Nano-scavengers for Blood Biomarker Discovery in Ovarian Carcinoma

Marilena Hadjidemetriou[1]*, Lana Papafilippou[1], Richard D. Unwin[2], Jane Rogan[3], Andrew Clamp[4] and Kostas Kostarelos[1,5]*

[1]*Nanomedicine Lab, Faculty of Biology, Medicine & Health, AV Hill Building, The University of Manchester, Manchester, , United Kingdom*

[2]*Division of Cardiovascular Sciences, Faculty of Biology, Medicine & Health, The University of Manchester, Manchester, United Kingdom*

[3]*Manchester Cancer Research Centre Biobank, The Christie NHS Foundation Trust, CRUK Manchester Institute, Manchester, United Kingdom*

[4]*Institute of Cancer Sciences and The Christie NHS Foundation Trust, Manchester Cancer Research Centre (MCRC), The University of Manchester, Manchester, United Kingdom*

[5]*Catalan Institute of Nanoscience and Nanotechnology (ICN2), Campus UAB, Bellaterra, 08193 Barcelona, Spain*

* Correspondence should be addressed to: marilena.hadjidemetriou@manchester.ac.uk
kostas.kostarelos@manchester.ac.uk

Abstract

The development and implementation of biomarker-based screening tools for ovarian cancer require novel analytical platforms to enable the discovery of biomarker panels that will overcome the limitations associated with the clinically used CA-125.The systematic discovery of protein biomarkers directly from human plasma using proteomics remains extremely challenging, due to the wide concentration range of plasma proteins. Here, we describe the use of lipid-based nanoparticles (NPs) as an 'omics' enrichment tool to amplify cancer signals in the blood and to uncover disease specific signatures. We aimed to exploit the spontaneous interaction of clinically-used liposomes (Caelyx®) with plasma proteins, also known as 'protein corona' formation, in order to facilitate the discovery of previously unreported differentially abundant molecules. Caelyx® liposomes were incubated with plasma samples obtained from advanced ovarian carcinoma patients and healthy donors and corona-coated liposomes were subsequently recovered. Comprehensive comparison between 'healthy' and 'diseased' corona samples by label-free proteomics resulted in the identification of multiple differentially abundant proteins. Moreover, immunoassay-based validation of selected proteins demonstrated the potential of nanoparticle-platform proposed to discover novel molecules with great diagnostic potential. This study proposes a nanoparticle-enabled workflow for plasma proteomic analysis in healthy and diseased states and paves the way for further work needed to discover and validate panels of novel biomarkers for disease diagnosis and monitoring.

Keywords: protein corona, biomarkers, ovarian cancer, liposomes, nanomedicine

Introduction

Much effort is currently focused on the development of robust and high-throughput 'omics' platforms for the discovery of minimally invasive molecular biomarkers to aid early and accurate cancer diagnosis, monitor tumour growth and response to therapies. Despite significant investment by major stakeholders, few protein cancer biomarkers have been validated and received FDA approval, raising concerns regarding the efficiency of the biomarker-development pipeline. It is noteworthy that of the FDA-approved biomarkers, the majority are used to monitor the progression of cancer, rather than enabling its early diagnosis.[1]

Proteins are the biological endpoints that govern most pathophysiological processes and they have therefore attracted most interest so far as biomarkers for cancer diagnostics.[2] Blood is frequently the biosample of choice for biomarker identification; however the discovery of tumour-derived protein signatures directly from blood is hindered by the wide concentration range of blood proteins, in addition to the preponderance of highly abundant proteins.[3] Despite significant improvement in the sensitivity of mass spectrometry-based proteomics, the issue of the high dynamic range of plasma protein abundances still remains unresolved and the diagnostic information blood can offer is partially inaccessible.[4]

Nanotechnology-based platforms hold great promise in addressing the above issues associated with biomarker discovery.[5] It should be emphasised however, that the vast majority of nanoparticle-based technologies developed so far have been designed to capture and quantify already known cancer-specific analytes,[6-9] enabling the verification and validation phases of biomarker development. The NP-enabled discovery of new plasma buried biomarkers has only been recently attempted.[10]

The fact that the surface of NPs is instantly covered by a wide range of adsorbed proteins and other biomolecules once in contact with blood, a self-assembly phenomenon known as 'protein' or 'biomolecule' corona formation, [11,12] makes NPs ideal biomarker discovery platforms. Biomolecule corona formation has become a popular line of research in the last decade and ongoing research is mainly focused on the proteomic analysis of corona profiles after the *ex vivo* and more recently the *in vivo* interaction of NPs with biofluids (mainly plasma).[13-17] Nanoparticle-protein interactions at the bio-nano interface not only can shed new light on the development of nanotechnologies but are now gradually being exploited as an engineering tool with therapeutic and diagnostic capabilities. [10,11,18-20]

The surface-capture of a complex blood proteome by NPs as well as the recently proposed concept of 'personalized corona' has sparked interest for utilizing the biomolecule corona

fingerprinting as a proteomic discovery platform.[10,18,21,22] We have recently demonstrated that the NP protein corona formed in the blood circulation of humans has the potential to be exploited as an enrichment and pre-fractionation tool that allows in depth coverage of the plasma proteome.[18] In a subsequent study, we employed two different tumour mouse models (a subcutaneous melanoma model and human lung carcinoma xenograft model) to demonstrate that intravenously injected lipid-based NP-scavengers (liposomes) surface-capture low MW, low abundant and disease-specific plasma proteins which cannot be detected by conventional plasma proteomic analysis.[10] Moreover, this study demonstrated that protein coronas, formed around intravenously injected NPs, differ both quantitatively and qualitatively in the presence and absence of a disease, allowing the uncovering of differentially abundant potential biomarker proteins.[10]

When animal models are employed for biomarkers discovery the exploitation of the molecularly richer *in vivo* protein corona is advantageous as opposed to its counterpart *ex vivo* corona.[16] However, hypothesis-free discovery proteomics often require the use of human clinical samples and therefore, in this study we aimed to explore the use of the *ex vivo* protein corona formed around the clinically used PEGylated liposomal doxorubicin formulation (Caelyx®), to identify disease-specific proteins directly from plasma samples, obtained from patients with recurrent ovarian carcinoma.

The work flow of this study is summarized in **Figure 1A** and involved the incubation of Caelyx® liposomes with plasma samples from patients with recurrent ovarian cancer and from healthy donors and the comprehensive comparison of the resultant protein coronas by label-free mass spectrometry. The above analysis led to in the discovery of 413 differentially abundant proteins between ‘healthy’ and ‘diseased’ corona samples, of which nine were quantified by immunoassays to further validate the potential use of the nanoparticle-protein corona technology for plasma proteomic analysis and biomarkers discovery.

.

Results

Recovery and purification of corona-coated liposomes from plasma samples obtained from ovarian carcinoma patients and healthy donors. To investigate the exploitation of the *ex vivo* formed NP protein corona for biomarker discovery, Caelyx® liposomes (20 μl of 1.5mM) were incubated with plasma samples (980 μl) obtained from patients with recurrent ovarian carcinoma about to commence the first cycle of Caelyx® as part of standard-of-care treatment (n=19) and age- matched female healthy donors (n=10). Patient clinical and basic blood analysis characteristics are summarized in **Tables S1** and **S2**. The physicochemical characteristics of the Caelyx® liposomes employed are summarized in **Table S3**. It should be noted that Caelyx® was employed because of its clinical use for the treatment of advanced ovarian cancer. The presence of the encapsulated doxorubicin has been shown not to affect the surface properties of liposomes and therefore corona formation.[15]

The *ex vivo* protein corona was allowed to form upon incubation of Caelyx® liposomes with plasma samples for 90 min at 37°C (**Figure 1A**). A purification protocol dependent on size exclusion chromatography was immediately performed to separate corona-coated liposomes from unbound plasma proteins. Membrane ultrafiltration was then used to concentrate the corona samples and to remove any large unbound or softly attached proteins, as previously optimised and described.[10,13-16,18] The above two-step purification process results in the complete elimination of unbound plasma proteins as demonstrated by plasma control experiments (**Figure S1**).

