



OPEN The role of GCNT1 mediated O-glycosylation in aggressive prostate cancer

Kirsty Hodgson^{1,6}, Margarita Orozco-Moreno^{1,6}, Emma Scott^{1,6}, Rebecca Garnham¹, Karen Livermore¹, Huw Thomas², Yuhan Zhou³, Jiepei He³, Abel Bermudez⁴, Fernando Jose Garcia Marques⁴, Kayla Bastian¹, Gerald Hysenaj¹, Emily Archer Goode¹, Rakesh Heer^{2,5}, Sharon Pitteri⁴, Ning Wang³, David J. Elliott¹ & Jennifer Munkley¹✉

Prostate cancer is the most common cancer in men and a major cause of cancer related deaths worldwide. Nearly all affected men develop resistance to current therapies and there is an urgent need to develop new treatments for advanced disease. Aberrant glycosylation is a common feature of cancer cells implicated in all of the hallmarks of cancer. A major driver of aberrant glycosylation in cancer is the altered expression of glycosylation enzymes. Here, we show that GCNT1, an enzyme that plays an essential role in the formation of core 2 branched O-glycans and is crucial to the final definition of O-glycan structure, is upregulated in aggressive prostate cancer. Using in vitro and in vivo models, we show GCNT1 promotes the growth of prostate tumours and can modify the glycome of prostate cancer cells, including upregulation of core 2 O-glycans and modifying the O-glycosylation of secreted glycoproteins. Furthermore, using RNA sequencing, we find upregulation of GCNT1 in prostate cancer cells can alter oncogenic gene expression pathways important in tumour growth and metastasis. Our study highlights the important role of aberrant O-glycosylation in prostate cancer progression and provides novel insights regarding the mechanisms involved.

Prostate cancer is a major cause of male cancer-related deaths around the world¹. The progression of prostate cancer is driven by the androgen receptor (AR) and first line treatments for advanced prostate cancer include AR-targeted therapies². However, in a significant number of cases, tumours progress to an aggressive state for which AR-targeted therapies become ineffective, known as metastatic castrate resistant prostate cancer (mCRPC)^{3–5}. Several second-generation agents targeting AR signalling, including enzalutamide⁶, abiraterone⁷, and darolutamide⁸, can be used to treat mCRPC, but nearly all affected men will also develop resistance to these agents^{3–5} and there is an urgent need to discover new ways to treat prostate cancer.

Glycosylation is the most common, complex, and dynamic post-translational modification of lipids and proteins and is essential for every biological process⁹. Aberrant glycosylation is an established hallmark of cancer cells^{10,11} and glycans hold huge potential for cancer research. In prostate cancer, glycan alterations are common and typically include increased tumour sialylation, truncated O-glycans, increased core fucosylation, and changes to PSA glycosylation^{12,13}. A key driver of aberrant glycosylation in cancer is the dysregulated expression of glycosyltransferase enzymes¹⁴. We previously revealed glycosylation is an androgen regulated process in prostate cancer cells and identified a set of glycosylation enzymes that are upregulated in clinical prostate tumour tissue^{15,16}. In this study, we focus on GCNT1 (also known as C2GNT1), a glycosyltransferase enzyme that plays an essential role in the formation of core 2 branched O-glycans, an early step in O-glycan synthesis that promotes chain branching and elongation and is crucial to the final definition of O-glycan structure^{17–19}.

Aberrant expression of GCNT1 has been related to aggressive disease in different types of cancers, including breast cancer²⁰, prostate cancer^{12,15,16,21–23}, chronic lymphocytic leukaemia (CLL)²⁴, endometrial cancer²⁵, colon

¹Newcastle University Centre for Cancer, Newcastle University Institute of Biosciences, Newcastle NE1 3BZ, UK. ²Newcastle University Centre for Cancer, Translational and Clinical Research Institute, Newcastle University, Paul O’Gorman Building, Newcastle upon Tyne NE2 4HH, UK. ³Department of Oncology and Metabolism, The Mellanby Centre for Musculoskeletal Research, The University of Sheffield, Sheffield, UK. ⁴Department of Radiology, Canary Center at Stanford for Cancer Early Detection, Stanford University, Palo Alto, CA 94304, USA. ⁵Department of Urology, Freeman Hospital, The Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne NE7 7DN, UK. ⁶These authors contributed equally: Kirsty Hodgson, Margarita Orozco-Moreno and Emma Scott. ✉email: Jennifer.munkley@ncl.ac.uk

cancer²⁶, testicular germ cell cancer²⁷ and bladder cancer²⁸. In prostate cancer, GCNT1 is upregulated at both the gene and protein level^{12,16,21–23}, and has been linked to recurrence after surgery²³ and the spread of cancer cells outside of the prostate gland²¹. Findings show that upregulation of GCNT1 enhances the growth of orthotopic prostate tumours²⁹, correlates with higher levels of core 2 O-SLe^x on PSA, PAP, and MUC1 proteins²², and can promote resistance to NK cell immunity³⁰. However, the impact of GCNT1 on the global glycosylation of prostate cancer cells and the molecular mechanisms underlying the role GCNT1 in prostate cancer are largely unknown.

Here, we analyse four independent cohorts of patient RNA samples and further confirm upregulation of GCNT1 gene expression in aggressive prostate cancer. Our findings reveal GCNT1 can alter the glycome of prostate cancer cells, specifically by upregulating the expression of core 2 O-glycan structures (including the cancer-associated glycan SLe^x) and modifying the O-glycosylation of secreted glycoproteins. Using both in vitro and in vivo studies, we show that GCNT1 promotes the growth of prostate tumours and regulates oncogenic gene expression pathways that are important for disease progression. Our study highlights the role of GCNT1 in aggressive prostate cancer and provides novel insights into the role of aberrant glycosylation in disease pathology.

