Original Research

Proteomic Study on Multi-Organ Metastases of Human Ovarian Clear Cell Carcinoma Cell Line in a Xenograft Mouse Model Based on a Novel Sequence-Specific Analysis Strategy

Nan Ye[1,2], Yijin Li[1], Xin Li[1,2], Tao Su[3], Sisheng Wang[3], Wen Zheng[3], Hao Yang[2], Jingqiu Cheng[1,2,*], Meng Gong[1,*]

[1]Laboratory of Clinical Proteomics and Metabolomics, Institutes for Systems Genetics, Frontiers Science Center for Disease-related Molecular Network, National Clinical Research Center for Geriatrics, West China Hospital, Sichuan University, 610041 Chengdu, Sichuan, China
[2]Key Laboratory of Transplant Engineering and Immunology of the National Health Commission, West China Hospital, Sichuan University, 610041 Chengdu, Sichuan, China
[3]Metabolomics and Proteomics Technology Platform, West China Hospital, Sichuan University, 610041 Chengdu, Sichuan, China
*Correspondence: jqcheng@scu.edu.cn (Jingqiu Cheng); gongmeng@scu.edu.cn (Meng Gong)
Academic Editor: Dario Rusciano
Submitted: 16 August 2022   Revised: 1 November 2022   Accepted: 9 November 2022   Published: 16 March 2023

Abstract

Background: To investigate the gene regulation of tumor cells in the process of different organ metastasis on a xenograft mouse model and screen the genes involved in the organ-target metastasis of tumor cells. **Methods**: A multi-organ metastasis model was constructed with a human ovarian clear cell carcinoma cell line (ES-2) based on a severe immunodeficiency mouse strain (NCG). Differentially expressed tumor proteins among multi-organ metastases were successfully characterized by microliter liquid chromatography-high-resolution mass spectrometry, sequence-specific data analysis and multivariate statistical data analysis. Liver metastases were selected as typical for subsequent bioinformatic analysis. Selected liver metastasis-specific genes in ES-2 cells were validated by sequence-specific quantitation including high resolution-multiple reaction monitoring quantification at protein level and quantitative real-time polymerase chain reaction at mRNA level. **Results**: From the mass spectrometry data, a total of 4503 human proteins were identified using the sequence-specific data analysis strategy. Of them, 158 proteins were selected as specifically regulated genes in liver metastases for subsequent bioinformatics studies. Based on Ingenuity Pathway Analysis (IPA) pathway analysis and sequence-specific quantitation, Ferritin light chain (FTL), lactate dehydrogenase A (LDHA) and long-chain-fatty-acid–CoA ligase 1 (ACSL1) were finally validated as specifically upregulated proteins in liver metastases. **Conclusions**: Our work provides a new approach to analyze gene regulation in tumor metastasis in xenograft mouse model. In presence of a large number of mouse protein interference, we validated the up-regulation of human ACSL1, FTL and LDHA in ES-2 liver metastases, which reflects the adaptive regulation of tumor cells to the liver microenvironment through metabolic reprogramming.

Keywords: tumor organ-specific metastasis; severe immunodeficiency mouse; ES-2 cell line; data independed acquisition; sequence-specific data analysis

1. Introduction

Metastasis is a process in which tumor cells disperse from the primary site. The tumor cell disseminates to distant areas through blood vessels, lymph vessels or direct invasion [1,2]. In other tissues or organs, secondary tumors form with a histological type similar to that of the primary tumor [3]. Early-stage tumor metastasis-related studies mostly focus on a series of molecular events occurring in the early and middle stages of metastasis from tumor growth in situ to the transport survival stage of the vasculature [4]. However, the mechanism of extravasation to select a secondary organ and colonize distant sites during the late steps is still poorly understood [5]. The earliest understanding of the tendency of organs with tumor metastasis can be traced back to the "seed-soil" doctrine proposed by surgeon Stephen Paget in 1889 [6]. It was called the three landmark theories in tumor metastasis research along with Ewing's

"theory of anatomical mechanisms" [7] and the "metastatic waterfall theory" proposed by Bross and Blumenson [8]. In recent years, it has been gradually recognized that the interaction between the tumor microenvironment and tumor cells through bidirectional signaling regulation is an important cause and key rate-limiting step for the occurrence of molecular events in the later stage of tumor metastasis and the final formation of metastatic lesions in specific organs [9].

There is evidence that single or clusters of cells can spread from the primary tumor and that after extravasation to the vasculature, successful metastatic cells must be selected and survive the new tissue microenvironment conducive to survival [3]. Thus, metastatic colonization is not only a product of primary tumor dissemination but also through a complex interaction between the disseminated tumor cells and the tissue microenvironment of the whole or-
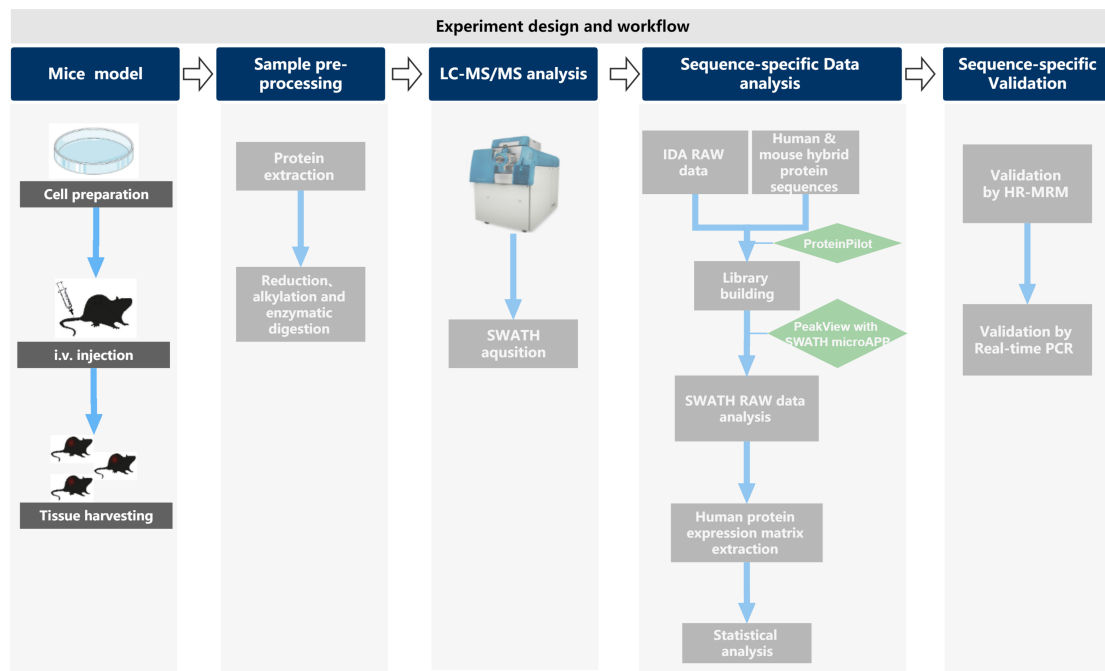
Fig. 1. Experiment design and workflow.

ganism [10]. In this process, tumor cells exist in a continuous phenotypic state, triggering their intrinsic programs through their interaction with the extracellular microenvironment, conferring their metastatic properties through epigenetic regulation, and overcoming various obstacles to eventually form metastatic lesions. It has been reported that metastatic lesions in different tumor types are often detected in different specific organs (liver, lung, bone and brain) [11]. However, little is known about the tropism of different organs in cancer subtypes. Understanding the complex interplay between the complex molecular-level behavior of tumor cells and the tumor microenvironment leading to metastasis requires the integration of massive molecular signature data collected from *in vitro* and *in vivo* experimental models, processing high-throughput data using different computational methods, and integrating information to characterize the metastatic process of tumor cells. However, to date, our understanding of this dynamic complex process remains incomplete.

