RESOURCE
https://doi.org/10.1038/s41556-018-0040-4

nature cell biology

Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric flow field-flow fractionation

Haiying Zhang[1*], Daniela Freitas[1,2,3,4,24], Han Sang Kim[1,5,24], Kristina Fabijanic[6], Zhong Li[7], Haiyan Chen[1,8], Milica Tesic Mark[9], Henrik Molina[9], Alberto Benito Martin[1], Linda Bojmar[1], Justin Fang[6], Sham Rampersaud[6], Ayuko Hoshino[1], Irina Matei[1], Candia M. Kenific[1], Miho Nakajima[1], Anders Peter Mutvei[10], Pasquale Sansone[1], Weston Buehring[1], Huajuan Wang[1], Juan Pablo Jimenez[11], Leona Cohen-Gould[11], Navid Paknejad[12], Matthew Brendel[12], Katia Manova-Todorova[12], Ana Magalhães[2,3], José Alexandre Ferreira[2,3,13], Hugo Osório[2,3,14], André M. Silva[15], Ashish Massey[1], Juan R. Cubillos-Ruiz[16], Giuseppe Galletti[17], Paraskevi Giannakakou[17], Ana Maria Cuervo[18], John Blenis[10], Robert Schwartz[19], Mary Sue Brady[20], Héctor Peinado[1,21], Jacqueline Bromberg[19,22], Hiroshi Matsui[6], Celso A. Reis[2,3,4,14] and David Lyden[1,23*]

The heterogeneity of exosomal populations has hindered our understanding of their biogenesis, molecular composition, biodistribution and functions. By employing asymmetric flow field-flow fractionation (AF4), we identified two exosome subpopulations (large exosome vesicles, Exo-L, 90–120 nm; small exosome vesicles, Exo-S, 60–80 nm) and discovered an abundant population of non-membranous nanoparticles termed 'exomeres' (~35 nm). Exomere proteomic profiling revealed an enrichment in metabolic enzymes and hypoxia, microtubule and coagulation proteins as well as specific pathways, such as glycolysis and mTOR signalling. Exo-S and Exo-L contained proteins involved in endosomal function and secretion pathways, and mitotic spindle and IL-2/STAT5 signalling pathways, respectively. Exo-S, Exo-L and exomeres each had unique *N*-glycosylation, protein, lipid, DNA and RNA profiles and biophysical properties. These three nanoparticle subsets demonstrated diverse organ biodistribution patterns, suggesting distinct biological functions. This study demonstrates that AF4 can serve as an improved analytical tool for isolating extracellular vesicles and addressing the complexities of heterogeneous nanoparticle subpopulations.

Exosomes are nanosized extracellular membrane vesicles of endosomal origin secreted by most cell types, including cancer cells[1–3]. Proteins, genetic material (for example, mRNAs, miRNAs, lncRNAs, DNA), metabolites and lipids are selectively recruited and packaged into exosomes, which horizontally transfer their cargo to recipient cells, thereby acting as vehicles of intercellular communication under both physiological and pathological conditions[4–7]. Harnessing this knowledge, translational researchers have focused on developing exosome-based diagnostic/prognostic biomarkers and therapeutic strategies.

Although our understanding of the biology, function and translational potential of exosomes is expanding rapidly, the heterogeneous

[1]Children's Cancer and Blood Foundation Laboratories, Departments of Pediatrics, and Cell and Developmental Biology, Drukier Institute for Children's Health, Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA. [2]i3S - Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal. [3]Institute of Molecular Pathology and Immunology of University of Porto, Ipatimup, Porto, Portugal. [4]Instituto de Ciências Biomédicas de Abel Salazar (ICBAS), University of Porto, Porto, Portugal. [5]Yonsei Cancer Center, Division of Medical Oncology, Departments of Internal Medicine, and Pharmacology, Yonsei University College of Medicine, Seoul, Korea. [6]Department of Chemistry and Biochemistry, City University of New York, Hunter College, New York, NY, USA. [7]Metabolomics Center, University of Illinois, Urbana, IL, USA. [8]Department of Surgical Oncology, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang, China. [9]Proteomics Resource Center, The Rockefeller University, New York, NY, USA. [10]Department of Pharmacology, Meyer Cancer Center, Weill Cornell Medical College, New York, NY, USA. [11]Microscopy & Image Analysis Core Facility, Weill Cornell Medicine, New York, NY, USA. [12]Molecular Cytology Core Facility, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [13]Experimental Pathology and Therapeutics Group, Portuguese Institute of Oncology, Dr. António Bernardino de Almeida, Porto, Portugal. [14]Medical Faculty, University of Porto, Al. Prof. Hernâni Monteiro, Porto, Portugal. [15]LAVQ-REQUIMTE/Department of Chemistry and Biochemistry, Faculty of Sciences, University of Porto, Porto, Portugal. [16]Microbiology & Immunology in Obstetrics and Gynecology, Obstetrics and Gynecology, Weill Cornell Medicine, New York, NY, USA. [17]Pharmacology in Medicine, Joan and Sanford I. Weill Department of Medicine, Weill Cornell Medicine, New York, NY, USA. [18]Department of Developmental & Molecular Biology, Albert Einstein College of Medicine, Jack and Pearl Resnick Campus, Bronx, NY, USA. [19]Department of Medicine, Weill Cornell Medicine, New York, NY, USA. [20]Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [21]Microenvironment and Metastasis Laboratory, Department of Molecular Oncology, Spanish National Cancer Research Center (CNIO), Madrid, Spain. [22]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [23]Department of Pediatrics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [24]These authors contributed equally: Daniela Freitas and Han Sang Kim. *e-mail: haz2005@med.cornell.edu; dcl2001@med.cornell.edu

nature of nanovesicles and the technical limitations in efficiently separating exosomal subpopulations have hindered the characterization of their molecular composition and biogenesis. The state-of-the-art asymmetric flow field-flow fractionation (AF4) technology[8] exhibits a unique capability to separate nanoparticles and has been widely used to characterize nanoparticles and polymers in the pharmaceutical industry and to examine various biological macromolecules, protein complexes and viruses[8,9], but it has rarely been tested for extracellular vesicle (EV) analysis[10–14]. Using AF4, nanoparticles are separated based on their density and hydrodynamic properties by two perpendicular flows: forward laminar channel flow and variable crossflow.

Here, we establish and optimize AF4 parameters and protocols, followed by rigorous biophysical and molecular characterization of small EV (sEV) fractions isolated from numerous cancer and normal cells. Through our modified AF4 protocols, we identify a distinct nanoparticle we term 'exomere', as well as two exosome subpopulations that demonstrate distinct biophysical and molecular properties.

Results

Identification of a distinct nanoparticle population and subsets of exosomes. We first fractionated B16-F10 melanoma-derived sEVs by AF4 (see Methods). A linear separation of the sEV mixture was achieved based on hydrodynamic radius (black dots, *y* axis) along the time course (*x* axis) (Fig. 1a), and the hydrodynamic radius of the particles was determined using an online quasi-elastic light scattering (QELS) monitor for real-time dynamic light scattering (DLS) measurements (red trace). UV absorbance (blue trace) was measured to assess protein concentration and the abundance of particles at specific time points for different particle sizes. Particles with a 35–150 nm diameter were successfully separated by AF4 (Fig. 1a). We identified five peaks (P1–P5) corresponding to the time and particle size at which the most abundant particles were detected. P1 represented the void peak, a mixture of all types of nanoparticles. P5 was composed of individual or aggregated particles and protein aggregates with much larger sizes, which are outside the separation range of the current AF4 protocol and eluted when crossflow dropped to zero (Supplementary Fig. 1a). The hydrodynamic diameters of peaks P2, P3 and P4 were 47 nm, 62 nm and 101 nm, respectively. To infer the hydrodynamic radius, correlation functions were fitted to single exponentials (Fig. 1b, representative P3 fraction graph).

Individual fractions were measured by NanoSight tracking analysis (NTA), validating a consistent particle size for each fraction between 60 nm and 140 nm (Supplementary Fig. 1b). DLS combined with AF4 showed a broader dynamic range than NTA for those particles with a smaller (~70 nm) or larger (~160 nm) particle size (Supplementary Fig. 1c). Moreover, NTA of each individual fraction in the range of 60–160 nm revealed a monomodal profile with a peak of ~77 nm (Supplementary Fig. 1d).

Transmission electron microscopy (TEM) with negative staining of AF4 input and representative fractions across the full dynamic range revealed three populations of particles (P2, P3 and P4; Fig. 1a) with distinct morphology and size (Fig. 1c). P2 represented a distinct population of nanoparticles not previously described; these were smaller than 50 nm (~35 nm) and clearly lacked an external membrane structure (Fig. 1c), so we named these structures 'exomeres'. The other two nanoparticle subpopulations we refer to as small exosomes (Exo-S; 60–80 nm, P3) and large exosomes (Exo-L; 90–120 nm, P4) (Fig. 1c). All three particle types were readily detected in the input TEM image (Fig. 1c). Western blot analysis confirmed exosome markers Tsg101 and Alix for Exo-S and Exo-L, and heat shock protein 90 (Hsp90) for exomeres (Fig. 1d). The sizes of each particle type measured in batch mode showed consistent results (Fig. 1e).

In summary, a single run of AF4 can efficiently discern exomeres and two distinct exosome subpopulations in a robust and highly reproducible manner (Supplementary Fig. 1e,f). Freeze–thawing of samples led to inconsequential differences (Supplementary Fig. 1g). However, changes in culture conditions led to differences in the relative abundance of each particle type (Supplementary Fig. 1h,i).

Importantly, only a minor peak eluted in the time range similar to exomeres in a blank medium control compared to conditioned media (CM) of B16-F10 and MDA-MB-231-4175 when processed in parallel (Supplementary Fig. 1j,k), thereby confirming that exomeres are indeed actively secreted by cultured cells and not mere aggregates present in media.

Using AF4, we detected distinct particles with diameters corresponding to exomeres and Exo-S/L in more than 20 cell lines analysed (Supplementary Table 1 and Supplementary Fig. 2a), and our findings were confirmed by TEM analysis of pooled fractions from selected cell lines (Supplementary Fig. 2b). Based on UV absorbance and TEM analysis, all cells secreted higher amounts of exomeres relative to Exo-S/L, except for B16-F10 and B16-F1, where Exo-S were relatively more abundant (Supplementary Fig. 2a and Fig. 1a,b). Measurement of the hydrodynamic diameter of each of these particles using Zetasizer (Malvern) revealed sizes similar to the B16-F10 preparations (Fig. 1e).

We also detected exomeres and Exo-S/L in AF4-fractionated sEVs from CM of human melanoma tumour explants by TEM (Fig. 1f, arrows; Supplementary Fig. 3a). Exomere and Exo-S size, measured in batch mode using Zetasizer, was comparable to results from tumour cell lines (Fig. 1g). AF4 profiling and TEM imaging analysis showed that normal mouse tissue explants (mammary fat pad and lung) also secreted exomeres and Exo-S/L nanoparticles (Supplementary Fig. 3b).

Biophysical characterization of exomeres and exosome subpopulations. Given the structural differences between exomeres and Exo-S/L, we examined their biophysical properties, including zeta potential and stiffness. Measuring zeta potential, an average surface charge, using Zetasizer, revealed that all particles were negatively charged, with exomeres being the weakest negatively charged (−2.7 mV to −9.7 mV), Exo-L the strongest (−12.3 mV to −16.0 mV) and Exo-S intermediate (−9.0 mV to −12.3 mV) (Fig. 2a).

To assess particle stiffness, atomic force microscopy (AFM) was performed in solution (see Methods). Exomeres demonstrated the highest stiffness (~145–816 MPa) and Exo-L the lowest (~26–73 MPa), with Exo-S stiffness being intermediate (~70–420 MPa).

AFM analysis of exomeres derived from B16-F10, MDA-MB-231-4175 and AsPC-1 cell lines demonstrated exomere structural heterogeneity and average exomere heights of 5.9 nm, 7.0 nm and 5.8 nm, respectively (Fig. 2c,d).

Collectively, these findings demonstrate the diverse biophysical properties exhibited by exomeres versus distinct exosome subpopulations. How size, charge and mechanical properties influence the differential stability, trafficking and uptake of the nanoparticles in vivo requires further investigation[15,16].

Distinct proteomic content and cellular functions of exomeres and exosome subpopulations. To characterize the molecular composition of exomeres and distinct exosome subpopulations we conducted proteomic profiling of nanoparticles derived from B16-F10, Pan02, 4T1, AsPC-1 and MDA-MB-231-4175 cells using label-free MS. A range of 165–483 proteins were identified in exomeres, 433–1,004 proteins in Exo-S and 247–1,127 proteins in Exo-L. Moreover, unique proteins were detected in each nanoparticle subtype (Fig. 3a), suggesting exomeres are unique entities released by cells rather than debris or fragments of exosomes.

Examination of the subcellular localization annotation of proteins revealed the specific enrichment of Exo-S/L in membrane-associated proteins, which were relatively depleted in exomeres (Supplementary Table 3), consistent with our structural studies