To confirm corona formation and to assess the morphology of Caelyx® liposomes before and after corona formation, Transmission Electron Microscopy (TEM) was performed. A well-dispersed liposome suspension was observed before and after incubation with plasma samples and purification. Corona-coated Caelyx® liposomes retained their size and spherical structure, while the occurrence of the proteins attached onto their surface revealed protein corona formation (**Figure 1B**).

To quantitatively compare 'healthy' and 'diseased' protein coronas, we calculated the total amount of protein associated with each μmole of lipid (Protein binding value; Pb). As shown in **Figure1C,** the average Pb value for ovarian carcinoma patients was 4 times higher than the average Pb value observed for healthy controls. These results are in agreement with our previous investigations in preclinical mouse models showing that protein corona fingerprints quantitatively differ in the absence and presence of tumorigenesis.[10]

Proteins associated with Caelyx® liposomes were separated by SDS-PAGE and visualized by Imperial Protein stain, as illustrated in **Figure 1D.** Despite the higher total amount of protein observed in the 'diseased' coronas, well distinct protein bands even at the low MW region were

observed, demonstrating the ability of the NP enrichment platform technology to minimise the noise of highly abundant proteins, such as albumin, and allow the interaction with low abundant proteins. The extensive purification processes applied to retrieve the corona-coated liposomes and purify them from the unbound proteins, worked as fractionation tool allowing the uncovering of the low MW plasma proteome (**Figure 1D**).

Proteomic comparison of 'healthy' and 'diseased' protein coronas. The goal of the proteomic discovery experiment was to comprehensively compare the 'healthy' and 'diseased' coronas in order to identify differentially abundant proteins.

In order to assess the reproducibility of the analysis of protein corona by LC-MS/MS, we isolated corona-coated liposomes from 6 aliquots of the same plasma sample (obtained from one healthy donor). The results demonstrated that the purification and quantification of the nanoparticle-bound protein fraction was reproducible, with 73% of proteins being measured with <30% CV (**Figure S2A**). To further validate the experimental reproducibility, two of the above replicated samples were analyzed in triplicates to evaluate the repeatability of the LC:MS/MS platform and the results demonstrated high analytical precision with approximately 95% of identified proteins with <30% CV (**Figure S2B & S3**). To assess the linearity of protein adsorption we incubated the same concentration of Caelyx® liposomes in full and diluted plasma. Interestingly, the total amount of corona proteins adsorbed onto the surface of liposomes was directly proportional to the total protein concentration in the incubation medium (**Figure S4**).

To compare 'healthy' and 'diseased' coronas, equal amounts of total protein from each corona sample were digested and subsequently analysed by LC-MS/MS. It should be emphasized that albumin and immunoglobulins were not depleted from corona samples prior to proteomic analysis. Processing of the raw data generated by LC-MS/MS analysis with Progenesis QI (version 3.0; Nonlinear Dynamics) software tool was carried out to statistically compare the abundance of proteins present in the 'healthy' and 'diseased' liposomal coronas. The Relative Protein Expression (fold change) and the reliability of measured differences (ANOVA, *p value*) were calculated. **Figures 2A & 2B** highlight the subset of differentially abundant proteins that met our confirmation criteria (see Experimental Section for further details). Out of 1187 identified proteins, 413 were found to be differentially abundant between the two groups with a *p value <0*.05, of which 171 were upregulated and 242 downregulated **(Figure 2B &Table S4)**. Considering the importance of achieving high confidence in the discovered proteins, we applied even more stringent criteria and interestingly out of the above 413 differentially abundant proteins 303 had a *p value*<0.01 (FDR of 2%) and a fold change> 2, which represents 25.5% of all proteins identified (**Figure 2B**).

The majority of highly specific cancer biomarkers are low MW intracellular proteins released from the tumour microenvironment into the blood circulation by leakage or secretion, however, their detection by plasma proteomic analysis remains challenging due to their extremely low concentration in the ng/ml to pg/ml range.[23] Interestingly, classification of the differentially abundant proteins identified in this study according to their cellular localization, demonstrated the enrichment of 189 intracellular proteins (present in the cytoplasm or nucleus) onto the surface of liposomes (**Figure S5**). In addition, ~50% of the differentially abundant proteins discovered had a MW<60 kDa (**Figure S6)**.

The above observation prompted us to investigate whether the differentially abundant corona proteins have been previously associated with ovarian carcinoma pathways. Disease and function IPA search revealed the association of 335 and 60 corona proteins with solid tumour pathways and metastasis processes, respectively. Interestingly, 72 proteins have been previously associated with ovarian cancer pathways, of which 15 have been described in the literature as potential biomarkers for ovarian cancer (n=8 for diagnosis; n=5 for unspecified applicability, n=1 for safety; and n=2 for efficacy), (**Figure 3**).

The plasma-incubated liposomes also surface-captured the clinically used blood biomarkers CA 125 (MUC 16), Transthyretin (TTR) and Apolipoprotein A1 (APOA1), all included in the FDA approved OVA1 diagnostic test which is used to evaluate ovarian masses for cancer prior to planned surgery.[24] In agreement with OVA1 test, CA125 and B2M were found to be upregulated in the 'diseased' corona whereas APOA1 was found to be downregulated. This suggests that the abundance of corona proteins, as calculated by LC-MS/MS analysis, directly reflects their concentration in blood.

Overall, the above data suggest that analysis of protein coronas formed after the *ex vivo* incubation of liposomes with plasma samples obtained from cancer patients and healthy controls can be used to uncover differentially abundant proteins, otherwise buried under the overwhelming signal of albumin.

Validation of the nanoparticle protein corona technology. To verify that the level of fold change observed by proteomic analysis of the 'healthy' and 'diseased' coronas is representative of the plasma proteome in healthy and diseased states, we performed ELISA experiments using plasma samples obtained from the same ovarian carcinoma patients and healthy controls. Distribution of the differentially abundant proteins identified by statistical significance and magnitude of change revealed that the majority exhibited fold change values much higher than the clinically used biomarkers CA125, TTR, and APOA1 (**Figure 2B & 4A**).

Nine corona proteins were selected to be validated and were divided in three groups: a) clinically used biomarkers for ovarian cancer (CA125, APOA1 and TTR; shown in orange); b) proteins mapped by IPA software to be associated with ovarian carcinoma pathways (THBS1, ENO1 and TGF-b1; shown in green) and c) proteins that have not been previously associated with ovarian carcinoma pathways but exhibited very promising fold change and *p values* (NME1, PDIA4 and PRKCSH; shown in blue). The fold change and *p values* of the nine selected proteins (as calculated by LC-MS/MS analysis) are illustrated in (**Figures 4A & S7**).

The plasma concentration profiles of selected proteins and their respective ROC curves, as calculated by ELISA experiments are shown in **Figure 4B**. In agreement with LC-MS/MS data (**Figures 4A & S7**), APOA1 and TTR proteins were found to be downregulated while CA125, THBS1,ENO1, TGF-b1, NME1, PDIA4 and PRKCSH proteins were found to be upregulated in ovarian carcinoma patients **(Figure 4B)**. This indicates that changes in the plasma proteome are directly reflected in the protein corona composition.

As illustrated in **Figure 4B**, the second group of proteins showed greater specificity and sensitivity than the clinically used biomarkers with AUC values ranging between 97.6% and 99.4%. Strikingly, an AUC value of 100% was observed for PDIA4 and PRKCSH proteins. The above ELISA validation data provide the first experimental evidence that *ex vivo* corona proteomic profiling provides allows increased penetration into the plasma proteome and has the potential to allow the discovery of candidate biomarkers.