In recent years, quantitative proteomics has become a powerful tool to study the changes in the protein profiles in cancer metastasis [12]. Mass spectrometry-based proteomics technology is a large-scale, high-throughput proteomics technique that enables the accurate analysis of complex protein mixtures [13]. This technique enables the scale-up search of multiple sets of protein or protein populations of interest and illustrates the complex dynamics of proteins in cells at the overall level through both qualitative and quantitative analysis of proteins [14]. Mass spectrometry, by combining proteomics and bioinformatics technologies, allows for a quantitative and precise systematic anal-

ysis of thousands of proteins, making it possible to analyze and characterize the changes in the protein spectrum involved in the metastatic colonization process of tumor cells [15].

In this research, to explore the phenomenon of gene regulation of tumor cells during metastatic colonization in different organs, a novel research strategy was established based on a sequence-specific proteomics analysis and a mouse model of multiple organ metastasis of human tumor cells. Herein, the ES-2 cell line, an ovarian clear cell carcinoma (OCCC) cell line, was applied to severely immunodeficient mice (NOD/ShiltJGpt-Prkdcem26Cd52 Il2rgem26Cd22/Gpt, NCG strain) to form multiple organ metastases, and an analytical method based on microliter liquid chromatography tandem high resolution mass spectrometry and a sequence-specific data analysis strategy was used to study the differences in protein expression among various organ metastases (Fig. 1) [16]. Through this strategy, we screened out a series of differential proteins for the organ tropism of ES-2 cells, and specific regulatory proteins in liver metastasis were presented as typical in this study. Our results indicated that this sequence-specific proteomics analysis strategy can provide a powerful tool to explore the gene regulation of xenografts or hosts from proteomes by avoiding quantitative interference of homologous proteins. To the best of our knowledge, this is the first study of multiple organ metastasis from one tumor cell line by sequence-specific proteomics analysis. It is of great significance for us to comprehensively understand the molecular mechanism of organ-specific metastasis of tumors.

2. Materials and Methods

2.1 Cell Preparation

The ES-2 human ovarian clear cell carcinoma cell line and HEK293T cell line were purchased from the American Culture Collection (ATCC) and cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco, Grand Island, NY, USA) containing 10% Fetal Bovine Serum (FBS, Hy-Clone, Logan, UT, USA) and 1% penicillin/streptomycin (P/S) (Gibco, Grand Island, NY, USA) at 37 °C with 5% CO_2. Lentiviral packaging plasmids (Pspax2, PMD2G and ZsGREEN1 GFP, Thermo Fisher, Waltham, MA, USA) were transfected into HEK293T cells according to the manufacturer's instructions to generate lentivirus, and a stable ES-2 cell line expressing GFP (ES-2-GFP) was obtained by monoclonal screening in 96-well plates after lentiviral transduction.

2.2 Construction of a Mouse Multiorgan Metastasis Model

NCG (NOD/ShiltJGpt-Prkdcem26Cd52Il2rgem26-Cd22/Gpt) strain mice, the severe immunodeficient strain mice lacking mature T cells, B cells, and NK cells used in the experiment (female, SPF grade, >8 weeks old) were purchased from JiCui Pharmaceutical Co., Ltd. All NCG mice were housed in a specific pathogen-free (SPF) animal room with free access to food and water. All animal experiments were performed and approved by the Animal Care and Use Committee of West China Hospital, Sichuan University. Exponentially growing ES-2-GFP cells were trypsinized to prepare a cell suspension in PBS at a concentration of 1×10^7 cells/mL, and the multiorgan metastatic model was induced in NCG mice by i.v. injection of 1×10^6 ES-2-GFP cells suspension per mouse.

2.3 Tissue Harvesting and Fluorescence Imaging

The mice were killed and left open by cervical dislocation between 18 and 20 days postinoculation. The distribution of metastatic lesions in each organ of 12 mice was observed and then validated by fluorescence imaging. The static dissection of each organ was loaded into a frozen storage tube and frozen in liquid nitrogen.

2.4 Sample Preprocessing

ES-2 cells were washed twice with cold Phosphate buffered saline (PBS) and lysed in 1000 μL SDT lysis buffer containing 2% sodium dodecyl sulfate, 100 mM Tris/HCl pH 7.4, 100 mM dithiothreitol (Sigma–Aldrich, St. Louis, MO, USA) and 1 × protease cocktail inhibitors (Roche Diagnostics, Sandhofer Strasse, Mannheim, Germany). After lysis, the mixture was transferred to a 1.5 mL centrifuge tube using cell scraping and sonicated on ice for one minute (5/5 s, 6 cycles, output 60 W, frequency 45 ±, SCIENTZ-IID, Ningbo, China). The supernatants were collected after centrifugation at 14,000 g and 4 °C for 10 min. Approximately 100 mg of tissue was extracted from each metastasis sample, homogenized in 800 μL SDT lysis buffer, sonicated

and centrifuged to collect supernatants under the same conditions. The protein concentration of each sample was determined by the tryptophan fluorescence emission method at 365 nm. Proteins in each sample (150 μg) were precipitated with cold acetone for 3 h at –20 °C. After centrifugation of the precipitated protein, the liquid layer was removed, and the cabinet was dried for 5 min. The protein precipitate was dissolved by the addition of 7 M guanidine hydrochloride and then incubated with 45 mM dithiothreitol for 30 min at 37 °C for reduction and alkylation and 106 mM iodoacetamide for 30 min at 25 °C in the dark. Then, the samples were transferred to a Vivacon 500 30 kDa Molecular weight cutoff (MWCO) centrifugation device (Sartorius, Gottingen, Germany), washed twice with an additional 100 μL 8 M urea solution and sequentially washed twice with 100 μL 50 mM ammonium bicarbonate. The samples were first digested with sequencing-grade trypsin (Promega, Madison, WI, USA) at a trypsin to protein ratio of 1:50 for 16 hours at 37 °C. After digestion, the peptides were dried in a vacuum and stored at –80 °C until use. All water used in the experiments was purified from a Milli-Q system (Millipore, Bedford, MA, USA), and other highest grade reagents and chemicals were purchased from Sigma–Aldrich or Thermo Fisher Scientific.

2.5 LC–MS/MS Analysis Based on the SWATH Acquisition Method

Sequential window acquisition of all theoretical mass spectra (SWATH-MS) analysis was performed on a quadrupole time-of-flight spectrometer (TripleTOF5600, AB/Sciex, Framingham, MA, USA) coupled to a micro-HPLC (NanoLC 415, Eksigent, Framingham, MA, USA). The LC flow phase consisted of mobile phase A (2% acetonitrile and 0.1% FA in ultrapure water) and mobile phase B (98% acetonitrile and 0.1% FA in ultrapure water) (all purchased from Thermo Fisher Scientific, Waltham, MA, USA). Samples were first captured using a ChromXP c18 precolumn (5 μm particle size, 120 Å, 75 μm id × 5 mm). Impurities and salt were removed for 6 min at a flow rate of 100% buffer a for 6 L/min. The peptides were then eluted onto a ChromXP c18 analysis column (3 μm particle size, 120 Å, 75 μm id × 15 cm). Mobile phases A and B were used to create the separation gradient of the eluted peptides in the column by increasing the solvent B concentration. A flow rate of 5 L/min was maintained throughout the flow phase, and the following gradient elution procedure was run: 5% B for 0.5 min; 5% to 25% B for 45 min; 25% to 35% B for 10 min; 35% to 55% B for 5 min; 55% to 80% B for 0.5 min; 80% B for 4.5 min; 80% to 5% B for 0.5 min; 5% B for 4.5 min. The key ion source parameters for data acquisition in the SWATH-MS mode are as follows: ion spray voltage float (ISVF) at 5500 V, curtain gas (CUR) at 30 psi, ion source gas 1 (GS1) at 17 psi, ion source gas 2 (GS2) at 13 psi, temperature (TEM) at 330 °C, and decluster potential (DP) at 100 V. The mass spec-

trometer was operated in cycle product ion mode. Full-MS scans were performed in the mass range of 350–1500 m/z in positive-ion mode. One hundred MS/MS acquisition windows with an overlap range of 1 m/z were constructed using a variable window mode. The coverage mass range was 100–1500 m/z in high sensitivity mode with a CES of 15 V. The maximum fill time for the MS scan was 250 ms and 30 ms for the MS/MS scan, with a final cycle time of 3.04 s. Peptide samples were prepared as 0.5 g/L, mixed with 12.5 fmol/L peptides from trypsin-digested bovine serum albumin (BSA), and analyzed for 5 L for each sample.