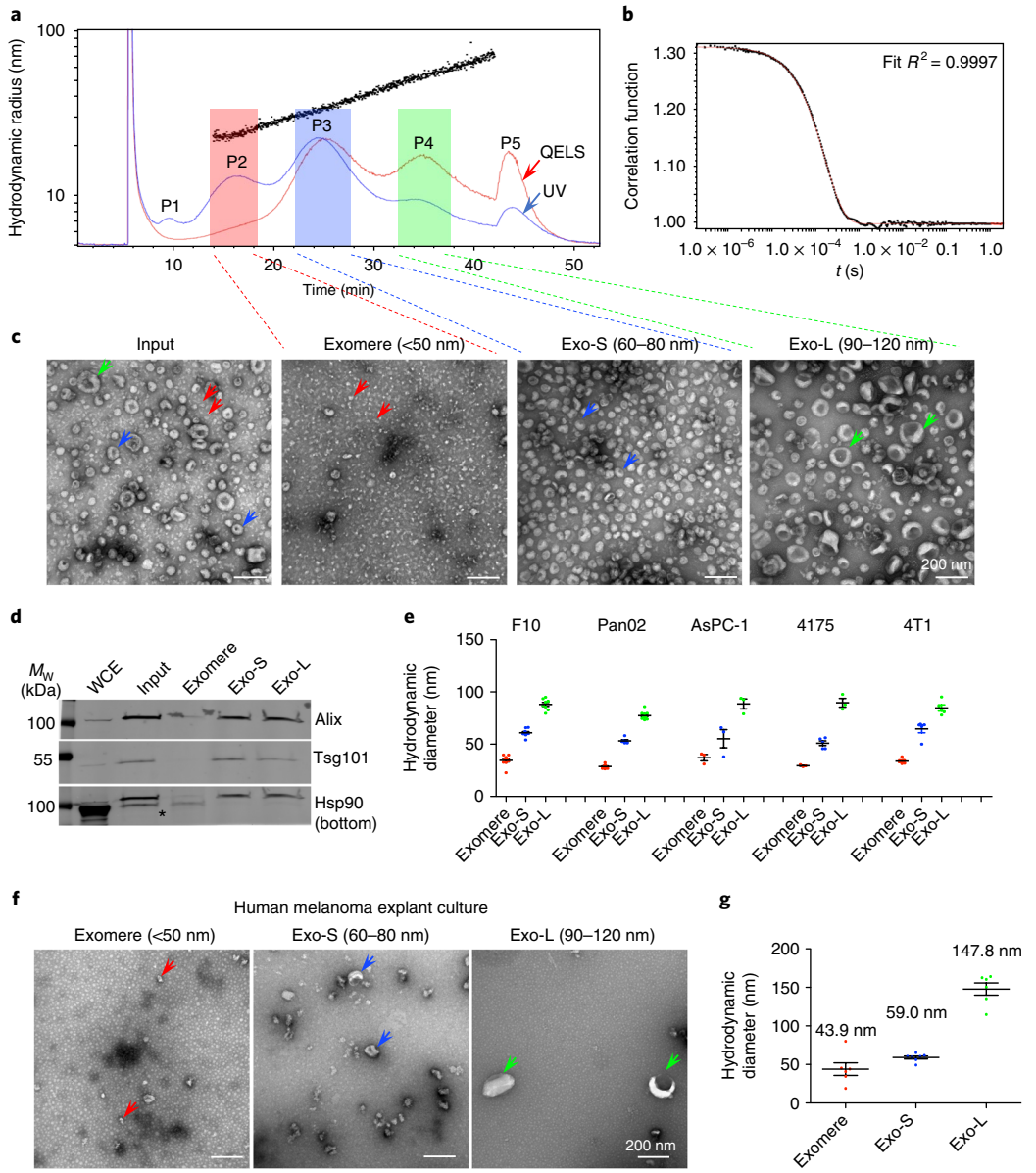

Fig. 1 | Identification, by AF4 and EM imaging analysis, of exomeres and two distinct subpopulations of exosomes released by tumour cells. **a**, A representative AF4 fractionation profile of B16-F10-derived exosomes. Black dots, hydrodynamic radius (nm); red and blue lines illustrate the QELS (DLS) intensity and UV absorbance (shown on a relative scale), respectively. P1–P5 mark the peaks detected based on UV absorbance. Fractions were pooled for exomeres (hydrodynamic diameter <50 nm), Exo-S (60–80 nm) and Exo-L (90–120 nm). **b**, Representative correlation function at peak 3 (P3), $t = 25.1$ min. For **a** and **b**, the experiment was repeated independently 50 times with similar results. **c**, TEM imaging analysis of exosome input mixture (pre-fractionation) and fractionated exomere, Exo-S and Exo-L subpopulations. Arrows indicate exomeres (red), Exo-S (blue) and Exo-L (green). Scale bars, 200 nm. This experiment was repeated seven times independently with similar results. **d**, Western blotting analysis of exosomal marker proteins in fractionated samples. Whole cell extract (WCE, 100 μg) and 10 μg of exosome and exomere mixture input and each subset were analysed. This experiment was performed once. **e**, Measurement of hydrodynamic diameters of exomeres, Exo-S and Exo-L derived from representative cell lines (B16-F10 (F10), AsPC-1, Pan02, MDA-MB-231-4175 (4175) and 4T1) in batch mode using Zetasizer after pooling fractions were collected for each subset of nanoparticles from an individual AF4 fractionation. Data are presented as mean ± s.e.m. in the order of exomere, Exo-S and Exo-L: B16-F10 ($n = 10$, 9 and 8 independent measurements, respectively); Pan02 ($n = 11$, 6, 11); AsPC-1 ($n = 5$, 5, 5); 4175 ($n = 3$, 5, 3); 4T1 ($n = 5$, 5, 5). **f**, TEM imaging analysis of fractions collected from explant culture of fresh human melanoma tissue. Scale bars, 200 nm. This experiment was performed with two independent specimens with similar results. **g**, Batch-mode measurement of the hydrodynamic diameters of fractions shown in **f**. Data are presented as mean ± s.e.m. (exomeres and Exo-L, $n = 6$; Exo-S, $n = 7$ independent measurements). Statistical source data are provided in Supplementary Table 8. Unprocessed blots are provided in Supplementary Fig. 7.

identifying Exo-S/L as membrane-encapsulated particles and exomeres as non-encapsulated particles. The endosomal sorting complexes required for transport (ESCRT)- and the soluble *N*-ethylmaleimide-sensitive factor activating protein receptor (SNARE)-related proteins, which are involved in vesicle budding, membrane fusion and exosome biogenesis[17,18], were identified within Exo-S/L. In particular, proteins associated with endosomes, multivesicular bodies, vacuoles and phagocytic vesicles were enriched in Exo-S. Plasma membrane, cell–cell contact/junction, late-endosome and trans-Golgi network proteins were enriched in

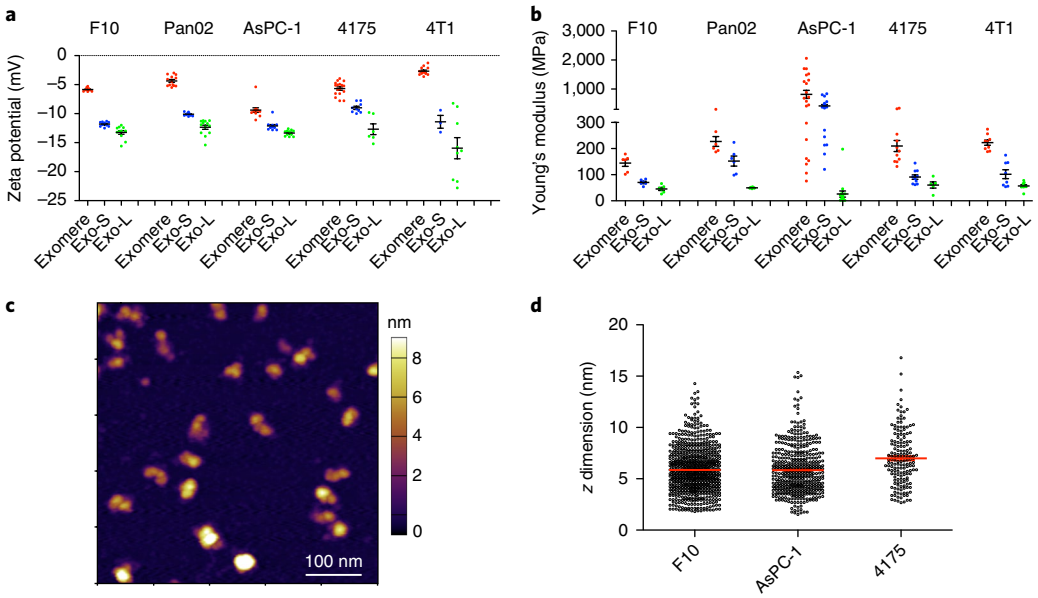

Fig. 2 | Characterization of physical and mechanical properties of exomeres and exosome subpopulations. a,**b**, Zeta potential (**a**) and stiffness (**b**) of exomeres and exosome subpopulations derived from various cancer cells were measured using Zetasizer and AFM indentation, respectively. Young's modulus was used to express particle stiffness. At least three and five replicates for each group of particles were measured for zeta potential and stiffness, respectively. Data are presented as mean ± s.e.m. For **a**, in the order of exomere, Exo-S and Exo-L: B16-F10 (n = 8, 10 and 12 independent measurements, respectively); Pan02 (n = 13, 11, 13); AsPC-1 (n = 12, 12, 12); 4175 (n = 17, 9, 6); 4T1 (n = 13, 3, 9). For **b**, B16-F10 (n = 6, 6, 6 particles measured); Pan02 (n = 6, 6, 6); AsPC-1 (n = 21, 19, 16); 4175 (n = 11, 10, 5); 4T1 (n = 9, 8, 9). **c**, Representative AFM image of exomeres derived from B16-F10. This experiment was repeated with samples derived from three different cell lines with similar results. **d**, AFM imaging analysis of the height (z dimension) of exomeres derived from B16-F10 (n = 754 particles analysed), AsPC-1 (n = 475) and MDA-MB-231-4175 (n = 160). Mean ± s.e.m. is depicted. Statistical source data are provided in Supplementary Table 8.

Exo-L. Notably, proteins associated with the extracellular matrix and space, proteasome accessory complex, endoplasmic reticulum (ER), mitochondrion and microtubule/cytoskeleton were packaged in exomeres. These findings imply possible fundamental differences in exomere, Exo-S and Exo-L biogenesis.

Principal component analysis (PCA) demonstrated a closer correlation of protein expression for Exo-S and Exo-L compared to exomeres from the same cell type (Supplementary Fig. 4a). According to PCA and consensus clustering analysis, exomeres from different cell types exhibited a higher degree of similarity to each other than to Exo-S and Exo-L from the same cell type (Fig. 3b,c).

To identify the signature proteins in each particle subset, we performed statistical analysis on the expression levels of proteins identified in these data sets. We pinpointed 64 proteins for exomeres and 99 proteins for Exo-S/L (Supplementary Table 4), with a false discovery rate (FDR) of <0.05, positive enrichment in each particle subset of interest, and a detection frequency of >80% (that is, a particular protein was positively enriched in at least four of five samples for each subtype of nanoparticles derived from five different cell lines). Remarkably, exomeres were significantly enriched in proteins involved in metabolism (see gene set enrichment analysis (GSEA) below), including MAT1A, IDH1, GMPPB, UGP2, EXT1 and PFKL. The sialoglycoprotein galectin-3-binding protein (LGALS3BP) and key proteins controlling glycan-mediated protein folding control (CALR)[19] and glycan processing (MAN2A1, HEXB, GANAB)[20–22] are also enriched in exomeres, suggesting exomere cargo may mediate the targeting of recipient cells through specific glycan recognition, and modulate glycosylation in recipient cells. Among proteins uniquely represented in Exo-S/L were annexins, ESCRT components (charged multivesicular body proteins (CHMPs), vacuolar protein-sorting proteins, HGS, Alix1/PDCD6IP and Tsg101), Hsp40 (DnaJ) family proteins, signalling transducer G protein subunits, integrins, Rab proteins and solute carrier family members. Members of key signalling pathways, such as JAK1, TGFBR2 and MET, were also enriched in Exo-S/L. To evaluate the unique markers of Exo-S and Exo-L subpopulations, we compared protein expression between these two sample sets and exomeres separately using a t-test. A second set of filters (protein intensity/area of $>1\times10^8$ and fold change of ≥5.0) was applied to the identified signature for exomeres and Exo-L, but not for Exo-S (Fig. 3d). Fewer signature proteins were identified for Exo-S than for exomeres and Exo-L, most probably due to the similarity of Exo-S to the other particles. Representative signature proteins identified by proteomics in each subset were validated by western blot analysis (Fig. 3e).

We further mined these proteomic data sets for conventional exosome markers, including flotillins, CD9, CD63, CD81, Alix1, Tsg101, HSC70 (HSPA8) and Hsp90 (Fig. 3f). Among the five cell lines, flotillins (FLOT1 and FLOT2) represented bona fide markers of Exo-S, while HSP90AB1 was preferentially associated with exomeres. Although CD9, CD63 and CD81 all demonstrated specific association with Exo-S/L subsets, they all showed a cell type- and particle-dependent preferential expression. Consistent with ref. [23], combining CD63, CD9 or CD81 will be necessary to isolate/label exosomes.

We found numerous Rab proteins in Exo-S/L subsets, but few of them in exomeres (Supplementary Fig. 4b), suggesting critical roles of Rab proteins for Exo-S/L formation and trafficking, but not for exomere biogenesis.

Next, we examined the most abundant proteins in each subset of nanoparticles. Haemoglobin, histones, cytoskeleton proteins (actins and tubulins), peptidylprolyl isomerase A (PPIA) and HSP ranked as the most abundant top 50 proteins in all three nanoparticle subpopulations (Supplementary Table 4). Hsp40/DnaJ family (HSP70 co-chaperones) members were also found in the top 50 proteins for Exo-L. Interestingly, HSP90AB1 was preferentially packaged