Discussion

In the UK, 55-58% of ovarian carcinoma patients are diagnosed at stage III or IV and 42-45% are diagnosed at stage I or II. Survival for ovarian cancer is strongly related to the stage of the disease at diagnosis (99% of patients diagnosed at stage I survive their disease for at least one year, versus 51% of patients diagnosed at stage IV).[25] The lack of disease-specific symptoms, in addition to the limited performance of the clinically used CA-125 serum biomarker, indicates the need for new biomarker-based tools to accurately detect ovarian cancer and to monitor disease progression.[26]

Label-free proteomics profiling of blood is a powerful tool to detect molecular biomarkers that are differentially expressed between healthy and disease states.[27] The concentration of a complex network of proteins that regulate tumorigenic pathways is often altered in the blood circulation of cancer patients; however their identification is hampered by the wide dynamic range of plasma proteins. Due to the limited analytical sensitivity, currently available mass spectrometry-based

approaches predominantly detect highly abundant proteins of limited diagnostic use and fail to detect low MW tumour-tissue derived proteins of lower abundance. To overcome the issue of albumin masking, plasma immunodepletion columns are extensively employed, however their use is limited by a significant loss of low MW proteins along with the highly abundant carrier plasma proteins. Therefore, the multifaceted process of tumorigenesis necessitates the development of 'omics-enrichment' platforms that will enable the discovery of biomarker panels with sufficient specificity and sensitivity.

We have recently proposed the use of nanoparticles (NPs) as protein-scavenging enrichment platforms to address the above fundamental issues associated with biomarker discovery in plasma. Our results demonstrated that intravenously administered lipid-based NPs, in mice and humans, were able to surface-capture and amplify low abundant tumour-released molecules that could not be detected by conventional plasma proteomics analysis.[10,18] This idea of using nanoparticles to allow an in depth-analysis of the blood proteome is based on the spontaneous and non-targeted adsorption of hundreds of proteins onto the nanoparticles surface once in contact with biological fluids, a phenomenon known as 'protein corona' formation.[11]

Herein, we aimed to further explore and validate the potential use of the nanoparticle-protein corona to discover novel disease-specific proteins from plasma samples obtained from ovarian carcinoma patients (**Figure 1A and Tables S1 & S2**). Even though analysis of the *in vivo* formed protein corona (after intravenous administration of NPs) has been shown to result in a richer sampling of the blood proteome,[16] exploitation of the *ex vivo* protein corona (after incubation of NPs with plasma samples) could be more easily incorporated in the discovery phase of the biomarker pipeline and deserves further investigation. It should be emphasized that unlike other nanoparticle-based technologies aiming to increase the sensitivity of detection of already known molecules,[8,28] the technology platform proposed here aims to identify previously unseen potential biomarker proteins that could potentially offer higher specificity and sensitivity than the clinically used biomarkers.

To assess the potential of the NP protein corona technology as a tool for biomarker discovery, we comprehensively compared protein coronas formed around a clinically used liposomal formulation, Caelyx®, upon incubation with plasma samples obtained from healthy donors (n=10) and from patients with recurrent ovarian cancer (n=19). Immediately after plasma incubations, corona-coated liposomes were purified from any unbound plasma proteins and only liposome-bound proteins were analysed by LC-MS/MS. It should be emphasized that albumin and immunoglobulins were not depleted from corona samples prior to LC-MS/MS analysis. Even though a fraction of highly abundant proteins interacts with the surface of liposomes, any unbound highly abundant proteins are removed by the purification process which addresses the 'signal-to-

noise' issue that biomarker discovery suffers from. We therefore propose that analysis of the nanoparticle protein corona can substitute plasma fractionation and immunodepletion methodologies.

In agreement with our previous studies in tumour-bearing mice,[10] we observed a significantly higher total amount of protein adhered onto the NPs surface in ovarian carcinoma patients in comparison to healthy controls (**Figure 1C**). This observation, not only confirms our hypothesis that protein corona is greatly affected by the ongoing tumorigenesis but it also paves the way for the development of diagnostic tests. More studies are needed to assess if the fluctuation in the amount of protein adsorbed onto the NPs surface can be used to indicate the onset of a disease or to monitor disease progression and response to the treatment.

Despite the higher amount of protein adhered onto the NPs surface after incubation with plasma samples obtained from ovarian carcinoma patients, gel electrophoresis experiments indicated that the analysis of the NP protein corona eliminates the issue of albumin masking and enables the analysis of a broad range of the plasma proteins including low MW proteins (**Figure 1D & Figure S5**). Subsequent analysis of the NPs protein coronas by LC-MS/MS revealed 413 proteins that were differentially abundant between ovarian carcinoma patients and healthy controls (with a *p value* < 0.05), of which 171 were under-expressed and 242 over-expressed (**Figures 2A & B**). Recent ongoing biomarker development efforts indicate that multiple markers, used individually or as part of a panel, are required to provide sufficient sensitivity and specificity. Noteworthy, although the majority of proposed cancer biomarkers are proteins found to be upregulated in the blood circulation of cancer patients, downregulated biomarkers are currently clinically used and should be taken into consideration.[26]

According to Ingenuity Pathway Analysis, out of 413 differentially abundant proteins discovered in this study 57 have been previously associated with ovarian carcinoma pathways and only 15 have been previously proposed as potential biomarkers for ovarian cancer (**Figure 3**). Moreover, distribution of the differentially abundant proteins identified by statistical significance and magnitude of change (**Figure 2B)** revealed that n=303 proteins had a *p value*<0.01 and a fold change> 2. It should be noted that even though the clinically used biomarkers CA125, TTR and APOA1 were found to interact with the surface of liposomes, the majority of differentially abundant proteins identified, exhibited higher fold change values with higher statistical significance in comparison to the clinically used biomarkers (**Figure 3B**). These results suggest that analysis of the NP protein corona can unveil previously unseen disease-specific molecules, otherwise buried under the overwhelming signal of albumin and immunoglobulins. More studies are needed to prove the ability of this nanoplatform to enrich disease-specific molecules at the early asymptomatic

stages of cancer. Moreover, pathway analysis of the 'diseased' nanoparticle corona could provide valuable information about the ongoing pathological pathways and the mechanism of cancer initiation and progression and could potentially lead to discovery of novel therapeutic strategies.

The identification of previously unknown disease-specific proteins prompted us to verify the differences observed between 'healthy' and 'diseased' coronas by commercially available ELISA kits, using plasma samples obtained from the same ovarian carcinoma patients and healthy controls. We chose to validate and compare 3 groups of proteins: a) clinically used biomarkers for ovarian cancer (CA125, APOA1 and TTR), b) proteins mapped by IPA software to be associated with ovarian carcinoma pathways (THBS1, ENO1 and TGF-b1) and c) proteins that have not been previously associated with ovarian carcinoma pathways (NME1, PDIA4 and PRKCSH). ELISA data were found to directly reflect the changes observed by LC-MS/MS analysis of 'healthy' and 'diseased' corona samples (**Figure 4**), indicating that protein corona composition mirrors the concentration fluctuations of the plasma proteome in the presence of a disease. Interestingly, ELISA quantification of the last two groups of proteins revealed higher performance than the clinically used biomarkers with AUC values between 97.6% and 99.4%, while an AUC value of 100% was shown for PDIA4 and PRKCSH proteins. Given the same plasma source was used for the corona formation and ELISA experiments, the above data represent solely an orthogonal validation of the LC-MS/MS data. Clearly, more validation studies will be required to prove the ability of the proposed proteins to discriminate between ovarian carcinoma patients and healthy controls.

Although the samples employed in this study were obtained from advanced ovarian cancer patients and have limited value for the discovery of screening biomarkers, the above validation data provide the first experimental evidence of the exploitation of the NP protein corona, formed *ex vivo* in human clinical samples, for the discovery of potential biomarker proteins. Considering the low number of samples used in this study, it should be emphasized that the clinical utility of the differentially abundant proteins identified as screening or monitoring biomarkers will require validation in much larger and well-defined patient cohorts and with the appropriate control groups (i.e. benign gynaecological conditions).