2.6 LC–MS/MS Validation Based on the HR-MRM Acquisition Method

For HR-MRM experiments, LC–MS/MS was performed on a quadrupole time-of-flight spectrometer (TripleTOF5600, AB/Sciex, Framingham, MA, USA) coupled to a micro-HPLC (NanoLC 415, Eksigent, Framingham, MA, USA). The LC flow phase consisted of mobile phase A (2% acetonitrile and 0.1% FA in ultrapure water) and mobile phase B (0.1% FA and 2% ultrapure water in acetonitrile) (all purchased from Thermo Fisher Scientific, Waltham, MA, USA). The sample was first loaded onto a ChromXP c18 precolumn (5 μm particle size, 120 Å, 75 μm id × 5 mm). Impurities and salts were removed at a flow rate of 6 μL/min in 100% buffer A for 6 minutes. The peptides were then eluted onto a ChromXP c18 analysis column (3 μm particle size, 120 Å, 75 μm id × 15 cm). Mobile phases A and B were used to create the separation gradient for eluting peptides in the column at an increasing concentration of solvent B. A flow rate of 5 μL/min was used throughout the flow phase, and the following gradient elution procedure was run: 5% to 8% B for 0.5 min; 8% to 30% B for 45 min; 30% to 50% B for 5 min; 50% to 80% B for 0.5 min; 80% B for 4.5 min; 80% to 5% B for 0.5 min; 5% B for 4.5 min. The key ion source parameters for data acquisition in the HR-MRM mode were as follows: ion spray voltage float (ISVF) at 5500 V, curtain gas (CUR) at 30 psi, ion source gas 1 (GS1) at 18 psi, ion source gas 2 (GS2) at 15 psi, temperature (TEM) at 350 °C, and cluster potential (DP) at 100 V. Individual candidate peptide TOF-MS scans were set at their m/z value followed by a MS/MS product ion scan from 100 to 1500 Da. The final cycle time was 2.16 s. Data acquisition was performed using Analyst software (version 1.6.3, Sciex-Foster, CA, USA).

2.7 Data Processing

Mass spectrometry acquisition data were processed using Protein-Pilot software (AB/Sciex, v.5.0.1, Framingham, MA, USA). The nonredundant human UniProtKB/Swiss-Prot protein database (https://www.uniprot.org/uniprotkb?facets=reviewed%3Atrue&query=proteome%3AUP000005640), BSA peptide sequence and mouse UniProtKB/Swiss-Prot protein database (https://www.uniprot.org/uniprotkb?facets=reviewed%3Atrue&query=proteome%3AUP000000589) were combined to form a heterozygous library (https://www.uniprot.org/uniprotkb/P02769/entry) for IDA data alignment to identify proteins. Search parameters were set as follows: sample type: identification; Cys alkylation: iodoacetamide; digestion: trypsin; instrument: TripleTOF 5600; special factors: none; species: Homo sapiens; search effort: thorough ID; results quality: 0.05 (detected protein threshold) performed Paragon searches. The false discovery rate (FDR) for proteins and peptides was set to 0.01 (minimum length set to seven amino acids) for analysis. Data from all IDA method identification runs performed using ProteinPilot software were merged into one batch and used for SWATH assay library construction. After database retrieval, the mouse polypeptides were removed by using the self-compiled R language software package (R 3.6.0 programming language, R Foundation for Statistical Computing, Vienna, Austria) to establish a human protein quantification database. Protein quantification information was extracted using PeakView software 2.2 and the swath plugin, where the BSA peptides were used to calibrate the retention time calibration. To obtain better quantitative results using the SWATH acquisition method, the following parameters were set: peptide filter: 15 peptides per protein; six transitions per peptide; 95% peptide confidence; 1% FDR threshold; exclusion of modified peptides; XIC option: 8 min XIC extraction window; 50 ppm XIC width. Missing values in the results are automatically populated by the PeakView software (AB/Sciex, version 2.2, Framingham, MA, USA) built-in SWAP plugin. Subsequent statistical analysis of the data was performed using R statistical software, version 3.6.0. The median method was chosen to normalize the data, and the p value of protein expression was calculated with the Kruskal–Wallis test. A Benjamini–Hochberg adjustment p value < 0.05 was used to screen for differentially expressed proteins (DEPs).

Bioinformatics analysis was performed using the Gene Ontology (GO) database (http://geneontology.org/) and Kyoto Encyclopedia of Genes and Genome (KEGG) (http://www.genome.jp/kegg/pathway.html) based on the DEPs. In addition, Ingenuity Pathway Analysis (IPA) (http://www.ingenuity.com/products/ipa) was used to determine the canonical pathways among DEPs. Fisher's exact test was used to calculate the p values.

2.8 Human-Specific Primer Synthesis

Considering the inevitable mixing of human tumor cells with murine organ tissue during tumor sampling, we designed primers for differential bases in human and mouse mRNAs, which is used to solve the problem that the human metastasis samples cannot be completely separated from the tissues of NCG mice due to surgical operation during the RT–PCR experiments. Primer synthesis was commissioned
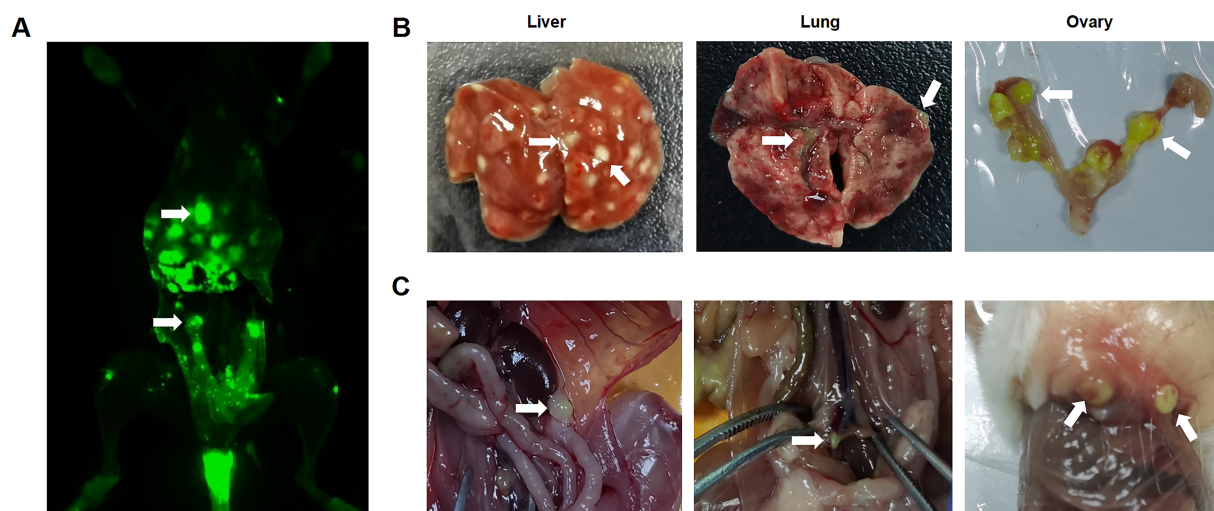
Fig. 2. Tissue harvesting and fluorescence imaging of ES-2 cell line. (A) Fluorescent imaging of the metastasis formed by ES-2 cells. (B) *In vivo* optical imaging of liver, lung, and ovary metastases formed by ES-2 cells. (C) *In vivo* optical imaging of connective tissue metastases formed by ES-2 cells. The white arrowhead indicates metastases. The connective tissues in different organs are collectively referred to as Other.

by Beijing Qingke Biological Company. The PCR primer sequences are shown in **Supplementary Table 1**.

