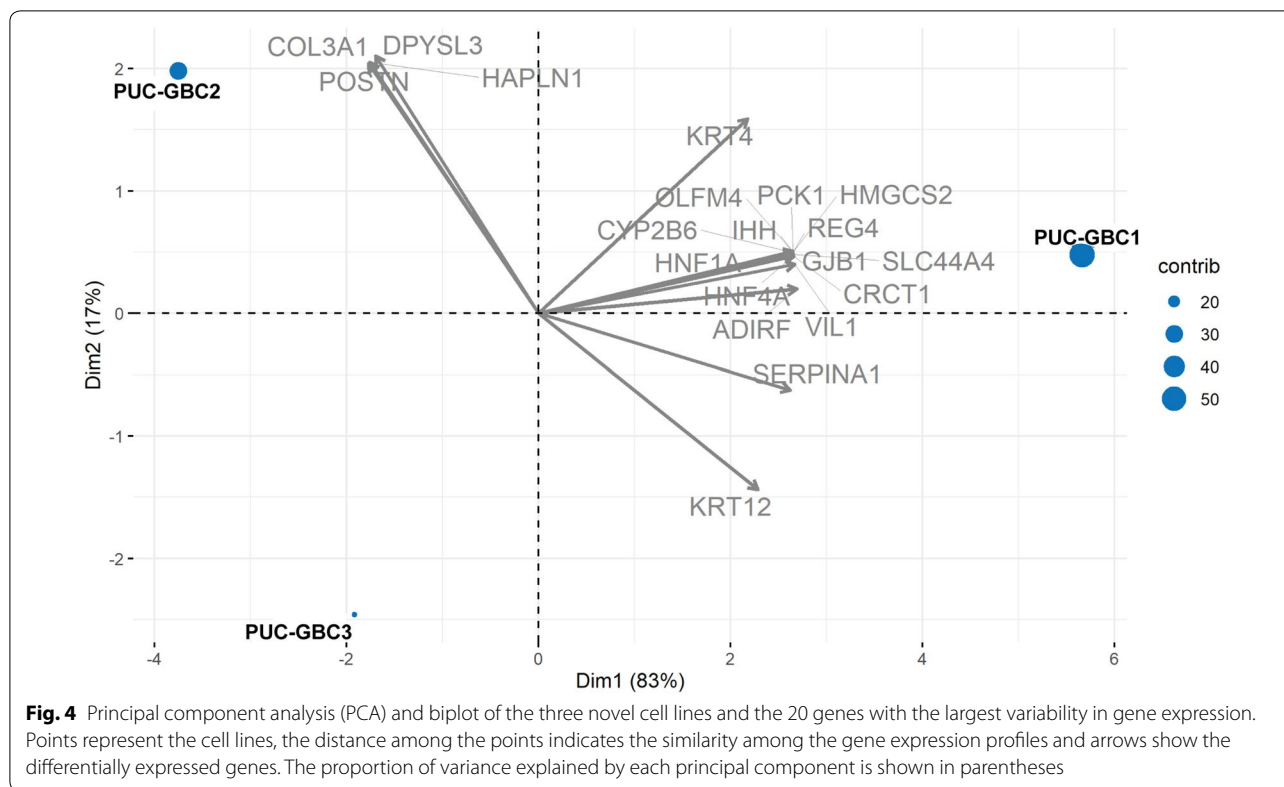
Fig. 3 Karyotype analysis of cancer cell lines reveals chromosomal heterogeneity

cell metabolism (*CYP2B6* and *SLC44A4*) were overrepresented in PUC-GBC1. Additional file 2 provides the expression values for the three newly established GBC cell lines as normalized transcripts per million. We also performed a gene ontology (GO) enrichment analysis, processing genes in terms of their associated molecular function and biological process. The results showed a similar enrichment of GO terms among the three cell lines (Fig. 5). Binding was the largest sub-category in the molecular function category, followed by catalytic activity, while the major biological processes were cellular process, biological regulation, and regulation of biological process.

According to the MSK-IMPACT database [26], mutations in at least 10% of primary GBC tumors are expected in ten genes: *TP53*, *ATM*, *SMAD4*, *ARID1A*, *CTNNB1*, *NF1*, *NOTCH3*, *PTPRD*, *KEAP1* and *ARID1B*. Among these ten candidate genes, RNA sequencing revealed a novel non-synonymous mutation in *NF1* shared by the three GBC cell lines (chr17:31232174, exon25:c.C3299T). According to SIFT (Sorting Intolerance from Tolerance) and Polyphen-2, this mutation has a damaging

Table 1 Short tandem repeat (STR) DNA profiling of ascites-derived gallbladder cancer cell lines

STR marker	Allele(s)		
	PUC-GBC1	PUC-GBC2	PUC-GBC3
AMEL	XY	XY	XY
CSF1PO	10 11	10 11	10 11
D13S317	11 12	11 12	11 12
D16S539	10 12	10 12	10 12
D21S11	33 2	33 2	33 2
D5S818	7 11	7 11	7 11
D7S820	7 10	7 10	10
TH01	7	7	7
TPOX	8	8	8
vWA	16	16	16



amino acid impact (p.S1100L). Additional file 3 lists the number of identified exonic mutations and the annotated genetic variants in the three GBC cell lines. Synonymous variants were more common ($n=5923$) than non-synonymous ($n=4368$) and frameshift ($n=208$), with multiple overlapping variants between the different cell lines (Additional file 4). Considering that these cell lines derived from a metastatic site, genomic alterations were expected in multiple genes, many of them involved in relevant oncogenic signaling pathways. These included non-synonymous single-nucleotide variations (nsSNV) in *CDKN1A* (rs1801270), *MDM4* (rs4252716), *PIK3CA* (rs2230461), *ERBB2* (rs1058808) and *ERBB3* (rs55699040), a frameshift deletion in *CDKN2A* (NA), among others (see Additional file 3).

In vitro characterization

Growth curves were examined for the three cell lines. The population doubling time was 60 h for PUC-GBC1, 36 h for PUC-GBC2 and 44 h for PUC-GBC3. Representative dose–response curves to gemcitabine, cisplatin and fluorouracil (5-FU), along with IC_{50} values are shown in Fig. 6a. Comparisons with calculated IC_{50} from five commercially available GBC cell lines are shown in Fig. 6b. Our three newly established cell lines showed higher sensitivity to gemcitabine and cisplatin than the

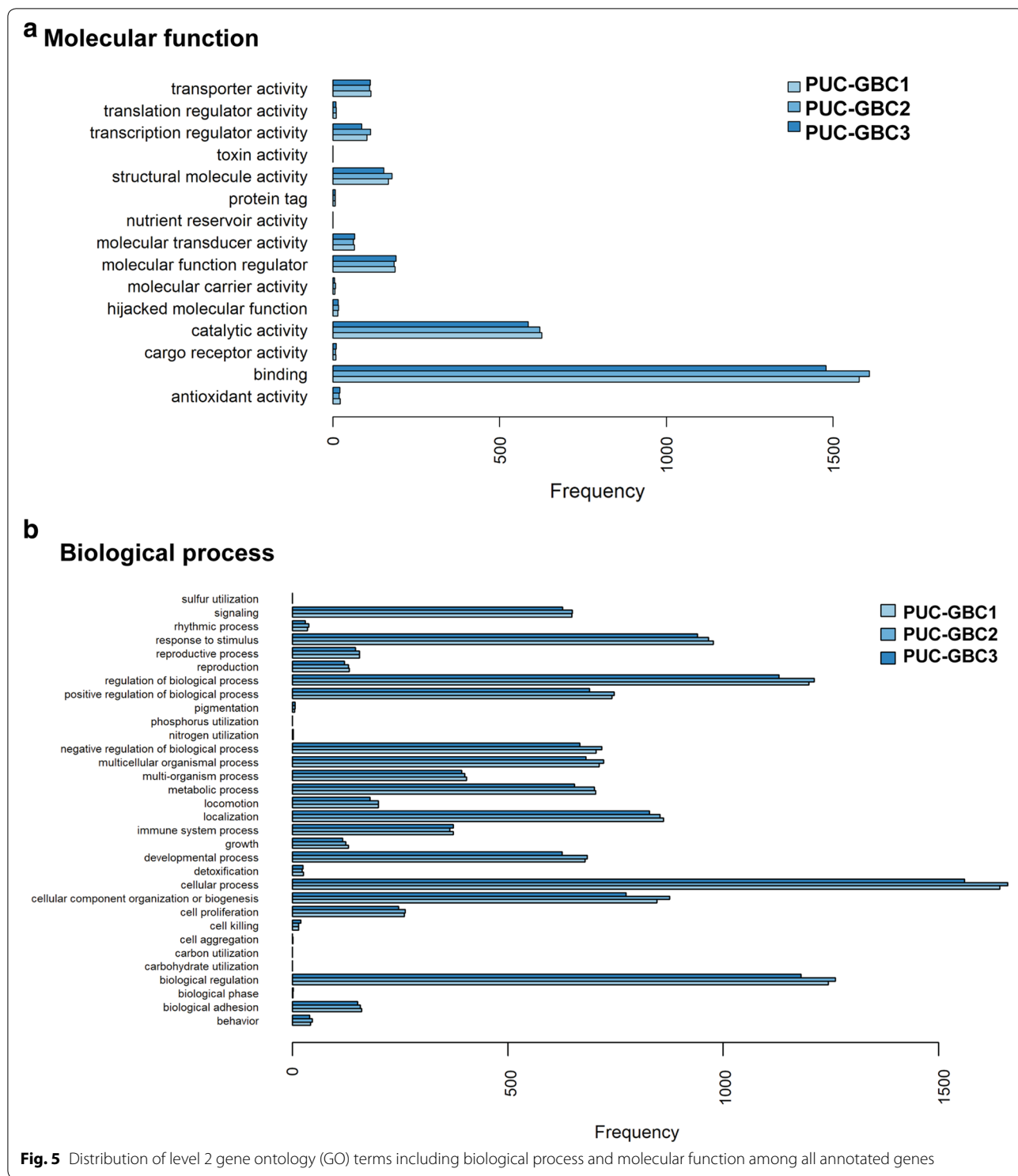
commercially available ones. Sensitivity to 5-FU was similar to that shown by GB-d1 and NOZ.

