Novel Plasma Proteomic Biomarkers for Early Identification of Induction Chemotherapy Beneficiaries in Locoregionally Advanced Nasopharyngeal Carcinoma

Shan-Qiang Zhang†, Su-Ming Pan‡, Shu-Zhen Lai§, Hui-Jing Situ*, Jun Liu†, Wen-Jie Dai†, Si-Xian Liang†, Li-Qing Zhou†, Qi-Qi Lu†, Pei-Feng Ke†, Fan Zhang†, Hai-Bin Chen§ and Ji-Cheng Li†,§,*

1 Medical Research Center, Yuebei People’s Hospital, Shantou University Medical College, Shaoguan, China, 2 Department of Radiation Oncology, Yuebei People’s Hospital, Shantou University Medical College, Shaoguan, China, 3 Department of Histology and Embryology, Shantou University Medical College, Shantou, China, 4 Institute of Cell Biology, Zhejiang University School of Medicine, Hangzhou, China

Background: Induction chemotherapy (IC) can alleviate locoregionally advanced nasopharyngeal carcinoma (LA-NPC), but effectiveness differs between patients, toxicity is problematic, and effective blood-based IC efficacy predictors are lacking. Here, we aimed to identify biomarkers for early identification of IC beneficiaries.

Methods: Sixty-four pairs of matched plasma samples collected before and after IC from LA-NPC patients including 34 responders and 30 non-responders, as well as 50 plasma samples of healthy individuals, were tested using data-independent acquisition mass spectrometry. The proteins associated with clinical traits or IC benefits were investigated by weighted gene co-expression network analysis (WGCNA) and soft cluster analysis. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes functional annotations were performed to determine the potential function of the identified proteins. The area under the receiver operating characteristic curve (AUC) was used to evaluate the performance of candidate biomarkers in predicting IC beneficiaries.

Results: Compared with healthy individuals, 1027 differentially expressed proteins (DEPs) were found in the plasma of LA-NPC patients. Based on feedback from IC outcomes, 463 DEPs were identified in the pre-IC plasma between responders and non-responders. A total of 1212 DEPs represented the proteomic changes before and after IC in responders, while 276 DEPs were identified in post-IC plasma between responders and non-responders. WGCNA identified nine protein co-expression modules correlated with clinical traits. Soft cluster analysis identified four IC benefits-related protein clusters. Functional enrichment analysis showed that these proteins may play a role in IC via immunity, complement, coagulation, glycosaminoglycan and serine. Four proteins differentially expressed in all group comparisons, paraoxonase/arylesterase 1 (PON1), insulin-like growth factor-binding protein 3 (IGFBP-3), rheumatoid factor D5 light chain...
INTRODUCTION

Nasopharyngeal carcinoma (NPC) is an aggressive type of head and neck malignant disease with a relatively high rate of morbidity and metastasis (1). In 2020, the World Cancer Report released by the International Agency for Research on Cancer (IARC) estimated 133,354 new cases of NPC and 80,008 NPC-related deaths worldwide (2). Within China, the corresponding numbers for such cases in 2020 were 62,444 and 34,810, accounting for 46.83% and 42.72% of the global total, respectively (3). Although progress in medical technology has made it possible to cure patients with early-stage NPC, more than 70% of patients were diagnosed as locally advanced (LA)-NPC at initial presentation (4). Treatment fails for ~30–40% of LA-NPC patients, and tumour relapse or distant metastasis reduces the overall survival rate of these patients to 20% (5, 6). Therefore, improving the clinical treatment of LA-NPC is crucial for effectively controlling disease progression and reducing mortality.

In recent years, data from several large-scale clinical trials demonstrated that induction chemotherapy (IC) prior to radiotherapy of LA-NPC leads to improved survival outcomes (7–9). When employing IC as a National Comprehensive Cancer Network (NCCN) guideline recommendation for the treatment of LA-NPC, clinical analysis revealed that the short-term tumour response to IC varies significantly among individuals; symptoms of ~30% of LA-NPC patients receiving IC were not effectively alleviated, and IC only achieved 8% absolute survival benefit, indicating that numerous patients underwent ineffective treatment (8, 10, 11). Furthermore, the incidence of acute adverse grade 3 or 4 events triggered by IC was as high as 75% (7). Hematotoxicity, weight loss, oral mucositis and salivary gland dysfunction could postpone or interrupt follow-up treatment, thus increasing the risk of cancer progression and chemoresistance (12). Therefore, identifying in advance patients likely to have a favourable response to IC would be of great importance for individualised treatment of LA-NPC. Currently, clinical indicators cannot provide strategies for identifying patients suitable for IC before treatment, because existing evidence of Epstein-Barr virus (EBV) only supports its use in the diagnosis and prognosis of NPC, while the Response Evaluation Criteria in Solid Tumours (RECIST) value is limited to the determination of short-term efficacy after IC (13). Therefore, there is an urgent need to establish a biomarker-driven laboratory method for more precise IC application.