2.9 Real-Time Polymerase Chain Reaction (Real-Time PCR)

Total RNA from ES-2 cells and metastatic foci of various organs was extracted using TRIzol reagent (Gibco, Life Technologies, Carlsbad, CA, USA) and reverse-transcribed into cDNA using a RevertAid™ First Strand cDNA Synthesis Kit with a DNase I Reverse Transcription Kit (ThermoFisher, USA). PCR amplification was performed using the TransStart Taq DNA Polymerase PCR Kit (TransGen Biotech, Inc., Beijing, China). Fluorescence quantitative PCR experiments were performed using a CFX96 dual-channel real-time PCR instrument (Bio-Rad, Hercules, CA, USA) with $2\times$ TSINGKE® Master qPCR Mix (SYBR Green I) (TSINGKE, Beijing, China). The cycling conditions were as follows: 50 °C for 2 min, followed by 35 cycles of 95 °C for 2 min, 95 °C for 15 s, and 60 °C for 30 s. Data were analyzed by Bio-Rad CFX Manager software. The relative expression levels of the target genes versus the reference gene (GAPDH) were calculated using the $2^{-\triangle\triangle CT}$ method. All experiments were performed in triplicate, providing three technical repeats.

2.10 Statistical Analysis

Quantitative data are represented as the mean ± standard deviation. Data analysis was performed using the one-way ANOVA method in GraphPad Prism 8.0.2 software (GraphPad Software Inc., San Diego, CA, USA). A value of $p < 0.05$ was considered statistically significant.

3. Results

3.1 Establishment of the ES-2 (Human Ovarian Clear Cell Carcinoma Cell Line) NCG Mouse Multiorgan Metastasis Model

To facilitate observation and verification of metastases, ES-2 cells were labeled using lentiviral transfection with green fluorescent protein (GFP), and distinctly fluorescent green was observed under fluorescence microscopy. At 18 to 22 weeks after cell injection, mice were sacrificed, and multiorgan metastases of ES-2 cells, such as in the liver, lung, ovary and peritoneum, were clearly observed in NCG mice. In addition, fluorescence imaging proved that the number of metastases formed by the ES-2 cell lines labeled with GFP varied significantly in different organs (Fig. 2A–C).

3.2 Proteomics Results and Data Analysis

From the data acquired by the SWATH strategy, a total of 6023 proteins were identified from 32,602 different unique peptides by matching to the heterozygous library containing human and mouse protein sequences and bovine serum albumin peptide sequences. After excluding 17,809 mouse source peptides, the number of remaining human peptides was 14,793, from which a total of 4503 human proteins were identified. Ultimately, a human protein quantitative database was established. *p* values were calculated using the Kruskal–Wallis test, which were corrected using the Benjamini–Hochberg strategy with an FDR <0.1 (*p*. adjusted), and a total of 2868 proteins with a *p* value < 0.05 after correction were screened.

To understand the gene regulation and protein profile expression changes from circulating ES-2 cells in blood
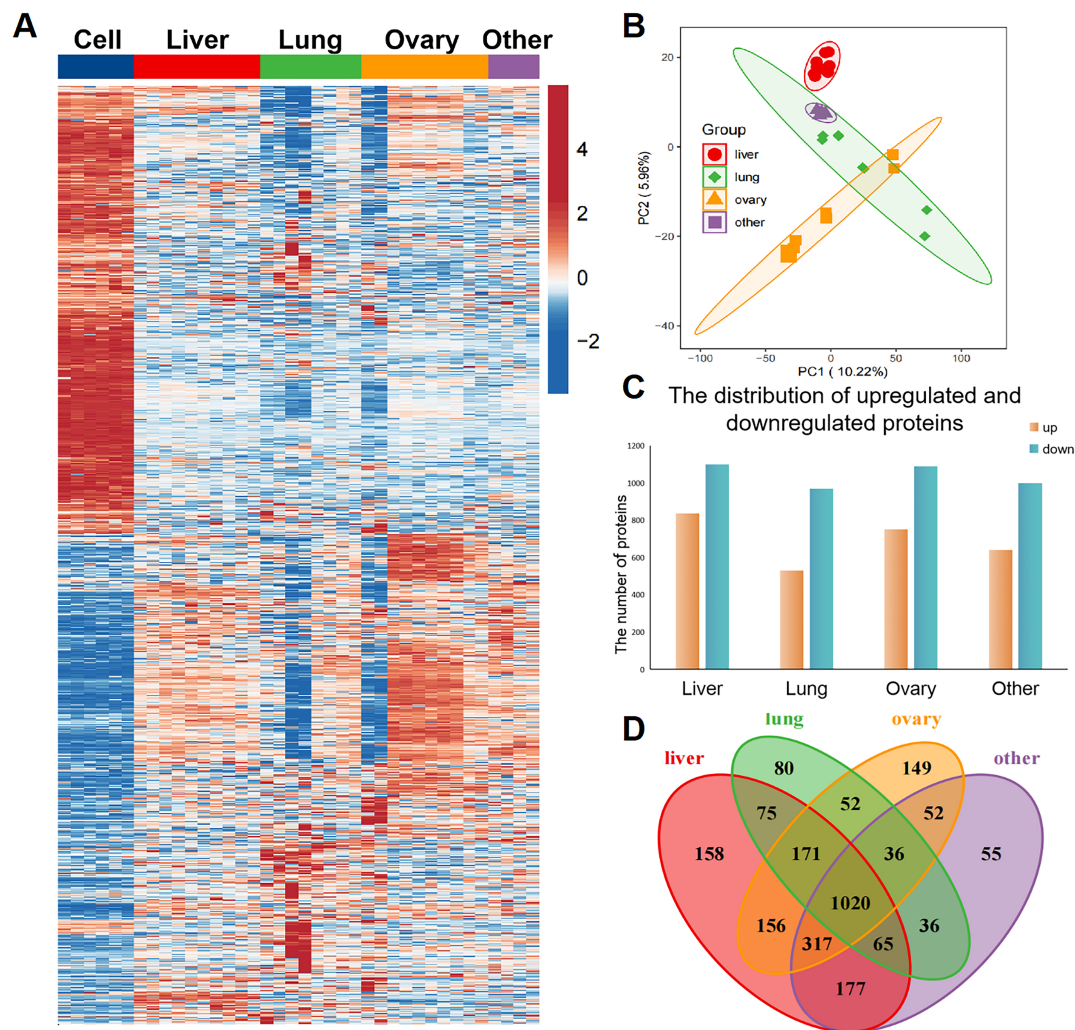
Fig. 3. Multivariate statistical analysis of the total protein expressed by different organs metastasis and ES-2 cell lines. (A) Hierarchical cluster analysis of the 2868 DEPs. Heatmap showed that clustering of differential proteins between ES-2 cells and different organ metastases. (B) Principal Component Analysis of the remanent DEPs from metastases in different organs after removing the proteins expressed by the cells. PCA analysis showed that there was a significant outlier trend of proteins expressed by liver metastases. The p-value was calculated using the Kruskal-Wallis test. A Benjamini-Hochberg adjustment p-value < 0.05 (FDR <0.1) was used to screen for DEP. (C) Distribution of different protein quantities expressed by different organ metastases. (D) According to the screening criteria of FDR <0.02 and fold change (FC) >1.5, Veen plot shows the intersectional relationship of proteins expressed by different organ metastases.

vessels to different organs in colonization and metastasis, we first performed a hierarchical clustering analysis (HCA) on protein data from ES-2 cell lines and the metastases of different organs. From the heatmap (Fig. 3A), the tumors metastasized to organs showed significant differences in protein expression profiles compared with the ES-2 cell lines. Moreover, various profiles were also clearly observed among the metastases from different organs. Subsequently, a principal component analysis (PCA) of the protein expression data of different metastasis groups was carried out, and the results showed obvious distinctions in the

distribution of these metastatic samples, each with significant organ-dependent separation. Interestingly, the protein spectrum of liver metastases showed the most pronounced deviation compared with metastases from other organs (Fig. 3B). Next, we performed statistical analysis of differential proteins between the ES-2 cell line and each organ metastatic tumor group separately. According to the screening criteria of FDR <0.02 and fold change (FC) >1.5 (volcano plots are shown in **Supplementary Fig. 1**), compared with the original ES-2 cell line, we obtained 1938 differential proteins in the liver metastasis group, 1499 differ-

Table 1. High Resolution Multi-Response Identifies 6 Proteins Upregulated in Liver Metastases.