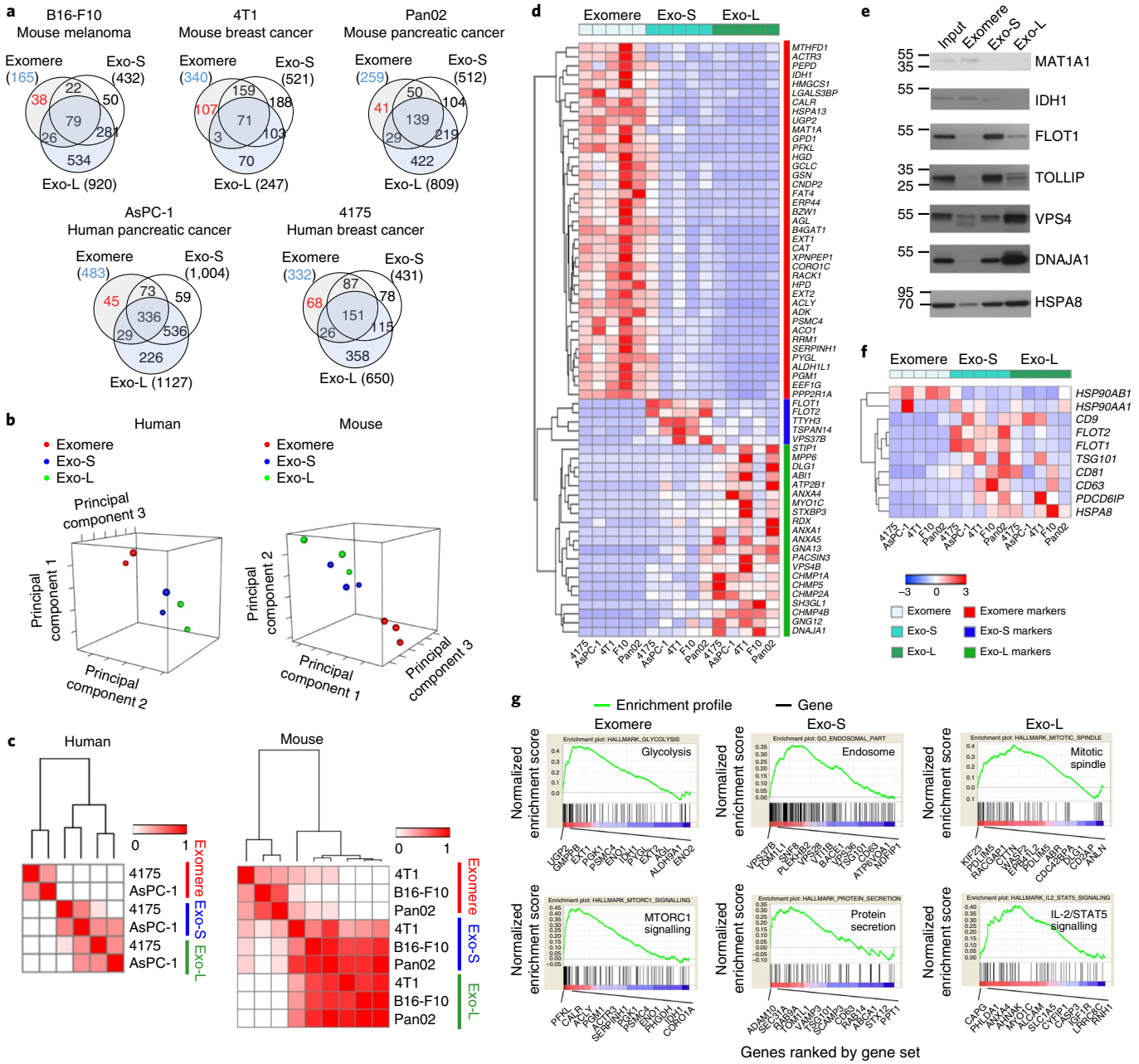

Fig. 3 | Proteomic profiling of exomeres and exosome subpopulations derived from various cancer cells. a, Venn diagram of proteins identified in each subset of particles. **b**,**c**, Principal component analysis (**b**) and consensus clustering analysis (**c**) of normalized proteomic MS data sets from human (MDA-MB-231-4175 and AsPC-1) and mouse (B16-F10, 4T1 and Pan02) cell lines. **d**, Heatmap illustration of unique proteins specifically associated with exomeres, Exo-S and Exo-L. Scale shown is intensity (area) subtracted by mean and divided by row standard deviation (that is, Δ(area − mean)/s.d.). **e**, Western blot analysis of representative signature proteins in fractionated samples. Equal amounts (10 μg) of exosome and exomere input mixture and each subset were analysed. This experiment was carried out once. **f**, Heatmap illustration of the relative abundance of conventional exosome markers in exomeres, Exo-S and Exo-L. Scale shown is intensity (area) subtracted by mean and divided by row standard deviation (that is, Δ(area − mean)/s.d.). **g**, Identification of top candidate gene sets enriched in exomere, Exo-S and Exo-L populations by GSEA. Proteins in each subset of nanoparticles are ranked by GSEA based on their differential expression level. Whether a pre-specified pathway is significantly over-represented towards the top or bottom of the ranked gene list in each subset of nanoparticle is evaluated using the normalized enrichment score (green line). Black vertical lines mark positions where members of a particular pathway appear in the ranked list of genes. Proteins that contributed most to the enrichment score are listed below the plot. For all proteomic analysis (**b**–**d**, **f**,**g**), a total of 30 samples (three nanoparticle subtypes derived from five different cell lines and two independent biological replicates for each nanoparticle sample) were subjected to statistical analysis. Unprocessed blots are provided in Supplementary Fig. 7.

in exomeres, while HSP70 members (HSPA8, HSPA2 and HSPA5) were more abundant in Exo-S/L. Other proteins relatively enriched in exomeres included inter-α-trypsin inhibitor heavy chain family members (ITIH), gelsolin (GSN), talin 1 (TLN1), WD repeat domain 1 (WDR1) and proteins involved in metabolism, such as phosphoglycerate kinase 1 (PGK1), pyruvate kinase muscle (PKM) and enolase 1 (ENO1). Consistent with the analysis above, SDCBP, PDCD6IP/Alix, tetraspanins (CD9, CD63, CD81 and others),

G protein family proteins and integrins were highly represented in both Exo-S and Exo-L. Tetraspanins were preferentially enriched in Exo-S, while G proteins and integrins were more prominent in Exo-L. Eukaryotic translation elongation factor 1 alpha 1 (EEF1A1) was most often present in exomeres and Exo-L. Other proteins preferentially associated with Exo-S included immunoglobulin superfamily member 8 (IGSF8) and its paralogue prostaglandin F2 receptor inhibitor (PTGFRN), milk fat globule-EGF factor 8 protein (MFGE8) and components of the ESCRT-I complex. Notably, annexins and S100 proteins were only represented in the top 50 proteins of Exo-L.

Furthermore, to exclude the possibility of lipoprotein contamination in exomeres, we examined proteins that are typically associated with purified lipoprotein particles (high-, low- and very-low-density lipoproteins, that is, HDL, LDL and VLDL) by proteomic MS analysis and then evaluated their presence in exomeres and Exo-S/L. Much fewer proteins were found in lipoproteins (Supplementary Table 5), and only some of these proteins were detected in exomeres and Exo-S/L, suggesting that most nanoparticle proteins are distinct from lipoproteins. A rough estimation showed that the lipoprotein-associated proteins account for 0–8% of total nanoparticle proteins (Supplementary Fig. 4c). Moreover, TEM analysis revealed that lipoprotein morphology/structure was clearly distinct from exomeres and Exo-S/L (Supplementary Fig. 4d). Taken together, these analyses ruled out the possibility that exomeres were mere lipoprotein contaminants.

The possible contamination of exomeres with other types of protein complexes with high molecular weight was also ruled out when exomere proteins were surveyed for subunits of known complexes (using the comprehensive resource of mammalian protein complexes: http://mips.helmholtz-muenchen.de/corum/#; 02.07.2017 CORUM current release). The coexistence of multiple subunits of protein complexes of similar size to exomeres was not detected (Supplementary Table 5) except for 10 out of 59 subunits of parvulin-associated pre-rRNP complex in 4T1 exomeres, 17 subunits of ribosomes in AsPC-1 exomeres and 7 out of 16 subunits of kinase maturation complex 1 in MDA-MB-231-4175 exomeres. However, these proteins account for only 1.8, 2.1 and 1.8% of total exomere proteins in each case, respectively, suggesting that their contribution diminishes the purity of exomeres by a negligible ~2%.

To gain insight into the function of these particle subsets, we conducted GSEA using the Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Hallmark databases (Supplementary Table 6). Strikingly, GSEA demonstrated that exomere-specific proteins were selectively enriched in metabolic processes, including carbohydrate metabolism, protein synthesis and small molecules. At least 36 of the top 50 'GO-biological processes' pathways identified metabolic processes associated with exomeres, in contrast to no metabolic processes associated with Exo-S/L (Supplementary Table 6). Genes encoding proteins involved in hypoxia, microtubules and coagulation were identified in exomeres (Supplementary Fig. 4e). Exo-S were enriched in membrane vesicle biogenesis and transport, protein secretion and receptor signalling gene sets. For Exo-L, enriched gene sets included mitotic spindle, IL-2/Stat5 signalling, multi-organism organelle organization and G-protein signalling. Profiles of top-rank gene sets enriched in exomeres (glycolysis and mTORC1 signalling), Exo-S (endosome and protein secretion) and Exo-L (mitotic spindle and IL-2/Stat5 signalling) are presented in Fig. 3g.

Collectively, these bioinformatic analyses of the proteomic content of each particle subset revealed the predominant link between exomere-associated proteins and metabolism and the link between Exo-S/L-associated proteins and multiple signalling transduction pathways, including biogenesis-related ESCRT complexes.

Distinct *N*-glycan profiles of exomeres and exosome subpopulations. Aberrant glycosylation is involved in pathological processes, including cancer[24]. Here, we aimed to determine the *N*-glycan profiles of each particle subset in three cell lines by conducting lectin blotting analysis (Fig. 4a) and glycomic MS.

Phaseolus vulgaris erythroagglutinin (E-PHA) recognizing bisected *N*-glycans detected a major band at ~75 kDa in both Exo-S and Exo-L of B16-F10 and AsPC-1, with faint detection in exomeres across the three cell lines and Exo-S of MDA-MB-231-4175. E-PHA detected a high-molecular-weight glycoprotein (240 kDa) in MDA-MB-231-4175 exomeres and a high-molecular-weight glycoprotein (150 kDa) in AsPC-1 and MDA-MB-231-4175 exomeres. *Phaseolus vulgaris* leucoagglutinin (L-PHA) recognizing branched *N*-glycans detected a predominant band at 75 kDa in both Exo-S and Exo-L of B16F-10 and AsPC-1. Multiple bands ranging from 50 to 70 kDa were also detected in all exomeres (especially MDA-MB-231-4175). Using *Aleuria aurantia* lectin (AAL), analysis of structures related to fucosylation (fucose-linked α-1,6) to GlcNAc or fucose-linked (α-1,3) to GlcNAc-related structures revealed two abundant glycoproteins between 70 and 100 kDa in both Exo-S and Exo-L of B16-F10 and AsPC-1. Exomeres across all three cell lines and Exo-S of MDA-MB-231-4175 displayed strong fucosylation on high-molecular-weight glycoproteins (200–280 kDa). *Sambucus nigra* lectin (SNA), recognizing α-2,6-linked sialic acid, detected the presence of high-molecular-weight α-2,6-sialylated glycoproteins (200–250 kDa) in all exomeres. Moreover, a low-molecular-weight protein (~60 kDa) displaying α-2,6-linked sialic acid modification was uniquely detected in Exo-L (but not Exo-S) from B16-F10. For AsPC-1, exomeres were the major carriers of sialylated glycoproteins, while these sialylated structures were almost absent in Exo-L. Lectin-binding profiles did not overlap with the most abundant proteins in the SDS–PAGE gel, indicating the specificity of lectin recognition independently of protein abundance (Supplementary Fig. 5a). Therefore, Exo-S and Exo-L versus exomeres display distinct *N*-glycosylation patterns. Notably, exomere and Exo-S/L-associated *N*-glycan profiles vary by cell type. Future studies will address the identity of these glycoproteins via glycoproteomic approaches.

We next aimed to identify profiles of the glycan structures enriched in each particle subset by matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) mass spectrometry (MS). Two independent, semi-quantitative MS analyses were conducted on B16-F10-derived exomeres and Exo-S/L (Fig. 4b). Figure 4c depicts the quantification of the top six most abundant *N*-glycan structures detected in one of the representative experiments. We observed the ubiquitous expression of certain complex *N*-glycans in all subsets, corresponding to peaks at *m/z* 2,209.8, 2,223.7, 2,237.7 and 2,365.5. Specifically, a complex *N*-glycan at *m/z* 2,015.7 and a hybrid *N*-glycan at *m/z* 2,404.8 were enriched in exomeres. Moreover, four of these six *N*-glycans contained sialic acid, and three of six were fucosylated. Similarly, the ions at *m/z* 2,015.7 and 2,404.8 were enriched in exomeres from MDA-MB-231-4175 (Supplementary Fig. 5b,d). The ion at *m/z* 2,404.8 was slightly enriched in AsPC-1 exomeres, but the ion at *m/z* 2,015.7 was not detected in AsPC-1 samples (Supplementary Fig. 5b and c). Instead, the ion at *m/z* 2,012.7 was strongly detected in AsPC-1 exomeres and Exo-S. Two other ions, at *m/z* 2,117.7 and 2,389.9, demonstrating Exo-S enrichment, were detected in AsPC-1 only (Supplementary Fig. 5b–d).

High-resolution MS analysis allowed further structural characterization of certain *N*-glycans, namely ions at m/z 2,223.7 and 2,015.7 (corresponding to extracted ion chromatogram m/z 1,111.39 (2-) and 1,007.38 (2-) in Supplementary Fig. 5e–j). In addition, the combination of collision-induced dissociation tandem mass spectrometry (CID-MS/MS) de novo sequencing and porous graphitized carbon liquid chromatography (PGC-LC) relative retention times for extracted ion chromatogram *m/z* 1,111.39 (2-) revealed

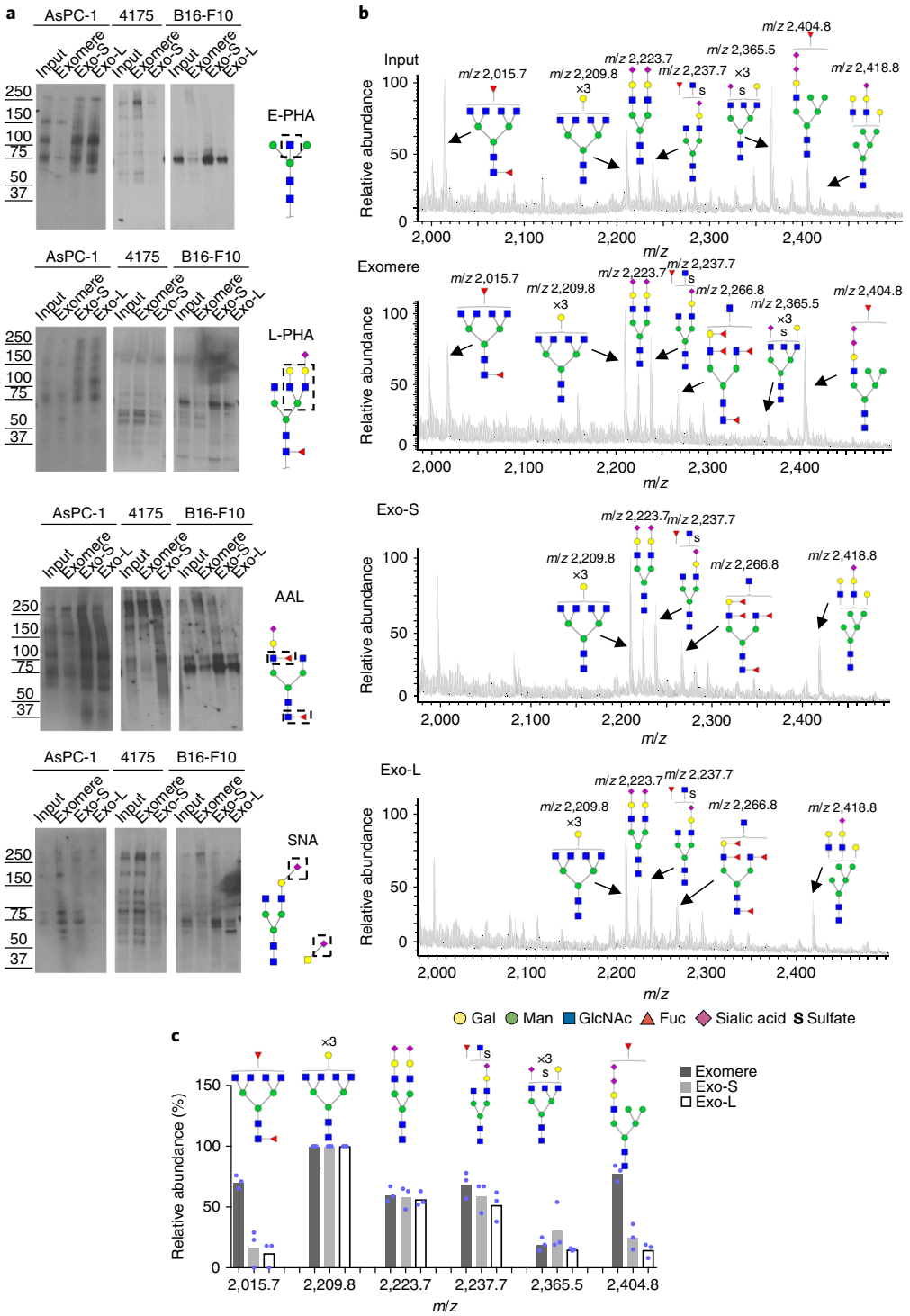

Fig. 4 | Characterization of *N*-glycosylation of proteins associated with exomere, Exo-S and Exo-L. a, Lectin blotting analysis of *N*-glycan profile of proteins associated with exomeres versus exosome subpopulations Exo-S and Exo-L. *P. vulgaris* erythroagglutinin (E-PHA) and *P. vulgaris* leucoagglutinin (L-PHA) recognize bisected and branched *N*-glycans, respectively. *A. aurantia* lectin (AAL) recognizes Fucα6GlcNAc and Fucα3GlcNAc. *S. nigra* lectin (SNA) recognizes α-2,6-linked sialic acid. All experiments were repeated independently twice with similar results except for AAL and E-PHA blotting for B16-F10 and 4175, which were done once. **b**, MS analysis of *N*-glycans of glycoproteins present in exomeres, Exo-S and Exo-L subsets of B16-F10. One representative experiment of two biologically independent replicates is shown. **c**, Comparison of the relative abundance of the top six most abundant *N*-glycan structures among exomere, Exo-S and Exo-L of B16-F10. The assignments (*m/z*) (charge; neutral exchange) for MALDI-MS and nanoLC-ESI-MS/MS are as follows: (2015.8 [-H; 0]; 1,007.4[a] [-2H; 0]), (2,209.8 [-H; 0]; 1,104.4[a] [-2H; 0]), (2,237.7[b] [-H; Na-H]; 732.57[a] [-3H; 0]), (2,365.5[b] [-H;4K-4H]; 783.9[a] [-3H; 0] and 1,182.4[a,b] [-2H; 4K-4H]) and (2,404.8[b] [-H; 2K-2H]; 1,201.9[b] [-2H; 2K-2H]). Data shown were quantified and normalized to the most abundant structure in the sample. Results are represented as average of three independent analytical measurements of one representative experiment. Statistical source data are provided in Supplementary Table 8 and unprocessed blots are provided in Supplementary Fig. 7. [a]The product ion spectra for this species did not allow a complete structural assignment. [b]Assignments admit neutral exchanges of protons with cations in sialoglycans, including the presence of potassium and sodium.

that this *N*-glycan from exomeres contained both α-2,3-linked and α-2,6-linked sialic acids, whereas the glycan from Exo-S contained exclusively α-2,3-linked sialic acids. The unique presence of *m/z* 1,007.38 (2-) in exomeres was also further confirmed.