In vitro migration assays showed that the relative migration rate after 24 h was highest for PUC-GBC2 and lowest for PUC-GBC3 ($P=0.0073$) (Fig. 7).

Tumorigenic potential

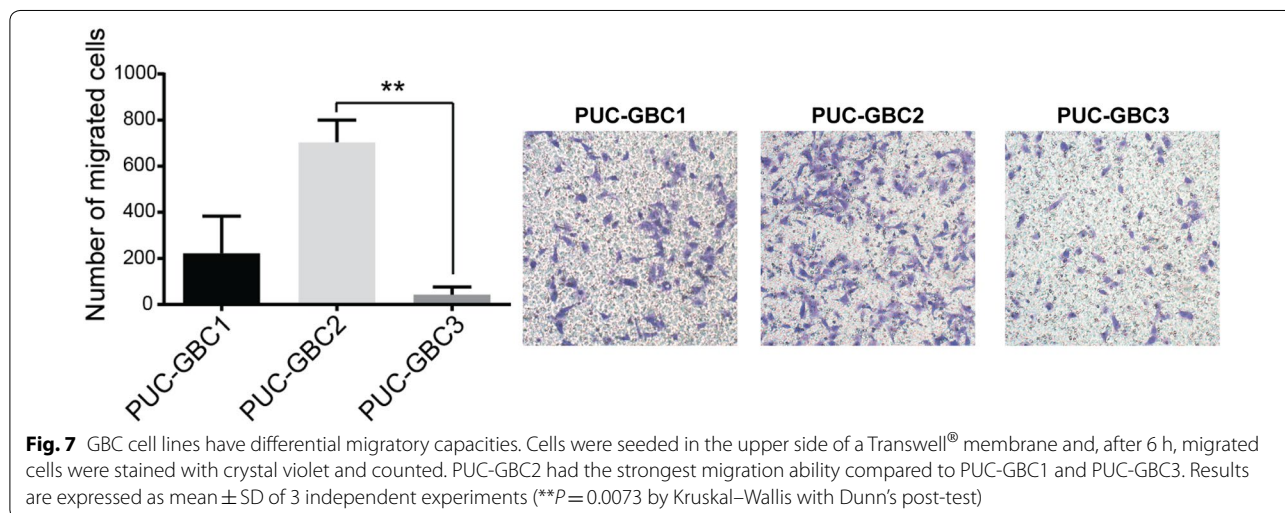
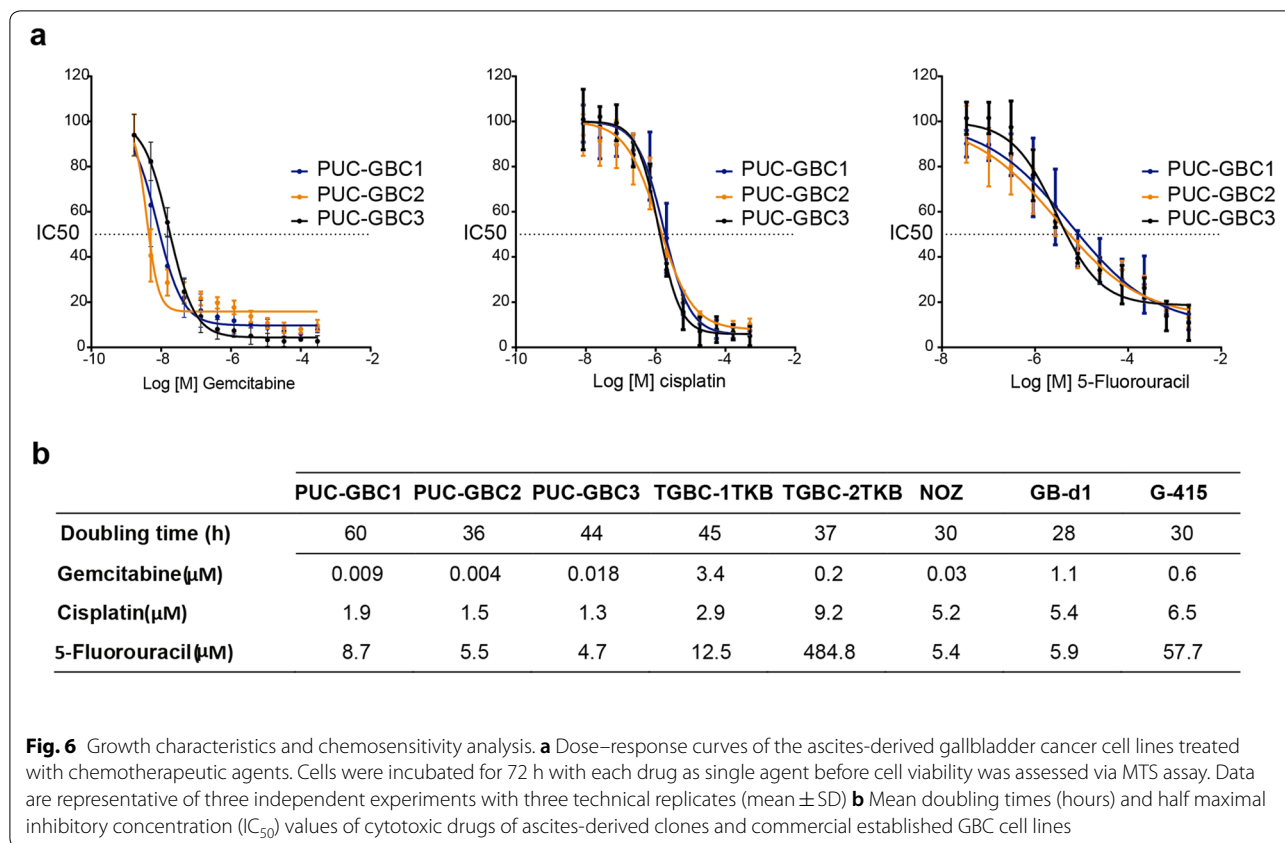
In vivo tumorigenicity assays revealed palpable tumors in all mice within 15 days. As shown in Fig. 8, PUC-GBC3 exhibited the highest growth kinetics, followed by PUC-GBC1, whereas PUC-GBC2 had the lowest tumor-growing potential comparing to PUC-GBC3 ($P=0.0002$) and PUC-GBC1 ($P=0.031$) at day 38.

The hematoxylin- and eosin-stained tissue sections of the mice tumor xenografts showed histological features of adenocarcinoma, with high mitotic index, anisokaryosis and anisocytosis. Particularly, PUC-GBC1 and PUC-GBC2 xenografts were histologically similar and characterized by the presence of tubular structures constituted by cells with enlarged, hyperchromatic nuclei, frequent mitotic activity, abundant apoptotic bodies, and murine stroma dominating the tumor mass. In contrast, PUC-GBC3 tumors had a more aggressive appearance, with a more solid growth and less gland formation; also, isolated signet ring cells and scarce infiltration of the host stroma were noted (Fig. 9).



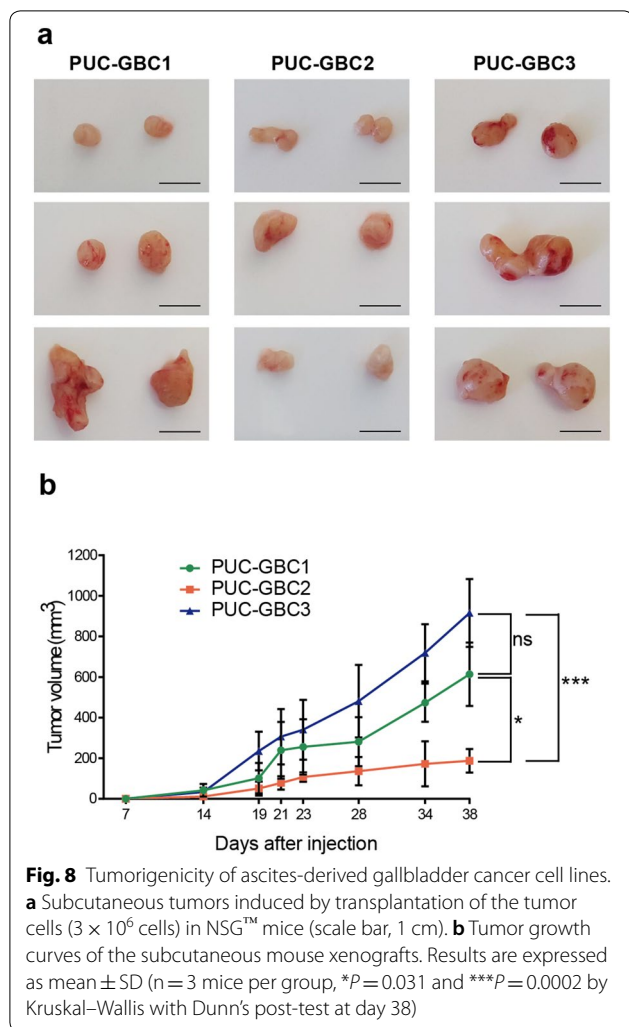
Immunohistochemistry (IHC) showed a strong positivity of the epithelial markers CK7, CK19 and E-cadherin (3+, 100% of the cells), which were also expressed in the primary tumor (Fig. 9 and Additional file 2 with

RNA sequencing expression results). Nuclear positive p53 expression was observed in more than 70% of the cells in the xenografts in agreement with the expression observed in the primary tumor (Fig. 9). We also



decided to evaluate the expression of human epidermal growth factor receptor 2 (HER2) because there is a subset of HER2-positive patients who might be candidates for an anti-HER2 therapy [27, 28]. We did not observe HER2 expression in the tissue section of the primary

tumor or in the xenograft model of PUC-GBC1. However, PUC-GBC2 and PUC-GBC3 showed moderate and high HER2 intensity, and RNA sequencing revealed an increased *ERBB2* expression for all GBC cell lines, and particularly PUC-GBC1 and PUC-GBC2 (Additional file 2).