Proteins control all physiological processes including immunity, material exchange, metabolism and signal pathway regulation (14). During cancer progression, protein synthesis and the decomposition of tumour tissue are enhanced, which affects the expression and composition of circulating proteins. These abnormalities are accompanied by dynamic changes in tumour cell apoptosis or proliferation during treatment. Therefore, circulating proteins have been studied for their utility as biomarkers in a wide range of malignancies and disorders (15, 16). High-throughput technology has been employed to develop biomarkers for early diagnosis, chemoresistance, and prognosis of NPC in the field of proteomics (17–19). However, the circulating protein profiles of LA-NPC patients with different IC efficacy, and the proteomic biomarkers for predicting the short-term efficacy of IC have not been reported.

In the present study, we investigated the plasma protein profiles of LA-NPC patients with different IC efficacy using data-independent acquisition (DIA) mass spectrometry, and conducted in-depth screening for differentially expressed proteins (DEPs), as well as the correlative analysis between pre-treatment plasma proteins, clinical traits and IC benefits. Then, we explored whether pre-treatment plasma proteins have the potential as biomarkers for early identification of IC beneficiaries.

MATERIAL AND METHODS

Participants, Samples and Grouping

We recruited 50 healthy controls (HC group) and 64 patients with non-metastatic stage III–IVa LA-NPC patients (LA-NPC group) at Yuebei People’s Hospital, Shantou University Medical College (Shaoguan, China), between May 2020 and January 2021. The detailed descriptions of participants are shown in Table 1. The distributions of age, gender, clinical stage, T- and N-classification, and EBV DNA load between patients showed no statistical differences. All patients were diagnosed under the guidance of the 8th edition of the American Joint Committee on Cancer (AJCC) Staging Manual (20). Patients participating in this study were not
suffering from any other major diseases at the time of diagnosis, and all were initially diagnosed and did not receive any anti-tumour therapy before blood sampling. Histopathological diagnosis was non-keratinising undifferentiated carcinoma in all cases.

To ensure that the identified proteins are relevant to IC efficacy, we only enrolled patients who received IC alone and who were administrated with the same regimen before radiotherapy. For IC, 75 mg/m² docetaxel and 80 mg/m² cisplatin (DP regimen) were given once every 3 weeks for two cycles. Prior to radiotherapy, we evaluated the tumour response to IC using RECIST (version 1.1) (21). Based on the measurement of gross tumour volume (GTV) using an Elekta Unity dedicated treatment planning system (TPS) and Monaco (v5.40.01) (22), we defined patients with complete and partial responses (CR/PR) as the effective (EFF) cohort, and patients with stable and progressive disease (SD/PD) as the non-effective (NEFF) cohort.

Plasma sampling was performed following the Human Proteome Organization (HUPO) recommendations (23). A total of 64 pairs of fasting blood samples were collected at two timepoints before and after IC. For plasma separation, peripheral blood was centrifuged under anticoagulant conditions at 1300 g for 10 min at 4°C, and the supernatant was transferred to an Eppendorf tube and stored at -80°C. Based on sampling times, we further divided the EFF cohort into pre-treatment effective (pre-EFF) and post-treatment effective (post-EFF) groups, and the NEFF cohort into pre-treatment non-effective (pre-NEFF) and post-treatment non-effective (post-NEFF) groups. This study was conducted in compliance with the ethical guidelines of the Helsinki Declaration, and was approved by the Institutional Review Board of Yuebei People’s Hospital, Shantou University Medical College (KY-2020-020). All participants gave written informed consent.

**Sample Preparation**

Briefly, high abundance and low abundance proteins were separated from the plasma pools using a Multiple Affinity Removal LC Column-Human 14 (Hu14; Agilent Technologies), and their components were desalted and concentrated using a 5 kDa cut-off ultrafiltration device (Sartorius Corp). Samples were mixed with SDT buffer comprising 4% sodium dodecyl sulphate (SDS), 100 mM dithiothreitol (DTT) and 150 mM Tris-HCl pH 8.0, boiled for 15 min, centrifuged at 14,000 g for 20 min, and the supernatant was quantified using a BCA Protein Assay Kit (Bio-Rad Corp). Samples were stored at -80°C for subsequent experiments.

For digestion, 200 µg of proteins was repeatedly ultrafiltered with UA buffer (8 M Urea,150 mM Tris-HCl pH 8.0) according to the filter-aided sample preparation (FASP) procedure to remove detergent and low-molecular-weight components. Next, 100 µL of 100 mM iodoacetamide (IAA) in UA buffer was added to block reduced cysteine residues and samples were incubated for 30 min in darkness. Filters were washed with 100 µL UA buffer three times then with 100 µL 25 mM NH₄HCO₃ buffer twice. The obtained protein suspensions were digested with 4 µg trypsin (Promega Corp) in 40 µL 25 mM NH₄HCO₃ buffer overnight at 37°C, and the resulting peptides were collected as a filtrate. The peptide content of each sample was estimated by UV light spectral density at 280 nm after desalting on an Empore SPE C18 cartridge (standard density, bed I.D. 7 mm, volume 3 mL; Sigma Corp), concentrated by vacuum centrifugation, and reconstituted in 40 µL of 0.1% (v/v) formic acid.