Taken together, our glycomics study demonstrated the prevalence of complex *N*-glycans in all particle subsets with relatively high levels of sialylation, consistent with previous findings of complex *N*-glycans and sialoglycoproteins in tumour microvesicles/exosomes[25–27]. Furthermore, our study revealed differences in *N*-glycan composition and structures among exomeres, Exo-S and Exo-L.

Distinct lipid composition among exomeres and exosome subpopulations. To investigate the lipid composition of each subset of particles, we performed quantitative lipidomics on these nanoparticles derived from B16-F10, MDA-MB-231-4175 and AsPC-1. By lipid MS, we found that Exo-S and Exo-L contained more lipids than exomeres for all cell lines (Fig. 5a, more than fivefold in all subpopulations, except for Exo-S of MDA-MB-231-4175, which was more than threefold).

Eighteen lipid classes were commonly identified in all samples (Supplementary Table 7 and Fig. 5b) and their relative frequency in each sample was compared. Phosphatidylcholine (PC) was the predominant lipid component in all subpopulations (~46–89%) except for AsPC-1 exomeres (13%) (Fig. 5b), which contained higher levels of diglyceride (DG, 38%) and triglyceride (TG, 26%) instead. Other phospholipids, including phosphatidylethanolamine (PE) and phosphatidylserine (PS), accounted for 2–6% of total lipids in Exo-S/L across all cell lines (Fig. 5b). However, PE and PS levels were lower in exomeres from MDA-MB-231-4175 and AsPC-1, but similar to Exo-S/L in B16-F10 (Fig. 5b,c). Phosphatidylinositol (PI) levels were lower than other phospholipids but had a pattern of distribution across nanoparticle subsets similar to those of PE and PS (Fig. 5b,c). Sphingomyelin (SM) accounted for 2–10% of the total lipid in all samples except for AsPC-1 Exo-S/L, which contained a

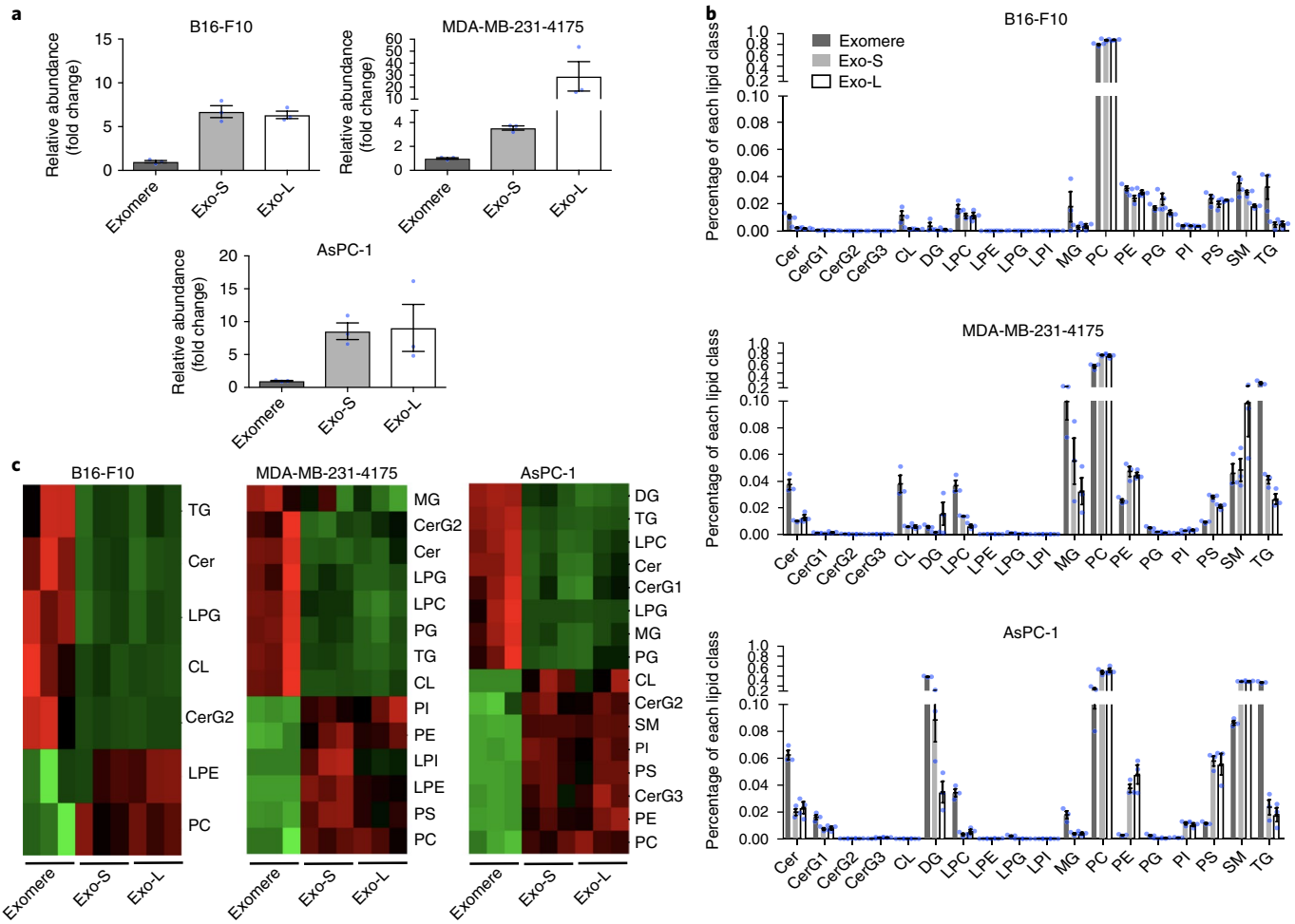

Fig. 5 | Characterization of lipid composition in exomeres and exosome subsets. a, Comparison of total lipid content of each subset of nanoparticles derived from different cell lines. Total signal intensity of each sample after normalization to sample weight and internal standards was compared to that of exomeres from the same set of samples (expressed as fold change). Data are presented as mean ± s.e.m. ($n = 3$ biologically independent samples). **b**, Relative abundance of each lipid class present in each subset of nanoparticles from different cell lines. Data are presented as mean ± s.e.m. ($n = 3$ biologically independent samples). For **a** and **b**, statistical source data are provided in Supplementary Table 8. **c**, Heatmap illustration of lipid classes specifically associated with exomeres, Exo-S and Exo-L (ANOVA test, $q < 0.05$). Statistical analysis was performed on a total of nine samples for each cell line (three different nanoparticle subtypes and three independent biological repeats for each nanoparticle sample). Cer, ceramide; CerG1–3, glucosylceramides; CL, cardiolipin; DG, diglyceride; LPC, lysophosphatidylcholine; LPE, lysophosphatidylethanolamine; LPG, lysophosphatidylglycerol; LPI, lysophosphatidylinositol; MG, monoglyceride; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphatidylinositol; PS, phosphatidylserine; SM, sphingomyelin; TG, triglyceride.

higher level of SM (28%, Fig. 5b,c) Cholesterol data were not collected in this study.

The relative levels of ceramide (Cer), TG and lysophosphatidylglycerol (LPG) varied significantly between exomeres and Exo-S/L across cell lines (analysis of variance (ANOVA) test, $q < 0.05$). Additionally, simple glycosphingolipid CerG2 and mitochondrion-specific cardiolipin (CL) were more abundant in exomeres of B16-F10 and MDA-MB-231-4175 compared to exosome subsets. In contrast, CerG2 and CL were more abundant in Exo-S/L compared to exomeres isolated from AsPC-1 cells. Monoglyceride (MG), phosphatidylglycerol (PG) and lysophosphatidylcholine (LPC) were more abundant in exomeres than in Exo-S/L from MDA-MB-231-4175 and AsPC-1, but present at equal levels in all three B16-F10 nanoparticle subsets. Lastly, lysophosphatidylethanolamine (LPE) was detected at higher levels in Exo-S/L from B16-F10 and MDA-MB-231-4175, but not from AsPC-1. Thus, our study revealed cell type-dependent differences in the total lipid content and composition among distinct nanoparticle subsets.

Distinct nucleic acid content among exomeres and exosome subpopulations. Because we previously detected dsDNA in tumour-derived exosomes[6], we determined the relative abundance of DNA in exomeres and Exo-S/L. DNA was detected in all three types of nanoparticles; however, the relative abundance varied by cell type (Fig. 6a). The relative amount of DNA was highest in exomeres derived from MDA-MB-231-4175 and in Exo-S from B16-F10 cells and AsPC-1. Bioanalyzer (Agilent) analysis revealed a distinct size distribution of DNA associated with each subset of nanoparticles (Fig. 6b and Supplementary Fig. 6). Exomere DNA was relatively evenly distributed in a broad range of sizes between 100 bp and 10 kb with a slight enrichment around 2 kb in several cases. In contrast, a strong enrichment between 2 kb and 4 kb was detected for Exo-S/L DNA, and the peak size of Exo-L DNA was slightly larger than that of Exo-S DNA. This phenomenon may be due to the structural capacity and different biogenesis mechanisms of each particle subset.

RNA was preferentially associated with Exo-S/L in both B16-F10 and AsPC-1 (Fig. 6c). RNA associated with exomeres and Exo-S showed a monomodal distribution (peak at 400 nt and 500 nt, respectively), whereas Exo-L RNA displayed a bimodal distribution (Fig. 6d) (additional peak at >4,000 nt). Specifically, 18S and 28S rRNAs were detected at very low levels in Exo-L, barely detected in Exo-S and absent in exomeres compared to cellular RNA. A strong small RNA peak (corresponding to tRNAs, microRNAs and other small RNAs) was detected in Exo-S and Exo-L, but not in exomeres. Remarkably, a unique RNA peak of unknown identity, of ~315 nt in size, was detected only in Exo-L.

Distinct organ biodistribution of exomeres and exosome subpopulations. Next, we investigated the organ biodistribution of B16-F10-derived nanoparticle subsets in naive mice. At 24 hours post-intravenous injection of near-infrared dye (NIR)-labelled exomeres, Exo-S and Exo-L into mice, organs were collected and analysed using the Odyssey imaging system (LI-COR Biosciences; Fig. 7). Interestingly, all nanoparticles were uptaken by haematopoietic organs, such as the liver (~84% of total signals), spleen (~14%) and bone marrow (~1.6%). The lungs (~0.23%), lymph nodes (~0.07%) and kidneys (~0.08%) showed less uptake of all nanoparticle subtypes. We did not detect particle uptake in the brain. Subsequently, the dynamic range of signal intensity in each organ was adjusted to compare the uptake of each subset of nanoparticles in the same organ (Fig. 7a). Punctate distribution patterns of nanoparticles were detected specifically in the lung and lymph nodes. This is in contrast to the homogeneous distribution pattern found for all nanoparticle subsets in the liver, spleen and bone marrow. Importantly, although exomeres and Exo-S/L were predominantly taken up in the liver, Exo-L displayed lymph node tropism. In addition, although not statistically significant, a trend of higher uptake of exomeres in the liver was observed. Quantification is shown in Fig. 7b. Distinct organ distributions indicate that nanoparticle subsets may be involved in different aspects of tumour progression and metastasis.

Discussion

Dissecting the heterogeneity of EV populations by differential ultracentrifugation, immuno-affinity capture, ultrafiltration and size-exclusion chromatography, polymer-based precipitation and microfluidics[28–33] in an attempt to separate nanoparticle populations has proven daunting. By employing state-of-the-art AF4 technology, we have succeeded in separating two discernible exosome subpopulations, Exo-S and Exo-L, and identified a distinct nanoparticle, the exomere, which differs in size and content from other reported particles. Unlike labour-intensive and time-consuming gradient methods, AF4 is highly reproducible, fast, simple, label-free and gentle. Moreover, we were able to efficiently resolve the exosome subpopulations and exomeres in a single AF4 run with real-time measurements of various physical parameters of individual particles.

Our analyses revealed that exomeres were selectively enriched in proteins involved in metabolism, especially 'glycolysis' and 'mTORC1' metabolic pathways, suggesting their potential roles in influencing the metabolic program in target organ cells, as well as in proteins associated with coagulation (for example, factors VIII and X) and hypoxia. Our proteomic analysis also showed that exomeres were enriched in key proteins controlling glycan-mediated protein folding control (CALR)[19] and glycan processing (MAN2A1, HEXB and GANAB)[20–22], suggesting exomere cargo may modulate glycosylation in distant recipient cells. Subcellular localization analysis of exomere-enriched proteins revealed their specific association with the ER, mitochondria and microtubules, demonstrating the potential roles of these proteins in exomere biogenesis and secretion.

Proteins unique to exosomes (Exo-L and Exo-S) versus exomeres were also identified. Multiple components of ESCRT complexes were specifically associated with Exo-S and Exo-L, but not observed within exomeres, suggesting a major role for ESCRT complexes in Exo-S/L but not exomere production. Other exosome-enriched proteins included Rab proteins, annexins, Hsp40 members and proteins involved in multiple signaling transduction pathways, such as integrins, G-proteins, JAK1 and TGFBRs.