Results

GCNT1 gene expression levels are upregulated in aggressive prostate cancer tissue

The glycosyltransferase GCNT1 has been previously identified as upregulated in prostate cancer and correlates with aggressive disease at both the gene and protein level^{12,15,16,21–23}. Here, we monitor GCNT1 gene expression in four additional prostate cancer clinical cohorts. Analysis of RNA sequencing data from The Cancer Genome Atlas Prostate Adenocarcinoma (TCGA PRAD) cohort^{31,32} reveals GCNT1 gene expression levels are 2.3-fold higher in prostate tumours relative to normal prostate tissue (n = 544, p = 0.021) (Fig. 1A). Using quantitative PCR, we further show the GCNT1 gene is 1.8-fold upregulated in prostate cancer tissue relative to benign prostate hyperplasia in an additional patient cohort (n = 20, p = 0.0248) (Fig. 1B). Next, we monitored GCNT1 gene levels in a molecular subgroup of prostate cancer patients with metastatic potential at presentation (previously published by^{33,34}). Within this dataset, GCNT1 is 2.6-fold upregulated in the ‘metastatic subgroup’ compared to the ‘non-metastatic’ sub-group, suggesting upregulation of the GCNT1 gene in primary prostate cancer patients presenting with metastatic biology (n = 20, p = 0.0494) (Fig. 1C). Furthermore, GCNT1 gene levels are 7.5-fold higher in metastatic CRPC compared to hormone naïve disease (n = 20, p = 0.0481) (Fig. 1D). Our findings confirm upregulation of GCNT1 gene levels in prostate cancer tissue compared to normal or benign prostate tissue, and suggest GCNT1 may be upregulated in patients with increased risk of metastasis and those developing relapse to castrate resistant disease. Next, we analysed GCNT1 gene expression in primary prostate cancer tissue samples from the TCGA PRAD clinical cohort (n = 493), and identified tumours with high (n = 124) and low (n = 123) GCNT1 expression^{31,35,36}. Fast gene set enrichment analysis (fgsea)³⁷ of genes differentially expressed by tumours with high GCNT1 expression vs low GCNT1 expression identified significant enrichment of 8 hallmark pathways in tumours with high GCNT1 expression (FDR q value < 0.05) (Fig. 1E). Of particular interest, when GCNT1 is highly expressed in primary prostate cancer tissue there is significant enrichment of the hallmark pathway ‘G2M checkpoint’, which is associated with the trajectory of prostate cancer progression³⁸ (NES 1.94, FDR q value 1.6e–02) (Fig. 1F).

Upregulation of GCNT1 modifies the prostate cancer glycome

GCNT1 catalyses the formation of core 2 branched O-glycan structures on polypeptides¹⁹ (Fig. 2A). In prostate cancer, GCNT1 has previously been linked to higher levels of core 2 O-linked Sialyl Lewis X (sLe^x) on PSA, PAP, and MUC1 proteins²², however how GCNT1 impacts global glycosylation in prostate cancer cells has not been previously studied. To test if upregulation of GCNT1 alters the cell surface glycosylation of prostate cancer cells, we monitored the recognition by lectins and glycan-binding antibodies in PC3 prostate cancer cells overexpressing GCNT1 (PC3 cells were chosen for these experiments as they have low levels of endogenous GCNT1 expression, Supplementary Fig. 1). Consistent with increased levels of core 2 branched O-glycans, lectin immunofluorescence revealed overexpression of GCNT1 correlates with enhanced binding of LEA lectin (which binds GlcNAc oligomers, polyLacNAc, and/or chitin) and RCA₁₂₀ lectin (which predominantly recognises terminal type 2 LacNAc)^{39,40}. However, in line with previous studies¹⁹ in addition to detecting changes to core 2 O-glycans, we also observed additional changes in the repertoire of glycans detected in GCNT1 overexpressing cells. GCNT1 overexpression also correlated with decreased binding of WGA lectin (which recognises terminal GlcNAc), and decreased binding of ConA lectin (which recognises mannose terminated and biantennary structures respectively)⁴⁰ (Fig. 2B and Supplementary Fig. 2). The formation of the core 2 branch is critical for the biosynthesis of sLe^x in O-glycans. To monitor expression of the sLe^x structure, we used a specific glycan binding antibody⁴¹ to assess prostate cancer cells overexpressing GCNT1⁴². Our findings suggest increased GCNT1 expression correlates with upregulation of the cancer-associated sLe^x glycan on the cell surface, which is a functional ligand for Selectins^{43–45} (Fig. 2C and Supplementary Fig. 2). Next, we analysed the O-glycoproteome in secretomes of DU145 prostate cancer cells with upregulated GCNT1 using mass spectrometry. This identified 60 glycopeptides as potential substrates for GCNT1 in prostate cancer cells, including a potential increase in core 2 O-glycan structures on PSAP and MUC16 (Fig. 2D). Together, these experiments suggest that GCNT1 may impact the phenotype of prostate cancer cells by modifying both cell surface glycosylation and the O-glycosylation of proteins secreted by prostate cancer cells.

GCNT1 promotes prostate tumour growth in vivo

We next investigated the impact of GCNT1 expression on the biology of prostate cancer cells using both in vitro and in vivo assays. Our in vitro studies show knockdown of GCNT1 suppresses prostate cancer cell proliferation and colony formation, whereas overexpression of GCNT1 had the opposite effect (Supplementary Fig. 3).