Collectively, our results suggest that NPs dispersed in biological fluids have the potential to be used as an enrichment 'omics' platform for biomarker discovery. It is now well established that the physicochemical properties of NPs directly affect the composition of protein corona[11] and more work is needed to investigate whether the use of other types of NPs and/or combinations of different NPs will further increase the range of plasma proteome detected.

Conclusion

In this study, we propose the use of lipid-based nanoparticles as ‘omics’ enrichment platforms to reveal disease specific signatures in the blood of ovarian carcinoma patients. We demonstrate that the molecular composition of protein corona, spontaneously formed around NPs upon incubation with plasma samples, reflects the concentration fluctuations of the blood proteome in the presence of tumorigenesis. Comprehensive comparison between the *ex vivo* formed ‘healthy’ and ‘diseased’ protein coronas by label-free proteomics (LC-MS/MS), revealed the discovery of 413 differentially abundant proteins. Subsequent immunoassay- based validation demonstrated the potential of the nanoparticle-platform proposed to identify novel potential biomarker proteins. This work is thought to pave the way for many more studies needed to allow the clinical exploitation of protein corona fingerprinting as a novel tool to track tumours over time and discover panels of novel biomarkers for early and accurate disease diagnosis.

Experimental

Ethical Approvals. This project was reviewed and approved by the Manchester Cancer Research Centre Biobank Sample Access Committee and all sample collection was conducted under the MCRC Biobank Research Tissue Bank Ethics (ref: 07/H1003/161+5).

Blood sample collection. Eligible cases for this study included women with recurrent ovarian cancer commencing Caelyx® chemotherapy as part of standard chemotherapeutic management for disease progression. Caelyx® contains 2mg/ml doxorubicin hydrochloride encapsulated in a PEGylated liposomal formulation (16 mg lipid content) and is indicated for the treatment of advanced ovarian cancer in women who have failed a first-line platinum-based chemotherapy. Plasma samples (before cycle 1 Caelyx® infusion) were collected into commercially available anticoagulant-treated tubes (K2 EDTA BD Vacutainer®). Plasma was then prepared by inverting the collection tubes 10 times to ensure mixing of blood with EDTA and subsequent centrifugation for 12 minutes at 1300 RCF at 4 °C. Following centrifugation supernatant was immediately collected into labelled Protein LoBind Eppendorf tubes and stored in -80^{0}C. Age-matched plasma samples from healthy female donors (n=2 Caucasian; n=4 Black; n=4 Hispanic) were purchased from Seralab UK (LOT BRH1221742-BRH1221751). Considering the impact of the anticoagulant agent on the formation of the protein corona,[29] healthy plasma samples contained the same anticoagulant agent (K2 EDTA BD Vacutainer® tubes) as that described above for the human clinical samples and were subjected to the same preparation protocol (centrifugation for 12 min at 1300 rpm at 4 ^{0}C). Healthy human plasma samples were received on dry ice and were stored in a -80^{0}C. Finally, samples were thawed only before the incubations.

***Ex vivo* protein corona formation.** To investigate the *ex vivo* protein corona, Caelyx® liposomes were incubated with plasma samples obtained from recurrent ovarian cancer patients and from healthy donors. Caelyx® liposomes (20 ul of 0.15 mM) were incubated with 980 ul of plasma for 90 min at 37°C in orbital shaker at 250 rpm. The *ex vivo* protein corona was allowed to form using the same liposome concentration (0.15 mM) as that extracted in 1 mL of plasma from intravenously injected patients.[18] This liposome concentration results in a final sample protein concentration (upon purification of corona-coated liposomes) that allows in gel digestion of 20ug of protein/sample and subsequent LC-MS/MS analysis. Our previous time evolution studies of the nanoparticle protein corona demonstrated that a complex protein corona forms as early as 10 min post-incubation and does not quantitatively change over time.[15] In the present study, we chose to incubate liposomes for 90 min, which reflects the time of Caelyx® infusion in ovarian carcinoma patients.

Separation of corona-coated liposomes from unbound and weakly bound proteins. Corona-coated liposomes were separated as we have previously described.[15,16] Briefly, *ex vivo* incubated liposomes were separated form unbound plasma proteins by size exclusion chromatography followed by membrane ultrafiltration. Immediately after incubation, samples (1ml) were loaded onto a Sepharose CL-4B (Sigma-Aldrich) column (15cm) equilibrated with HEPES buffer. Fractions containing liposomes (4,5,6) were then pooled together and concentrated to 500 µL using a Vivaspin 6 column (10 000 MWCO, Sartorious, Fisher Scientific) at 9000 rpm. Vivaspin 500 centrifugal concentrator (1 000k MWCO, Sartorious, Fisher Scientific) was then used at 9000 rpm, to further concentrate the samples to 100 µL and to ensure separation of protein-coated liposomes from the remaining large unbound proteins. Corona-coated liposomes were then washed 3 times with 100 µL HEPES buffer to remove weekly bound proteins. To validate the separation of corona-coated liposomes from unbound proteins, the same procedure was performed with controls of plasma samples (without prior incubation with liposomes (**Figure S1**).

Transmission Electron Microscopy (TEM). Bare and corona-coated liposomes were stained by uranyl acetate solution 1% and visualized with transmission electron microscopy (FEI Tecnai 12 BioTwin) before and after their *in vivo* interaction with plasma proteins. Samples were diluted to 0.5 mM lipid concentration and carbon Film Mesh Copper Grids (CF400-Cu, Electron Microscopy Science) were used.

SDS-PAGE electrophoresis. Proteins associated with 0.025 µM of liposomes were loaded onto a 4-20% NOVEX Tris-Glycine Protein Gel (ThermoFisher Scientific). The gel was run until the proteins neared the end

of the gel (25-40 minutes at 225V). Staining was performed with Imperial Gel Staining reagent (Sigma Life Science).

Quantification of adsorbed proteins. Proteins associated with recovered liposomes were quantified by BCA Protein assay kit according to manufacturer's instructions. To make sure that liposomes in solution do not interfere with the absorbance at 562 nm we measured the absorbance of corona-coated liposomes in HEPES buffer and subtracted it from the total absorbance, measured when corona-coated liposomes were mixed with the BCA reagent. Lipid concentration was quantified by Stewart assay and Protein binding ability; Pb values (μg of protein/μM lipid) were then calculated.

Mass Spectrometry. In-gel digestion of corona proteins (20ug/sample) was performed prior to LC-MS/MS analysis, as we have previously described.[15,16,30] Before SDS-PAGE samples were boiled for 5 min at 90°C in the presence of Tris-Glycine SDS Sample Buffer and NuPAGE reducing agent (ThermoFisher). Digested samples were analysed by LC-MS/MS using an UltiMate® 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to a Q Exactive™ Hybrid Quadrupole-Orbitrap™ (Thermo Fisher Scientific, Waltham, MA) mass spectrometer.

To assess the repeatability of the sample processing workflow we incubated Caelyx® liposomes with plasma obtained from a healthy donor and repeated the same protocol for 6 times/replicates (**Figure S2A**). To assess the analytical variation/repeatability of the LC:MS/MS platform two of the samples used above to determine the repeatability of the sample processing workflow were run in triplicates (**Figure S2B & S3**). To assess the linearity of the method we incubated the same concentration of Caelyx® liposomes (0.15 mM) in full and diluted plasma (**Figure S4**).