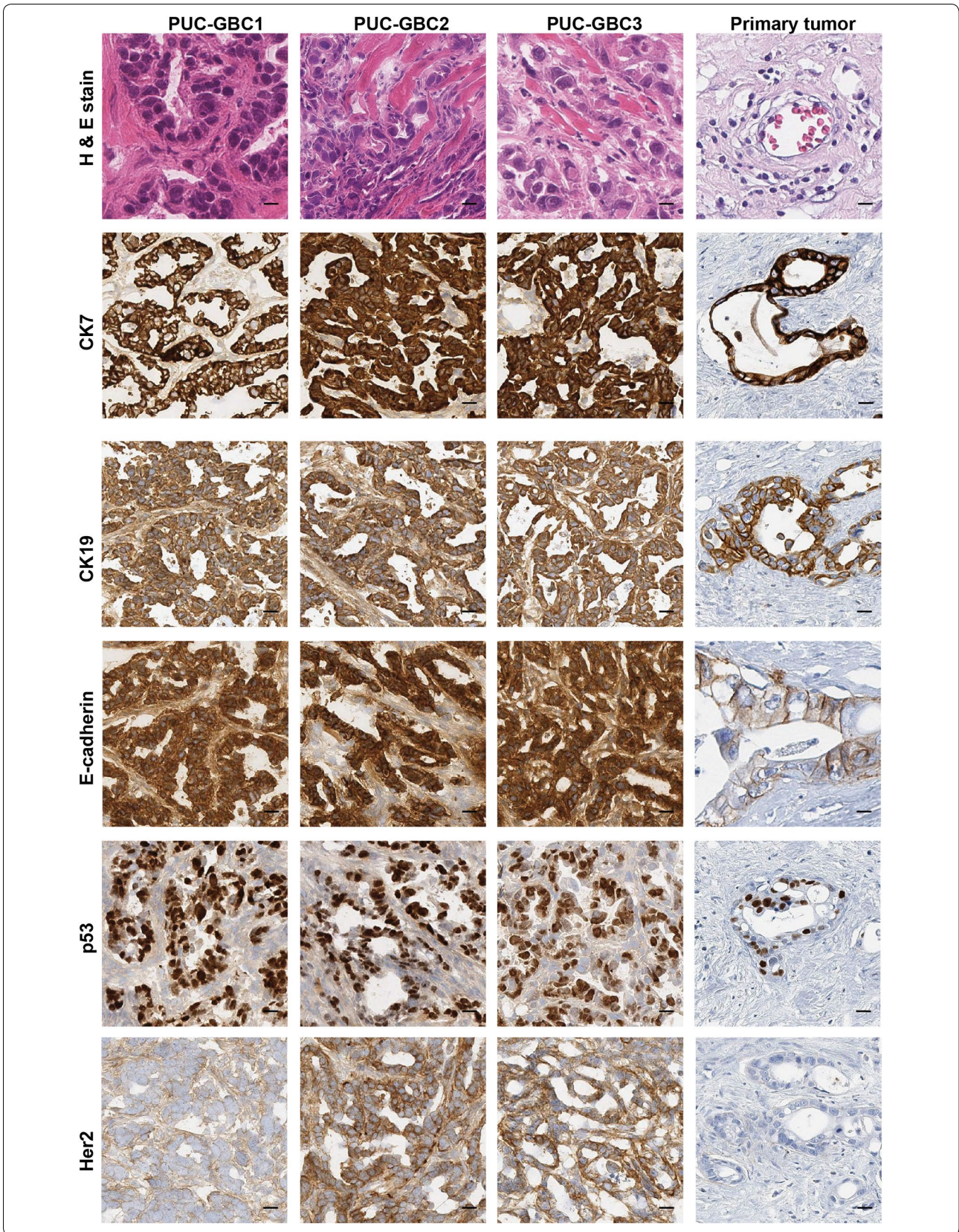
Discussion

We report in this study on the phenotypic, genomic and functional characteristics of three new GBC cell lines derived from the metastatic ascites of a Chilean male patient; these were named PUC-GBC1, PUC-GBC2 and PUC-GBC3. Mapuche Native American ancestry has been strongly associated with the risk of GBC, and we found that the percentage of Mapuche Native American ancestry of the male donor of the primary tumor cells was 36%. For comparison, the first (third) quartiles of Mapuche ancestry reported in a recent study that included 1805 Chileans were 28% (43%) [6], confirming the representativeness of

the established GBC cell lines for the Chilean population. The three new cell lines displayed an epithelial phenotype and showed abnormality in structure and number of chromosomes, with a tendency to triploidy. The doubling time varied from 36 to 60 h and was very similar to that observed for commercially available Asian cell lines. Evaluation of chemosensitivity indicated that our ascites-derived lines are more chemosensitive to gemcitabine and cisplatin than Asian GBC cell lines, which could be due to differences in molecular patterns. Furthermore, PUC-GBC1, PUC-GBC2 and PUC-GBC3 were tumorigenic when injected subcutaneously in NSGTM mice. Interestingly, the tumor volume of the PUC-GBC2 xenografts was the lowest despite PUC-GBC2 cells having the shortest doubling time and a higher migration capacity than PUC-GBC1 and PUC-GBC3 cells. In vivo, PUC-GBC3 cells showed a more aggressive behavior in agreement with the histopathological characteristics of the derived xenografts. The analysis of RNA sequencing data revealed differentially expressed genes that could be responsible for the phenotypic differences observed between the three cell lines. For instance, PUC-GBC2 showed elevated expression of *COL3A1*, which encodes an extracellular matrix (ECM) protein (collagen α -1(III)) associated with tumor progression [29, 30] and has been reported as a marker of poor prognosis [31–33]. At the functional level, recent reports have demonstrated that *COL3A1* promotes cell proliferation and migration [34, 35]. Similarly, high expression of *POSTN*, also overrepresented in PUC-GBC2, might explain the higher migratory ability of these cells. *POSTN* encodes for periostin, a multifunctional ECM protein that contributes to the remodeling of the tumor microenvironment during tumor progression [36]. In CCA and hepatoblastoma, overexpressed periostin was found to induce cell migration and epithelial-to-mesenchymal transition (EMT) features [37, 38]. On the other hand, PUC-GBC1 cells exhibited a slower growth rate and an enhanced in vivo tumorigenic capability. The transcriptome profile showed a high expression of genes related with cancer stem cell properties, such as *HNF1A*, a transcription factor recently reported to be enriched in pancreatic cancer stem cells [39] and a key player in the transcriptional reprogramming of colorectal cancer cells that promote liver metastasis [40]. RNA-Seq also confirmed the increased expression of cytokeratin and mucin genes in all three cell lines, showed an elevated *TP53* and *ERBB2* expression for PUC-GBC1 and PUC-GBC2, and revealed a novel exonic mutation in *NF1*.

(See figure on next page.)

Fig. 9 Xenograft tumors display high similarities with the original primary tumor. Xenografts showed histological features of adenocarcinoma, with tumor infiltration of muscle tissue. Expression of all markers was observed in the tumor cells of xenografts, except for HER2, which was present with different intensity only in cells derived from peritoneal metastasis. Scale bar, 20 μ m; magnification, $\times 20$



We confirmed the epithelial origin of the primary culture and the derived cell lines by evaluation of the immunocytochemical expression of CK7 and CK19 (Additional file 1 and Fig. 9). Keratins are a complex subclass of the intermediate filaments, made up of more than 20 different polypeptides. Among them, CK19 belongs to the type I group of cytokeratins and is specifically expressed in the periderm, whereas CK7 is a type II cytokeratin specifically expressed in the simple epithelia lining the cavities of the internal organs and in the gland ducts and blood vessels [41]. Both CKs have shown normal expression in gallbladder epithelium [42] and are considered markers of biliary tract tumors [43].

We also evaluated the expression of CA 19-9, which is a glycolipid synthesized by the pancreatic, biliary, gastric, colonic cells, as well as the endometrial and salivary epithelia [43]. We found elevated CA 19-9 levels in almost all the cells of the primary culture and of the clones PUC-GBC1 and PUC-GBC3. The intensity of CA 19-9 expression was also high in PUC-GBC2, but only in 20% of cells (Fig. 2). Recently, Barnett et al. reported that differential expression of CA 19-9 and sTRA (Sialyl-TRA), another carbohydrate antigen, made it possible to identify morphologically distinct subsets of pancreatic cancer cells. Subpopulations expressing both markers were part of well differentiated and mucin-secreting pancreatic adenocarcinomas, whereas those expressing just one were often poorly differentiated and vacuolated and never mucin-secreting. In addition, the authors observed cases displaying sometimes more than one subpopulation in the same tumor [44]. A heterogeneous distribution of CA 19-9 expression in GBC could explain our findings, especially considering that the clones were isolated from malignant ascites, characterized by containing a heterogeneous group of tumor cells. However, more research is needed to determine if cells with different patterns of CA 19-9 expression co-exist in gallbladder tumors.

Other tumor markers strongly expressed by the primary culture and individual clones were the mucins MUC1 and MUC16, particularly MUC1 according to RNA sequencing results (Additional file 2). MUC1 and MUC16 are transmembrane glycoproteins, located in 1q22 and 19p13 locus respectively, which contribute to forming physiological barriers and transmitting growth and survival signals to the cells [45]. High MUC1 immunoreactivity has been associated with vascular invasion and the malignant progression of many tumors [46–51], including GBC [52, 53]. Studies have reported that MUC1 expression is higher in GBC than in normal and inflammatory gallbladder tissue [54–57] and more frequently and extensively expressed in poorly differentiated adenocarcinoma [58]. The upregulation of MUC1 in cancer is a product of the aberrant O-linked glycosylation

that lead to an increased sialylation of this protein [59, 60]. Interestingly, the expression rate of stromal localization of sialylated MUC1 at the deepest invading sites of pT2 gallbladder carcinoma seems to be associated with a more frequent postsurgical peritoneal dissemination and lower survival rate of these patients [54]. Similar findings have reported that sialylated MUC1 mucin plays an important role in the progression of prostate cancer [61] and may be involved in the metastatic potential of pancreatic ductal adenocarcinoma [62].

On the other hand, the upregulation of MUC16 in primary tumor and ascites-derived clones could be associated with a biologically aggressive phenotype (Figs. 1, 2). Indeed, MUC16 (cancer antigen 125, CA 125) is used in clinics to diagnose and predict prognosis in ovarian cancer patients, but it is also upregulated in a large percentage of digestive tract adenocarcinomas and has been proposed as a prognostic marker for gastrointestinal cancers [63]. In a recent study, positive immunohistochemical staining of MUC16 (immunoreactivity in more than 10% of the tumor cells) was observed in more than 75% of extrahepatic pancreatobiliary tumors (n = 234 cases), including 37 gallbladder cancers [64]. Chaube et al. proposed MUC16 as a potential diagnostic marker because serum levels are increased in patients with GBC and benign disease can be differentiated from malignant disease with a sensitivity of 64% and a specificity of 90% [65].