We found further differences between Exo-S and Exo-L protein cargo. Flotillin 1, flotillin 2, tweety family member 3, tetraspanin 14 and ESCRT-I subunit VPS37B were specifically enriched in Exo-S. In contrast, levels of proteins such as annexin A1/A4/A5, charged multivesicular body protein 1A/2A/4B/5, vacuolar protein sorting 4 homologue B, DnaJ heat shock protein family (Hsp40) member A1 and myosin IC were relatively higher in Exo-L. Interestingly, tissue factor, a well-studied exosome protein[34], was enriched in Exo-L. It is thus plausible that exomeres and Exo-L cooperate to optimize the coagulation cascade in vivo.

Exo-S were predominantly enriched in proteins associated with endosomes, multivesicular bodies, vacuoles and phagocytic vesicles, while Exo-L were specifically enriched in plasma membrane, cell–cell contact/junction, late-endosome and trans Golgi network proteins. These data indicate that Exo-S are most likely bona fide/canonical exosomes (that is, derived from intraluminal vesicles of endosomal compartments), whereas Exo-L may represent non-canonical exosomes or probably sEVs of different subcellular origin (that is, plasma membrane budding).

Identifying specific exosome and exomere markers to better isolate and characterize these nanoparticles is critical to advancing our knowledge of EV biology. Because flotillin 1 and 2 were specifically associated with Exo-S, these proteins may represent reliable markers of conventionally defined exosomes. Other previously reported

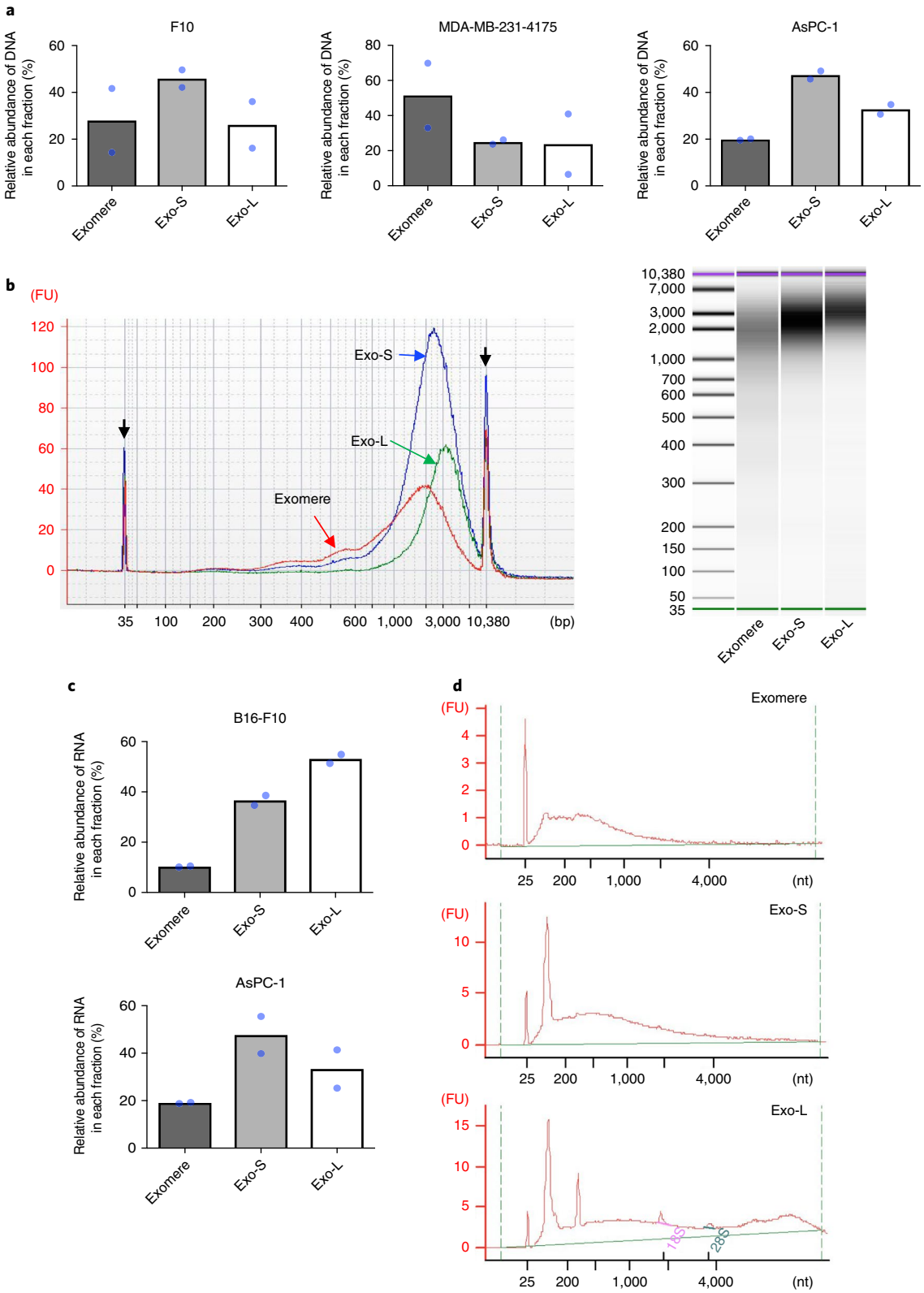

Fig. 6 | Characterization of nucleic acid association with exomere and exosome subsets. a, Relative abundance of DNA associated with each subpopulation of particles from representative fractionations of B16-F10, AsPC-1 and MDA-MB-231-4175. **b**, Agilent Bioanalyzer analysis of the size distribution of DNA associated with different subsets of particles. Data shown are the electropherograms (left) and electrophoresis images (right) from a representative of two independent experiments on AsPC-1-derived particles. Black arrows: internal standards (35 bp and 10,380 bp). Red line: exomeres; blue line: Exo-S; green line: Exo-L. FU, arbitrary fluorescent unit. **c**, Relative abundance of total RNA associated with each subpopulation of particles from representative fractionations of B16-F10 and AsPC-1. **d**, Size distribution of RNA isolated from different fractions of B16-F10. Shown are representative profiles from one of two independent experiments. For **a** and **c**, data shown are mean ($n = 2$ biologically independent samples). Statistical source data are provided in Supplementary Table 8.

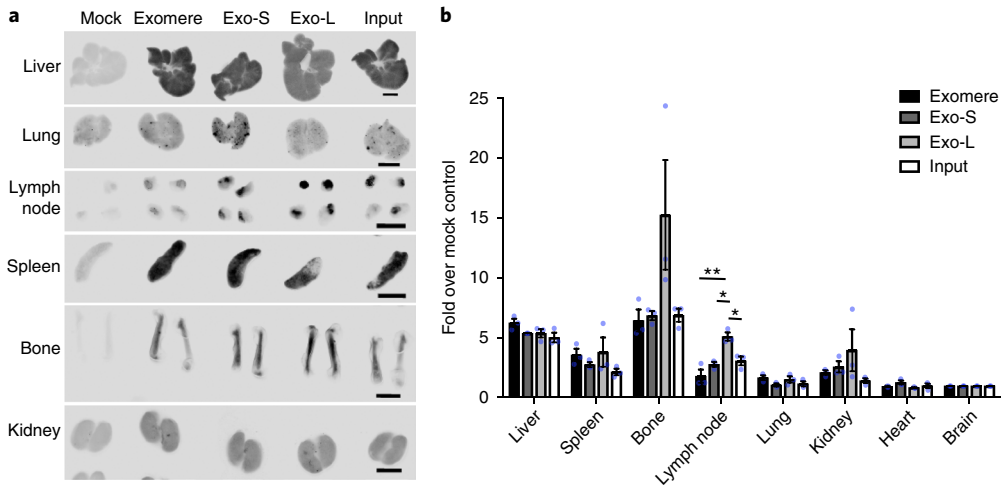

Fig. 7 | Organ biodistribution of B16-F10-derived exomeres and exosome subpopulations in syngeneic naive mice. a, Whole organ imaging of NIR dye-labelled B16-F10 exomeres, Exo-S and Exo-L from a representative experiment using the Odyssey imaging system (LI-COR Biosciences; $n = 4$ independent experiments). The dynamic range of signal intensity was adjusted for each organ so that the differences among these nanoparticle subsets can be easily recognized. Scale bars, 2.5 mm. **b**, Quantification of nanoparticle uptake in different organs in one representative experiment. This experiment was repeated independently four times with similar results. The signal intensity in each organ was acquired using Image Studio (LI-COR Biosciences) and normalized to the brain from the same animal due to undetectable uptake of nanoparticles in this organ. Fold changes (*y* axis) were then calculated for each organ between the experimental group (input, exomere, Exo-S and Exo-L) versus the mock control. $n = 3$ animals per group. Results shown are mean ± s.e.m. Statistical significance determined using one- way ANOVA (*$P < 0.05$; **$P < 0.01$; unmarked, not significant). For lymph nodes, the *P* values for comparisons input versus Exo-L, exomere versus Exo-L, and Exo-S versus Exo-L are 0.022, 0.001 and 0.01, respectively. Statistical source data are provided in Supplementary Table 8

exosome markers, including CD9, CD63, CD81, Tsg101 and Alix1, were present in Exo-S and/or Exo-L in a cell type-dependent manner and therefore would have to be combined with size exclusion to distinguish exosome subpopulations. Notably, Hsp90-β, highly represented in exomeres, could be a potential exomere marker, whereas several Hsp70 family members, such as HSC70/HSPA8, could serve as markers for Exo-S/L subpopulations.

Our glycomic, lipidomic and genomic studies also revealed additional distinct molecular signatures in exomeres and exosomes. Similar to the expression in metastatic tumour cells, exosome subsets were enriched in sialylated glycoproteins, supporting the role of these structures in exosome-mediated cellular recognition. One predominant sialoglycoprotein previously identified in exosomes[25,35], the galectin-3-binding protein (LGALS3BP), a modulator of cell communication and immune responses[36,37], was highly enriched in exomeres. This ligand could mediate the specific interaction of exomeres with target cells through proteins, such as collagens, fibronectin, nidogen, galectin-3 and integrin beta-1[38,39].

Interestingly, our lipidomic analyses revealed that exomeres contained fewer lipids compared to Exo-S and Exo-L. Phospholipids and SM, the major structural components of the plasma lipid bilayer membrane[40], ranked top in all nanoparticles examined. Such an observation is expected for Exo-S/L subsets due to their vesicular membrane structure; however, exomeres seem to lack external membrane structures. Yet, differences in several lipid classes distinguished exomeres from Exo-S and Exo-L. For instance, exomeres contained higher levels of triglycerides and ceramides than exosome subpopulations and thus may serve to transport these metabolites to recipient cells. Our study further revealed that DNA packaging in exomeres and exosomes varied by tumour type, while RNA was packaged in Exo-S and Exo-L independent of tumour classification.

Collectively, our findings demonstrate that proteins, glycans, lipids and nucleic acids are selectively packaged in exomeres, Exo-S and Exo-L, further supporting the idea that these are distinct nanoparticle subsets.

Our observation that nanoparticle subtypes have different organ biodistribution patterns suggests that they mediate the pleiotropic effects of cancer. The punctate pattern of Exo-L uptake and its apparent tropism for lymph nodes implicate this nanoparticle in facilitating metastasis of disseminated tumour cells. Exomeres, along with exosomes, were taken up by haematopoietic organs, including the liver, spleen and bone marrow. Interestingly, the predominant exomere uptake by the liver and the exomere enrichment in protein cargo involved in metabolism lead us to speculate that exomeres may specifically target the liver for metabolic reprogramming during tumour progression. Our data indicate that the size of nanoparticles, in addition to their specific cargo, may influence metastatic patterning and the systemic effects of cancer.

Our identification of exomeres highlights the diversity of EVs and particles secreted by cells. Elucidating their biogenesis will be essential to unravel their roles in cellular and organ function. Target cells and the functional outcomes exerted by each nanoparticle subset in organs need to be further delineated to advance our understanding of the collective, systemic effects of nanoparticles in the metastasis process. Undoubtedly, these discoveries will open avenues for translational studies of EVs and particles in diagnostic, prognostic and therapeutic applications.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41556-018-0040-4.

Received: 18 October 2016; Accepted: 12 January 2018;
Published online: 19 February 2018

References

1. Thery, C., Zitvogel, L. & Amigorena, S. Exosomes: composition, biogenesis and function. *Nat. Rev. Immunol.* **2**, 569–579 (2002).
2. El Andaloussi, S., Mager, I., Breakefield, X. O. & Wood, M. J. Extracellular vesicles: biology and emerging therapeutic opportunities. *Nat. Rev. Drug. Discov.* **12**, 347–357 (2013).