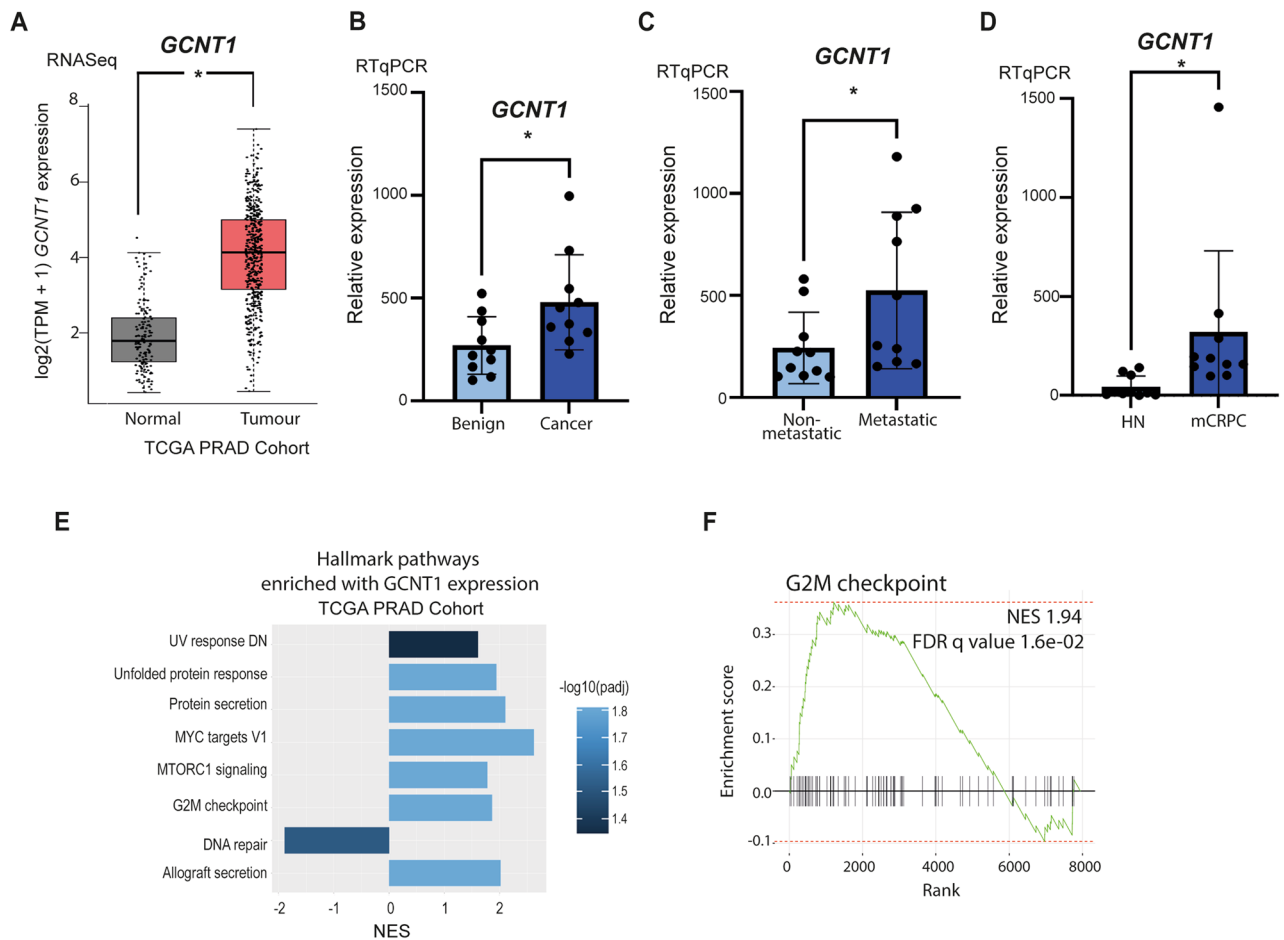


Figure 1. *GCNT1* is upregulated in aggressive prostate cancer tissue. Analysis of *GCNT1* gene levels in clinical prostate tissue. (A) *GCNT1* mRNA levels in the TCGA PRAD cohort³¹ are significantly higher in prostate cancer tissue relative to normal prostate tissue (n = 544, p = 0.021). (B) Real-time PCR analysis of *GCNT1* in RNA samples extracted from FFPE prostate tissue. *GCNT1* is significantly upregulated in cancer relative to benign tissue (n = 20, unpaired t test, p = 0.0248). (C) Analysis of *GCNT1* mRNA levels in samples from the Walker et al.³³ cohort. *GCNT1* is significantly upregulated in patients presenting with a ‘metastatic’ phenotype relative to patients with a ‘non-metastatic’ phenotype (n = 20, unpaired t test, p = 0.0494). (D) Real-time PCR analysis of *GCNT1* mRNA in hormone naïve prostate cancer and mCRPC tissue samples. *GCNT1* is significantly upregulated in mCRPC tissue samples relative to hormone naïve prostate cancer (n = 20, unpaired t test, p = 0.0481). (E,F) In the TCGA PRAD cohort (n = 493)³¹, *GCNT1* mRNA levels correlate with 8 hallmark response pathways, including ‘G2M checkpoint’ (NES 1.94, FDR q value 1.6e–02).

Consistent with this, knockdown of *GCNT1* using sub-cutaneous in vivo mouse models significantly reduces the growth of CWR22RV1 tumours (Fig. 3A), whereas overexpression of *GCNT1* significantly increases the growth of PC3 tumours (Fig. 3B). These findings are consistent with previous studies showing that upregulation of *GCNT1* enhances the growth of orthotopic LNCaP tumours²⁹. The effect of *GCNT1* levels on prostate cancer cell adhesion was also investigated using cell adhesion assays. The results show that knockdown of *GCNT1* in CWR22RV1 cells significantly decreased cell adhesion to both collagen and fibronectin coated plates, whereas overexpression of *GCNT1* in PC3 cells significantly increased cell adhesion (Fig. 3C).