These newly established cell lines, in particular PUC-GBC1 and PUC-GBC2, showed overexpression of p53, which was also observed in the primary tumor and xenografts (Fig. 9). In GBC, most studies have reported a frequency over 50% of p53 overexpression [20, 23, 66–68]. In addition, there is a progressive increase in p53 positivity from dysplastic lesions to carcinoma in situ and invasive carcinoma [23, 66, 69], and a high correlation between protein overexpression and the presence of *TP53* mutations [20, 21]. The cell lines described here conserved the p53 overexpression observed in the primary tumor, although our analysis did not show any *TP53* mutation. It has been generally accepted that mutations in *TP53* may result in the synthesis of a functionally defective p53 protein that exhibits an extended half-life, which tends to accumulate in the cell nucleus and can be detected by immunohistochemistry [70, 71]. However, studies in other tumors have described immunohistochemical overexpression of p53 in absence of gene mutations due to the deregulation of upstream factors of p53 pathway, which lead to protein stabilization and accumulation in the nucleus [72, 73]. Based on our data, it is highly probable that the p53 signaling pathway, along with other relevant oncogenic pathways, is dysfunctional in all three cell lines. For instance, genomic analysis revealed a nsSNV in *TP53* (dbSNP identifier: rs121912651) in all

cell lines, which has not been reported as an oncogenic mutation but likely possesses a pathogenic role according to the ClinVar database (<https://www.ncbi.nlm.nih.gov/clinvar/>). Additionally, we found a nsSNV in *MDM4* and a frameshift deletion in *CDKN2A*, both genes encoding proteins related with regulation of p53 (Additional file 3). However, further studies are needed in order to determine the level of p53 activity in these novel cell lines, as well as the functional and clinical significance of those genetic variants in gallbladder cancer.

We also assessed the expression of HER2 in primary tumors and xenografts. Overexpression of this protein has been reported in 14–48% of GBC cases [27, 28, 74–78], and its amplification and mutation has been observed in 5–20% and 10% of the cases, respectively [27, 75–77, 79]. The newly established GBC cell lines did not present any *ERBB2* mutation. HER2 upregulation has been associated with poor prognosis in GBC [77, 78], and experimental tumor models have suggested that HER2 signaling is involved in gallbladder carcinogenesis [80, 81]. In this study, immunohistochemistry for HER2 was negative for the primary tumor and PUC-GBC1 xenografts but was 2+ for PUC-GBC2 and 3+ for PUC-GBC3 (Fig. 9). *ERBB2* was particularly highly expressed in PUC-GBC1 and PUC-GBC2, and showed a somewhat lower expression in PUC-GBC3. This is not an unexpected finding considering that intratumor heterogeneity is a common event in cancer, and genetic heterogeneity of HER2 has been reported for other tumors [82–88]. For instance, focal or patchy positivity of HER2 is a pattern encountered in primary gastric tumors, which is consequence of the intratumor heterogeneity and could explain discordances observed in HER2 status between the primary tumor and metastatic sites [82, 85]. Recently, Toshiba et al. investigated the status of HER2 in a large cohort of patients with GBC and observed intratumor heterogeneity in 51% (20/39) of the cases with IHC scores of 2+ and 3+. Among them, around 80% showed amplification of HER2/neu gene as determined by fluorescent in situ hybridization (FISH) [27]. Based on these observations, it would be relevant to perform the immunohistochemical test for HER2 in primary and metastatic sites if clinical trials with anti-HER2 therapies are considered for patients with this disease. In this regard, it would be even more advisable to implement the technology of liquid biopsies and take advantage of their clinical value for the identification of heterogeneous subclonal populations of tumor cells.

Conclusions

In summary, we established and characterized three new GBC cell lines derived from a patient with metastatic disease, who represents well the Chilean population with a

36% percentage of Mapuche Native American ancestry. The new cell lines share some biological characteristics, but also show genomic and phenotypic differences that reflect intratumor heterogeneity. The established GBC cell lines provide a new experimental model for future research, including the study of cellular and molecular mechanisms involved in metastasis, the identification of new tumor-associated markers, and the evaluation of responses to new therapeutic agents.

Materials and methods

Cell lines and growth conditions

A panel of five cell lines of Asian origin was used for comparison of doubling times and drug sensitivity. GB-d1 [14] was provided by Dr. Anirban Maitra (Department of Pathology, Division of Pathology and Laboratory Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA); NOZ was obtained from the Health Science Research Resources Bank (Osaka, Japan; No JCRB1033); and G-415, TGBC-1TKB and TGBC-2TKB were purchased from RIKEN BioResource Center (Ibaraki, Japan; No RCB2640, RCB1129 and RCB1130). G-415 and GB-d1 were grown in RPMI 1640 medium (Thermo Scientific HyClone, Logan, UT, USA) supplemented with 10% fetal bovine serum (FBS), 10 units/mL penicillin and 10 mg/mL streptomycin (1% penicillin/streptomycin, Thermo Scientific HyClone). NOZ, TGBC-1TKB and TGBC-2TKB were grown in Dulbecco's Modified Eagle Medium (DMEM high glucose; Corning, New York, NY, USA) supplemented with 5% FBS and 1% penicillin/streptomycin.

Isolation and establishment of human gallbladder cancer cell lines

The primary cell culture was established from the ascites of a 60-year-old male patient with GBC. Histological examination classified this tumor as moderately differentiated tubular adenocarcinoma with a T3NxM1 stage. There was evidence of vascular and lymphatic system invasion and peritoneal carcinomatosis. Ascitic fluid (50 mL) was obtained from the patient, delivered to the laboratory and centrifuged at 1000 rpm for 10 min, and the pellet was rinsed twice with sterile phosphate-buffered saline (PBS; Corning, New York, NY, USA) containing antibiotics (1% penicillin–streptomycin solution, P/S). The supernatant was removed and saved as ascitic fluid supplement at -20 °C, and the remaining cells were resuspended in DMEM/F12 medium (HyClone, GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) supplemented with 5% FBS, 50% of ascitic fluid (HyClone, GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) and 1% P/S, seeded into 6-well culture plates, and incubated at 37 °C in a humidified atmosphere of 5% CO₂ in the air.

The growth medium was replaced every 2–3 days, and the plate was regularly checked for epithelial cells and fibroblast outgrowth. The ascitic fluid supplement was maintained in the culture and reduced 10 times every week until the cells reached confluency. After the first subculture, the cells were grown in complete culture medium alone (without added ascitic fluid) and submitted to immunocytochemistry analysis to confirm the epithelial origin and evaluate the expression of tumor markers. This analysis was repeated at passage 17. Primary culture cells were subcultured until they reached more than 20 passages. Then, three individual clones were obtained by serial dilution of the primary cell culture in a 96-well plate. Once established, all cell lines were cultured in DMEM supplemented with 5% FBS and 1% P/S. All cultures were tested for mycoplasma contamination with a commercial PCR kit (EZ-PCR Mycoplasma Test Kit, Biological Industries, Cromwell, CT, USA).

Immunostaining procedures

Primary antibodies and working dilutions used for immunostaining analysis were mouse monoclonal anti-cytokeratin 7 (CK7; Cat. No M7018; Clone OV-TL 12/30) and mouse monoclonal anti-p53 (Cat. No IS616; Clone DO-7) from Dako (Agilent Technologies, Santa Clara, CA, USA); mouse monoclonal anti-mesothelin (Cat. No ACI 3175; Clone MSLN-15C11) from Biocare Medical, LLC (Pacheco, CA, USA); mouse monoclonal anti-CA 15-3 (MUC1; Cat. No MU323-UC; Clone BGX323A) from Biogenex (Fremont, NE, USA); mouse monoclonal anti-cytokeratin 19 (CK19; Cat. No 760-4281; Clone A53-B/A2.26), mouse monoclonal anti-CA 19-9 (Cat. No 760-2609; Clone 121SLE), mouse monoclonal anti-CA 125 (MUC16; Cat. No 760-2610; Clone OC125), rabbit monoclonal anti-calretinin (Cat. No 790-4467; Clone SP65), mouse monoclonal anti-vimentin (Cat. No 790-2917; Clone V9), mouse monoclonal anti-E-cadherin (Cat. No 790-4497; Clone 36) and rabbit monoclonal anti-HER2 (Cat. No 790-2991; Clone 4B5) from Roche Tissue Diagnostics (Oro Valley, AZ, USA).

The epithelial and tumor origin of the primary culture from the ascites and the derived cell lines was evaluated by immunocytochemistry. Cells were seeded at 2×10^5 per well in 24-well plates and allowed to attach for 24 h. Following fixation in 4% paraformaldehyde for 15 min, PBS wash and permeabilization with 0.1% Triton X-100, cells were incubated with primary antibodies at 1:100 dilution for 1 h. Then, cells were washed four times with 0.2% BSA (bovine serum albumin) in PBS and incubated for 30 min with the SuperPicture™ Polymer Detection Kit (Thermo Fisher Scientific Inc., Waltham, MA, USA). The bound antibody was visualized with a red chromogen (ImmPACT™ NovaRED™ Substrate; Vector Laboratories

Inc., Burlingame, CA, USA). Antibody for HER2/neu was immunostained on the Benchmark XT automated stainer (Ventana Medical Systems, Tucson, AZ, USA) according to the manufacturer's instructions. Images were acquired using the EVOS XL Core Imaging System (Thermo Fisher Scientific Inc., Waltham, MA, USA).

The primary tumor and xenografts were analyzed immunohistochemically on formalin-fixed paraffin-embedded (FFPE) sections (2 μ m thickness) using the automated immunostainer BenchMark ULTRA and the ultraVIEW universal DAB detection kit (Ventana Medical Systems Inc., Roche Group, Tucson, AZ, USA). Digital images were captured using the Aperio AT2 Digital Pathology Scanner (Leica Biosystems, Nussloch, Germany). The expression was evaluated by the pathologist (J.C.R.) considering the intensity of the staining (1+, weak; 2+, moderate; 3+, strong) and the percentage of labeled cells. Immunohistochemical scoring of HER-2 expression was based on the system proposed for gastric and gastroesophageal junction cancer [89].