3. Raposo, G. & Stoorvogel, W. Extracellular vesicles: exosomes, microvesicles, and friends. *J. Cell. Biol.* **200**, 373–383 (2013).
4. Balaj, L. et al. Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. *Nat. Commun.* **2**, 180 (2011).
5. Choi, D. S., Kim, D. K., Kim, Y. K. & Gho, Y. S. Proteomics, transcriptomics and lipidomics of exosomes and ectosomes. *Proteomics* **13**, 1554–1571 (2013).
6. Thakur, B. K. et al. Double-stranded DNA in exosomes: a novel biomarker in cancer detection. *Cell. Res.* **24**, 766–769 (2014).
7. Tetta, C., Ghigo, E., Silengo, L., Deregibus, M. C. & Camussi, G. Extracellular vesicles as an emerging mechanism of cell-to-cell communication. *Endocrine* **44**, 11–19 (2013).
8. Fraunhofer, W. & Winter, G. The use of asymmetrical flow field-flow fractionation in pharmaceutics and biopharmaceutics. *Eur. J. Pharm. Biopharm.* **58**, 369–383 (2004).
9. Yohannes, G., Jussila, M., Hartonen, K. & Riekkola, M. L. Asymmetrical flow field-flow fractionation technique for separation and characterization of biopolymers and bioparticles. *J. Chromatogr. A* **1218**, 4104–4116 (2011).
10. Oh, S. et al. Miniaturized asymmetrical flow field-flow fractionation: application to biological vesicles. *J. Sep. Sci.* **30**, 1082–1087 (2007).
11. Sitar, S. et al. Size characterization and quantification of exosomes by asymmetrical-flow field-flow fractionation. *Anal. Chem.* **87**, 9225–9233 (2015).
12. Petersen, K. E. et al. A review of exosome separation techniques and characterization of B16-F10 mouse melanoma exosomes with AF4-UV-MALS-DLS-TEM. *Anal. Bioanal. Chem.* **406**, 7855–7866 (2014).
13. Ashby, J. et al. Distribution profiling of circulating microRNAs in serum. *Anal. Chem.* **86**, 9343–9349 (2014).
14. Agarwal, K. et al. Analysis of exosome release as a cellular response to MAPK pathway inhibition. *Langmuir* **31**, 5440–5448 (2015).
15. Beningo, K. A. & Wang, Y. L. Fc-receptor-mediated phagocytosis is regulated by mechanical properties of the target. *J. Cell. Sci.* **115**, 849–856 (2002).
16. Key, J. et al. Soft discoidal polymeric nanoconstructs resist macrophage uptake and enhance vascular targeting in tumors. *ACS Nano* **9**, 11628–11641 (2015).
17. Colombo, M., Raposo, G. & Thery, C. Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. *Annu. Rev. Cell. Dev. Biol.* **30**, 255–289 (2014).
18. Hessvik, N. P. & Llorente, A. Current knowledge on exosome biogenesis and release. *Cell. Mol. Life Sci.* **75**, 193–208 (2017).
19. Molinari, M. & Helenius, A. Chaperone selection during glycoprotein translocation into the endoplasmic reticulum. *Science* **288**, 331–333 (2000).
20. Fukuda, M. N., Masri, K. A., Dell, A., Luzzatto, L. & Moremen, K. W. Incomplete synthesis of N-glycans in congenital dyserythropoietic anemia type II caused by a defect in the gene encoding α-mannosidase II. *Proc. Natl Acad. Sci. USA* **87**, 7443–7447 (1990).
21. Yang, W. H. et al. An intrinsic mechanism of secreted protein aging and turnover. *Proc. Natl Acad. Sci. USA* **112**, 13657–13662 (2015).
22. Martiniuk, F., Ellenbogen, A. & Hirschhorn, R. Identity of neutral alpha-glucosidase AB and the glycoprotein processing enzyme glucosidase II. Biochemical and genetic studies. *J. Biol. Chem.* **260**, 1238–1242 (1985).
23. Kowal, J. et al. Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. *Proc. Natl Acad. Sci. USA* **113**, E968–977 (2016).
24. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: mechanisms and clinical implications. *Nat. Rev. Cancer* **15**, 540–555 (2015).
25. Escrevente, C. et al. Sialoglycoproteins and N-glycans from secreted exosomes of ovarian carcinoma cells. *PLoS ONE* **8**, e78631 (2013).
26. Batista, B. S., Eng, W. S., Pilobello, K. T., Hendricks-Munoz, K. D. & Mahal, L. K. Identification of a conserved glycan signature for microvesicles. *J. Proteom. Res.* **10**, 4624–4633 (2011).
27. Saraswat, M. et al. N-linked (N-) glycoproteomics of urinary exosomes. *Mol. Cell. Proteom.* **14**, 263–276 (2015).
28. Thery, C., Amigorena, S., Raposo, G. & Clayton, A. Isolation and characterization of exosomes from cell culture supernatants and biological fluids. *Curr. Protoc. Cell Biol.* **3**, 22.1–22.29 (2006).
29. Merchant, M. L. et al. Microfiltration isolation of human urinary exosomes for characterization by MS. *Proteom. Clin. Appl.* **4**, 84–96 (2010).
30. Lasser, C., Eldh, M. & Lotvall, J. Isolation and characterization of RNA-containing exosomes. *J. Vis. Exp.* **59**, e3037 (2012).
31. Chen, C. et al. Microfluidic isolation and transcriptome analysis of serum microvesicles. *Lab. Chip* **10**, 505–511 (2010).
32. Jorgensen, M. et al. Extracellular vesicle (EV) array: microarray capturing of exosomes and other extracellular vesicles for multiplexed phenotyping. *J. Extracell. Vesicles* **2**, 20920 (2013).
33. Tauro, B. J. et al. Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. *Methods* **56**, 293–304 (2012).
34. Gardiner, C. et al. Extracellular vesicles, tissue factor, cancer and thrombosis—discussion themes of the ISEV 2014 Educational Day. *J. Extracell. Vesicles* **4**, 26901 (2015).
35. Liang, Y. et al. Complex N-linked glycans serve as a determinant for exosome/microvesicle cargo recruitment. *J. Biol. Chem.* **289**, 32526–32537 (2014).
36. White, M. J., Roife, D. & Gomer, R. H. Galectin-3 binding protein secreted by breast cancer cells inhibits monocyte-derived fibrocyte differentiation. *J. Immunol.* **195**, 1858–1867 (2015).
37. Laubli, H. et al. Lectin galactoside-binding soluble 3 binding protein (LGALS3BP) is a tumor-associated immunomodulatory ligand for CD33-related Siglecs. *J. Biol. Chem.* **289**, 33481–33491 (2014).
38. Hellstern, S. et al. Functional studies on recombinant domains of Mac-2-binding protein. *J. Biol. Chem.* **277**, 15690–15696 (2002).
39. Sasaki, T., Brakebusch, C., Engel, J. & Timpl, R. Mac-2 binding protein is a cell-adhesive protein of the extracellular matrix which self-assembles into ring-like structures and binds beta1 integrins, collagens and fibronectin. *EMBO J.* **17**, 1606–1613 (1998).
40. van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9**, 112–124 (2008).

Acknowledgements

The authors acknowledge technical support from Wyatt Technology and especially J. Champagne. The authors also acknowledge the Genomics Resource Core facility (WCM) for their high-quality service. The authors thank C. Ghajar and J. Weiss for feedback on the manuscript and members of the Lyden laboratory for discussions. Our study was supported by the National Cancer Institute (U01-CA169538 to D.L.), the National Institutes of Health (NIH; R01-CA169416 to D.L. and H.P.; R01-CA218513 to D.L. and H.Z.), the US Department of Defense (W81XWH-13-10249 to D.L.), W81XWH-13-1-0425 (to D.L., J.Br.), the Sohn Conference Foundation (D.L., I.M., H.P. and H.Z.), the Children's Cancer and Blood Foundation (D.L.), The Manning Foundation (A.H. and D.L.), The Hartwell Foundation (D.L.), The Nancy C. and Daniel P. Paduano Foundation (D.L.), The Starr Cancer Consortium (H.P. and D.L.; D.L. and H.Z.), the Pediatric Oncology Experimental Therapeutic Investigator Consortium (POETIC; D.L.), the James Paduano Foundation (D.L. and H.P.), the NIH/WCM CTSC (NIH/NCATS: UL1TR00457 to H.M. and H.Z.; UL1TR002384 to D.L., H.M. and H.Z.), the Malcolm Hewitt Wiener Foundation (D.L.), the Champalimaud Foundation (D.L.), the Thompson Family Foundation (D.L., R.S.), U01-CA210240 (D.L.), the Beth Tortolani Foundation (J.Br.), the Charles and Marjorie Holloway Foundation (J.Br.), the Sussman Family Fund (J.Br.), the Lerner Foundation (J.Br.), the Breast Cancer Alliance (J.Br.), the Manhasset Women's Coalition Against Breast Cancer (J.Br.), the National Institute on Minority Health and Health Disparities (NIMHD) of the NIH (MD007599 to H.M.), NIH/NCATS (UL1TR00457 to H.M.). C.R., A.M., D.F., A.F., A.S. and H.O. acknowledge FEDER (Fundo Europeu de Desenvolvimento Regional funds through COMPETE 2020) POCI, Portugal 2020 (NORTE-01-0145-FEDER-000029) and FCT – Fundação para a Ciência e a Tecnologia in the framework of the project 'Institute for Research and Innovation in Health Sciences' (POCI-01-0145-FEDER-007274) and the FCT project POCI-01-0145-FEDER-016585 (PTDC/BBB-EBI/0567/2014). The authors acknowledge FCT for grants to A.M. (SFRH/BPD/75871/2011) and A.F. (SFRH/BPD/111048/2015). D.F. acknowledges FCT (SFRH/BD/110636/2015), the BiotechHealth PhD Programme (PD/0016/2012) and the American Portuguese Biomedical Research Fund.

Author contributions

H.Z. designed the experimental approach, performed the experimental work, analysed the data, coordinated the project and wrote the manuscript. K.F., S.R., J.F. and H.M. performed zeta potential and stiffness measurements. C.A.R., D.F. A.Mag., J.A.F., A.M.S. and H.O. conducted the glycomics analysis. M.T.M. and H.M. performed and analysed exosome MS. Z.L. conducted and analysed lipidomic MS. H.S.K. conducted proteomic data analysis. H.C., L.B., A.B.M., M.N., A.P.M., P.S., W.B., H.W., A.Mas., G.G. and J.R.C.-R. facilitated exosome isolation/fractionation and other experimental work. J.P.J. and L.C.-G. performed electron microscopy. N.P., M.B. and K.M.-T. performed atomic force microscopy. A.H., I.M. and C.K. contributed to manuscript reading, editing and providing feedback. P.G., A.M.C., J.Bl., R.S., H.P. and J.Br. discussed the hypothesis and contributed to data interpretation and manuscript writing. M.S.B. contributed to studies involving human specimens. D.L. conceived the hypothesis, led the project, interpreted the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41556-018-0040-4.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to H.Z. or D.L.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Cell lines and cell culture. B16-F10, B16-F1, 4T1, MDA-MB-231 series (parental, -1833, -4175 and -831, gifts from J. Massagué), LLC, SW620, HCT116 (Horizon Discovery), PANC-1, AsPC-1, Pan02 (purchased from the National Cancer Institute Tumor Repository) and NIH3T3 cells were cultured in DMEM. Human melanoma cells (SK-Mel103, A375M and A375P, obtained from Memorial Sloan Kettering Cancer Center (MSKCC)), human prostatic carcinoma cell lines PC3 and DU145, as well as BXPC-3, HPAF-II, PC-9, ET2B (gift from P. Gao and J. Bromberg), K-562 (DSMZ) and NB-4 (DSMZ) cells were cultured in RPMI, supplemented with penicillin (100 U ml^{-1}) and streptomycin (100 µg ml^{-1}) and 10% FBS. Cell lines were obtained from American Type Culture Collection, if not otherwise mentioned, and authenticated using short tandem repeat (STR) profiling by commercial providers. All cells were maintained in a humidified incubator with 5% CO_2 at 37 °C and routinely tested and confirmed to be free of mycoplasma contamination. When collecting conditioned media for exosome isolation, FBS (Gibco, Thermo Fisher Scientific) was first depleted of exosomes by ultracentrifugation at 100,000*g* for 90 min. Cells were cultured for 3 days before supernatant collection.

Human specimens and processing. Fresh human tumour tissues were obtained from subjects with stage 1–3 melanoma at MSKCC and who had histologically confirmed melanoma. All individuals provided informed consent for tissue donation according to a protocol approved by the institutional review board of MSKCC (IRB# 11-033A, MSKCC; IRB#0604008488, Weill Cornell Medicine (WCM)) and the study is compliant with all relevant ethical regulations regarding research involving human participants. Tissues were cut into small pieces and cultured for 24 h in serum-free RPMI supplemented with penicillin/streptomycin. Conditioned medium was processed for exosome isolation and AF4 fractionation as described in the following.

Exomere and exosome isolation and NTA. sEV were prepared using differential ultracentrifugation methods[41] and resuspended in phosphate buffered saline (PBS, pH7.4) for subsequent analysis and AF4 fractionation. Isolated samples were quantified using the bicinchoninic acid (BCA) assay (Pierce, Thermo Fisher Scientific). NTA analysis of exosome size and particle number was performed using an LM10 or NS500 NanoSight system (Malvern Instruments) equipped with a blue laser (405 nm) following the manufacturer's instructions.

AF4 fractionation. The detailed step-by-step AF4 fractionation protocol, including sample preparation, AF4 setting parameters and running method, data collection and analysis, and fraction collection and characterization, is provided on ProtocolExchange[42].

TEM and AFM. For negative-stain TEM analysis, 5 µl of sample solution was placed on a formvar/carbon-coated grid and allowed to settle for 1 min. The sample was blotted and negative-stained with four successive drops of 1.5% (aq.) uranyl acetate, blotting between each drop. Following the last drop of stain, the grid was blotted and air-dried. Grids were imaged with a JEOL JSM 1400 transmission electron microscope operating at 100 kV. Images were captured on a Veleta 2k×2k charge-coupled device camera (Olympus-SIS).

For AFM, dilutions were made for each sample and then plated on freshly cleaved mica substrate (SPI) for ~2 min before washing with 10 ml of molecular biology grade H_2O (Fisher BP2819-1) and being blown dry with nitrogen gas. Imaging was performed using an MFP-3D-BIO AFM (Asylum Research), with an Olympus AC240TS-R3 AFM probe (Asylum Research) in tapping mode at room temperature. Images were captured at 1 µm × 1 µm. Image analysis was performed using a custom-written ImageJ/FIJI (NIH) code.

Zeta potential measurement. Fractionated samples were diluted in PBS (0.01 M phosphate buffer, 0.0027 M KCl, 0.137 M NaCl; pH 7.4 tablets, Sigma) for ζ potential analysis using a Zetasizer Nano ZS90 (Malvern Instruments). Samples were freshly prepared prior to loading onto the instrument at a 90° angle (with respect to the light source). All experiments were performed at a constant temperature of 25 °C.

Stiffness measurement. Freshly cleaved mica coverslips were first coated with poly-L-lysine (0.1% wt/vol in H_2O) for 30 min and then incubated with samples on the mica surface for 45 min. The samples were then rinsed with 1 ml of MilliPure water, washed three times with PBS buffer, then submerged in a drop of PBS on the mica surface. A stand-alone MFP-3D atomic force microscope (Asylum Research) was used to perform the analysis. The spring constant of the cantilever was determined as 559.73 pN nm^{-1} by the thermal noise method[43]. The radius of curvature of the cantilever was ~15 nm, and a resonant frequency of 325 kHz was used for the stiffness analysis (indentation of the cantilever) and imaging. Force measurements were performed with an approximate force distance of 300 nm and velocity of 500 nm s^{-1}.

Western blot analysis. Whole cell extract (WCE) and exosome fractions were lysed directly with SDS sample buffer and lysates were cleared by centrifugation at 14,000*g* for 10 min. WCE (100 µg) and 10 µg of input and each nanoparticle subset were separated on a Novex 4–12% Bis-Tris Plus Gel (Life Technologies) and transferred onto a PVDF membrane (Millipore). Membranes were blocked for 1 h at room temperature (RT) followed by primary antibody incubation overnight at 4 °C. The following antibodies were used for western blot analysis: anti-Tsg101 (Santa Cruz sc-7964), anti-Alix1 (Cell Signaling 2171), anti-Hsp90 (Stressgen ADI-SPA-830-F), anti-MAT1A1 (Abcam ab174687), anti-IDH1 (Proteintech 23309-1-AP), anti-FLOT1 (BD Biosciences 610820), anti-TOLLIP (Abcam ab187198), anti-VPS4B (Santa Cruz sc-32922), anti-DNAJA1 (Abcam ab126774) and anti-HSPA8/HSC70 (LifeSpan Biosciences LS-C312344-100). All primary antibodies were used at 1:1,000 dilution. IRDye 800 CW Goat-anti-mouse IgG (LI-COR Biosciences P/N 926-32210, 1:15,000 dilution), horseradish peroxidase (HRP)-linked sheep-anti-mouse IgG (GE Healthcare Life Sciences NA931, 1:2,500 dilution) and HRP-linked donkey-anti-rabbit IgG (GE Healthcare Life Sciences NA934, 1:2,500 dilution) were used as secondary antibody. The blot was analysed either using the Odyssey imaging system (LI-COR Biosciences) or enhanced chemiluminescence substrates (Thermo Fisher Scientific).