**TABLE 1** Demographics and baseline characteristics of enrolled participants.

|                          | HC              | LA-NPC          | P value |
|--------------------------|-----------------|-----------------|---------|
| Population size          | 50              | 34              | 30      | 0.680  |
| Age                      | 50.18 ± 10.69   | 48.74 ± 11.35   | 51.30 ± 13.68 | 0.240  |
| Gender-no. (%)           |                 |                 |         |       |
| Male                     | 37 (74)         | 22 (64.71)      | 25 (83.33) | 0.518  |
| Female                   | 13 (26)         | 12 (35.29)      | 5 (16.67)  | 0.563  |
| Clinical stage-no. (%)   |                 |                 |         |       |
| III                      | 20 (58.82)      | 20 (66.67)      |         | 0.563  |
| IVA                      | 14 (41.18)      | 10 (33.33)      |         | 0.563  |
| T Classification-no. (%) |                 |                 |         |       |
| T1                       | 0 (0)           | 1 (3.33)        |         | 0.289  |
| T2                       | 10 (29.41)      | 9 (30.01)       |         | 0.289  |
| T3                       | 11 (32.35)      | 12 (40.00)      |         | 0.289  |
| T4                       | 13 (38.24)      | 8 (26.67)       |         | 0.289  |
| N Classification-no. (%) |                 |                 |         |       |
| N0                       | 2 (5.88)        | 1 (3.33)        |         | 0.559  |
| N1                       | 6 (17.65)       | 1 (3.33)        |         | 0.559  |
| N2                       | 23 (67.65)      | 25 (83.33)      |         | 0.559  |
| N3                       | 3 (8.82)        | 3 (10.00)       |         | 0.559  |
| EBV DNA load (copies/mL) |                 |                 |         |       |
| <1000                    | 26 (76.47)      | 21 (70.00)      |         | 0.559  |
| >1000                    | 8 (23.53)       | 9 (30.00)       |         | 0.559  |

no. (%), number; EBV, Epstein-Barr virus.
Digested pool peptides were subsequently fractionated using a High pH Reversed-phase Peptide Fractionation Kit (Thermo Scientific Corp) into 10 fractions. Each fraction was concentrated by vacuum centrifugation and reconstituted in 15 µL of 0.1% (v/v) formic acid. Collected peptides were desalted and reconstituted successively with a C18 cartridge and 40 µL of 0.1% (v/v) formic acid. Finally, iRT-Kits (Biognosys Corp) were added to correct the relative retention time differences between runs at a volume ratio of 1:3 for iRT standard peptides versus sample peptides.

**Data-Dependent Acquisition Mass Spectrometry**

Peptide fractions for DDA library generation were analysed using a Q Exactive HF X mass spectrometer (Thermo Scientific Corp) coupled to an Easy nLC 1200 chromatography system (Thermo Scientific Corp). Peptide samples (1.5 µg) were first loaded onto an EASY-Spray C18 Trap column (P/N 164946, 3 µm, 75 µm × 2 cm; Thermo Scientific Corp), then separated on an EASY-Spray C18 LC Analytical Column (ES802, 2 µm, 75 µm × 25 cm; Thermo Scientific Corp) over a 120 min gradient from buffer B (0.1% formic acid and 84% acetonitrile). The column flow rate was maintained at 250 nL/min. MS detection was performed in positive ion mode, full scans were performed between 300 and 1800 m/z, the resolution for the MS1 scan was 60,000 at 200 m/z, the automatic gain control (AGC) target for the MS scan was set to 3e6, and the maximum injection time (IT) was 25 ms. The dynamic exclusion was set to 30s. Each full mass spectrometry (MS)-selected ion monitoring (SIM) scan followed 20 ddMS2 scans. The MS2 scan was performed at 15,000 resolution, the AGC target was 5e4, the maximum IT was 25 ms, and the collision energy was 30 eV.

**DIA Mass Spectrometry**

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) running in DIA mode (Shanghai Applied Protein Technology Co., Ltd) was used for analysis of sample peptides. Each DIA cycle was comprised of one full MS-SIM scan, and 30 DIA scans covered a mass range of 350–1,800 m/z. For full scans, the resolution was 120,000 at 200 m/z, the AGC target was 3e6, and the maximum IT was 50 ms. DIA scans were performed at 15,000 resolution, and the AGC target and maximum IT were set to 3e6 and auto, respectively. The collision energy was 30 eV and the runtime was 120 min, with a linear gradient of buffer B (84% acetonitrile and 0.1% formic acid) at a flow rate of 250 nL/min. MS performance monitoring was conducted by injection of QC acetonitrile and 0.1% formic acid) at a flow rate of 250 nL/min. The collision energy was 30 eV and auto, respectively. The maximum IT was 50 ms. DIA scans were performed at a mass range of 350–1,800 m/z. For full scans, the resolution was 120,000 at 200 m/z, the AGC target was 3e6, and the maximum IT was 25 ms. The dynamic exclusion was set to 30s. Each full mass spectrometry (MS)-selected ion monitoring (SIM) scan followed 20 ddMS2 scans. The MS2 scan was performed at 15,000 resolution, the AGC target was 5e4, the maximum IT was 25 ms, and the collision energy was 30 eV.