Short tandem repeat (STR) DNA profiling

The genetic profiling of the cell lines derived from the primary cell culture was established using the polymorphic STR loci detection service offered by the Department of Molecular Medicine at Aarhus University Hospital (IdentiCell service for human cell line authentication). DNA of each cell line was purified using the PureLink® Genomic DNA mini kit (Thermo Fisher Scientific Inc., Waltham, MA, USA).

Chromosome analysis

The cytogenetic karyotyping analysis was performed using a standard air-dried method. Briefly, cells in an exponential growth phase were treated with a final concentration of 0.1 μ g/mL colcemid (Colcemid® Solution, Irvine Scientific, Santa Ana, CA, USA) for 45 min at 37 °C. Then, cells were harvested to arrest metaphases and treated with the hypotonic solution (0.075 M KCl) for 10 min at 37 °C. After two changes in the fixative (3:1, methanol: glacial acetic acid), the cell suspension was incubated overnight at -20 °C. The next day, the fixative was changed another three times and two drops of the cell suspension were dropped from a distance of about 50 cm onto clean dry slides tilted at an angle of about 45°, allowing the cells to roll across the slide. After air drying, one slide was stained with Giemsa (3:1 ratio of Gurr Buffer and Giemsa Stain) and analyzed under a light microscope at 10 \times and 100 \times magnification. If the metaphase cells were abundant and well spread, the remaining slides were used for chromosome analysis using trypsin G-banding. To determine ploidy, chromosomes were counted from a minimum of 50 metaphase spreads using

GenASIs Karyotyping (Applied Spectral Imaging Inc. Carlsbad, CA, USA).

Cell growth characteristics

The doubling time was calculated by seeding the cells at a density of 20 000 cells/well in 12-well plates and counting them by Trypan blue dye exclusion in a Neubauer camera every 24 h for 15 days (the medium was replaced every 3 days). The doubling times were determined from the growth curves using the data generated from 3 independent experiments, each with three technical replicates.

Chemosensitivity assay

The cytotoxic drugs used for in vitro chemosensitivity assays were gemcitabine (Sigma-Aldrich, St. Louis, MO, USA), cisplatin (Calbiochem, Merck group, Darmstadt, Germany) and Fluorouracil (Laboratorios Kampar, Santiago, Chile). These drugs are the main chemotherapy agents used to treat GBC. Cells were seeded into 96-well plates at a density of 5×10^3 cells/well in 100 μ L cell culture medium. After an overnight attachment period, cells were treated with either gemcitabine (starting at 300 μ M), cisplatin (starting at 350 μ M) or fluorouracil (starting at 2 mM) with a serial dilution of 1/3. Following 72 h of drug incubation, cell viability was assessed by incubating the cells for 2 h at 37 °C with a MTS-PMS colorimetric solution (Promega Corp., Madison, WI, USA), and absorbance of each well was read at a wavelength of 490 nm using a multiwell plate reader (Epoch Microplate Spectrophotometer, BioTek Instruments Inc., Winooski, VT, USA). The half maximal inhibitory concentration was calculated from the dose response curves (IC_{50} : dose of drug required to reduce final cell number or optical density in MTS assay to 50% of control). Three independent experiments were performed, each with three technical replicates.

Cell migration assay

Migration assays were performed using Transwell® 24-well plates containing polycarbonate filters with an 8-micron pore size (BD Biosciences, Bedford, MA, USA). Complete medium was placed in the lower chamber to act as a chemoattractant, and 5×10^4 cells were seeded with serum-free medium into the upper chamber. After 24 h, cells were fixed in methanol for 15 min and stained with 0.05% crystal violet in 25% methanol/PBS for 15 min. Cells on top of the membrane were removed using a cotton swab, and filters were washed with PBS. Cells on the underside of filters were viewed and counted under a microscope in 10 randomly selected fields. Three independent experiments were performed, each with three technical replicates.

Tumorigenicity assay

The in vivo tumorigenicity assay was performed using 7 to 8-week-old male NOD *scid* gamma (NSG™) mice obtained from The Jackson Laboratory (Bar Harbor, ME, USA). The animals ($n=9$) were randomized into three groups, anesthetized with isoflurane (3% for induction of anesthesia and 1.5% for maintenance) and 3×10^6 cells suspended in 150 μ L PBS/matrigel (Matrigel™ Basement Membrane Matrix, BD Biosciences, Bedford, MA, USA) were injected bilaterally and subcutaneously on the back of the anesthetized mice. Animals were examined every week for the development of tumors. Tumor volumes were estimated twice a week from caliper measurements ($\text{volume} = 0.52 \times (\text{width})^2 \times \text{length}$). When tumors had grown to 1.5–2.0 cm^3 after 4 weeks, mice were euthanized by carbon dioxide (CO_2) and the tumor xenografts were removed, weighted and photographed. Then, tumors were fixed in 10% neutral buffered formalin, and processed for routine histopathological examination.

Variant and gene expression analysis from RNA sequencing, and estimation of genetic ancestry proportions

Total RNA was extracted and sequenced on an Illumina Hi-Seq 2000. The quality of the sequencing reads was assessed using FastQC (0.11.2) (<http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>). PRINSEQ (0.20.3) [90] and cutadapt (1.9) [91] were used to filter the reads based on quality parameters and to remove adapters. Variant calling was performed following the GATK guidelines for RNA sequencing data, including a two-pass alignment with STAR (2.5.2b) [92], towards the GRCh38 version of the human genome, removal of PCR duplicates with picard tools (2.1.0) (<http://broadinstitute.github.io/picard/>) and variant calling with GATK 3.5 [93]. Thresholds for the detection of mutations were read depth > 20, allelic depth > 10, mapping quality > 40 and Fisher's strand < 60. Variants were annotated with Annovar [94]. Gene expression values were calculated using feature Counts from the subread package (1.6.4) [95] and normalized as transcripts per million. More than 30 M reads were generated per each cell line sample. Additional file 5 provides the sequencing statistics for the three newly established GBC cell lines. Gene ontology analyses were conducted with the R package goProfiles.

The identified variants with an allele frequency > 5% according to the 1000 Genome Project (more than 5000 variants in total) were used to infer the genetic ancestry proportions of the donor of the primary GBC tumor cells [96]. Surrogates of African and European ancestry were 87 Yorubans in Ibadan, Nigeria, and 80 Utah residents with Northern and Western European ancestry from the

1000 Genome Project. Nine Mapuche and nine Aymara individuals were selected to represent the two largest indigenous peoples in Chile based on the three following criteria: four grandparental Mapuche or Aymara surnames, estimated Native American proportion of at least 74% for Mapuche and at least 99% for Aymara reference individuals, and mitochondrial DNA haplogroups consistent with Mapuche (haplogroup C or D) or Aymara (haplogroup B) descent [6]. The ADMIXTURE software was used for supervised estimation of the African, European, Mapuche Native American and Aymara Native American ancestry components relying on the above-mentioned identified common variants, reference individuals and genotype data from a recent study that included 1805 Chileans [6, 97].

Statistical analysis

Statistical analyses were performed using R programming environment in Rstudio© version 1.0 (Rstudio, Inc., Boston, MA). A Kruskal–Wallis test was used to compare multiple groups, followed by Dunn’s post hoc multiple comparisons test. Probability values smaller than 0.05 were considered statistically significant.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s40659-020-00282-7>.

Additional file 1. Representative micrographs of immunocytochemical staining for CK7, CK19, Mesothelin, Vimentin and Calretinin in the three clones isolated from the ascites-derived primary culture (magnification, 40×).

Additional file 2. Gene expression of normalized transcripts per million in Chilean GBC cell lines.

Additional file 3. Exonic mutations in Chilean GBC cell lines.

Additional file 4. Venn diagrams depicting unique and shared (overlapping circles) variants in the three cell lines. (a) non-synonymous; (b) synonymous; and (c) frameshift. The green circle depicts the variants identified in PUC-GBC1, the orange circle depicts the variants identified in PUC-GBC2, and the blue circle depicts the variants identified in PUC-GBC3.

Additional file 5. Sequencing statistics.

Abbreviations

GBC: Gallbladder cancer; CK7: Cytokeratin 7; CK19: Cytokeratin 19; CA 19-9: Carbohydrate antigen 19-9; MUC1: Mucin 1; MUC16: Mucin 16; TP53: Tumor Protein P53; STR: Short tandem repeat; GO: Gene ontology; nsSNV: Nonsynonymous single-nucleotide variation; SIFT: Sorting intolerance from tolerance; 5-FU: Fluorouracil; HER2: Human epidermal growth factor receptor 2; ERBB2: Erb-B2 Receptor Tyrosine Kinase 2; IHC: Immunohistochemistry; CKs: Cytokeratins; sTRA: Sialyl-TRA; CA 125: Cancer antigen 125; FISH: Fluorescent in situ hybridization; DMEM: Dulbecco’s Modified Eagle Medium; FBS: Fetal bovine serum; PBS: Phosphate-buffered saline; P/S: Penicillin-streptomycin solution; BSA: Bovine serum albumin; FFPE: Formalin-fixed paraffin-embedded; NSG: NOD scid gamma.

Acknowledgements

Not applicable.

Authors’ contributions

PG, CB and LR contributed equally to this work. PG, CB, LR, JAE and JCR conceived and designed the experiments. PG, CB, LR, FA and CL-A performed the in vitro experiments and characterized the primary culture and derived clones. JC-I, MS and VPM performed the in vivo experiments. JL-B, MD-L and FB performed the gene expression analysis from RNA sequencing and the estimation of genetic ancestry proportions. FA, ZP and ML provided help with the karyotyping. SK provided help with the primary culture establishment. DR assisted with tissue processing and imaging. PG, CB, LR, JAE, HW, PL, JL-B and JCR analyzed the data and drafted the manuscript. All authors read and approved the final manuscript.