Analysis of proteomic profiling data. Protein MS analyses of fractionated exosomes were performed at the Rockefeller University Proteomics Resource Center as described previously[44,45], and conducted on two independent biological replicates for each sample (exomere, Exo-S and Exo-L) derived from five different cell lines (B16-F10, 4T1, Pan02, AsPC-1 and MDA-MB-231-4175). Raw data are provided in Supplementary Table 2.

For proteomic data processing and principal component analysis (PCA), the proteomic expression data were processed using the 'Limma' package of the R program (https://www.r-project.org, v3.2.5). Proteomic expression data were imported and normalized using the 'normalizeBetweenArrays' function (method=quantile)[46]. PCA was performed for data reduction, simplifying data sets to three dimensions for plotting purposes using the 'princomp()' function with default options, and illustrated using the 'rgl' package and 'plot3d()' function.

For clustering and marker selection, consensus clustering analysis, marker selection for each fraction, and heatmap generation were conducted using GENE-E software (https://software.broadinstitute.org/GENE-E/). Consensus clustering was conducted to assess whether proteomic expression differed between fractions[47]. To identify fraction-specific markers, the probe (based on UniProt ID) values were collapsed to protein level using maximum probe. Only proteins detected in both replicates of a sample were included for further analysis. Proteins were sorted by the signal-to-noise statistic $(\mu_A - \mu_B)/(\alpha_A + \alpha_B)$ where μ and α are the mean and standard deviation of proteomic expression, respectively, for each class[48]. Next, the signal-to-noise marker selection tool from GENE-E was used to identify fraction-specific markers with 1,000 permutations. The cutoff to select fraction-specific markers was a fold change of ≥5, a false discovery rate (FDR) of <0.05 and mean protein expression of $\geq 1 \times 10^8$ with positivity in ≥80% (that is, at least four of five samples from five cell lines for each nanoparticle subset) of the corresponding fraction. Heatmaps for visualization of differential protein expression patterns were generated for 65 markers (39 exomere-specific markers, 5 Exo-S markers and 21 Exo-L markers) using GENE-E with relative colour scheme (by subtracting each mean protein expression, and dividing by each standard deviation for each row).

For GSEA we used the entire proteomic expression data set[49]. Gene sets from the Molecular Signatures database (MSigDB, http://www.broadinstitute.org/gsea/msigdb/index.jsp) v5.1 were used for GSEA (H: 50 hallmark gene sets; C2:KEGG: 186 canonical pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database; C5: 825 gene sets based on gene ontology (GO) term)[50]. Default parameters were used to identify significantly enriched gene sets (FDR $q < 0.25$).

Glycoprotein extraction and lectin blotting. Nanoparticles were lysed with RapiGest SF (Waters) containing 1 mM sodium orthovanadate and protease inhibitor cocktail (Roche) for 30 min on ice, then centrifuged at 16,000*g* for 20 min. For lectin blotting, 0.5 µg of total protein extracts were separated using 4–15% gradient gels (Biorad) and transferred onto nitrocellulose membranes. Samples were incubated at RT for 1 h with the following biotinylated lectins: *Aleuria aurantia* lectin (AAL; Fucα6GlcNAc and Fucα3GlcNAc), *Sambucus nigra* lectin (SNA; Neu5Acα6(Gal or GalNAc)), *Phaseolus vulgaris* leucoagglutinin (L-PHA; Galβ4GlcNAcβ6(GlcNAcβ2Manα3)Manα3) and *Phaseolus vulgaris* erythroagglutinin (E-PHA; Galβ4GlcNAcβ2Manα6(GlcNAcβ4)(GlcNAcβ4Manα3)Manβ4) (Vector Laboratories, 1:2,000 dilution, except 1:1,000 for L-PHA). A Vectastain Elite ABC HRP kit (Vector Laboratories) was used for signal detection with an ECL enhanced chemiluminescence technique (GE Healthcare Life Sciences). The total protein profile of the samples was assessed in parallel on a silver-stained gel (Supplementary Fig. 5a). (Abbreviations: Fuc, fucose; GlcNAc, *N*-acetylglucosamine; Man, mannose; Neu5Ac, neuraminic acid; Gal, galactose; GalNAc, *N*-acetylgalactosamine.)

Glycomics analysis. Glycoprotein extracts from the different fractions were reduced, alkylated and digested with sequencing-grade, modified trypsin (Promega) using a standard proteomics protocol[51]. The *N*-glycans were analysed based on a modification of ref. [52]. Briefly, *N*-linked glycans were released with

PNGase F (*Elizabethkingia* meningoseptica; Sigma), deaminated, and partially purified using porous graphitized carbon solid-phase extraction cartridges (PGC-SPE, HyperSep-96-Hypercarb, 25 mg, Thermo Scientific) as described previously[53].

Glycan profiling and characterization were performed by MALDI–TOF/TOF MS (4800 Plus, SCIEX) using α-cyano-4-hydroxycinnamic acid (CHCA; 10 mg ml^{-1} in 50% acetonitrile (ACN)), operated in reflector negative mode (mass range of *m/z* 1,000–5,000) with external calibration (TOF/TOF calibration mixture, SCIEX). Three independent analytical measurements were performed. NanoHPLC high-resolution MS (HRMS) was used to validate the presence of most discriminative ions in MALDI–MS spectra using a nanoHPLC system (Dionex, 3000 Ultimate RSLCnano) coupled on-line to an LTQ-Orbitrap XL mass spectrometer (Thermo Scientific) equipped with a nano-electrospray ion source (Thermo Scientific, EASY-Spray source). *N*-Glycan chromatographic separation using porous graphitized carbon (PGC) was adapted from a procedure described previously[53]. We combined in series a nanoflow PGC column (Hypercarb, 150 mm × 75 µm inner diameter (ID), 3 µm particle size, Thermo Scientific) and a reversed-phase C18 column (EASY-Spray C18 PepMap, 100 Å, 150 mm × 75 µm ID and 3 µm particle size, Thermo Scientific). This allowed a better separation of carbohydrates and remaining tryptic peptides, while minimizing salt precipitation events encountered when a nanospray emitter was used directly after the PGC column. The mass spectrometer was operated in negative ion mode.

The monosaccharide compositions for the glycan precursors on MALDI–MS spectra were predicted using the GlycoMod tool (http://www.expasy.ch/tools/glycomod) considering mass accuracies bellow 10 ppm. The possibility of neutral exchanges with Na^+ and K^+ was considered for sialoglycans. The glycan structures were assigned based on nanoHPLC–PGC–HRMS analysis considering (1) molecular monoisotopic mass, (2) CID-MS/MS de novo sequencing and (3) PGC-LC relative retention times. In particular, α-2,3-linked and α-2,6-linked sialylated *N*-glycans were differentiated based on retention time (α-2,6 < α -2,3)[52]. For further validation, MS/MS fragmentation profiles were matched to glycosidic fragments calculated in silico on GlycoWorkBench (http://relax.organ.su.se/eurocarb/gwb/home.action)[54]. General understanding of mammalian *N*-glycosylation was used to determine some structural aspects, yet some structural ambiguity remained in a subset of the reported *N*-glycans, as indicated with brackets above select glycan structures. A semiquantitive approach was used to compare glycan compositions based on MALDI–MS assignments, taking into account the monoisotopic peak intensity. Glycan standards and negative controls were analysed in parallel. These results were validated based on the intensity of each species on nanoHPLC–HRMS ion chromatograms (EIC) ($m/z \pm 0.01$).

Lipidomics, sample preparation, MS and data analysis. Equal amounts of each sample (based on BCA quantification) were subjected to lipidomic analysis. Samples were first sonicated with a Model Q700 QSonica sonicator equipped with an Oasis 180 Chiller (4 °C; amplitude, 95; process, 5 min; pulse on 30 s; plus off 55 s), centrifuged at 14,800 r.p.m. for 10 min at 4 °C, then 50 µl of the extract supernatant was spiked with 2 µl, 50 µg ml^{-1} internal standard mixture (Cer 18:1/12:0; PC 12:0/12:0; PE 14:0/14:0; PG 14:0/14:0; PS 14:0/14:0). Subsequently, the samples were analysed by using the Thermo Q-Exactive MS system in the Metabolomics Laboratory of the Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign. The software Xcalibur 3.0.63 was used for data acquisition and analysis. A Dionex Ultimate 3000 series HPLC system (Thermo) was used, and LC separation was performed on a Thermo Accucore C18 column (2.1 × 150 mm, 2.6 µm) with mobile phase A (60% acetonitrile: 40% H_2O with 10 mM ammonium formate and 0.1% formic acid) and mobile phase B (90% isopropanol: 10% acetonitrile with 10 mM ammonium formate and 0.1% formic acid) and a flow rate of 0.4 ml min^{-1}. The linear gradient was as follows: 0 min, 70% A; 4 min, 55% A; 12 min, 35% A; 18 min, 15% A; 20–25 min, 0% A; 26–33 min, 70% A. The autosampler was set to 15 °C and the column was kept at 45 °C. The injection volume was 10 µl. Mass spectra were acquired under both positive (sheath gas flow rate, 50; auxiliary gas flow rate, 13; sweep gas flow rate, 3; spray voltage, 3.5 kV; capillary temperature, 263 °C; auxiliary gas heater temperature, 425 °C) and negative (sheath gas flow rate, 50; auxiliary gas flow rate, 13; sweep gas flow rate, 3; spray voltage, −2.5 kV; capillary temperature, 263 °C; auxiliary gas heater temperature, 425 °C) electrospray ionization. The full scan mass spectrum resolution was set to 70,000 with a scan range of *m/z* ~230–1,600 and the automatic gain control (AGC) target was 1E6 with a maximum injection time of 200 ms. For MS/MS scans, the mass spectrum resolution was set to 17,500 and the AGC target was 5E4 with a maximum injection time of 50 ms. The loop count was 10. The isolation window was 1.0 *m/z* with normalized collision energy (NCE) of 25 and 30 eV. For data analysis, LipidSearch (v.4.1.30, Thermo) was used for lipid identification. The lipid signal responses were normalized to the corresponding internal standard signal response. For those lipid classes without corresponding internal standard, positive lipid ion signals were normalized with the signal of internal standard Cer 18:1/12:0 and negative ion signals were normalized with the signal of internal standard PG 14:0/14:0. The percentage of lipid classes within a sample was calculated by adding that of each of the individual molecular species quantified within a specific lipid class, and the relative abundance was represented by the mean percentage of three replicates for each group of samples. Differences among different subpopulations of particles derived from the same cell line were analysed using ANOVA test ($q < 0.05$).

Nucleic acid analysis. DNA was extracted from nanoparticles using the AMPure XP beads (Agencourt) following the manufacturer's protocol. Equal volumes of nanoparticles in PBS and lysis buffer AL (QIAGEN) were mixed and incubated with proteinase K (20 µg ml^{-1}, QIAGEN) at 56 °C for 10 min. The mixture was mixed with one volume each of AMPure beads, isopropanol and PEG solution (Beckman) and incubated for 5 min at RT. DNA bound to the beads was then separated from the solution/supernatant on a magnet for 5 min at RT. The supernatant was removed by pipetting, and bead-bound DNA was washed twice with freshly prepared 80% ethanol, then air-dried for 5 min. Finally, DNA was eluted from beads with nuclease-free water and quantified using a QuBit assay (Life Technology). DNA extraction was performed for two independent biological replicates of each sample.

RNA was extracted using the Ambion mirVarna kit (Life Techonology) following the manufacturer's protocol with one modification: one volume of nanoparticles in PBS was first lysed with seven volumes of lysis buffer. The samples were analysed using Agilent Total RNA Pico kits. RNA extraction was performed for two independent biological replicates of each sample.

Biodistribution assessment. Fractionated nanoparticles were first labelled with the near-infrared dye CellVue NIR815 (eBioscience) following the manufacturer's protocol, followed by washing with 20 ml of PBS and pelleting by ultracentrifugation at 100,000*g* for 70 min at 10 °C. Labelled nanovesicles (10 µg) resuspended in 100 µl of PBS, or an equivalent volume of mock reaction mixture, were retro-orbitally injected into naive mice (6-week-old female C57BL/6 mice purchased from Jackson Labs). At 24 h post-injection, tissues were collected and analysed using the Odyssey imaging system (LI-COR Biosciences). Two independent experiments with three animals per group were performed. No statistical method was used to predetermine sample size. The experiments were neither randomized nor blinded. All animal experiments were performed in compliance with ethical regulations and in accordance with WCM institutional, IACUC and AAALAS guidelines, approved for animal protocol 0709-666A.

Code availability. Custom-written ImageJ/FIJI (NIH) code for AFM image analysis is fully available from the corresponding authors upon reasonable request.

Statistics and reproducibility. Error bars in graphical data represent means ± standard error of the mean (s.e.m.). Statistical significance was determined using one-way ANOVA. $P < 0.05$ was considered statistically significant. Statistical analyses were performed using GraphPad Prism software. For lipid class analysis, the ANOVA test ($q < 0.05$) was performed using Qlucore Omics Explorer (Sweden). For proteomic analysis, proteins were sorted by the signal-to-noise statistic $(\mu_A - \mu_B)/(\alpha_A + \alpha_B)$, where μ and α represent the mean and standard deviation of proteomic expression, respectively. The cutoff to select fraction-specific markers was a fold changeof ≥5, an FDR of <0.05 and mean protein expression of $\geq 1 \times 10^8$ with positivity in ≥80% (at least four out of five samples from five different cell lines for each subset of nanoparticles) of the corresponding fraction. For GSEA, the Kolmogorov–Smirnov statistic was calculated to evaluate whether proteins from a predetermined pathway are significantly over-represented towards the top or bottom of the ranked gene list (FDR $q < 0.25$).

Multiple AF4 analyses were performed for each cell line studied in this work: B16-F10, >50× (repeated times); AsPC-1, 9×; Pan02, 16×; MDA-MB-231-4175 (4175), 17×; 4T1, 10×. TEM imaging analysis of fractionated particles was conducted for B16-F10, 7×; AsPC-1, 3×; Pan02, 2×; 4175, 1×; 4T1, 4×. Four independent human melanoma specimens were analysed using AF4, and two of them were analysed by TEM. Proteomic profiling of exomeres, Exo-S and Exo-L was performed on two biologically independent samples of each particle derived from five different cell lines (B16-F10; AsPC-1; Pan02; 4175; 4T1). Western blotting validation of specific signature proteins of each particle subtype was done once (noted in the legend for Fig. 1d). For *N*-glycan study, lectin blotting was repeated independently twice except for AAL and E-PHA blotting for B16-F10 and 4175, which were performed once (Fig. 4a). Glycomic MS was performed on two biologically independent B16-F10 samples and one sample of AsPC-1 and 4175 (Supplementary Fig. 5b–d). Quantification of the top six most abundant glycans was based on three independent analytical measurements of one experiment (Fig. 4c and Supplementary Fig. 5c,d). Silver-stained PAGE analysis was repeated independently twice for B16-F10 and 4175 and once for AsPC-1 (Supplementary Fig. 5a). NanoHPLC-PGC-HRMS was carried out once (Supplementary Fig. 5e–i). Lipidomic analysis was conducted on three biologically independent samples. DNA and RNA analyses of each particle subtype were repeated twice. Organ biodistribution analysis of each particle subtype was repeated four times independently. NTA analysis was conducted using three biologically independent samples. TEM analysis was repeated three times for AF4 peaks P1 and P5 and once for HDL, LDL and VLDL (Supplementary Fig. 4d). AF4 analyses of B16-F10 sEVs collected from technical and biological replicates, and samples kept at either 4 °C or −80 °C, were $^{\text{repeated independently}}$ three times, cells of different passage numbers twice, and under hypoxic versus normoxic conditions were repeated with three

different cell lines independently. AF4 and TEM analysis of particles isolated from the blank medium control and CM of B16-F10 and 4175 was carried out once (Supplementary Fig. 1j,k).