Mass Spectrometry data analysis. To statistically compare the abundance of proteins identified in the 'healthy' and 'diseased' coronas MS peak intensities were analyzed using Progenesis LC-MS software (version 3.0; Nonlinear Dynamics). RAW files were imported into Progenesis LC-MS software (version 3.0; Nonlinear Dynamics) with automatic feature detection enabled. The Progenesis QI default method of normalisation was applied ('Normalise to all proteins') to compensate for experimental variations. A representative reference run was selected automatically, to which all other runs were aligned in a pair-wise manner. Automatic processing was selected to run with applied filters for peaks charge state (maximum charge 5). Protein quantitation method was selected to be the relative quantitation using Hi-N with N=3 peptides to measure per protein. The resulting MS/MS peak lists were exported as a single Mascot generic file and upload onto a local Mascot Server (version 2.3.0; Matrix Science, UK). The spectra were searched against the UniProt database using the following parameters: tryptic enzyme digestion with one missed cleavage allowed, peptide charge of +2 and +3, precursor mass tolerance of 15 mmu, fragment mass tolerance of 8 ppm, oxidation of methionines as variable modifications and carbamidomethyl as fixed modifications, with decoy database search disabled and ESI-QUAD-TOF the selected instrument. Each search produced an XML file from Mascot and the resulted peptides (XML files) were imported back into Progenesis LC-MS to assign peptides to features. The peptide intensities were compared between groups by one way analysis of variance. Subsequently data were exported in Excel format. Finally, results were filtered to present a mean normalized abundance of more than 50,000 in at least one of the two groups.

Mass Spectrometry data were also analysed with QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity). Diseases and functions IPA tool was used to identify proteins involved in ovarian carcinoma pathways. The biomarker overlay IPA tool was then used to identify proteins described in the literature as potential biomarkers for ovarian cancer.

Enzyme-linked Immunosorbent Assay (ELISA). ELISA kits for human CA125 (MUC16, ab195213, Abcam, UK), apolipoprotein AI (APOAI, ab189576, Abcam, UK), prealbumin (Transthyretin TTR, ab108895, Abcam, UK), thrombospondin 1 (THBS1, ab193716, Abcam, UK), alpha-enolase (ENO1, ab181417, Abcam, UK), transforming growth factor beta 1 (TGF-β1, DB100B, R&D Systems Europe, LTD.), nucleoside diphosphate kinase A (NME1, orb406403, Biorbyt Ltd., UK), glucosidase 2 subunit beta (PRKCSH, EH2259, Wuhan Fine Biotech Co., Ltd.) and disulphide-isomerase A4 (PDIA4, abx250438, Abbexa Ltd., UK) were purchased for the quantitative measurement of each human protein in plasma. Experiments were performed according to manufacturer's instructions

Statistical analysis. Statistical analysis of the data was performed using GraphPad Prism software. Mann-Whitney t-test was used for the quantification of the total amount of protein adsorbed (Pb values of **Figure 1C**) and for ELISA experiments (**Figure 4B**).

Author contributions

M.H. initiated, designed and performed the experiments, analyzed all data and took responsibility for planning and writing the manuscript. L.P. contributed to the analysis of the mass spectrometry data. R.U. provided guidance on the proteomic data analysis. J.R. and A.C. contributed to the clinical design, provided oversight of the ethical approval process and were responsible for access and storage of the samples in the MCRC Biobank. K. K. provided intellectual input throughout the study and contributed to the writing of the manuscript.

Data availability

The raw/processed data required to reproduce these findings are all included in the supporting information.

Acknowledgements

This research was partially funded by the CRUK Pioneer Award (C54921/A25435). We would like to acknowledge the CRUK/NIHR Manchester Experimental Cancer Medicine Centre for providing research nurse funding and their clinical infrastructure used for the blood sampling. This work was partially funded by the CRUK Centre (C147/A18083) and Major Centre Award (C147/A25254) awarded to the Manchester Cancer Research Centre. We would like to acknowledge the CRUK/NIHR Manchester Experimental Cancer Medicine Centre for providing research nurse funding and their clinical infrastructure used for the blood sampling. Authors wish to acknowledge Manchester Cancer Research (MCRC) Biobank team for their assistance in the collection of blood samples. The authors also would like to thank the Faculty of Life Sciences EM Facility at the University of Manchester for their assistance in Electron Microscopy imaging. In addition, Biological Mass Spectrometry (Bio-MS) Facility staff at the University of Manchester for their support. We would also like to thank the patients for their kind donation of blood samples to support this research.

References

1 Fuzery, A. K., Levin, J., Chan, M. M. & Chan, D. W. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. *Clin Proteom* **10**, doi:Artn 13 10.1186/1559-0275-10-13 (2013).

2 Anderson, N. L. & Anderson, N. G. The human plasma proteome: history, character, and diagnostic prospects. *Molecular & cellular proteomics : MCP* **1**, 845-867 (2002).

3 Hanash, S. M., Pitteri, S. J. & Faca, V. M. Mining the plasma proteome for cancer biomarkers. *Nature* **452**, 571-579, doi:10.1038/nature06916 (2008).

4 Rifai, N., Gillette, M. A. & Carr, S. A. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat biotechnol* **24**, 971-983, doi:10.1038/nbt1235 (2006).

5 Liotta, L. A., Ferrari, M. & Petricoin, E. Clinical proteomics: written in blood. *Nature* **425**, 905, doi:10.1038/425905a (2003).

6 Tam, J. O. *et al.* A comparison of nanoparticle-antibody conjugation strategies in sandwich immunoassays. *J Immunoassay Immunochem* **38**, 355-377, doi:10.1080/15321819.2016.1269338 (2017).

7 Huo, Q. *et al.* A facile nanoparticle immunoassay for cancer biomarker discovery. *J Nanobiotechnology* **9**, 20, doi:10.1186/1477-3155-9-20 (2011).

8 Luchini, A. *et al.* Smart hydrogel particles: biomarker harvesting: one-step affinity purification, size exclusion, and protection against degradation. *Nano letters* **8**, 350-361, doi:10.1021/nl072174l (2008).

9 Moro, L., Turemis, M., Marini, B., Ippodrino, R. & Giardi, M. T. Better together: Strategies based on magnetic particles and quantum dots for improved biosensing. *Biotechnol Adv* **35**, 51-63, doi:10.1016/j.biotechadv.2016.11.007 (2017).

10 Hadjidemetriou, M., Al-Ahmady, Z., Buggio, M., Swift, J. & Kostarelos, K. A novel scavenging tool for cancer biomarker discovery based on the blood-circulating nanoparticle protein corona. *Biomaterials* **188**, 118-129, doi:10.1016/j.biomaterials.2018.10.011 (2019).

11 Hadjidemetriou, M. & Kostarelos, K. Nanomedicine: Evolution of the nanoparticle corona. *Nat Nanotechnol* **12**, 288-290, doi:10.1038/nnano.2017.61 (2017).

12 Cedervall, T. *et al.* Understanding the nanoparticle-protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. *P Natl Acad Sci USA* **104**, 2050-2055, doi:DOI 10.1073/pnas.0608582104 (2007).

13 Al-Ahmady, Z. S., Hadjidemetriou, M., Gubbins, J. & Kostarelos, K. Formation of protein corona in vivo affects drug release from temperature-sensitive liposomes. *J Control Release* **276**, 157-167, doi:10.1016/j.jconrel.2018.02.038 (2018).

14 Garcia-Alvarez, R., Hadjidemetriou, M., Sanchez-Iglesias, A., Liz-Marzan, L. M. & Kostarelos, K. In vivo formation of protein corona on gold nanoparticles. The effect of their size and shape. *Nanoscale* **10**, 1256-1264, doi:10.1039/c7nr08322j (2018).

15 Hadjidemetriou, M., Al-Ahmady, Z. & Kostarelos, K. Time-evolution of in vivo protein corona onto blood-circulating PEGylated liposomal doxorubicin (DOXIL) nanoparticles. *Nanoscale* **8**, 6948-6957, doi:10.1039/c5nr09158f (2016).

16 Hadjidemetriou, M. *et al.* In Vivo Biomolecule Corona around Blood-Circulating, Clinically Used and Antibody-Targeted Lipid Bilayer Nanoscale Vesicles. *ACS nano* **9**, 8142-8156, doi:10.1021/acsnano.5b03300 (2015).

17 Chen, F. *et al.* Complement proteins bind to nanoparticle protein corona and undergo dynamic exchange in vivo. *Nat Nanotechnol* **12**, 387-393, doi:10.1038/nnano.2016.269 (2017).