**Bioinformatics and Statistical Analyses**

All statistical testing was performed using R version 4.1.1 (R Foundation for Statistical Computing) software packages and SPSS Statistics version 22.0 (IBM Corp.). To mine the proteomic differences between IC responders and non-responders, proteins with either increased or decreased relative abundance in the inter-group comparisons were defined as DEPs according to fold change ≥1.2 or ≤0.833 and p-value <0.05. A weighted protein co-expression network was constructed to identify hub modules related to clinical traits among the identified proteins using the Weighted Gene Co-Expression Network Analysis (WGCNA) package for R software. The fuzzy C-means algorithm, a software package termed Mfuzz (http://mfuzz.sysbiolab.eu/) based on the open-source statistical language R, was used for soft clustering analysis to observe changes in the expression of plasma proteins in the EFF cohort before and after IC. For DEPs and proteins related to clinical traits or IC benefits, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional annotation were performed to explore the enriched functions and pathways using the clusterProfiler R package (24). Receiver operating characteristic (ROC) curve analysis was conducted to test the discriminative ability of the prediction models for IC beneficiaries.

Shapiro-Wilk normality tests were used to confirm whether data obeyed a normal distribution. For statistical comparison, one-way analysis of variance or Student t-tests were used for parameter data, chi-square tests were used for composition ratio, mean ± standard deviations were used to describe continuous variables, and p <0.05 for two-tailed tests was considered significant.

**RESULTS**

**IC Outcomes and Quality Control Data**

The detailed roadmap for this study is shown in Figure 1. Based on the IC outcomes of the 64 enrolled patients, we classified one patient with CR and 33 patients with PR as the EFF cohort, and 29 patients with SD and 1 patient with PD as the NEFF cohort. Figure 2 shows the typical changes in GTV before and after IC for 6 patients (PR = 3, PD = 1, SD = 2) assessed by magnetic resonance imaging, and GTV measurements for all patients are provided in Supplementary Table 1. Quality control data are shown in Supplementary Figures 1A–D. For protein identification, the FDR method was applied for multiple testing.
FIGURE 1 | Roadmap for identifying biomarkers for early identification of IC beneficiaries. HC, healthy control; LA-NPC, locoregionally advanced nasopharyngeal carcinoma; IC, induction chemotherapy; pre-EFF, pre-effective; pre-NEFF, pre-non-effective; post-EFF, post-effective; post-NEFF, post-non-effective; DDA, data-dependent acquisition; DIA, data-independent acquisition; DEPs, differentially expressed proteins; WGCNA, weighted gene co-expression network analysis; ROC, receiver operating characteristic.

FIGURE 2 | Typical MRI images for GTV changes before and after IC. MRI images of three patients with PR (EFF1-3), one patient with PD (NEFF1), and two patients with SD (NEFF2-3) at baseline (pre-IC) and after two cycles of IC (post-IC). The area within the red line is the primary focus. MRI, magnetic resonance imaging; GTV, gross tumour volume; IC, induction chemotherapy; EFF, effective; NEFF, non-effective.
correction, and the quantitative results with a Q value <0.01 indicate high reliability. The intensity jitter distribution plot includes a few outlier samples, and the strength of the correlation between quality control samples was greater than 0.9 in all cases. A total of 2512 proteins were identified across all groups, indicating high repeatability and quantitative stability.

**Differences in Protein Profiles Between IC Responders and Non-Responders**

To comprehensively understand the differences and changes of plasma protein profiles between responders and non-responders before and after IC, screening of DEPs was carried out in tiers. A total of 1027 proteins were found to be differentially expressed in the LA-NPC group compared with the HC group, of which 324 were upregulated and 703 were downregulated (Figure 3A). Subsequently, 463 DEPs, including 370 upregulated and 93 downregulated, were identified between the pre-EFF group and the pre-NEFF group (Figure 3B). Figure 3C shows 1212 DEPs (380 upregulated and 832 downregulated) identified between pre-EFF and post-EFF groups, while Figure 3D shows 276 DEPs (86 upregulated and 190 downregulated) between post-EFF and post-NEFF groups.