Funding

This research was supported by grants from the Fondo Nacional de Desarrollo Científico y Tecnológico, FONDECYT No. 1170893 (JCR), 1171463 (CB), 1201734 (PL), 11180987 (HW) and 1160800 (SK); Dirección de Investigación Medicina UC-Pontificia Universidad Católica de Chile, Project No 16-240 (JCR); and Instituto Milenio IMIL P09/016-F (JCR). Lorena Rosa is a student of the Ph.D Program “Doctorado en Ciencias mención Biología Celular y Molecular Aplicada”, Universidad de La Frontera and was supported by CONICYT (#21140027). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials

All data generated and/or analyzed in this study are included in this published article and its additional files. Additionally, they are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

The present study was approved by the Scientific Ethics Committee of the Faculty of Medicine, Pontificia Universidad Católica de Chile, and was conducted in accordance with the tenets of the 1964 Helsinki declaration and its subsequent amendments. The clinical sample used to develop and establish the primary culture was obtained after the patient signed the written informed consent form. Mouse husbandry and animal experiments were conducted in accordance with the recommendations of the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. For injections of cancer cells, mice were anesthetized under 3% isoflurane for induction of anesthesia and 1.5% for maintenance. At the end of the assay, mice were euthanized using carbon dioxide (CO₂) overdose. The protocol was approved by the Committee on Animal Experimentation, Pontificia Universidad Católica de Chile (Protocol No CBB-224/2014).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Pathology, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile. ² Applied Molecular and Cellular Biology PhD Program, Universidad de La Frontera, Temuco, Chile. ³ Science for Life Laboratory, Division of Genome Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden. ⁴ Center of Excellence in Translational Medicine (CEMT) and Scientific and Technological Bioresource Nucleus (BIOREN), Universidad de La Frontera, Temuco, Chile. ⁵ Department of Hematology Oncology; Cellular and Molecular Biology, Pontificia Universidad Católica de Chile, Santiago, Chile. ⁶ Department of Hematology Oncology, Pontificia Universidad Católica de Chile, Santiago, Chile. ⁷ Statistical Genetics Research Group, Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Germany. ⁸ Bioinformatics Core Facility, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. ⁹ Cytogenetics Laboratory, Complejo Asistencial Dr. Sótero del Río, Santiago, Chile. ¹⁰ Division of Obstetrics and Gynecology, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile. ¹¹ Department of Clinical Laboratory, Faculty of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile. ¹² Department of Pathology, Faculty of Medicine, Millennium Institute of Immunology and Immunotherapy, Pontificia Universidad Católica de Chile, Santiago, Chile.

Received: 11 October 2019 Accepted: 6 April 2020
Published online: 15 April 2020

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. <https://doi.org/10.3322/caac.21492> (Epub 2018/09/13).
- Hundal R, Shaffer EA. Gallbladder cancer: epidemiology and outcome. *Clin Epidemiol*. 2014;6:99–109. <https://doi.org/10.2147/clep.s37357> (Epub 2014/03/19).
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015;136(5):E359–86. <https://doi.org/10.1002/ijc.29210> (Epub 2014/09/16).
- Series y Gráficos de Mortalidad. Ministerio de Salud. <http://www.deis.cl/series-y-graficos-de-mortalidad/>. Accessed 05 Feb 2019.
- Villanueva L. Cancer of the gallbladder-Chilean statistics. *E Cancer Med Sci*. 2016;10:704. <https://doi.org/10.3332/ecancer.2016.704> (Epub 2017/01/21).
- Lorenzo Bermejo J, Boekstegers F, Gonzalez Silos R, Marcelain K, Baez Benavides P, Barahona Ponce C, et al. Subtypes of Native American ancestry and leading causes of death: Mapuche ancestry-specific associations with gallbladder cancer risk in Chile. *PLoS Genet*. 2017;13(5):e1006756. <https://doi.org/10.1371/journal.pgen.1006756> (Epub 2017/05/26).
- Andren-Sandberg A. Diagnosis and management of gallbladder cancer. *N Am J Med Sci*. 2012;4(7):293–9. <https://doi.org/10.4103/1947-2714.98586> (Epub 2012/08/07).
- Lazcano-Ponce EC, Miquel JF, Munoz N, Herrero R, Ferrecio C, Wistuba II, et al. Epidemiology and molecular pathology of gallbladder cancer. *CA Cancer J Clin*. 2001;51(6):349–64.
- Randi G, Franceschi S, La Vecchia C. Gallbladder cancer worldwide: geographical distribution and risk factors. *Int J Cancer*. 2006;118(7):1591–602. <https://doi.org/10.1002/ijc.21683>.
- Compton CC, Byrd DR, Garcia-Aguilar J, Kurtzman SH, Olawaiye A, Washington MK, editors. *AJCC Cancer Staging Atlas*. 2nd ed. New York: Springer-Verlag; 2012.
- Goldin RD, Roa JC. Gallbladder cancer: a morphological and molecular update. *Histopathology*. 2009;55(2):218–29. <https://doi.org/10.1111/j.1365-2559.2008.03192.x> (Epub 2009/06/06).
- Homma S, Hasumura S, Nagamori S, Kameda H. Establishment and characterization of a human gall bladder carcinoma cell line NOZ. *Hum Cell*. 1988;1(1):95–7 (Epub 1988/03/01).
- Koyama S, Yoshioka T, Mizushima A, Kawakita I, Yamagata S, Fukutomi H, et al. Establishment of a cell line (G-415) from a human gallbladder carcinoma. *Gan*. 1980;71(4):574–5 (Epub 1980/08/01).
- Shimura H, Date K, Matsumoto K, Nakamura T, Tanaka M. Induction of invasive growth in a gallbladder cancer cell line by hepatocyte growth factor in vitro. *Jpn J Cancer Res*. 1995;86(7):662–9 (Epub 1995/07/01).
- Ghosh M, Koike N, Yanagimoto G, Tsunoda S, Kaul S, Hirano T, et al. Establishment and characterization of unique human gallbladder cancer cell lines. *Int J Oncol*. 2004;24(5):1189–96 (Epub 2004/04/07).
- Yamada N, Chung Y, Ohtani H, Ikeda T, Onoda N, Sawada T, et al. Establishment and characterization of a new human gallbladder carcinoma cell line (OCUG-1) producing TA-4. *Int J Oncol*. 1997;10(6):1251–5 (Epub 1997/06/01).
- Ku JL, Yoon KA, Kim JJ, Kim WH, Jang JY, Suh KS, et al. Establishment and characterization of six human biliary tract cancer cell lines. *Br J Cancer*. 2002;87(2):187–93. <https://doi.org/10.1038/sj.bjc.6600440> (Epub 2002/07/11).
- Wang JH, Li LF, Yu Y, Li B, Jin HJ, Shen DH, et al. Establishment and characterization of a cell line, EH-GB2, derived from hepatic metastasis of gallbladder cancer. *Oncol Rep*. 2012;27(3):775–82. <https://doi.org/10.3892/or.2011.1570> (Epub 2011/12/03).
- Liu ZY, Xu GL, Tao HH, Yang YQ, Fan YZ. Establishment and characterization of a novel highly aggressive gallbladder cancer cell line, TJ-GBC2. *Cancer Cell Int*. 2017;17:20. <https://doi.org/10.1186/s12935-017-0388-8> (Epub 2017/02/15).
- Itoi T, Watanabe H, Yoshida M, Ajioka Y, Nishikura K, Saito T. Correlation of p53 protein expression with gene mutation in gall-bladder carcinomas. *Pathol Int*. 1997;47(8):525–30 (Epub 1997/08/01).
- Roa I, Melo A, Roa J, Araya J, Villaseca M, de Aretxabala X. P53 gene mutation in gallbladder cancer. *Rev Med Chil*. 2000;128(3):251–8 (Epub 2000/08/30).
- Quan ZW, Wu K, Wang J, Shi W, Zhang Z, Merrell RC. Association of p53, p16, and vascular endothelial growth factor protein expressions with the prognosis and metastasis of gallbladder cancer. *J Am Coll Surg*. 2001;193(4):380–3 (Epub 2001/10/05).
- Ghosh M, Sakhuja P, Singh S, Agarwal AK. p53 and beta-catenin expression in gallbladder tissues and correlation with tumor progression in gallbladder cancer. *Saudi J Gastroenterol*. 2013;19(1):34–9. <https://doi.org/10.4103/1319-3767.105922> (Epub 2013/01/16).
- Kim K, Kim DH, Chae SW, Shin JH, Kim HJ, Do SI, et al. Expression of cell cycle-related proteins, p16, p53 and p63 as important prognostic markers in gallbladder adenocarcinoma. *Pathol Oncol Res*. 2014;20(2):409–15. <https://doi.org/10.1007/s12253-013-9710-5> (Epub 2013/11/02).
- Singh A, Mishra PK, Saluja SS, Talikoti MA, Kirtani P, Najmi AK. Prognostic significance of HER-2 and p53 expression in gallbladder carcinoma in North Indian patients. *Oncology*. 2016;91(6):354–60. <https://doi.org/10.1159/000450999> (Epub 2016/10/27).
- Zehir A, Benayed R, Shah RH, Syed A, Middha S, Kim HR, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med*. 2017;23(6):703–13. <https://doi.org/10.1038/nm.4333> (Epub 2017/05/10).
- Yoshida H, Shimada K, Kosuge T, Hiraoka N. A significant subgroup of resectable gallbladder cancer patients has an HER2 positive status. *Virchows Arch*. 2016;468(4):431–9. <https://doi.org/10.1007/s00428-015-1898-1> (Epub 2016/01/14).
- Halder S, Kundu S, Chakraborty J, Chakrabarti S. Significance of HER2 and Ki-67 in preneoplastic lesions and carcinoma of gallbladder. *J Gastrointest Cancer*. 2018. <https://doi.org/10.1007/s12029-018-0162-8> (Epub 2018/08/30).
- Tapper J, Kettunen E, El-Rifai W, Seppala M, Andersson LC, Knuutila S. Changes in gene expression during progression of ovarian carcinoma. *Cancer Genet Cytogenet*. 2001;128(1):1–6. [https://doi.org/10.1016/s0165-4608\(01\)00386-7](https://doi.org/10.1016/s0165-4608(01)00386-7) (Epub 2001/07/17).
- Yue H, Wang J, Chen R, Hou X, Li J, Lu X. Gene signature characteristic of elevated stromal infiltration and activation is associated with increased risk of hematogenous and lymphatic metastasis in serous ovarian cancer. *BMC Cancer*. 2019;19(1):1266. <https://doi.org/10.1186/s12885-019-6470-y> (Epub 2020/01/01).
- Yuan L, Shu B, Chen L, Qian K, Wang Y, Qian G, et al. Overexpression of COL3A1 confers a poor prognosis in human bladder cancer identified by co-expression analysis. *Oncotarget*. 2017;8(41):70508–20. <https://doi.org/10.18632/oncotarget.19733> (Epub 2017/10/21).
- Shi S, Tian B. Identification of biomarkers associated with progression and prognosis in bladder cancer via co-expression analysis. *Cancer Biomark*. 2019;24(2):183–93. <https://doi.org/10.3233/cbm-181940> (Epub 2019/01/29).
- Engqvist H, Parris TZ, Kovacs A, Nemes S, Werner Ronnerman E, De Lara S, et al. Immunohistochemical validation of COL3A1, GPR158 and PITHD1 as prognostic biomarkers in early-stage ovarian carcinomas. *BMC Cancer*. 2019;19(1):928. <https://doi.org/10.1186/s12885-019-6084-4> (Epub 2019/09/20).
- Gao YF, Zhu T, Chen J, Liu L, Ouyang R. Knockdown of collagen alpha-1(III) inhibits glioma cell proliferation and migration and is regulated by miR128-3p. *Oncol Lett*. 2018;16(2):1917–23. <https://doi.org/10.3892/ol.2018.8830> (Epub 2018/07/17).
- Carpino G, Overi D, Melandro F, Girmaldi A, Cardinale V, Di Matteo S, et al. Matrisome analysis of intrahepatic cholangiocarcinoma unveils a peculiar cancer-associated extracellular matrix structure. *Clin Proteomics*. 2019;16:37. <https://doi.org/10.1186/s12014-019-9257-x> (Epub 2019/11/07).
- Ma H, Wang J, Zhao X, Wu T, Huang Z, Chen D, et al. Periostin promotes colorectal tumorigenesis through integrin-FAK-Src Pathway-mediated YAP/TAZ activation. *Cell Rep*. 2020;30(3):793–806. <https://doi.org/10.1016/j.celrep.2019.12.075> (Epub 2020/01/23).
- Chen L, Tian X, Gong W, Sun B, Li G, Liu D, et al. Periostin mediates epithelial-mesenchymal transition through the MAPK/ERK pathway