18 Hadjidemetriou, M. *et al.* The Human In Vivo Biomolecule Corona onto PEGylated Liposomes: A Proof-of-Concept Clinical Study. *Adv Mater*, e1803335, doi:10.1002/adma.201803335 (2018).

19 Ke, P. C., Lin, S., Parak, W. J., Davis, T. P. & Caruso, F. A Decade of the Protein Corona. *ACS nano* **11**, 11773-11776, doi:10.1021/acsnano.7b08008 (2017).

20 Westmeier, D. *et al.* Nanomaterial-microbe cross-talk: physicochemical principles and (patho)biological consequences. *Chem Soc Rev* **47**, 5312-5337, doi:10.1039/c6cs00691d (2018).

21 Miotto, G. *et al.* Protein corona as a proteome fingerprint: The example of hidden biomarkers for cow mastitis. *Colloids Surf B Biointerfaces* **140**, 40-49, doi:10.1016/j.colsurfb.2015.11.043 (2016).

22 Magro, M. *et al.* Analysis of hard protein corona composition on selective iron oxide nanoparticles by MALDI-TOF mass spectrometry: identification and amplification of a hidden mastitis biomarker in milk proteome. *Anal Bioanal Chem* **410**, 2949-2959, doi:10.1007/s00216-018-0976-z (2018).
23 Schiess, R., Wollscheid, B. & Aebersold, R. Targeted proteomic strategy for clinical biomarker discovery. *Mol Oncol* **3**, 33-44, doi:10.1016/j.molonc.2008.12.001 (2009).
24 Nolen, B. M. & Lokshin, A. E. Biomarker testing for ovarian cancer: clinical utility of multiplex assays. *Mol Diagn Ther* **17**, 139-146, doi:10.1007/s40291-013-0027-6 (2013).
25 *Office for National Statistics. Cancer survival by stage at diagnosis for England (experimental statistics)* , 2016).
26 Nolen, B. M. & Lokshin, A. E. Protein biomarkers of ovarian cancer: the forest and the trees. *Future Oncol* **8**, 55-71, doi:10.2217/fon.11.135 (2012).
27 Sandin, M., Chawade, A. & Levander, F. Is label-free LC-MS/MS ready for biomarker discovery? *Proteomics Clin Appl* **9**, 289-294, doi:10.1002/prca.201400202 (2015).
28 Longo, C. *et al.* Core-shell hydrogel particles harvest, concentrate and preserve labile low abundance biomarkers. *PloS one* **4**, e4763, doi:10.1371/journal.pone.0004763 (2009).
29 Schottler, S., Klein, K., Landfester, K. & Mailander, V. Protein source and choice of anticoagulant decisively affect nanoparticle protein corona and cellular uptake. *Nanoscale* **8**, 5526-5536, doi:10.1039/c5nr08196c (2016).
30 Al-Ahmady, Z. S., Hadjidemetriou, M., Gubbins, J. & Kostarelos, K. Formation of protein corona in vivo affects drug release from temperature-sensitive liposomes. *Journal of Controlled Release* **276**, 157-167, doi:10.1016/j.jconrel.2018.02.038 (2018).

FIGURE LEGENDS

Figure 1*: Protein corona formation after the ex vivo* incubation of PEGylated, doxorubicin-encapsulated liposomes (Caelyx®) with plasma samples obtained from healthy controls (n=10) and ovarian carcinoma patients (n=19). (A) Schematic description of the experimental design. Caelyx® liposomes were incubated *ex vivo* with plasma samples obtained from patients with recurrent ovarian carcinoma (n=19) and from healthy donors (n=10) for 90 min at 37°C. Corona-coated liposomes were isolated and purified from unbound proteins by size exclusion chromatography and membrane ultrafiltration. 'Healthy' and 'diseased' protein coronas were comprehensively characterized and compared by label-free mass spectrometry to identify differentially abundant potential biomarker proteins. Selected potential biomarker proteins were further validated by commercially available ELISA kits. **(B)** Negative stain TEM of liposomes before and after corona formation. All scale bars are 100nm. **(C)** The total amount of protein adsorbed onto the surface of liposomes recovered from plasma samples obtained from healthy and ovarian carcinoma patients expressed as Pb values (µg of protein/µM lipid). Individual biological replicates are shown. Error bars indicate mean +/− SEM. (Mann-Whitney t-test; ****indicates *p<0.0001*). **(D)** Imperial stained SDS-PAGE gel of representative 'healthy' and 'diseased' corona samples.

Figure 2: Proteomic comparison of the liposomal protein coronas formed in plasma samples obtained from ovarian carcinoma patients (n=19) and healthy controls (n=10). MS peak intensities were analyzed using Progenesis LC-MS software (version 3.0; Nonlinear Dynamics). Results were filtered to present a mean normalized abundance of more than 50,000 in at least one of the two groups. The peptide intensities were compared between groups by one way analysis of variance (ANOVA). **(A)** Heatmap of Normalized Abundance (NA) values of proteins found to be differentially expressed between 'healthy' and 'diseased' coronas. Only proteins with *p value < 0.05* are shown (n=413). Proteins are classified from highest to the lowest max fold change. Average abundance of each protein for each group is also shown. The full list of differentially abundant corona proteins and their respective mean normalized abundance, *p value* and max fold change are shown in **Table S4. (B)** Volcano plot displays the relationship between fold change and significance between the two groups. The y-axis depicts the negative log10 of p-values and the x-axis is the difference in expression between the two groups as log10 fold changes. Only proteins with at least 2-fold change and a p value<0.01 *value* are highlighted (n=303; downregulated proteins are shown in red and upregulated proteins in blue). **Figure 3**: **Ingenuity Pathway Analysis (IPA) of potential biomarker corona proteins.** Out of 413 potential biomarker proteins n=15 (shown in orange) were previously reported as potential biomarkers for ovarian cancer and n=57(shown in green) have been previously associated with ovarian carcinoma pathways. The name of proteins illustrated in the diagram and their respective gene symbols are shown in **Table S5.**

Figure 4: **ELISA validation of the nanoparticle protein corona technology. (A)** Scatter plot displays the relationship between fold change and significance of the nine potential biomarker proteins selected to be further validated. The y-axis depicts the negative log10 of p-values and the x-axis is the difference in expression between the two groups as log10 fold changes. Clinically used biomarkers are shown in orange, proteins previously associated with ovarian carcinoma pathways (according to IPA) are shown in green and proteins that have not been previously associated with ovarian carcinoma pathways are shown in blue. **(B)** Plasma concentration profiles of selected potential biomarker proteins in healthy controls (n=10) and ovarian carcinoma patients (n=15-17) and their respective ROC curves based on ELISA assays. AUC values are also shown; Mann-Whitney t-test; * indicates p<0.05, *** indicates p < 0.001, **** indicated p<0.0001.

(CA125=mucin 16, APOA1=apolipoprotein A1, TTR=transthyretin, THBS1=thrombospondin 1, ENO1=enolase 1, TGF-b1= transforming growth factor b1, NME1= nucleoside diphosphate kinase A, PRKCSH=glucosidase 2 subunit beta and PDIA4= protein disulfide-isomerase A4).