Enrichment analysis of DEPs is shown in Figures 4, 5. GO annotation showed that the highly enriched terms of biological process (BP), cellular component (CC), and molecular function (MF) categories in all batch data are roughly the same, while some MF terms, such as glycosaminoglycan binding, sulfur compound binding, heparin binding, serine-related enzyme activities, complement binding, structural constituents of the epidermis, and opsonin binding were enriched between pre-EFF and pre-NEFF groups, pre-EFF and post-EFF groups, and post-EFF and post-NEFF groups, respectively. In addition, KEGG analysis indicated that the complement and coagulation cascade was the most significantly enriched pathway for all batch data.

**Correlation Between Pre-Treatment Plasma Proteins and Clinical Traits**

To investigate correlations between the identified proteins and clinical traits, a weighted protein co-expression network was constructed from 2297 proteins after filtering out proteins expressed in fewer than half of the samples. The power of $\beta = 5$ was selected for soft thresholding to ensure a scale-free topology (Figure 6A). Among the 19 co-expression modules, proteins contained in the grey module were not attributed to any modules (Figure 6B). Clinical traits, such as patient’s age, gender, T classification, N classification, clinical stage and IC outcome were used for correlation analysis. The results presented in Figure 6C indicate that the pink module was negatively correlated with gender. The modules negatively correlated with T classification and clinical stage are brown, while the positive correlations are coloured black and pink. The midlightblue module represents a positively correlation with N classification. Only the cyan module is negatively correlated with IC outcome, while greenyellow, yellow, blue and turquoise modules are positively correlated with IC outcome.

Modules significantly correlated with tumour stage and IC outcome were integrated separately using the criteria correlation $>0.3$ and $p < 0.001$ to perform enrichment analysis.
GO annotation showed that the overrepresented BP, CC and MF terms for the two batches were roughly the same, and similar to the annotation results for DEPs. KEGG analysis indicated that the complement and coagulation cascade was still the most significantly enriched pathway for the two batches of data.

**Correlation Between Pre-Treatment Plasma Proteins and IC Benefit**

To determine whether the identified proteins are correlated with IC benefits, nine clusters that tended to exhibit significant changes were identified by soft clustering analysis to reveal proteomic changes in the EFF cohort around the time of IC (Figure 7). In cluster 1, expression of proteins in the pre-EFF group was significantly higher than that in the HC group, and it was decreased significantly in the post-EFF group. Similarly, proteins in cluster 2 were also highly expressed in the pre-EFF group compared with the HC group, but tended to be normal in the post-EFF group. By contrast, protein expression levels in clusters 7 and 8 in the pre-EFF group were significantly lower than that in the HC group, but in the post-EFF group, protein expression levels for cluster 7 were significantly higher than that in the HC group, while protein expression levels for cluster 8 tended to be normal. These expression changes appear to be associated with tumor regression (i.e., PR and CR, respectively).

For the other clusters, the changes in protein expression after IC in clusters 3, 4 and 6 did not show a strong correlation with the benefits of IC. Although the changes in protein expression in cluster 5 showed a similar trend to clusters 7 and 8, a considerable proportion of proteins could be found whose expression did not change significantly after IC. In addition, the protein expression in the pre-EFF group in cluster 9 did not show significant differences compared to the HC group, and their expression in the post-EFF group was significantly lower than that in the HC group. Therefore, we did not include the proteins in these clusters in further analyses.

We performed enrichment analysis of proteins contained in clusters 1, 2, 7 and 8 (Figure 8). GO annotation showed that the most significantly enriched terms were similar to those identified in the DEPs annotation results. KEGG analysis indicated that the complement and coagulation cascade was the most enriched pathway.

**Identification of Predictive Biomarkers for IC Beneficiaries**

DEPs with a given gene name in all datasets overlapped, and four proteins were identified (Figure 9A); paraoxonase/arylesterase 1...
(PON1), insulin-like growth factor-binding protein 3 (IGFBP3), rheumatoid factor D5 light chain (v-kappa-3), and RNA helicase (DDX55). Their relative protein abundances are shown in Figure 9B. Based on the results of the above analyses, the brown module to which PON1 belongs was negatively correlated with clinical stage and T classification (Figure 6C). For v-kappa-3, the turquoise module in which it is located was correlated with IC outcome, and its expression change in the EFF cohort around the time of IC is correlated with IC benefits, as observed for cluster 2 (Figures 6C, 7). The blue module containing DDX55 was also correlated with IC outcome (Figure 3C). The expression changes for DDX55 in the EFF cohort around the time of IC presented a positive correlation with IC benefits, as shown for cluster 8 (Figure 7).

To explore the potential of the four proteins as biomarkers, we performed multivariate logistic regression analysis and ROC curve analysis. When applying the four proteins to differentiate HC and LA-NPC groups, the AUC values for PON1, IGFBP3, v-kappa-3 and DDX55 were 0.82, 0.66, 0.81 and 0.65, respectively (Figure 9C). When the four proteins were combined into a diagnostic model, the AUC value of the model integrating all four proteins was 0.97 (Figure 9H). Furthermore, the results presented in Figures 9I, J show that the four proteins could readily discriminate between post-EFF and post-NEFF groups (PON1, AUC = 0.71; IGFBP3, AUC = 0.87; DDX55, AUC = 0.94), especially the model combining all four proteins (AUC = 0.94).