Upregulation of *GCNT1* alters oncogenic signalling pathways in prostate cancer cells

To further investigate the role of *GCNT1* in prostate cancer cells, we used RNA-sequencing to search for gene expression pathways that change with upregulation of *GCNT1*. Bioinformatic analyses identified 2445 differentially expressed genes when *GCNT1* is overexpressed in PC3 cells (adjusted p value < 0.05, Log₂FC 0.58) (Fig. 4A and Supplementary Table 1) with enrichment in ‘cell adhesion’ and ‘cell communication’ biological processes (Fig. 4B) and downregulation in the ‘epithelial mesenchymal transition (EMT)’ hallmark signature (Fig. 4C). Visualisation of significantly differentially expressed EMT hallmark genes in a heatmap confirmed that *GCNT1* overexpression results in downregulation of the EMT pathway (Fig. 4D). Of particular interest in the EMT gene set, the *LGALS1* gene, which encodes Galectin-1 (a carbohydrate binding protein implicated in modulating cell–cell and cell–matrix interactions^{46–53}), was significantly downregulated when *GCNT1* is overexpressed (adjusted p value < 0.01, Log₂FC – 0.7). Validation at the protein level confirmed that upregulation of *GCNT1* in prostate cancer cells promotes loss of Galectin-1 (Fig. 4E) and analysis of The Cancer Genome Atlas Prostate