- in hepatoblastoma. *Cancer Biol Med.* 2019;16(1):89–100. <https://doi.org/10.20892/j.issn.2095-3941.2018.0077> (Epub 2019/05/24).
38. Sonongbua J, Siritungyong S, Thongchot S, Kamolhan T, Utispan K, Thuwajit P, et al. Periostin induces epithelialto mesenchymal transition via the integrin alpha5beta1/TWIST2 axis in cholangiocarcinoma. *Oncol Rep.* 2020;43(4):1147–58. <https://doi.org/10.3892/or.2020.7485> (Epub 2020/02/06).
 39. Abel EV, Goto M, Magnuson B, Abraham S, Ramanathan N, Hotaling E, et al. HNF1A is a novel oncogene that regulates human pancreatic cancer stem cell properties. *Elife.* 2018. <https://doi.org/10.7554/elife.33947> (Epub 2018/08/04).
 40. Teng S, Li YE, Yang M, Qi R, Huang Y, Wang Q, et al. Tissue-specific transcription reprogramming promotes liver metastasis of colorectal cancer. *Cell Res.* 2020;30(1):34–49. <https://doi.org/10.1038/s41422-019-0259-z> (Epub 2019/12/08).
 41. Chougule P, Sumitran-Holgersson S. Cytokeratins of the Liver and Intestine Epithelial Cells During Development and Disease. In: Hamilton G, editor. *Cytokeratins tools in oncology*. London: IntechOpen; 2012. p. 15–32.
 42. Uhlen M, Bjorling E, Agaton C, Szzygarto CA, Amini B, Andersen E, et al. A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol Cell Proteomics.* 2005;4(12):1920–32. <https://doi.org/10.1074/mcp.m500279-mcp200> (Epub 2005/08/30).
 43. Malaguarnera G, Giordano M, Paladina I, Rando A, Uccello M, Basile F, et al. Markers of bile duct tumors. *World J Gastrointest Oncol.* 2011;3(4):49–59. <https://doi.org/10.4251/wjgo.v3.i4.49> (Epub 2011/04/30).
 44. Barnett D, Liu Y, Partyka K, Huang Y, Tang H, Hostetter G, et al. The CA19-9 and Sialyl-TRA antigens define separate subpopulations of pancreatic cancer cells. *Sci Rep.* 2017;7(1):4020. <https://doi.org/10.1038/s41598-017-04164-z> (Epub 2017/06/24).
 45. McGuckin MA. Introduction to mucins in cancer. In: Taylor-Papadimitriou J, Burchell JM, editors. *Mucins and Cancer*. London: Future Medicine Ltd; 2013. p. 6–18.
 46. Xu F, Liu F, Zhao H, An G, Feng G. Prognostic significance of mucin antigen MUC1 in various human epithelial cancers: a meta-analysis. *Medicine.* 2015;94(50):e2286. <https://doi.org/10.1097/md.0000000000002286> (Epub 2015/12/20).
 47. Wang XT, Kong FB, Mai W, Li L, Pang LM. MUC1 immunohistochemical expression as a prognostic factor in gastric cancer: meta-analysis. *Dis Markers.* 2016;2016:9421571. <https://doi.org/10.1155/2016/9421571> (Epub 2016/05/18).
 48. Lin X, Gu Y, Kapoor A, Wei F, Aziz T, Ojo D, et al. Overexpression of MUC1 and genomic alterations in its network associate with prostate cancer progression. *Neoplasia.* 2017;19(11):857–67. <https://doi.org/10.1016/j.neo.2017.06.006> (Epub 2017/09/21).
 49. Huang X, Sun Q, Chen C, Zhang Y, Kang X, Zhang JY, et al. MUC1 overexpression predicts worse survival in patients with non-small cell lung cancer: evidence from an updated meta-analysis. *Oncotarget.* 2017;8(52):90315–26. <https://doi.org/10.18632/oncotarget.19861> (Epub 2017/11/23).
 50. Ozcan HEA, Anuk T, Ozden O. Expression profile and cellular localizations of mucin proteins, CK7, and cytoplasmic p27 in Barrett's esophagus and esophageal adenocarcinoma. *Adv Med Sci.* 2018;63(2):296–300. <https://doi.org/10.1016/j.advms.2018.04.002> (Epub 2018/05/29).
 51. Saltos A, Khalil F, Smith M, Li J, Schell M, Antonia SJ, et al. Clinical associations of mucin 1 in human lung cancer and precancerous lesions. *Oncotarget.* 2018;9(86):35666–75. <https://doi.org/10.18632/oncotarget.26278> (Epub 2018/11/28).
 52. Toba T, Kijima H, Hakamada K, Igarashi Y. Histological phenotype is correlated with the wall-invasion pattern of gallbladder adenocarcinoma. *Biomed Res.* 2014;35(5):295–302 (Epub 2014/10/31).
 53. Hiraki T, Yamada S, Higashi M, Hatanaka K, Yokoyama S, Kitazono I, et al. Immunohistochemical expression of mucin antigens in gallbladder adenocarcinoma: MUC1-positive and MUC2-negative expression is associated with vessel invasion and shortened survival. *Histol Histopathol.* 2017;32(6):585–96. <https://doi.org/10.14670/hh-11-824> (Epub 2016/09/28).
 54. Kawamoto T, Shoda J, Irimura T, Miyahara N, Furukawa M, Ueda T, et al. Expression of MUC1 mucins in the subserosal layer correlates with postsurgical prognosis of pathological tumor stage 2 carcinoma of the gallbladder. *Clin Cancer Res.* 2001;7(5):1333–42 (Epub 2001/05/15).
 55. Chang HJ, Kim SW, Lee BL, Hong EK, Kim WH. Phenotypic alterations of mucins and cytokeratins during gallbladder carcinogenesis. *Pathol Int.* 2004;54(8):576–84. <https://doi.org/10.1111/j.1440-1827.2004.01666.x> (Epub 2004/07/21).
 56. Ghosh M, Kamma H, Kawamoto T, Koike N, Miwa M, Kapoor VK, et al. MUC 1 core protein as a marker of gallbladder malignancy. *Eur J Surg Oncol.* 2005;31(8):891–6. <https://doi.org/10.1016/j.ejso.2005.03.008> (Epub 2005/06/01).
 57. Xiong L, Yang Z, Yang L, Liu J, Miao X. Expressive levels of MUC1 and MUC5AC and their clinicopathologic significances in the benign and malignant lesions of gallbladder. *J Surg Oncol.* 2012;105(1):97–103. <https://doi.org/10.1002/jso.22055> (Epub 2011/08/05).
 58. Yamato T, Sasaki M, Watanabe Y, Nakanuma Y. Expression of MUC1 and MUC2 mucin core proteins and their messenger RNA in gall bladder carcinoma: an immunohistochemical and in situ hybridization study. *J Pathol.* 1999;188(1):30–7. [https://doi.org/10.1002/\(sici\)1096-9896\(199905\)188:1%3c30:aid-path291%3e3.0.co;2-q](https://doi.org/10.1002/(sici)1096-9896(199905)188:1%3c30:aid-path291%3e3.0.co;2-q) (Epub 1999/07/09).
 59. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer.* 2015;15(9):540–55. <https://doi.org/10.1038/nrc3982> (Epub 2015/08/21).
 60. Beatson R, Tajadura-Ortega V, Achkova D, Picco G, Tsourouktsoglou TD, Klausung S, et al. The mucin MUC1 modulates the tumor immunological microenvironment through engagement of the lectin Siglec-9. *Nat Immunol.* 2016;17(11):1273–81. <https://doi.org/10.1038/ni.3552> (Epub 2016/10/21).
 61. Arai T, Fujita K, Fujime M, Irimura T. Expression of sialylated MUC1 in prostate cancer: relationship to clinical stage and prognosis. *Int J Urol.* 2005;12(7):654–61. <https://doi.org/10.1111/j.1442-2042.2005.01112.x> (Epub 2005/07/28).
 62. Masaki Y, Oka M, Ogura Y, Ueno T, Nishihara K, Tangoku A, et al. Sialylated MUC1 mucin expression in normal pancreas, benign pancreatic lesions, and pancreatic ductal adenocarcinoma. *Hepatogastroenterology.* 1999;46(28):2240–5 (Epub 1999/10/16).
 63. Streppel MM, Vincent A, Mukherjee R, Campbell NR, Chen SH, Konstantopoulos K, et al. Mucin 16 (cancer antigen 125) expression in human tissues and cell lines and correlation with clinical outcome in adenocarcinomas of the pancreas, esophagus, stomach, and colon. *Hum Pathol.* 2012;43(10):1755–63. <https://doi.org/10.1016/j.humpath.2012.01.005> (Epub 2012/05/01).
 64. Fernandez Moro C, Fernandez-Woodbridge A, Alistair D'souza M, Zhang Q, Bozoky B, Kandaswamy SV, et al. Immunohistochemical typing of adenocarcinomas of the pancreatobiliary system improves diagnosis and prognostic stratification. *PLoS ONE.* 2016;11(11):e0166067. <https://doi.org/10.1371/journal.pone.0166067> (Epub 2016/11/10).
 65. Chaube A, Tewari M, Singh U, Shukla HS. CA 125: a potential tumor marker for gallbladder cancer. *J Surg Oncol.* 2006;93(8):665–9 (Epub 2006/05/26).
 66. Roa I, Villaseca M, Araya J, Roa J, de Aretxabala X, Melo A, et al. p53 tumour suppressor gene protein expression in early and advanced gallbladder carcinoma. *Histopathology.* 1997;31(3):226–30 (Epub 1997/11/14).
 67. Wang SN, Chung SC, Tsai KB, Chai CY, Chang WT, Kuo KK, et al. Aberrant p53 expression and the development of gallbladder carcinoma and adenoma. *Kaohsiung J Med Sci.* 2006;22(2):53–9. [https://doi.org/10.1016/s1607-551x\(09\)70221-9](https://doi.org/10.1016/s1607-551x(09)70221-9) (Epub 2006/03/30).
 68. Hidalgo Grau LA, Badia JM, Salvador CA, Monso TS, Canaletta JF, Nogues JM, et al. Gallbladder carcinoma: the role of p53 protein overexpression and Ki-67 antigen expression as prognostic markers. *HPB.* 2004;6(3):174–80. <https://doi.org/10.1080/13651820410025110> (Epub 2008/03/12).
 69. Wistuba II, Gazdar AF, Roa I, Albores-Saavedra J. p53 protein overexpression in gallbladder carcinoma and its precursor lesions: an immunohistochemical study. *Hum Pathol.* 1996;27(4):360–5 (Epub 1996/04/01).
 70. Yemelyanova A, Vang R, Kshirsagar M, Lu D, Marks MA, Shih IM, et al. Immunohistochemical staining patterns of p53 can serve as a surrogate marker for TP53 mutations in ovarian carcinoma: an immunohistochemical and nucleotide sequencing analysis. *Mod Pathol.* 2011;24:1248. <https://doi.org/10.1038/modpathol.2011.85>.