DISCUSSION

For LA-NPC, the main advantages of IC are eliminating micrometastases and shrinking primary tumours to improve the outcome of follow-up radiotherapy (7-10). However, despite
this treatment, the short-term tumour response rate and long-term survival rate are not improved for some patients, and serious systemic toxicity can occur following IC (11, 12). Therefore, early identification of patients suitable for IC before treatment is the ideal strategy to improve the effective rate of clinical treatment of LA-NPC. In other words, patients who do not respond to one IC regimen could be rapidly subjected to other regimens. Currently, no clear omics data were available on the molecular basis of IC-sensitivity and -resistance for LA-NPC. To address this issue, we comprehensively evaluated the plasma protein profiles of responders and non-responders around the time of IC using a DIA approach, and identified IC benefit-related proteins. To the best of our knowledge, this is the first study to apply a high-throughput proteomics approach to investigate the blood molecular characteristics of LA-NPC patients with different IC efficacy.

In order to mining the potential correlation between plasma proteins and the short-term efficacy of IC as much as possible, we explored various elements. IC comprises combinations of different drugs, usually from two or more drug classes. Considering that existing chemotherapeutics work through various mechanisms and sites of action, we unified the medication regimen and treatment cycle to avoid differences in efficacy caused by different regimens or dose intensities (25). Moreover, previous studies on biomarkers predicting the efficacy of IC included patients receiving concurrent chemoradiotherapy, while radiotherapy was proved to cover the efficacy of IC (13, 19, 26). Therefore, we only enrolled patients who received IC alone prior to radiotherapy in the present study, and focused on the short-term efficacy of IC by sampling before radiotherapy. Eventually, compared with other studies using only pre-treatment samples to identify predictive biomarkers of chemotherapy efficacy, we also utilised matched post-IC samples to confirm the correlation between pre-treatment plasma proteins and IC benefits, and to verify the potential of pre-treatment plasma proteins to develop efficacy predictive biomarkers (13, 15, 19).

Based on this design, we observed different protein profiles in the pre-treatment plasma between IC responders and non-responders, and the protein profiles of the two remained
differences after changes were produced under the intervention of IC. In addition, a number of proteins were found to be related to clinical traits or IC benefits. In the results of enrichment analysis, GO annotation revealed that many functional terms were repeatedly highlighted all batch data. Thus, we infer those proteins enriched in these terms may jointly participate in the occurrence and development of NPC. However, in terms of immune response, the results indicate that chemotherapy drugs may not only inhibit tumour cells and activate immune effectors, but also promote the release of some inflammatory and angiogenesis-related cytokines, which may explain why chemotherapy has no effect on some tumours (27). It is notable that the location of some functional terms in the data may imply their potential relevance to IC benefits. For example, glycosaminoglycan assembly in the Golgi is sensitive to drug treatment and a variety of environmental factors, which can alter the structure of glycosaminoglycans and thereby promote or inhibit different signalling pathways (28). In addition, serine synthesis can help hypoxic cancer stem cells to regulate redox state (29). Serine restriction and inhibiting its signalling pathway can reduce the proliferative activity of cancer cells (30). KEGG analysis revealed that complement activation and coagulation functions can enhance the malignant biological behaviours of cancer, such as inflammation, epithelial-mesenchymal transition (EMT) and angiogenesis (31). Some genes in the complement and coagulation cascade signalling pathway are considered predictors of chemosensitivity (32).

The above analysis indicates that LA-NPC circulating proteins are involved in a series of complex mechanisms in the process of IC, including protein-protein interactions and drug–protein target interactions. Our results lay a foundation for further revealing the precise molecular mechanism and new therapeutic targets. The differences in plasma protein profiles between responders and non-responders after IC reflect the different sensitivity of various subpopulations of NPC cells to chemotherapeutic drugs caused by tumor heterogeneity (33). Based on the efficacy feedback, we confirmed the potential correlation between pre-treatment plasma proteins and short-term efficacy of IC by screening of DEPs and...
bioinformatics analysis. Although some of the proteins we identified may not be NPC-specific, their expression may be regulated by pro- or anti-oncogenes, as well as other tumour-associated genes and their products, which can indirectly reflect the state of the tumor (34).