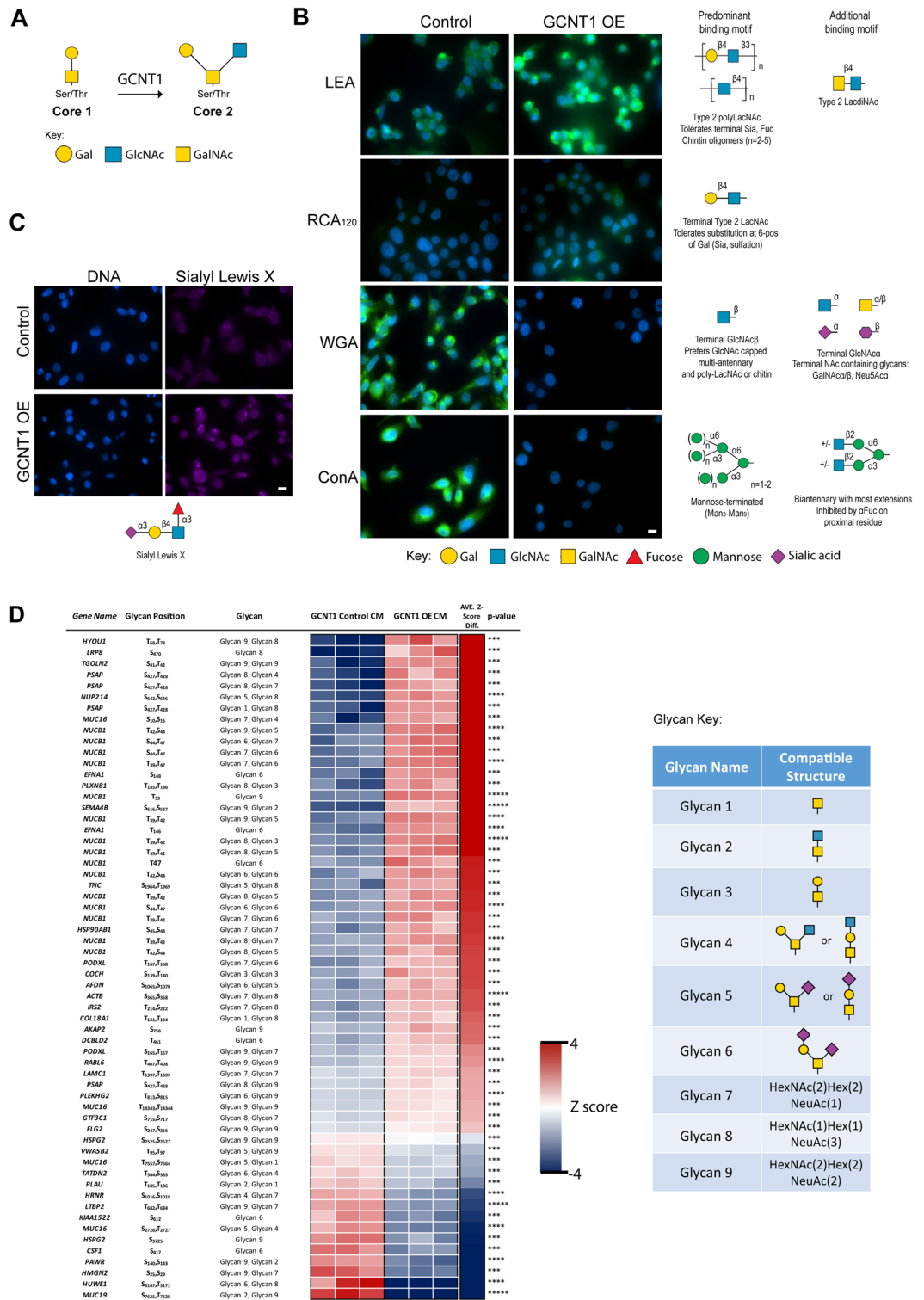


Figure 2. GCNT1 modifies the prostate cancer glycome. (A) GCNT1 is a key enzyme for the formation of core 2 linked O-glycans. In O-glycan synthesis GalNAc is transferred to Ser and Thr residues in the polypeptide, which can then be further extended with various carbohydrates. Four common O-glycan structures are expressed in mammalian tissues, core 1 through core 4. GCNT1 adds GlcNAc to GalNAc to form the core 2 branch. The core 2 branch is a scaffold for the production of lactosamine disaccharide repeats (specifically poly-N-acetylglucosamine) on O-glycans. (B) Lectin immunofluorescence shows PC3 cells with upregulated GCNT1 have increased levels of core 2 branched O-glycans (detected by enhanced binding of LEA and RCA₁₂₀ lectins^{39,40}), reduced levels of mannose terminated and biantennary structures respectively (detected using ConA lectin⁴⁰) and reduced levels of terminal GlcNAc (detected using WGA lectin⁴⁰). (C) Detection of SLe^x in PC3 cells using a glycan specific binding antibody⁴¹ indicates upregulation of GCNT1 promotes increased expression of the SLe^x antigen in prostate cancer cells. (D) Glycoproteomics to map GCNT1 glycosylation sites on specific proteins in DU145 prostate cancer cells with upregulated GCNT1 identified 60 glycopeptides as statistically significant (p < 0.001) specific substrates of GCNT1. Putative glycan structures in this figure were created using the DrawGlycan-SNFG tool⁸⁵.

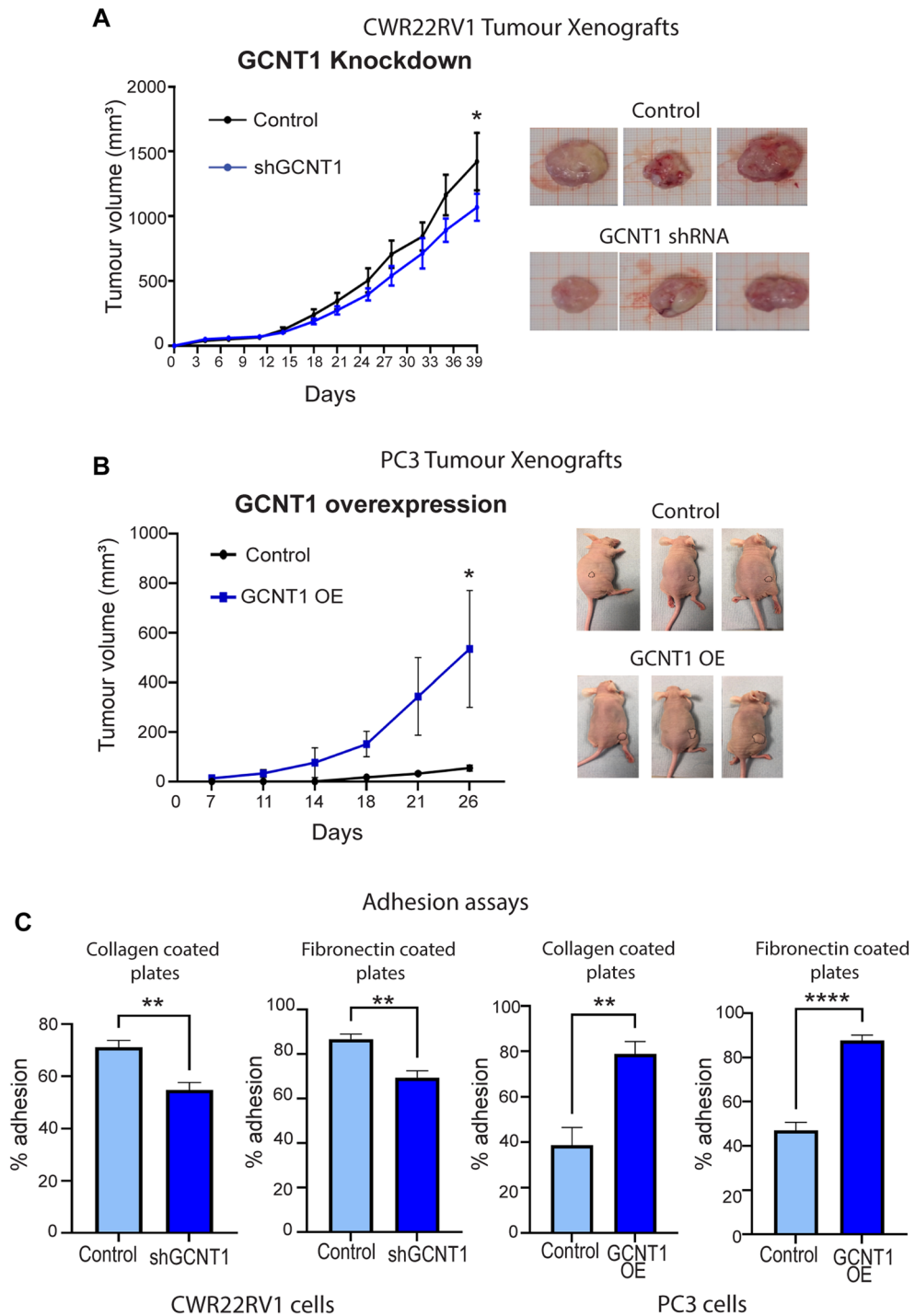


Figure 3. GCNT1 promotes prostate tumour growth in vivo. **(A)** Knockdown of GCNT1 using shRNA significantly reduces the growth of CWR22RV1 tumours xenografts in a subcutaneous xenograft model. **(B)** Upregulation of GCNT1 in PC3 cells significantly increases the growth of subcutaneous xenograft tumours. **(C)** Knockdown of GCNT1 significantly reduces the adhesion of CWR22RV1 cells to collagen and fibronectin coated plates, whereas overexpression of GCNT1 in PC3 cells has the opposite effect.