Independent measurements of hydrodynamic diameters of exomeres, Exo-S and Exo-L derived from different cell lines in batch mode were repeated as follows (in the order of exomere, Exo-S and Exo-L): B16-F10 ($n=10$, 9 and 8 independent measurements, respectively); Pan02 ($n=11$, 6, 11); AsPC-1 ($n=5$, 5, 5); 4175 ($n=3$, 5, 3); 4T1 ($n=5$, 5, 5). For the zeta potential, independent measurements were repeated as follows: B16-F10 ($n=8$, 10, 12); Pan02 ($n=13$, 11, 13); AsPC-1 ($n=12$, 12, 12); 4175 ($n=17$, 9, 6); 4T1 ($n=13$, 3, 9). For stiffness independent measurements were repeated as follows: B16-F10 ($n=6$, 6, 6); Pan02 ($n=6$, 6, 6); AsPC-1 ($n=21$, 19, 16); 4175 ($n=11$, 10, 5); 4T1 ($n=9$, 8, 9). For AFM imaging analysis of the height of exomeres independent measurements were repeated as follows: B16-F10 ($n=754$ particles analysed); AsPC-1 ($n=475$); 4175 ($n=160$). AFM imaging of exomeres was repeated with samples derived from three different cell lines.

For all experiments described above, all attempts at replication were successful, with similar results.

Data availability. The data sets for proteomic analysis of exomeres, Exo-S and Exo-L subpopulations derived from various cancer cell lines (Supplementary Table 2) have been deposited at https://figshare.com/s/302419bafecaae26b653.

Proteins that are uniquely associated with or are among the top 50 most abundant proteins in exomere, Exo-S and Exo-L derived from different cancer cell lines (Supplementary Table 4) have been deposited at https://figshare.com/s/5081b49c6716bbc8d630. Proteomics analysis of lipoprotein particles (Supplementary Table 5) has been deposited at https://figshare.com/s/031571ce9dd63aca4529. GSEA results for proteins associated with exomeres, Exo-S and Exo-L derived from various cancer cell lines (Supplementary Table 6) have been deposited at https://figshare.com/s/633ffe2120e23acc076d. Lipid classes identified in exomeres and exosome subsets derived from different cell lines (raw data and normalized data, Supplementary Table 7) have been deposited at https://figshare.com/s/0573bf5335bb46ee895e. Source data for Figs. 1e, 1g, 2a,b,d, 4c, 5a,b, 6a,b, 7b and Supplementary Figs. 1b,d, 4c, 5c,d have been provided as Supplementary Table 8. All other data supporting the findings of this study are available from the corresponding authors upon reasonable request.

References

41. Peinado, H. et al. Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat. Med.* **18**, 883–891 (2012).
42. Zhang, H. & Lyden, D. A protocol for asymmetric-flow field-flow fractionation (AF4) of small extracellular vesicles. *Protocol Exchange* https://doi.org/10.1038/protex.2018.002 (2018).
43. Langlois, E. D., Shaw, G. A., Kramar, J. A., Pratt, J. R. & Hurley, D. C. Spring constant calibration of atomic force microscopy cantilevers with a piezosensor transfer standard. *Rev. Sci. Instrum.* **78**, 093705 (2007).
44. Costa-Silva, B. et al. Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat. Cell Biol.* **17**, 816–826 (2015).
45. Hoshino, A. et al. Tumour exosome integrins determine organotropic metastasis. *Nature* **527**, 329–335 (2015).
46. Bolstad, B. M., Irizarry, R. A., Astrand, M. & Speed, T. P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185–193 (2003).
47. Monti, S., Tamayo, P., Mesirov, J. & Golub, T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* **52**, 91–118 (2003).
48. Golub, T. R. et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
49. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
50. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
51. Ferreira, J. A. et al. Synthesis and optimization of lectin functionalized nanoprobes for the selective recovery of glycoproteins from human body fluids. *Anal. Chem.* **83**, 7035–7043 (2011).
52. Kolarich, D., Windwarder, M., Alagesan, K. & Altmann, F. Isomer-specific analysis of released N-glycans by LC-ESI MS/MS with porous graphitized carbon. *Methods Mol. Biol.* **1321**, 427–435 (2015).
53. Jensen, P. H., Karlsson, N. G., Kolarich, D. & Packer, N. H. Structural analysis of N- and O-glycans released from glycoproteins. *Nat. Protoc.* **7**, 1299–1310 (2012).
54. Ceroni, A. et al. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Prote. Res.* **7**, 1650–1659 (2008).

nature research

Corresponding author(s): David Lyden, MD, PhD and Haiying Zhang, PhD

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   No statistical method was used to predetermine sample size.

2. Data exclusions

   Describe any data exclusions.

   No data was excluded

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   Multiple AF4 analyses were performed for each cell line studied in this work: B16-F10, > 50x (repeated times); AsPC-1, 9x; Pan02, 16x; MDA-MB-4175 (4175), 17x; and 4T1, 10x. TEM imaging analysis of fractionated particles were conducted for B16-F10, 7x; AsPC-1, 3x; Pan02, 2x; 4175, 1x; and 4T1, 4x. Four independent human melanoma specimens were analyzed using AF4 and two of them were analyzed by TEM. Proteomic profiling of exomeres, Exo-S and Exo-L was performed on two biologically independent samples of each particle derived from five different cell lines (B16-F10; AsPC-1; Pan02; 4175; and 4T1). Western blotting validation of specific signature proteins of each particle subtype was done once (noted in the legend for Fig. 1d). For N-glycan study, lectin blotting was repeated independently twice except for AAL and E-PHA blotting for B16-F10 and 4175 which were done once (Fig. 4a). Glycomic MS was performed on two biologically independent B16-F10 samples and one sample of AsPC-1 and 4175 (Supplementary Fig. 5b-d). Quantification of top 6 most abundant glycans was based on 3 independent analytical measurements of one experiment (Fig. 4c, Supplementary Fig. 5c and d). Silver stained-PAGE analysis was repeated independently twice for B16-F10 and 4175 and once for AsPC-1 (Supplementary Fig. 5a). NanoHPLC-PGC-HRMS was done once (Supplementary Fig. 5 e-i). Lipidomic analysis was conducted on 3 biologically independent samples. DNA and RNA analyses of each particle subtype were repeated twice. Organ biodistribution analysis of each particle subtype was repeated 4x independently. NTA analysis was conducted using 3 biologically independent samples. TEM analysis was repeated 3 times for AF4 peaks P1 and P5 and once for HDL, LDL and VLDL (Supplementary Fig. 4d). AF4 analysis of B16-F10 sEVs collected from technical and biological replicates, and samples kept at either 4 oC or -80 oC were repeated independently 3 times, cells of different passage numbers twice, and under hypoxic versus normoxic conditions was repeated with 3 different cell lines independently. AF4 and TEM analysis of particles isolated from the blank media control and CM of B16-F10 and 4175 was done once (Supplementary Fig. 1 j and k).

   Independent measurements of hydrodynamic diameters of exomeres, Exo-S and Exo-L derived from different cell lines in batch mode were repeated (in the order of exomere, Exo-S and Exo-L): B16-F10 (n=10, 9, and 8 independent measurements, respectively); Pan02 (n=11, 6, 11); AsPC-1 (n=5, 5, 5); 4175 (n=3, 5, 3); 4T1 (n=5, 5, 5)). For zeta potential, independent measurements were repeated: B16-F10 (n=8, 10, 12); Pan02 (n=13, 11, 13); AsPC-1 (n=12, 12, 12); 4175 (n=17, 9, 6); 4T1 (n=13, 3, 9). For stiffness, B16-F10 (n=6, 6, 6); Pan02 (n=6, 6, 6); AsPC-1 (n=21, 19, 16); 4175 (n=11, 10, 5); 4T1 (n=9, 8, 9). For AFM imaging analysis of the height of exomeres: B16F10 (n=754 particles analyzed), AsPC1 (n=475) and 4175 (n=160). AFM imaging of exomeres was repeated with samples derived from 3

different cell lines.

For all experiments described above, all attempts at replication were successful with similar results.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

No methods of randomization were used due to the fact that no subjective scoring for data evaluation is required in this work. All quantifications were done automatically using specific instruments.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

The investigators were not blinded to allocation during experiments and outcome assessment. This is due to the fact no subjective scoring for data evaluation was required for this work. All quantifications were done automatically using specific instruments.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.

Prism 7 (version 7.02) was used for statistical analysis and graphical data presentation (Graphpad software). GENE-E software was used for heatmap generation (http://www.broadinstitute.org/software/gene-e). ImageJ (version 1.51) was used for image analysis. Gene Ontology (GO) analysis was performed using KOBAS 3.0 (KEGG Orthology Based Annotation System), For proteomic data processing and Principal Component Analysis, the proteomic expression data was processed using the 'Limma' package of the R program (https://www.r-project.org, v3.2.5). MS/MS spectra were extracted and searched against Uniprot complete human or mouse proteome databases (January 2013) concatenated with common contaminants67 using Proteome Discoverer 1.4 (Thermo Scientific) and Mascot 2.4 (Matrix Science).The monosaccharide compositions for the glycan precursors on MALDI-MS spectra were predicted using the GlycoMod tool (http://www.expasy.ch/tools/glycomod) .MS/MS fragmentation profiles were matched to glycosidic fragments calculated in silico on GlycoWorkBench (http://www.eurocarbdb.org/applications/ms-tools). For data analysis, LipidSearch (v.4.1.30, Thermo) was used for lipid identification. In vivo nanoparticle biodistribution was quantified using the ImageStudio software (LI-COR). Chemstation software (Agilent Technologies) with integrated Eclipse module (Wyatt Technology) was used to operate the AF4 flow and Astra 6 (Wyatt

Technology) was used for data acquisition and analysis. Image J (NIH) bundled with 64-bit Java 1.8.0_112 was used for AFM analysis.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	Materials used are available from the authors and standard commercial sources.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	The following antibodies were used for western blot analysis: anti-Tsg101 (Santa Cruz sc-7964); anti-Alix1 (Cell Signaling 2171); anti-Hsp90 (Stressgen ADI-SPA-830-F), anti-MAT1A1 (Abcam ab174687); anti-IDH1 (Proteintech 23309-1-AP); anti-FLOT1 (BD Biosciences 610820); anti-TOLLIP (Abcam ab187198);anti-VPS4B (Santa Cruz sc-32922); anti-DNAJA1 (Abcam ab126774); anti-HSPA8/HSC70 (LifeSpan Biosciences LS-C312344-100). IRDye 800 CW Goat-anti-mouse IgG (LI-COR Biosciences P/N 926-32210), HRP-linked Sheep-anti-Mouse IgG (GE Healthcare Life Sciences NA931), and HRP-linked Donkey-anti-Rabbit IgG (GE Healthcare Life Sciences NA934) were used as secondary antibody. All antibodies are commercially available and have been validated by the manufacturer. All primary antibodies were used at 1:1000x dilution. IRDye 800 CW -Goat-anti-mouse IgG was used at 1:15,000 dilution), and HRP-linked secondary antibodies were used at 1:2,500 dilution.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	The following cell lines were used: B16-F10, B16-F1, 4T1, PANC-1, AsPC-1, LLC, SW620, PC3, DU145, BXPC-3, HPAF-II, PC-9, and NIH3T3 cells were all obtained from ATCC. MDA-MB231 series (parental, -1833, -4175, and -831) were gifts from Dr. J. Massagué, MSKCC. ET2B cells were a gift from Dr. P. Gao and J. Bromberg. HCT116 were obtained from Horizon Discovery. Leukemia cell lines K-562 and NB-4 were obtained from DSMZ. Human melanoma cell lines SK-Mel103, A375M and A375P were obtained from Memorial Sloan-Kettering Cancer Center (MSKCC). The murine pancreatic adenocarcinoma PAN02 was purchased from the National Cancer Institute Tumour Repository. Authentication using STR profiling was performed by commercial providers (except for ET2B which was not authenticated)
b. Describe the method of cell line authentication used.	Cell lines were authenticated - using STR profiling
c. Report whether the cell lines were tested for mycoplasma contamination.	Cell lines were routinely checked for mycoplasma and were always found to be mycoplasma negative.
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	No cell lines listed as misidentified were used in this study.

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.	Mice used for biodistribution experiments were 6 week-old C57BL/6 females purchased from Jackson laboratories.

Policy information about studies involving human research participants

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Melanoma patient tumor explant and lymphatic fluid studies were conducted in accordance with Weill Cornell Medicine IRB-approved protocol (IRB Protocol# 0604008488, Dr. David Lyden) and MSKCC IRB protocol # 11-033A, led by Dr. Mary-Sue Brady.
Sample ID: Mel280LN
Age at surgery: 56
Gender: M
Genotype information: negative for BRAFV600E, ALK and NRASQ61R
Diagnosis at time of surgery: Metastatic Melanoma of unknown primary in one axillary lymph node.
Treatment Categories: Ipilimumab x 2 doses (1/12/17, 2/2/17)

Sample ID: Mel376LN
Age at surgery: 76
Gender: F
Genotype information: N/A
Diagnosis at time of surgery: Metastatic Melanoma in one inguinal lymph node
Treatment Categories: Pembro (04/20/17, 05/11/17)

Sample ID: Mel360
Age at surgery: 75
Gender: F
Genotype information: positive for BRAFV600E
Diagnosis at time of surgery: Metastatic Melanoma, involving serosa, muscularis and submucosa of small intestine
Treatment Categories: Pembro 05/02/16, 05/23/16, 06/27/2016, 07/08/2016, 08/08/2016, 08/29/2016, 09/26/2016

Sample ID: Mel5004
Age at surgery: 49
Gender: M
Genotype information: negative for AKT1, BRAF, EGFR, ERBB2, KRAS, MEK1, NRAS, ans PIK3CA
Diagnosis at time of surgery: Residual paracicatricial melanoma in situ (skin, upper arm); Metastatic melanoma (right axillary mass)
Treatment Categories: Pembro (5/23/17, 6/13/17, 7/5/17, 7/25/17, 8/15/17, 9/5/17, 9/26/17, 10/17/17, 11/7/17, 11/28/17)