Considering the above-mentioned problems of IC for LA-NPC, we further confirmed the potential clinical value of pre-treatment plasma proteins by identifying efficacy predictive biomarkers. The four candidate biomarkers identified in this study have been linked to physiological processes or chemotherapy resistance related to malignant diseases (35–43). PON1 is a calcium-dependent hydrolase protein mainly involved in the occurrence and development of cancers through the regulation of oxidative stress and inflammation (35). PON1 is expressed at low levels in the blood of patients with a variety of cancers, and levels are further decreased after treatment (35, 36). We observed similar trends in the present study. The low expression of PON1 in LA-NPC plasma may be caused by inhibition of its synthesis in liver and/or oxidative stress, and the continuous decrease in levels after IC may be caused by inhibition of PON1 synthesis in the liver by chemotherapeutic drugs inhibiting peroxisome proliferator-activated receptor delta (37). A previous study found that decreased serum PON1 was significantly correlated with tumour load, consistent with our findings that PON1 levels were negatively correlated with clinical stage and T classification (35). Enrichment analysis showed that PON1 was enriched in terms related to lipids and cholesterol, suggesting that preventing low-density lipoprotein oxidation, maintaining cholesterol homeostasis, and reducing chronic inflammation may be effective ways to prevent the further deterioration of LA-NPC.

IGFBP3 is an N-linked glycosylated, phosphorylated, secretory protein correlated with multiple cancers. In this study, we found that the expression of IGFBP3 in LA-NPC was lower than in the HC group, and even lower in the pre-NEFF group. It is interesting to note that IGFBP3 was further downregulated in the post-EFF group, but significantly upregulated in the post-NEFF group and higher than that in the HC group. The previous studies found that upregulation of IGFBP3 promotes the proliferation, invasion and metastasis of NPC cells, consistent with our finding that IC non-responders may have a poor prognosis (38, 39). Enrichment analysis found that IGFBP3 may play a role in IC through binding insulin-like growth factor I (IGF-1). IGF-1 can promote progression of the cell cycle, leading to mitosis and anti-apoptosis, and phosphorylation of IGFBP3 can enhance its binding to IGF-1 (40). In addition, IGFBP3 can enhance TGF-β1-mediated EMT by activating other cells against decapentaplegic homolog 2/3 (SMAD2/3) (38). EMT is known to involve both cisplatin and docetaxel resistance mechanisms, which further indicates the potential of IGFBP3 as an efficacy predictive biomarker (41).

FIGURE 8 | Functional enrichment analysis of IC benefits-related proteins. (A–D) Results corresponding to functional enrichment analysis of proteins in clusters 1, 2, 7 and 8, respectively.
Compared with the above two proteins, the reports of v-kappa-3 in cancer are very limited, which may be related to its unknown biological function. DDX55 is a member of the RNA helicase family, members of which perform crucial roles in cell metabolism. Abnormally high expression of DDX55 can be used as a reference indicator for poor prognosis of a variety of cancers (42, 43). Similar to these studies, we also found that the expression of DDX55 in the pre-NEFF group was higher than that in the HC group. In lung cancer, DDX55-associated genes were found to be enriched in proto-oncogene (MYC) targets, and MYC transcription can be regulated by cytokines (42). Considering that NPC cells can also stimulate the secretion of cytokines (42). Considering that NPC cells can also stimulate the secretion of cytokines, hence we speculated that differences in the expression of DDX55 in plasma from patients with differences in IC efficacy is related to the modulation of RNA binding and ATP-dependent helicase activity by small molecules, or disrupting certain protein-protein interactions implicated in cancer progression (43).

In the performance evaluation of candidate biomarkers, we found that the model composed of all four proteins showed satisfactory accuracy in diagnosing LA-NPC, predicting IC efficacy, confirming effective IC, and evaluating IC outcome (all AUC values ≥0.92). It is worth noting that continuous and dynamic expression detection is one of the necessary conditions to determine whether biomarkers have the potential to predict efficacy, we used post-IC samples to validate the efficacy predictive biomarkers identified in pre-IC samples (44). The advantage of this is that it can confirm the post-treatment detectability and IC benefit relevance of candidate biomarkers in a way that the use of additional pre-IC samples alone does not (45). In addition, diagnostic model was constructed with the aim of detecting LA-NPC correlation of biomarkers. Predictive model was primarily used to assess the precision of biomarkers for identifying potential IC beneficiaries. The two models constructed using post-IC samples were used to confirm whether the potential IC beneficiaries identified by the biomarkers were indeed treated effectively, and to further exclude false positive

**FIGURE 9** | Identification of candidate biomarkers. (A) Venn diagram showing the overlap of DEPs between the four compared datasets. (B) Relative protein abundances of PON1, IGFBP3, v-kappa-3 and DDX55. (C, D) ROC analysis of the four proteins and their co-constructed model for distinguishing between HC and LA-NPC groups. (E, F) ROC analysis of the four proteins and their co-constructed model for distinguishing between pre-EFF and pre-NEFF groups. (G, H) ROC analysis of the four proteins and their co-constructed model for distinguishing between post-EFF and post-NEFF groups. (I, J) ROC analysis of the four proteins and their co-constructed model for distinguishing between post-EFF and post-NEFF groups. ****p < 0.0001
results. Thus, the cohort design containing six groups, the four-tier in-depth screening of DEPs and multi model evaluation increased the reliability of our results and demonstrated the potential clinical value of the candidate biomarkers.