Entry name	Protein names	Gene names
FRIL_HUMAN	Ferritin light chain	FTL
LDHA_HUMAN	Lactate Dehydrogenase A	LDHA, PIG19
IF2A_HUMAN	Eukaryotic translation initiation factor 2 subunit 1	EIF2S1
CPSM_HUMAN	Carbamoyl-phosphate synthase	CPS1
ACSL1_HUMAN	long-chain-fatty-acid–CoA ligase 1	ACSL1, FACL1, FACL2, LACS, LACS1, LACS2
GSH1_HUMAN	Glutamate–cysteine ligase catalytic subunit	GCLC, GLCL, GLCLC

ential proteins in the lung metastasis group, 1850 differential proteins in the ovary metastasis group and 1641 differential proteins in the peritoneum metastasis group (Fig. 3C). To study the commonness and organ specificity of gene expression of ES-2 cells in the formation of different organ metastases, we built a Venn diagram to show the collection of proteins expressed in different organ metastases, shown as Fig. 3D. Notably, there are 1020 proteins in the intersection of multiple organ metastases. We chose the 227 proteins uniformly upregulated in various organ metastasis for GO enrichment analysis and KEGG pathway analysis and results showed these proteins are mainly located in the cytoplasm and nucleus, as well as in exocytic exosomes, mainly involved in the intracellular protein transport and hydrolysis process, and their molecular functions focus on the cell adhesion with ATP binding and cadherin participation (**Supplementary Fig. 2A**). At the same time, these differential proteins were mainly enriched in metabolism as well as oxidative phosphorylation pathways, including the active glycolysis of tumor cells in order to adapt to the metabolic stress in the microenvironment, as well as the nucleotide metabolic pathways (**Supplementary Fig. 2B**). These results implied the proteins in the intersection of multiple organ metastases might be associated with general gene expression regulation in solid metastatic tumor formation from tumor cells. On the other side, Venn diagram also reflected the organ specificity of gene expression of ES-2 cells in different organ metastases: the mounts of proteins specifically regulated in metastases of different organs are separately 158 for liver, 80 for lung, 149 for ovary and 55 for other connective tissues, and these proteins may be involved in tumor colonization in corresponding target organs.

3.3 Liver Metastases Were Selected as Study Subjects for Differential Protein Screening

Since in all metastases, liver metastases showed the most significant differences in protein expression compared to original ES-2 cells and showed obvious deviation from other organ metastases in PCA scoring, we focused on the 158 proteins specifically expressed in liver metastases (Listed in **Supplementary Table 2**) as representatives in subsequent bioinformatical studies. We first performed GO analysis and KEGG pathway analysis to understand the protein expression changes observed in liver metastases. GO

is used to describe gene functions and is divided into three categories: biological processes, cellular components, and molecular functions. GO enrichment analysis showed that 158 proteins specifically expressed in the liver were mainly located in the cytosol and nucleus and were mainly involved in DNA transcription, and molecular function focused on ATP binding (Fig. 4A). Further KEGG pathway analysis of the above 158 proteins showed that these proteins were mainly enriched in metabolic pathways (Fig. 4B).

Next, we introduced these protein genes into the IPA for classical overlapping canonical pathway analysis and screened genes related to the liver and ovarian cancer background as well as tumor colonization and metastasis. Pathway analysis showed that the genes specifically expressed by liver metastases were mainly concentrated in the oxidative phosphorylation and glutathione biosynthetic pathways (4C). Based on IPA pathway analysis results and SWATH quantitative data, we screened 17 proteins that were upregulated in liver metastases with functional importance and sufficient MS response intensity for further verification by targeted quantification using HR-MRM acquisition. Six proteins (Listed in Table 1), including ferritin light chain (FTL), lactate dehydrogenase A (LDHA), eukaryotic translation initiation factor 2 subunit 1 (EIF2S1), carbamoyl-phosphate synthase (CPS1), long-chain-fatty-acid–CoA ligase 1 (ACSL1) and glutamate–cysteine ligase catalytic subunit (GCLC), which could be effectively quantified by HR-MRM acquisition and showed exact upregulation in liver metastases, were screened out (Fig. 5A), and this result further narrowed the range of target proteins.

3.4 Quantitative PCR Validation Based on Human-Specific Primers

To distinguish between human and mRNAs, we designed specific primers aimed at differential bases for distinguishing human- and murine-derived cDNA and performed PCR preexperiments. PCR experiments demonstrated that this series of primers are targeted to human-derived cDNA (Fig. 5B). To examine the reliability of the proteomic data in the study, six proteins that were completely upregulated in liver metastases resulting from the HR-MRM validation were selected for RT–PCR validation. The results showed that the three differential protein gene expression levels derived from the six candidate target proteins were consistent with the proteomic data, including
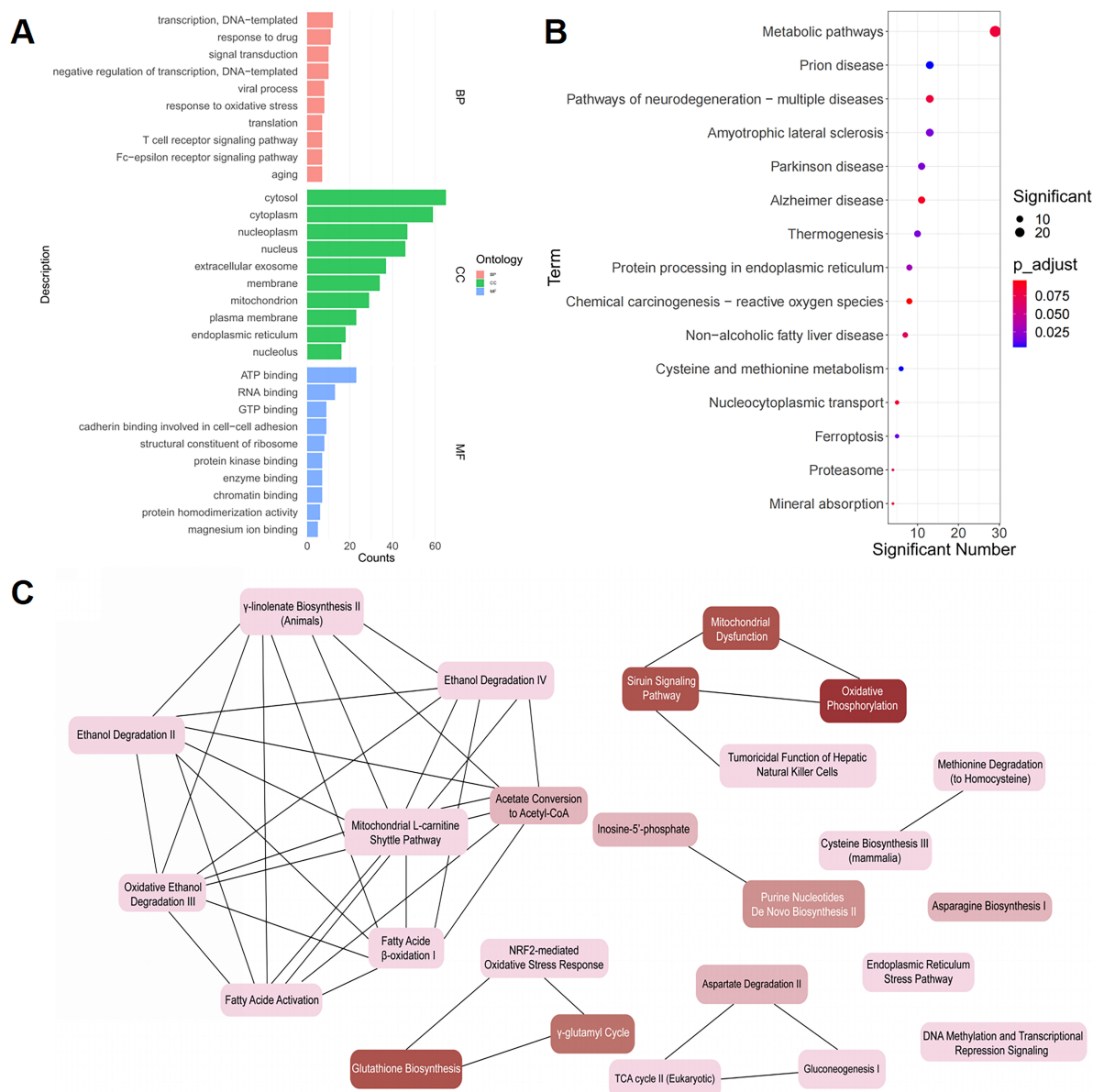
Fig. 4. Summary of the 158 proteins specifically expressed in liver metastases. (A) GO annotation of the 158 proteins specifically expressed in liver metastases. Only the top 10 words are shown according to the order of the *p*-adjusted. BP, biological process; CC, cellular component; MF, molecular function. (B) KEGG pathway enrichment analysis of 158 proteins specifically expressed by liver metastases. (C) Diagram of classical overlapping pathway analysis of 158 protein genes specifically expressed in liver metastases (Classical overlapping pathways share one or more genes by interconnection, where the deeper the red, the more significant the contribution).