Adenocarcinoma (TCGA PRAD) cohort^{31,35,36} revealed a correlation between the *GCNT1* and *LGALS1* genes in clinical prostate cancer tissue (Fig. 4F). Our data also reveals potential links between GCNT1 and the alteration of several other pathways important in prostate cancer, including MEK signalling which is a therapeutic target for mCRPC⁵⁴, ERBB2 signalling which is also believed to contribute to prostate cancer progression^{55,56}, and altered KRAS signatures which are a major player in promoting and maintaining tumorigenesis⁵⁷. Together,

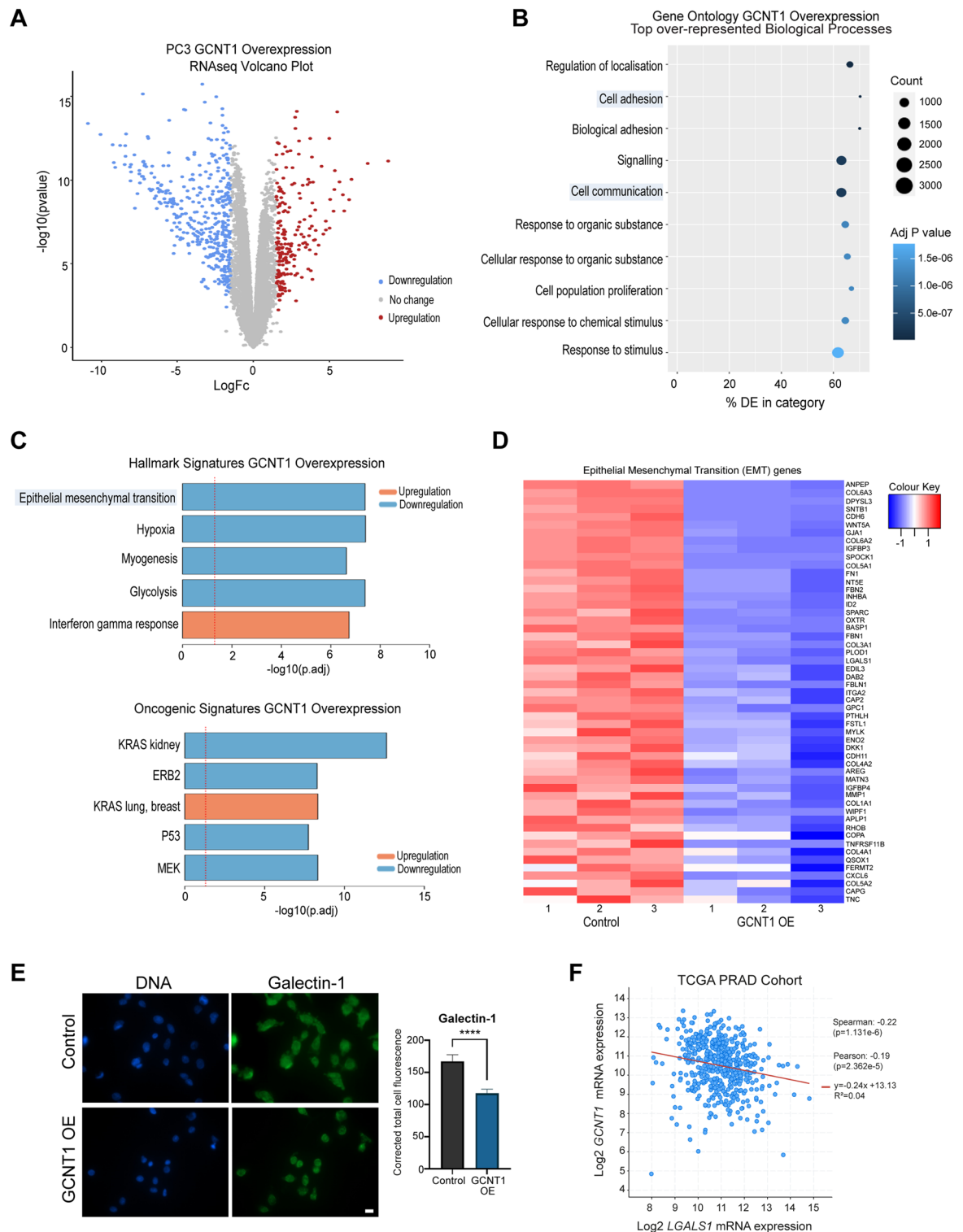


Figure 4. Upregulation of GCNT1 alters oncogenic gene expression pathways in prostate cancer cells. **(A)** RNA-seq analysis of PC3 cells with upregulated GCNT1 identified 2445 genes that are differentially expressed in response to GCNT1. **(B)** Gene Ontology analysis of genes regulated by GCNT1 shows an enrichment of genes with roles in ‘signalling’ and ‘cell communication’. **(C)** Ensemble Gene Set Enrichment Analysis of genes regulated by GCNT1 reveals downregulation in the ‘epithelial mesenchymal transition’ hallmark signature, and downregulation in ‘P53’ and ‘MEK’ oncogenic signatures. **(D)** Heatmap to illustrate 53 genes with roles in EMT that are significantly downregulated in GCNT1 overexpressing cells. **(E)** Validation at the protein level shows Galectin-1, which is encoded by the *LGALS1* gene, is downregulated when GCNT1 is overexpressed. **(F)** Analysis of the TCGA PRAD cohort ($n = 493$)^{31,35,36} shows a significant correlation between *GCNT1* and *LGALS1* mRNA levels in primary prostate cancer clinical tissue.

these findings reveal GCNT1 overexpression can alter oncogenic gene expression pathways in prostate cancer that are important in disease progression.