FIGURE 1

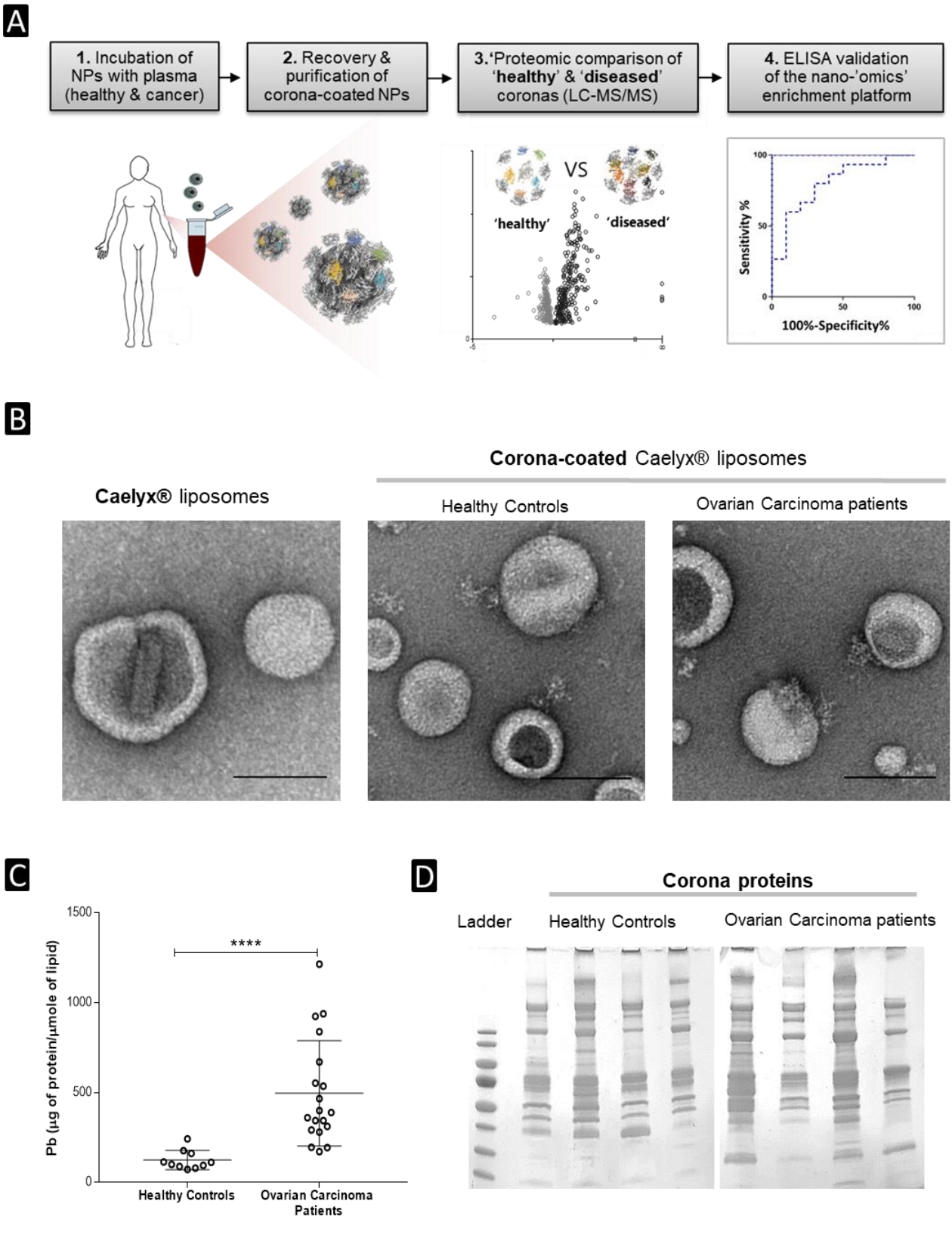
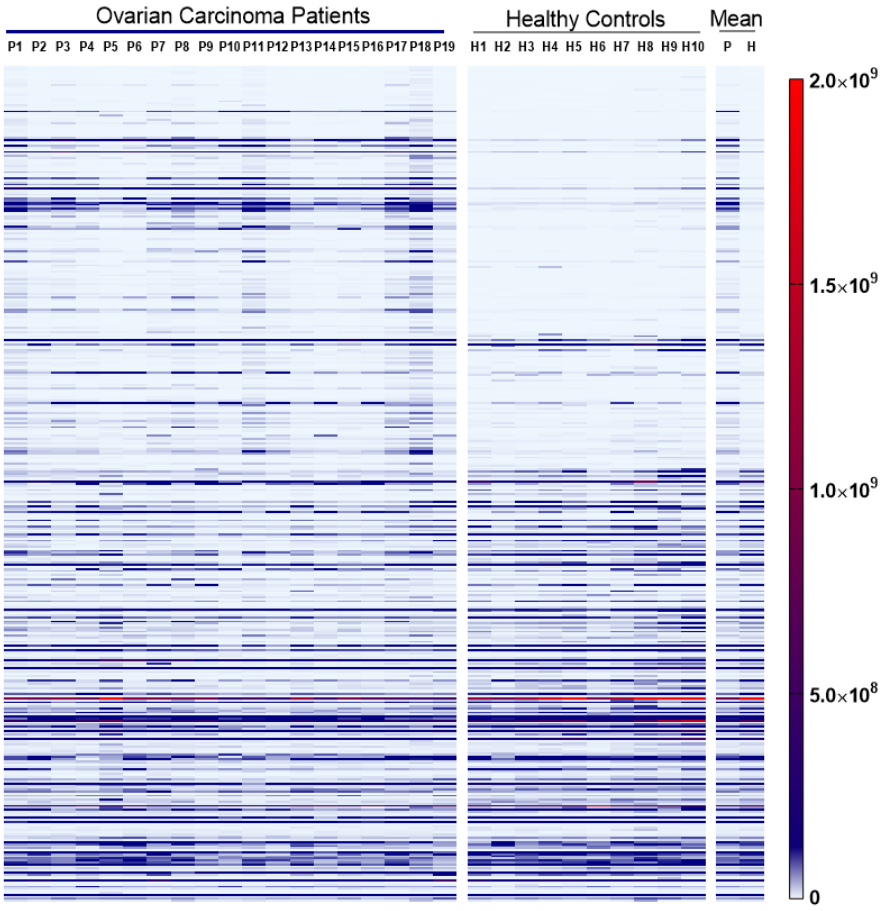

A
1. Incubation of NPs with plasma (healthy & cancer)
2. Recovery & purification of corona-coated NPs
3.'Proteomic comparison of 'healthy' & 'diseased' coronas (LC-MS/MS)
4. ELISA validation of the nano-'omics' enrichment platform
VS
'healthy'
'diseased'
Sensitivity %
100%-Specificity%
B
Caelyx® liposomes
Corona-coated Caelyx® liposomes
Healthy Controls
Ovarian Carcinoma patients
C
Pb (µg of protein/µmole of lipid)
****
Healthy Controls
Ovarian Carcinoma Patients
D
Corona proteins
Ladder
Healthy Controls
Ovarian Carcinoma patients

FIGURE 2

A


Ovarian Carcinoma Patients
Healthy Controls
Mean
P1 P2 P3 P4 P5 P6 P7 P8 P9 P10 P11 P12 P13 P14 P15 P16 P17 P18 P19
H1 H2 H3 H4 H5 H6 H7 H8 H9 H10
P H
2.0×10^9
1.5×10^9
1.0×10^9
5.0×10^8
0

B

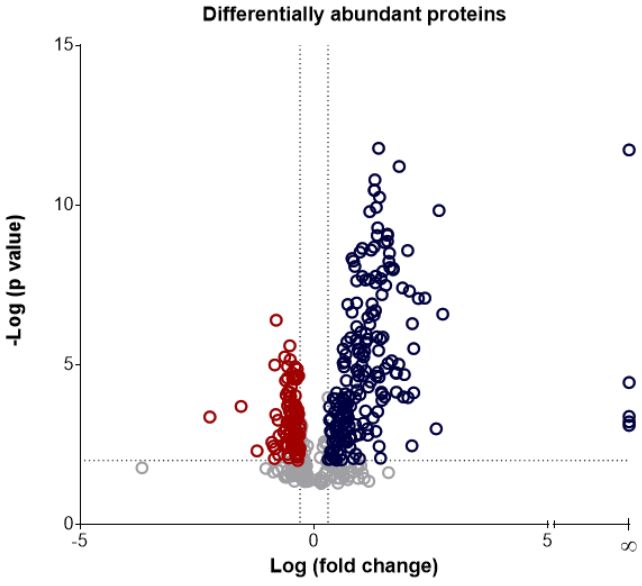

Differentially abundant proteins
-Log (p value)
15
10
5
0
-5
0
5
∞
Log (fold change)