ACSL1, LDHA, and FTL, which were significantly higher in liver metastases than in lung, ovary, and other metastases (Fig. 5C).

4. Discussion

Despite our deeper understanding of the process of tumor metastasis in recent years, distant metastasis remains the leading cause of cancer death [2]. Our understanding of the complex interaction of tumor metastasis leading to a series of molecular events between tumor cells and the microenvironment is still insufficient. This process per-

sists and is in dynamic changes, requiring the integration of large molecular data from *in vivo* and *in vitro* experimental models and the information therein to characterize the metastatic process of tumor cells. Here, we employed NCG mice, an immunodeficient mouse strain, to construct a human-derived tumor cell line xenograft model. NCG mice are a severely immunodeficient strain with defects in mature T cells, B cells, and NK cells, which makes NCG mice one of the best human-derived tumor cell line xenograft model animals to date [17]. In preliminary experiments, we found that ES-2 cell line can present very stable
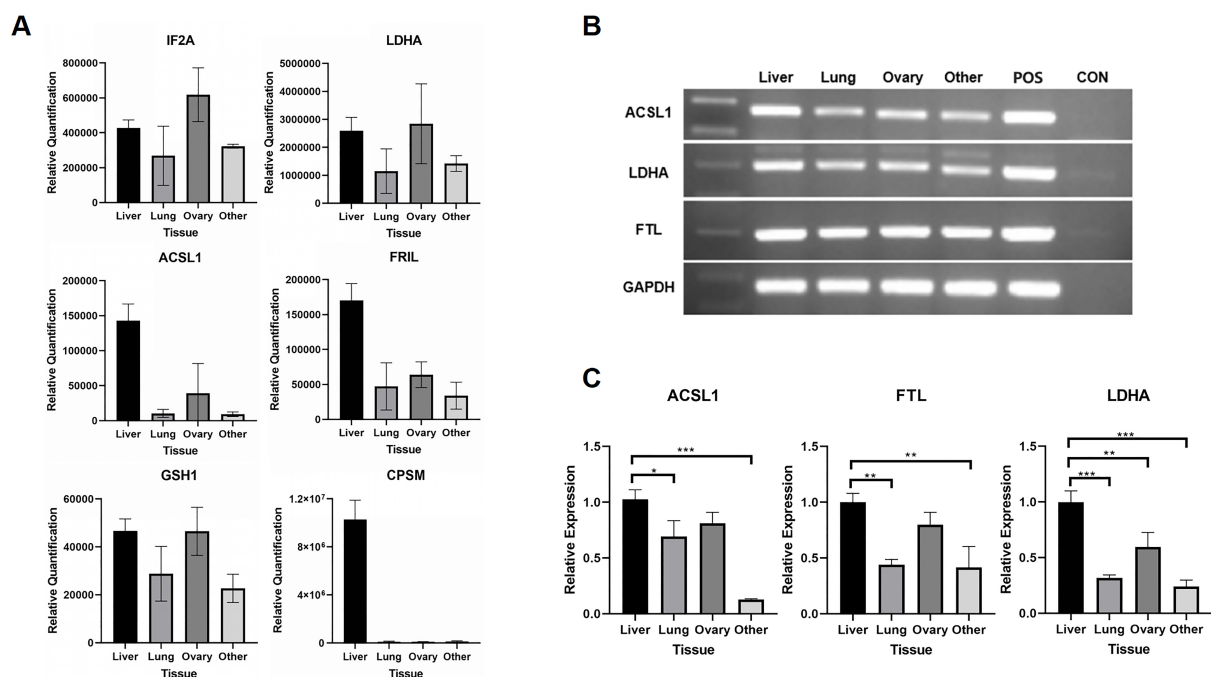
Fig. 5. RT-qPCR results. (A) Exact upregulation of six proteins quantified by HR-MRM acquisition. (B) PCR preliminary experiment of specific primer for target protein gene (ACSL1, LDHA and FTL), negative control for normal liver tissue of NCG mice. (C) Real-time PCR analysis of genes upregulated in liver metastases (ACSL1, LDHA, and FTL) expression in various organ metastases of NCG mice (n = 3). The positive control (POS) was the ES-2 cell line, and the negative control (CON) was the liver tissue of normal NCG mice, GAPDH was used as a reference gene. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

liver, lung, ovary, and subcutaneous multisite metastases in NCG mice, providing an excellent source of model samples for this study. At the same time, to solve the difficulty of separating human tumors from mouse tissues, we innovatively removed mouse peptides at the level of omics data processing and successfully constructed a protein quantitative database containing only human specificity. In the choice of methodology, we used a combination of untargeted and targeted protein quantification techniques to identify and verify the target proteins. This strategy integrates the advantages of targeted and nontargeted protein quantification techniques to make the final proteomics quantification results more realistic [18]. As a quantitative proteome technology based on a data-independent acquisition mode, SWATH has high throughput, high sensitivity and no bias and is more suitable for studies requiring the identification of large-scale unknown proteins [19]. To further validate the differentially expressed proteins obtained based on SWATH data screening, we chose HR-MRM quantification based on high-resolution mass spectrometry to validate the target protein quantified by SWATH acquisition mode [20,21]. It is worth mentioning that to detect the expression of human proteins specifically, quantitative strategies based on sequence specificity were used throughout this study. In SWATH analysis, we first formed a heterozygous library by combining human and mouse protein databases, and only unique peptides matched in this heterozygous li-

brary were chosen for protein quantification; thus, the coexisting peptides in human and mouse protein databases were excluded to minimize the interference between two species. This strategy made up for the disadvantage that the organ metastases obtained from the experimental operation cannot completely eliminate the mouse tissue and accurately distinguish the human and mouse proteins during the protein identification process, through which the subsequent data analysis could reflect the true regulation of protein expression in human tumor cells.