Discussion

In prostate cancer, aberrant glycosylation is closely linked to a malignant phenotype, but the mechanisms behind how altered glycosylation contributes to disease pathology are poorly understood. Here, we show that GCNT1, which encodes a key glycosyltransferase enzyme controlling O-glycan branching, is upregulated in aggressive prostate cancer tissue. In line with previous reports, our study shows that increased expression of GCNT1 promotes the growth of prostate tumours, and building on these findings we also reveal GCNT1 can increase cell adhesion. Furthermore, we find that upregulation of GCNT1 can alter the glycome of prostate cancer cells to increase levels of core 2 O-glycans and upregulate expression of sLe^x, and can impact oncogenic gene expression patterns.

Previous studies show that upregulation of GCNT1 correlates with higher levels of core 2 O-sLe^x on PSA, PAP, and MUC1 proteins²², but how GCNT1 modifies global glycosylation in prostate cancer cells has not been studied. This study shows overexpression of GCNT1 in prostate cancer cells modifies cell surface glycosylation and leads to increased levels of core 2 O-glycan structures, including upregulation of the poly-N-acetylglucosamine chain (which serves as a scaffold for terminal modifications such as the tumour-associated glycan sLe^x). We also find GCNT1 can upregulate the sLe^x structure, a Selectin ligand with a key role in metastasis^{43–45,58,59}. Using site specific O-glycoproteomics, we identified 60 potential substrates for GCNT1 in the prostate cancer secretome. It is interesting to speculate that the role of GCNT1 in prostate cancer tumour growth and control of gene expression patterns could be due to some of its protein substrates' roles or combined functions. Taken together, these studies provide insights into how upregulation of GCNT1 can modify the prostate cancer glycome, however additional studies will be needed to further identify and characterise the polypeptide substrates of GCNT1.

Several reports have correlated *GCNT1* expression in prostate cancer with aggressive disease, including recurrence after surgery²³ and the spread of cancer cells outside of the prostate gland²¹. It has also been demonstrated that GCNT1 enhances the growth of orthotopic prostate tumours²⁹. Consistent with previously published studies, the findings reported here also correlate *GCNT1* gene expression levels with more aggressive prostate cancer and show upregulation of GCNT1 increases the growth of prostate tumours. Our findings indicating upregulation of GCNT1 correlates with downregulation of the EMT hallmark pathway may provide insight into how GCNT1 can alter the adhesive properties of prostate cancer cells and potentially lead to a reduction in the seeding potential of prostate cancer cells during the metastatic process. Indeed, using an intra-cardiac injection model, we find that upregulation of GCNT1 in PC3 cells significantly reduces the number of metastatic tumours formed ($p = 0.0295$) (Supplementary Fig. 4). The complex biological process of EMT has been heralded as a key hallmark of cancer that confers metastatic properties upon cancer cells⁶⁰. There is considerable evidence to suggest that the transition of epithelial cells to a mesenchymal state via EMT promotes an invasive phenotype that contributes to prostate cancer metastasis^{61–63}. However, studies using pre-clinical models suggest the transition between epithelial and mesenchymal phenotypes represents a dynamic and complex set of events that play important roles in metastatic dissemination and colonisation⁶⁴ and that the transient induction of EMT and the subsequent reversion to an epithelial phenotype (known as mesenchymal to epithelial transition or MET) may be necessary for the formation of overt metastasis^{64–68}. Our findings indicating upregulation of GCNT1 correlates with downregulation of the EMT hallmark pathway may provide insight into how GCNT1 can alter the adhesive properties of prostate cancer cells and lead to a reduction in the seeding potential of prostate cancer cells during the metastatic process. Therefore, it is important to consider our data in the context of the complex and dynamic role EMT likely plays in aggressive prostate cancer.

Taken together, our findings reveal potential mechanisms to explain how upregulation of GCNT1 might contribute to the prostate cancer pathology, however further studies will be needed to further unpick the mechanisms involved. In particular, while the data presented here examines the role of GCNT1 in the formation of metastatic tumours by disseminated prostate tumour cells, to fully study the in vivo role of GCNT1 in prostate cancer metastasis it will be necessary to utilise a repertoire of mouse models encompassing different stages in the metastatic cascade. In addition, while the in vivo studies reported here were carried out using immunocompromised mouse models, moving forward, and as GCNT1 can upregulate sLe^x (a Selectin ligand linked to tumour metastasis which is recognised by endothelial cells and lymphocytes^{43–45,58,59}), it will be also important to study the role of GCNT1 in prostate cancer metastasis in the context of a functional immune system.

Methods

RT-qPCR

RNA extraction, cDNA synthesis and real-time PCR were as described previously³⁴.

Clinical RNA cohorts

The clinical RNA samples shown in Fig. 1B were collected from Castle Hill Hospital (Cottingham, Hull) and used with ethical approval by the local Research Ethics Committee (Ethics Number: 07/H1304/121). Patients gave informed consent and all patient samples were anonymised. The clinical RNA samples shown in Fig. 1C and D were obtained via the Prostate Cancer Biorepository Network (PCBN). Our study was peer reviewed and approved by PCBN. Samples were prepared and stored using standard protocols and all patients gave informed consent. All methods were performed in accordance with the relevant guidelines and regulations.