FIGURE 3

Corona proteins previously reported as **potential biomarkers** for ovarian cancer (n=15):

Corona proteins previously reported to be **associated with ovarian carcinoma pathways** (n=57):

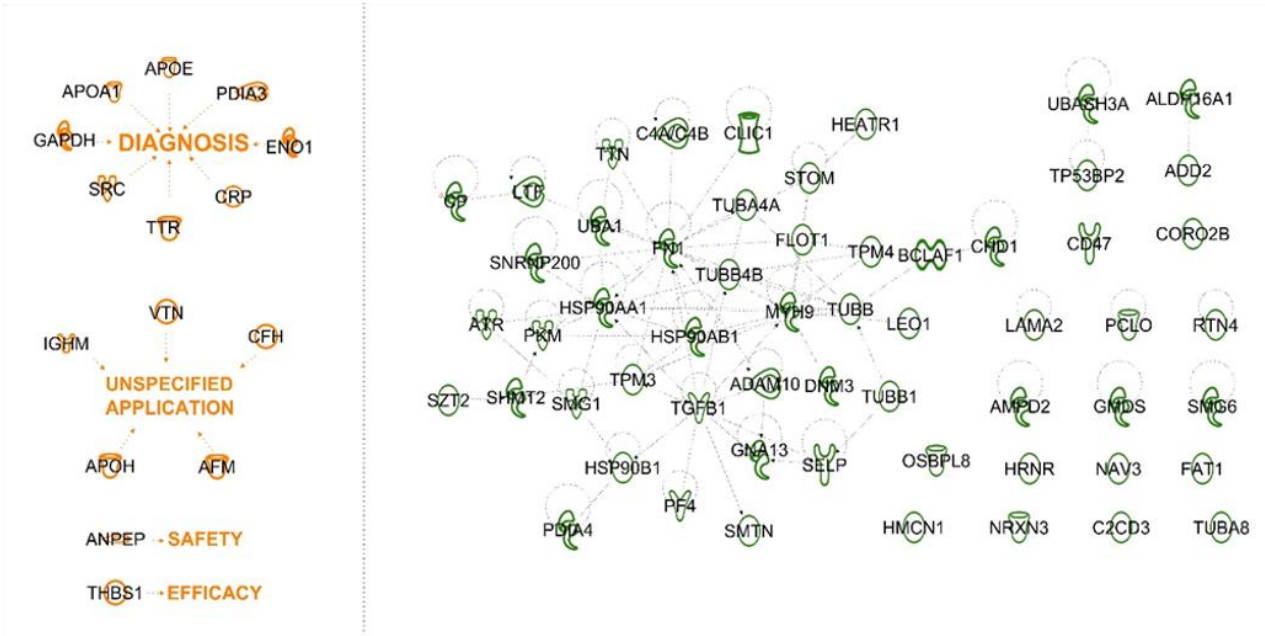
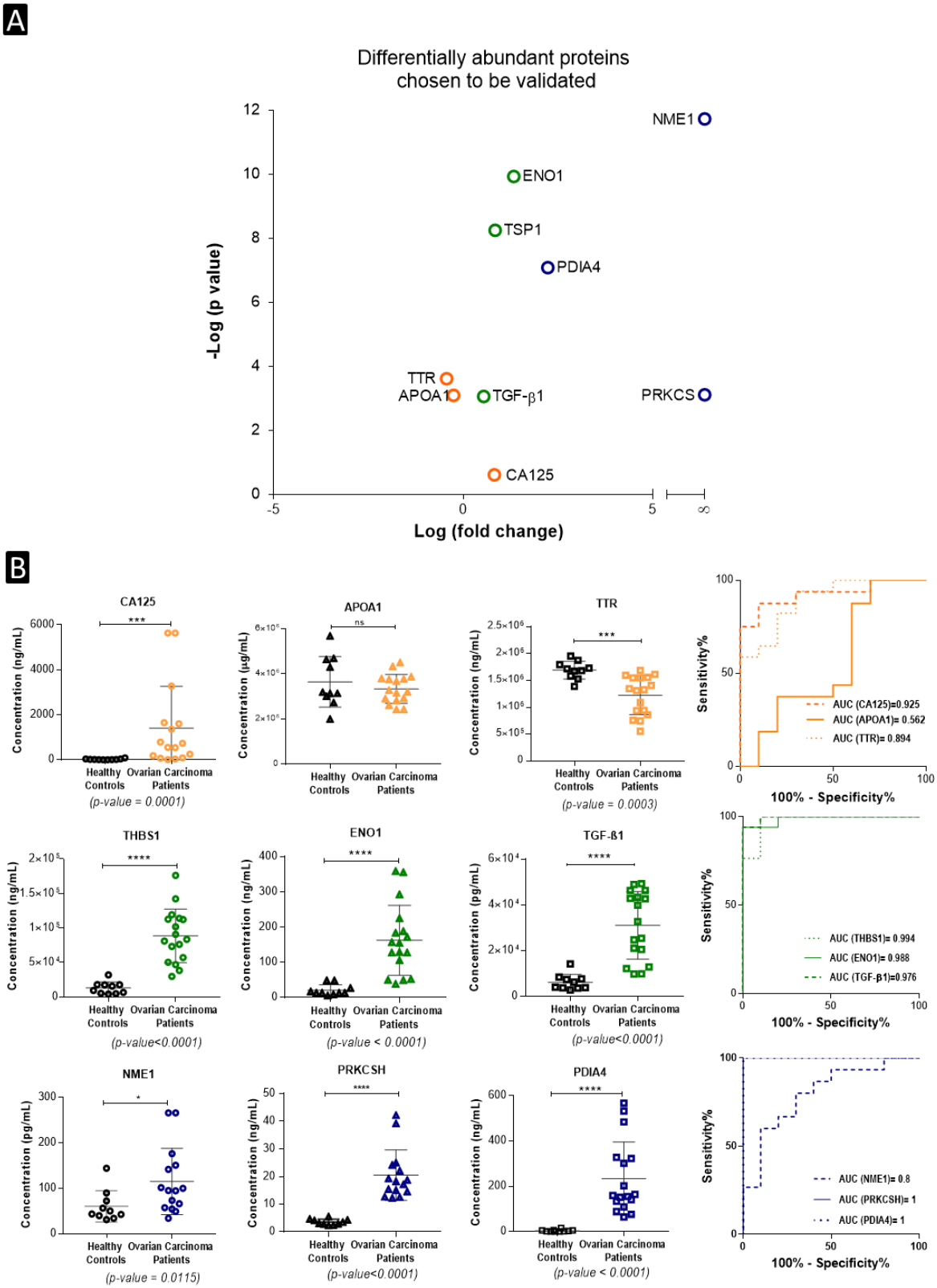


FIGURE 4


A
Differentially abundant proteins chosen to be validated
-Log (p value)
Log (fold change)
NME1
ENO1
TSP1
PDIA4
TTR
APOA1
TGF-β1
PRKCS
CA125
B
CA125
APOA1
TTR
THBS1
ENO1
TGF-ß1
NME1
PRKCSH
PDIA4
Concentration (ng/mL)
Concentration (µg/mL)
Concentration (pg/mL)
Healthy Controls
Ovarian Carcinoma Patients
(p-value = 0.0001)
(p-value = 0.0003)
(p-value<0.0001)
(p-value < 0.0001)
(p-value = 0.0115)
Sensitivity%
100% - Specificity%
AUC (CA125)=0.925
AUC (APOA1)= 0.562
AUC (TTR)= 0.894
AUC (THBS1)= 0.994
AUC (ENO1)= 0.988
AUC (TGF-β1)=0.976
AUC (NME1)= 0.8
AUC (PRKCSH)= 1
AUC (PDIA4)= 1