In this ES-2 metastasis model, compared with ES-2 cells cultured *in vitro*, the multi-organ metastatic tissues showed plenty of similar gene regulation, which was reflected by the 1020 proteins in the intersection of multiple organ metastases in Venn diagram. We analyzed 227 proteins that were consistently upregulated in various organ metastases and found a lot of proteins are involved in the cell adhesion with ATP binding and cadherin participation. For example, the upregulation of Integrin beta-4 (ITB4) and its downstream signaling molecules mitogen-activated protein kinase kinase 2 (MAPK2) confirmed that the Ras/Raf/MEK/ERK pathway associated with the tumor cell adhesion process is activated during the multiorgan metastasis of ovarian clear cell carcinoma [22]. It has been reported that the activation of the Ras/Raf/MEK/ERK pathway mediated by the upregulation of ITB4 is a switch for cancer cells to initiate a series of sequential metastatic steps,
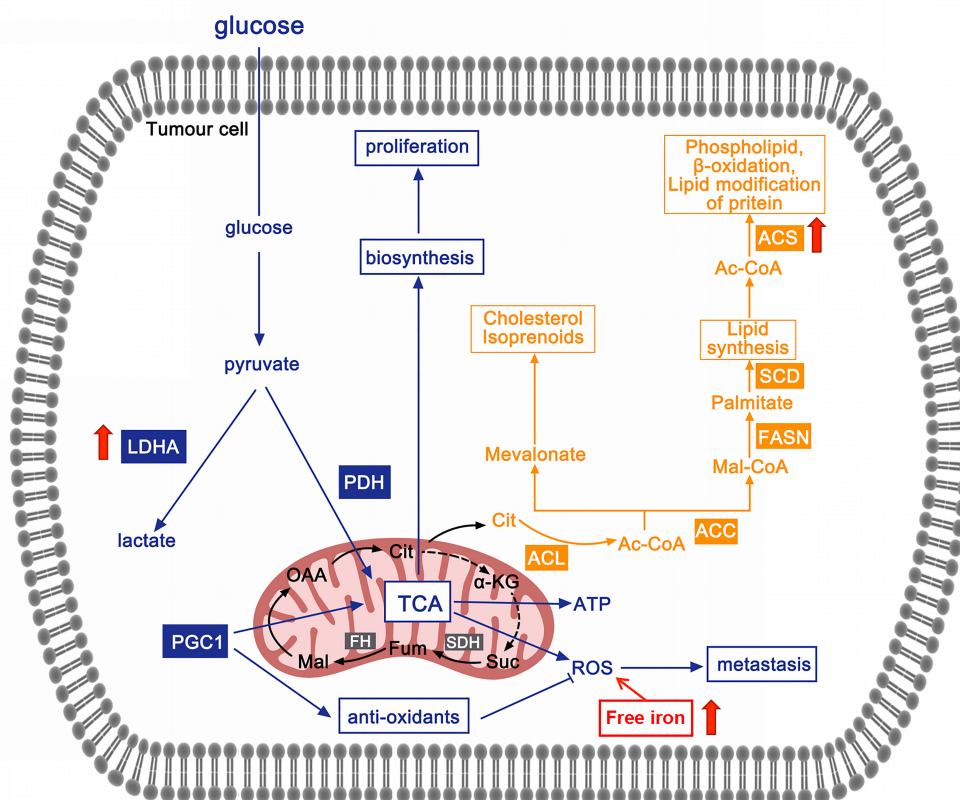
Fig. 6. Schematic representation of the metabolic features of the cancer cells in liver metastases. The figure illustrates the pathways of glycolysis, OXPHOS, and lipid metabolism in the tumor cells. Key enzymes involved in different metabolic pathways are marked in different colors: blue-glycolysis; orange-lipid metabolism. TCA, tricarboxylic acid cycle; OXPHOS, oxidative phosphorylation; ATP, Adenosine triphosphate; ROS, Reactive oxygen species; PGC1, PPARγcoactivator-1; PDH, Pyruvate dehydrogenase; ACL, ATP-citrate lyase; ACC, Acetyl-CoA carboxylase; FASN, Fatty acid synthase; SCD, Stearoyl-CoA Desaturase; ACS, long-chain-fatty-acid–CoA ligase.

including extracellular matrix remodeling and invasion, the active intracellular protein transport and the hydrolysis process [23,24]. In addition, insulin-like growth factor II mRNA-binding protein 3 (IGF2BP3), was also found to be upregulated in all organ metastases [25]. IGF2BP3 is known to be synthesized de novo in cancer, and it acts as oncogenes to stimulate the growth and migration of tumor cells in lung cancer, gastrointestinal cancer and ovarian cancer [26].

On the other hand, our result also reflected the obvious differential gene regulations among different organ metastases of ES-2 cells, especially, the protein expression profile of liver metastases showed a significant trend of outliers compared with that of other organ metastases. Therefore, liver metastases were chosen as typical research objects for differential protein screening in this study. Through multivariate statistical and bioinformatic analysis as well as sequence-specific validation, including HR-MRM quantification at the protein level and qRT–PCR at the mRNA level, LDHA (lactate dehydrogenase A), ACSL1 (long-chain fatty acid CoA ligase 1) and FTL (ferritin light chain)

were chosen as liver metastasis-specific genes of ES-2 cells. Notably, maybe due to the heterogeneity of tumor tissue colonizing in different mice, larger standard deviations were observed in some groups in HR-MRM assay (six biological replicates for each target protein), therefore, the expressions of target proteins were doubly verified by semi-quantitative PCR and real-time qPCR at mRNA level. These results clearly exhibited the truth of the regulation on target protein expression during liver metastasis of ES-2 cells. LDHA is known to be the A chain of LDH lactate dehydrogenase. During glycolysis, LDH (L-lactate dehydrogenase) is the key rate-limiting enzyme in the metabolic process of pyruvate to lactate [27,28]. LDH plays a crucial role in active glycolysis in cancer cells; moreover, the high expression of LDHA as a marker of tissue damage in OCCC has been demonstrated [29]. Furthermore, ACSL1 activates long-chain fatty acids by catalyzing the formation of long-chain fatty acyl-CoA and initiates fatty acid metabolism [30], and its high expression is closely related to the dysregulation of lipid metabolism and active fatty acid uptake [31], which is consistent with our finding that

genes specifically expressed in liver metastases are concentrated in metabolic pathways such as oxidative phosphorylation and glutathione biosynthesis (Fig. 6). A major function of ferritin is the storage of iron in a soluble and nontoxic state. Previous studies have shown that the particular microenvironment of persistent oxidative stress caused by high concentrations of free iron during the progression of its disease course has been shown to be important in the malignant degeneration of ovarian epithelial cells. The resulting metabolic reprogramming and the formation of a stress microenvironment are prominent features of OCCC [32]. It is well known that cellular metabolic reprogramming is an important marker of tumorigenesis [33]. By ingesting large amounts of nutrients from the microenvironment, tumor cells use aerobic glycolysis (Warburg effect) to provide more macromolecular intermediates while supporting their own proliferation and invasion through amino acids and lipid anabolites [15,34]. Interestingly, the changes in metabolite concentrations resulting from the metabolic reprogramming of tumor cells also stimulate the tumor microenvironment to secrete specific cytokines together with nutrients to maintain the malignant phenotype of cancer cells and to drive the invasion and metastasis of tumor cells through extracellular matrix remodeling [35]. It is well known that the organ tropism of metastatic tumors largely depends on the similarity of driver mutations in the metastatic site to the primary tumor [36], and studies have been reported predicting the properties of bone metastasis by the gene expression profile of primary breast cancer [37]. At the same time, similar expression trends have been reported of these three selected proteins in other cancer models, which may predict a homologous regulatory role in other multiple cancer models [38]. Secondary distal metastasis sites create a microenvironment similar to the primary tumor to support the metastatic colonization of the tumor cells through their interaction with the tumor cells. Actually, the liver is the most important organ for glucose and lipid metabolism and iron storage. On the one hand, its active metabolic microenvironment and colonization in the tumor cell complex interaction and metabolic coupling drive the metabolic conversion of tumor cells, making the key metabolic enzymes of anabolic pathway appear highly expressed and synthesizing more nutrients to maintain the abnormal proliferation of cancer cells and invasion and metastasis [39]. On the other hand, increased expression of ACSL1, FTL and LDHA in liver metastases reflects the adaptation of tumor cells to the liver microenvironment through metabolic reprogramming. The above results indicate that the liver has a microenvironment similar to that of OCCC due to metabolic reprogramming and oxidative stress, which might be one of the key reasons why OCCC cells tend to colonize the liver to form distal metastases.