71. Murthyak B, Hortobagyi T. Immunohistochemical correlates of TP53 somatic mutations in cancer. *Oncotarget*. 2016;7(40):64910–20. <https://doi.org/10.18632/oncotarget.11912> (Epub 2016/09/15).
72. Koga T, Hashimoto S, Sugio K, Yoshino I, Nakagawa K, Yonemitsu Y, et al. Heterogeneous distribution of P53 immunoreactivity in human lung adenocarcinoma correlates with MDM2 protein expression, rather than with P53 gene mutation. *Int J Cancer*. 2001;95(4):232–9. [https://doi.org/10.1002/1097-0215\(20010720\)95:4%3c232:aid-ijc1040%3e3.c0.co;2-5](https://doi.org/10.1002/1097-0215(20010720)95:4%3c232:aid-ijc1040%3e3.c0.co;2-5) (Epub 2001/06/16).
73. Wang YC, Lin RK, Tan YH, Chen JT, Chen CY, Wang YC. Wild-type p53 overexpression and its correlation with MDM2 and p14ARF alterations: an alternative pathway to non-small-cell lung cancer. *J Clin Oncol*. 2005;23(1):154–64. <https://doi.org/10.1200/jco.2005.03.139> (Epub 2004/12/31).
74. Kalekou H, Miliaras D. Immunohistochemical study of microvessel density, CD44 (standard form), p53 protein and c-erbB2 in gallbladder carcinoma. *J Gastroenterol Hepatol*. 2004;19(7):812–8. <https://doi.org/10.1111/j.1440-1746.2004.03357.x> (Epub 2004/06/24).
75. Nakazawa K, Dobashi Y, Suzuki S, Fujii H, Takeda Y, Ooi A. Amplification and overexpression of c-erbB-2, epidermal growth factor receptor, and c-met in biliary tract cancers. *J Pathol*. 2005;206(3):356–65. <https://doi.org/10.1002/path.1779> (Epub 2005/05/14).
76. Kawamoto T, Krishnamurthy S, Tarco E, Trivedi S, Wistuba II, Li D, et al. HER receptor family: novel candidate for targeted therapy for gallbladder and extrahepatic bile duct cancer. *Gastrointest Cancer Res*. 2007;1(6):221–7 (Epub 2007/11/01).
77. Harder J, Waiz O, Otto F, Geissler M, Olschewski M, Weinhold B, et al. EGFR and HER2 expression in advanced biliary tract cancer. *World J Gastroenterol*. 2009;15(36):4511–7 (Epub 2009/09/25).
78. Roa I, de Toro G, Schalper K, de Aretxabala X, Churi C, Javle M. Overexpression of the HER2/neu gene: a new therapeutic possibility for patients with advanced gallbladder cancer. *Gastrointest Cancer Res*. 2014;7(2):42–8 (Epub 2014/05/07).
79. Li M, Zhang Z, Li X, Ye J, Wu X, Tan Z, et al. Whole-exome and targeted gene sequencing of gallbladder carcinoma identifies recurrent mutations in the ErbB pathway. *Nat Genet*. 2014;46(8):872–6. <https://doi.org/10.1038/ng.3030> (Epub 2014/07/07).
80. Kiguchi K, Carbajal S, Chan K, Beltran L, Ruffino L, Shen J, et al. Constitutive expression of ErbB-2 in gallbladder epithelium results in development of adenocarcinoma. *Cancer Res*. 2001;61(19):6971–6 (Epub 2001/10/05).
81. Wu Q, Kiguchi K, Kawamoto T, Ajiki T, Traag J, Carbajal S, et al. Therapeutic effect of rapamycin on gallbladder cancer in a transgenic mouse model. *Cancer Res*. 2007;67(8):3794–800. <https://doi.org/10.1158/0008-5472.can-06-3214> (Epub 2007/04/19).
82. Kim MA, Lee HJ, Yang HK, Bang YJ, Kim WH. Heterogeneous amplification of ERBB2 in primary lesions is responsible for the discordant ERBB2 status of primary and metastatic lesions in gastric carcinoma. *Histopathology*. 2011;59(5):822–31. <https://doi.org/10.1111/j.1365-2559.2011.04012.x> (Epub 2011/11/19).
83. Seol H, Lee HJ, Choi Y, Lee HE, Kim YJ, Kim JH, et al. Intratumoral heterogeneity of HER2 gene amplification in breast cancer: its clinicopathological significance. *Mod Pathol*. 2012;25(7):938–48. <https://doi.org/10.1038/modpathol.2012.36> (Epub 2012/03/06).
84. Yoon HH, Shi Q, Sukov WR, Lewis MA, Sattler CA, Wiktor AE, et al. Adverse prognostic impact of intratumor heterogeneous HER2 gene amplification in patients with esophageal adenocarcinoma. *J Clin Oncol*. 2012;30(32):3932–8. <https://doi.org/10.1200/jco.2012.43.1890> (Epub 2012/09/19).
85. Cho EY, Park K, Do I, Cho J, Kim J, Lee J, et al. Heterogeneity of ERBB2 in gastric carcinomas: a study of tissue microarray and matched primary and metastatic carcinomas. *Mod Pathol*. 2013;26(5):677–84. <https://doi.org/10.1038/modpathol.2012.205> (Epub 2012/12/15).
86. Kurozumi S, Padilla M, Kurozumi M, Matsumoto H, Inoue K, Horiguchi J, et al. HER2 intratumoral heterogeneity analyses by concurrent HER2 gene and protein assessment for the prognosis of HER2 negative invasive breast cancer patients. *Breast Cancer Res Treat*. 2016;158(1):99–111. <https://doi.org/10.1007/s10549-016-3856-2> (Epub 2016/06/20).
87. Hou Y, Nitta H, Wei L, Banks PM, Portier B, Parwani AV, et al. HER2 intratumoral heterogeneity is independently associated with incomplete response to anti-HER2 neoadjuvant chemotherapy in HER2-positive breast carcinoma. *Breast Cancer Res Treat*. 2017;166(2):447–57. <https://doi.org/10.1007/s10549-017-4453-8> (Epub 2017/08/12).
88. Rye IH, Trinh A, Saetersdal AB, Nebdal D, Lingjaerde OC, Almendro V, et al. Intratumor heterogeneity defines treatment-resistant HER2+ breast tumors. *Mol Oncol*. 2018;12(11):1838–55. <https://doi.org/10.1002/1878-0261.12375> (Epub 2018/08/23).
89. Ruschoff J, Dietel M, Baretton G, Arbogast S, Walch A, Monges G, et al. HER2 diagnostics in gastric cancer—guideline validation and development of standardized immunohistochemical testing. *Virchows Arch*. 2010;457(3):299–307. <https://doi.org/10.1007/s00428-010-0952-2> (Epub 2010/07/29).
90. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27(6):863–4. <https://doi.org/10.1093/bioinformatics/btr026> (Epub 2011/02/01).
91. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *%J EMBnet journal*. 2011;17(1):3. <https://doi.org/10.14806/ej.17.1.200> (Epub 2011-08-02).
92. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21. <https://doi.org/10.1093/bioinformatics/bts635> (Epub 2012/10/30).
93. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–303. <https://doi.org/10.1101/gr.107524.110> (Epub 2010/07/21).
94. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164. <https://doi.org/10.1093/nar/gkq603> (Epub 2010/07/06).
95. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923–30. <https://doi.org/10.1093/bioinformatics/btt656> (Epub 2013/11/15).
96. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68–74. <https://doi.org/10.1038/nature15393> (Epub 2015/10/04).
97. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655–64. <https://doi.org/10.1101/gr.094052.109> (Epub 2009/08/04).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.