Although this study revealed differences in circulating protein profiles among LA-NPC patients with different IC outcomes, the identified potential biomarkers need to be further verified by more independent clinical cohorts due to the limitations of single-center nature. Furthermore, in addition to methods such as enzyme-linked immunosorbent assays, considering that IC usually involves two or more drugs which are administered to the patients at the same time, functional study of the potential biomarkers also needs to explore a reliable method to establish multidrug resistant cell line models.

CONCLUSIONS

In summary, the present study comprehensively investigated the plasma protein profiles of LA-NPC patients with different IC outcomes and demonstrated the potential correlation between pre-treatment plasma proteins and the short-term efficacy of IC. The identified four proteins, PON1, IGFBP3, v-kappa-3 and DDX55, can achieve the early identification of IC beneficiaries while diagnosing LA-NPC, and they could confirm whether the identified potential beneficiaries have been effectively treated after IC, and distinguish responders and non-responders. These findings may help to reveal new targets of IC and serve the individualised treatment of LA-NPC. Whether the identified potential biomarkers are specific to the DP regimen remains to be clarified.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

REFERENCES

1. Chen YP, Chan ATC, Le QT, Blanchard P, Sun Y, Ma J. Nasopharyngeal Carcinoma. Lancet (2019) 394(10192):64–80. doi: 10.1016/S0140-6736(19)30956-0
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin (2021) 71(3):209–49. doi: 10.3322/caac.21660
3. Cancer Today. Population Fact Sheets (2020). Available at: https://gco.iarc.fr/today/fact-sheets-populations (Accessed December 10, 2021).
4. Liu SL, Bian LJ, Liu ZX, Chen QY, Sun XS, Sun R, et al. Development and Validation of the Immune Signature to Predict Distant Metastasis in Patients With Nasopharyngeal Carcinoma. J Immunother Cancer (2020) 8(1):e000205. doi: 10.1136/jitc-2019-000205
5. Zhang W, Guo Q, Liu G, Zheng F, Chen J, Huang D, et al. NKILA Represses Nasopharyngeal Carcinoma Carcinogenesis and Metastasis by NF-kb Pathway Inhibition. PLoS Genet (2019) 15(8):e1008325. doi: 10.1371/journal.pgen.1008325
6. Lu S, Yu Z, Xiao Z, Zhang Y. Gene Signatures and Prognostic Values of M^a Genes in Nasopharyngeal Carcinoma. Front Oncol (2020) 10:875. doi: 10.3389/fonc.2020.00875
7. Zhang Y, Chen L, Hu GQ, Zhang N, Zhu XD, Yang KY, et al. Gemcitabine and Cisplatin Induction Chemotherapy in Nasopharyngeal Carcinoma. N Engl J Med (2019) 381(12):1124–35. doi: 10.1056/NEJMoa1905287
8. Sun Y, Li WF, Chen NY, Zhang N, Hu GQ, Xie FY, et al. Induction Chemotherapy Plus Concurrent Chemoradiotherapy Versus Concurrent Chemoradiotherapy Alone in Locoregionally Advanced Nasopharyngeal Carcinoma: A Phase 3, Multicentre, Randomised Controlled Trial. Lancet Oncol (2016) 17(11):1509–20. doi: 10.1016/S1470-2045(16)30410-7
9. Cao SM, Yang Q, Guo L, Mai HQ, Mo HY, Cao KJ, et al. Neoadjuvant Chemotherapy Followed by Concurrent Chemoradiotherapy Versus Concurrent Chemoradiotherapy Alone in Locoregionally Advanced Nasopharyngeal Carcinoma: A Phase III Multicentre Randomised Controlled Trial. Eur J Cancer (2017) 75:14–23. doi: 10.1016/j.ejca.2016.12.039
10. Liu J, Tang LQ, Chen QY, Zhang L, Guo SS, Guo L, et al. The Prognostic Value of Plasma Epstein-Barr Viral DNA and Tumor Response to Neoadjuvant Chemotherapy in Advanced-Stage Nasopharyngeal

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Institutional Review Board of Yuebei People’s Hospital, Shantou University Medical College (KY-2020-020). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conceptualization, J-CL, S-QZ and S-MP. Methodology, S-QZ. Software, JL. Formal analysis, all authors. Investigation, all authors. Writing—original draft preparation, S-QZ. Writing—review and editing, J-CL. Supervision, H-BC. Project administration, J-CL. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by grants from the Shaoguan Science and Technology Plan Projects in 2020, China (Grant no. 200812094530421).

ACKNOWLEDGMENTS

We thank the individuals who participated in this study and donated samples.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.889516/full#supplementary-material
Chemotherapeutic Biomarkers of Nasopharyngeal Carcinoma

Zhang et al.

Frontiers in Oncology | www.frontiersin.org June 2022 | Volume 12 | Article 889516
this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Zhang, Pan, Lai, Situ, Liu, Dai, Liang, Zhou, Lu, Ke, Zhang, Chen and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.