5. Conclusions

In summary, in this research, on a xenograft mouse model of a human-derived tumor cell line, the protein expression profile study of multiorgan metastases from ES-2 cells was realized by a sequence-specific proteomic analysis strategy. Our approach provides a new idea for the study of the mechanism of tumor target organ metastasis and colonization, by which the confusion of tumor cell proteins and target organ proteins can be effectively avoided and realize the analysis of the true protein expression profile of tumor cells.

Availability of Data and Materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

NY and MG designed this study. MG, JC, SW, HY and WZ processed, analyzed and interpreted data. NY, XL, YL and TS contributed to the experiments. NY and MG wrote the draft. JC and MG review and edited the draft. All authors read and approved the final version of the manuscript.

Ethics Approval and Consent to Participate

Study protocol obeyed the Declaration of Helsinki for all animal experimental investigations and was approved and filed by the Ethics Committee of the West China Hospital, Sichuan University (approval number: 20220222033).

Acknowledgment

Not applicable.

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/j.fbl2803053.

References

[1] Nguyen DX, Bos PD, Massagué J. Metastasis: from dissemination to organ-specific colonization. Nature Reviews Cancer. 2009; 9: 274–284.

[2] Steeg PS. Targeting metastasis. Nature Reviews Cancer. 2016; 16: 201–218.

[3] Suhail Y, Cain MP, Vanaja K, Kurywchak PA, Levchenko A, Kalluri R. Systems biology of cancer metastasis. Cell Systems. 2019; 9: 109–127.

[4] Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. Cell. 2011; 147: 275–292.

[5] Spill F, Reynolds DS, Kamm RD, Zaman MH. Impact of the physical microenvironment on tumor progression and metastasis. Current Opinion in Biotechnology. 2016; 40: 41–48.

[6] Paget S. The distribution of secondary growths in cancer of the breast. The Lancet. 1889; 133: 571–573.

[7] Stevens AR, Ewing J. Adenocarcinoma of the Testis in the Adult. Annals of Surgery. 1928; 88: 1074.

[8] Bross IDJ. Metastatic sites that produce generalized cancer: identification and kinetics of generalizing sites. Fundamental Aspects of Metastasis. 1976.

[9] Pradeep S, Kim SW, Wu SY, Nishimura M, Chaluvally-Raghavan P, Miyake T, et al. Hematogenous metastasis of ovarian cancer: rethinking mode of spread. Cancer Cell. 2014; 26: 77–91.

[10] Rohan TE, Xue X, Lin HM, D'Alfonso TM, Ginter PS, Oktay MH, et al. Tumor microenvironment of metastasis and risk of distant metastasis of breast cancer. Journal of the National Cancer Institute. 2014; 106: dju136.

[11] Minn AJ, Kang Y, Serganova I, Gupta GP, Giri DD, Doubrovin M, et al. Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors. The Journal of Clinical Investigation. 2005; 115: 44–55.

[12] Sethi N, Kang Y. Unravelling the complexity of metastasis—molecular understanding and targeted therapies. Nature Reviews Cancer. 2011; 11: 735–748.

[13] Ellis MJ, Gillette M, Carr SA, Paulovich AG, Smith RD, Rodland KK, et al. Connecting genomic alterations to cancer biology with proteomics: the NCI Clinical Proteomic Tumor Analysis Consortium. Cancer Discovery. 2013; 3: 1108–1112.

[14] Gillette MA, Carr SA. Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry. Nature Methods. 2013; 10: 28–34.

[15] Zhou S, Liu R, Yuan K, Yi T, Zhao X, Huang C, et al. Proteomics analysis of tumor microenvironment: Implications of metabolic and oxidative stresses in tumorigenesis. Mass Spectrometry Reviews. 2013; 32: 267–311.

[16] Cho SY, Kang W, Han JY, Min S, Kang J, Lee A, et al. An integrative approach to precision cancer medicine using patient-derived xenografts. Molecules and Cells. 2016; 39: 77.

[17] Cekanova M, Rathore K. Animal models and therapeutic molecular targets of cancer: utility and limitations. Drug Design, Development and Therapy. 2014; 8: 1911.

[18] Michalski A, Cox J, Mann M. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC−MS/MS. Journal of Proteome Research. 2011; 10: 1785–1793.

[19] Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Molecular & Cellular Proteomics. 2012; 11.

[20] Gallien S, Duriez E, Crone C, Kellmann M, Moehring T, Domon B. Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. Molecular & Cellular Proteomics. 2012; 11: 1709–1723.

[21] Vidova V, Spacil Z. A review on mass spectrometry-based quantitative proteomics: Targeted and data independent acquisition. Analytica Chimica Acta. 2017; 964: 7–23.

[22] Ricono JM, Huang M, Barnes LA, Lau SK, Weis SM, Schlaepfer DD, et al. Specific Cross-talk between Epidermal Growth Factor Receptor and Integrin αvβ5 Promotes Carcinoma Cell Invasion and Metastasis. Cancer Research. 2009; 69: 1383–1391.

[23] Roberts PJ, Der CJ. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. Oncogene. 2007; 26: 3291–3310.

[24] Friday BB, Adjei AA. Advances in targeting the Ras/Raf/MEK/Erk mitogen-activated protein kinase cascade with MEK inhibitors for cancer therapy. Clinical Cancer Research. 2008; 14: 342–346.

[25] Bell JL, Wächter K, Mühleck B, Pazaitis N, Köhn M, Lederer M, et al. Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? Cellular and Molecular Life Sciences. 2013; 70: 2657–2675.

[26] Jeng Y-M, Wang T, Lu S, Yuan R, Hsu HC. Prognostic significance of insulin-like growth factor II mRNA-binding protein 3 expression in gastric adenocarcinoma. Journal of British Surgery. 2009; 96: 66–73.

[27] Okamoto T, Mandai M, Matsumura N, Yamaguchi K, Kondoh E, Amano Y, et al. Hepatocyte nuclear factor-1β (HNF-1β) promotes glucose uptake and glycolytic activity in ovarian clear cell carcinoma. Molecular Carcinogenesis. 2015; 54: 35–49.

[28] Ji JX, Wang YK, Cochrane DR, Huntsman DG. Clear cell carcinomas of the ovary and kidney: clarity through genomics. The Journal of Pathology. 2018; 244: 550–564.

[29] Wong OG, Li J, Cheung ANY. Targeting DNA damage response pathway in ovarian clear cell carcinoma. Frontiers in Oncology. 2021; 11.

[30] Rossi Sebastiano M, Konstantinidou G. Targeting long chain acyl-CoA synthetases for cancer therapy. International Journal of Molecular Sciences. 2019; 20: 3624.

[31] Corn KC, Windham MA, Rafat M. Lipids in the tumor microenvironment: From cancer progression to treatment. Progress in Lipid Research. 2020; 80: 101055.

[32] Anglesio MS, George J, Kulbe H, Friedlander M, Rischin D, Lemech C, et al. IL6-STAT3-HIF Signaling and Therapeutic Response to the Angiogenesis Inhibitor Sunitinib in Ovarian Clear Cell CancerIL6-STAT3-HIF Signaling in Ovarian Clear Cell Cancer. Clinical Cancer Research. 2011; 17: 2538–2548.

[33] Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011; 144: 646–674.

[34] Warburg O. On the origin of cancer cells. Science. 1956; 123: 309–314.

[35] Orimo A, Gupta PB, Sgroi DC, Arenzana-Seisdedos F, Delaunay T, Naeem R, et al. Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. Cell. 2005; 121: 335–348.

[36] 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. Nature Medicine. 2017; 23: 703–713.

[37] Zhang XH-F, Jin X, Malladi S, Zou Y, Wen YH, Brogi E, et al. Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. Cell. 2013; 154: 1060–1073.

[38] Samaržija I. Site-specific and common prostate cancer metastasis genes as suggested by meta-analysis of gene expression data. Life. 2021; 11: 636.

[39] Martinez-Outschoorn UE, Pestell RG, Howell A, Tykocinski ML, Nagajyothi F, Machado FS, et al. Energy transfer in" parasitic" cancer metabolism: mitochondria are the powerhouse and Achilles' heel of tumor cells. Cell Cycle. 2011; 10: 4208–4216.