Cell culture and creation of stable cell lines

Cell culture and the cell lines used were as described previously³⁴. The stable cell lines used in the study were created using lentiviral transduction. For knockdown of GCNT1 the following shRNA lentiviral particles were purchased from Santa Cruz (GCNT1 shRNA sc-92945-V and Control shRNA sc-108080). Transductions were carried out according to the manufacturer's instructions using an MOI of 5. For overexpression of GCNT1 the following Lentitect Purified lentiviral particles were purchased from Tebu-Bio (GCNT1 LPP-Z6088-Lv242-050-S and Negative Control 217LPP-NEG-Lv242-025-C). Transductions were carried out according to the manufacturer's instructions using an MOI of 5. For the metastasis study, PC3 cells were transduced with Firefly Luciferase Lentivirus (BPS Bioscience, 79692-H) at a MOI of 1 in media containing 5 µg/ml polybrene. Stable cell lines were selected with 200 µg/ml hygromycin, and then further selected for GCNT1 overexpression using the methods described above.

Colony formation assays

Cells were plated at a density of 100 cells per 100 mm dish and maintained until colonies of more than 50 cells/colony had formed. Cells were fixed in 10% formalin for 10 min and stained with 0.5% crystal violet for 10 min at room temperature. Colony numbers were counted manually. Three biological repeats were conducted per cell line.

Adhesion assays

Cell adhesion assays were performed using the Vybrant Cell Adhesion Assay Kit (Fisher, V-13181) using both collagen coated plates (Sigma, CC304) and fibronectin coated plates (Sigma, S3815). 50,000 cells per well were allowed to adhere for two hours and assays were performed as per the manufacturer's instructions. Percentage adhesion was calculated by comparing the absorbance with that of an identical unwashed plate. Three biological repeats were conducted per cell line.

Mouse models

CWR22RV1 tumour xenografts

Male NMRI mice (Charles Rivers) were inoculated at 8 weeks of age with 1×10^7 CWR22RV1 cells with GCNT1 knockdown by unilateral subcutaneous injection into the flank. Cells were injected in a volume of 100 µL and Matrigel in a 1:1 mixture. Animals were weighed and tumour volumes monitored by caliper measurement three times a week until the first animal met a humane endpoint.

PC3 tumour xenografts

Male CD-1 Nude mice (Charles Rivers) were inoculated at 8 weeks of age with 1×10^7 PC3 cells with GCNT1 overexpression. Cells were injected in a volume of 50 µL of cell culture media and Matrigel in a 1:1 mixture. Animals were weighed and tumour volumes monitored by caliper measurement three times a week until the first animal met a humane endpoint.

PC3 tumour engraftment study

Six-week old male BALB/cAnNCrI immunocompromised (athymic nude) mice were purchased from Charles River (Kent, UK) and housed in a controlled environment in Optimice cages (Animal Care Systems, Colorado, USA) with a 12 h light/dark cycle at 22 °C with ad libitum water and 2018 Teklad Global 18% protein rodent diet containing 1.01% Calcium (Harlan Laboratories, UK). Mice were randomised into two groups to receive single-cell suspensions of 1×10^5 PC3 (Control or GCNT1 OE) cells/100 µL PBS via injection into the left cardiac ventricle of mice (intracardiac injection). Tumour progression was monitored weekly based on bioluminescence using the in vivo imaging systems (IVIS, PerkinElmer, Cambridge, UK) for 6 weeks and the images were blinded for data analysis.

Mice were euthanized through exsanguination under general anaesthesia, followed by cervical dislocation. All procedures complied with the UK Animals (Scientific Procedures) Act 1986 and were reviewed and approved by the local Research Ethics Committees of the University of Sheffield under Home Office project licence (PP3267943). The animal research performed in this study is reported in accordance with ARRIVE guidelines.

Immunofluorescence

Cells were cultured in Lab-Tek™II Chamber Slides (ThermoScientific, 154453) for 48 h. Cells were washed with PBS before permeabilization and fixation with ice-cold absolute methanol for 10 min at – 20 °C. Next, slides were washed with PBS and blocked with Carbo-Free™ Blocking solution (Vector Laboratories, SP-5040) for 1 h at room temperature. Slides were incubated overnight at 4 °C with FITC-conjugated LEA (Lectin from *Lycopersicon esculentum*, Sigma, L0401), RCA₁₂₀ Fluorescein (*Ricinus Communis* Agglutinin 120, Vector Laboratories, FLK-2100), WGA fluorescein (Wheat Grain Agglutinin, Vector Laboratories, FLK-2100), ConA fluorescein (Concanavalin A, Vector Laboratories, FLK-2100), CoraLite[®]488-conjugated Galectin-1 Monoclonal antibody (Proteintech, CL488-60223) at 1:200, and Sialyl Lewis X (Alexa Fluor 647 Rat Anti-Human Cutaneous lymphocyte antigen, BD Pharmagen, 563533) at 1:100. Finally, slides were washed with PBS and stained with Hoechst (ThermoScientific, 62249) for 15 min at room temperature. Images were acquired and processed with the ZEISS Axio Imager 3 microscope. The fluorescent images were analysed using NIH ImageJ software⁶⁹ (Version t1.51) by measuring the area, integrated density and mean grey value of one cell at a time ($n = 50$). The fluorescence intensity was calculated in Excel using the formula for corrected total cell fluorescence (CTCF) = integrated density – (area of selected cell × mean fluorescence of background readings).

Mass spectrometry-based glycoproteomics

Cell pellets were lysed using 500 μL of 1.5% SDS lysis solution containing 1X protease inhibitors (Sigma Aldrich) followed by sonication using a Branson probe sonicator (Fisher Scientific). Conditioned cell media was concentrated to a final volume of 500 μL using 10 kDa 15 mL spin filters (Amicon). Protein concentration was quantified with using a standard BCA protein assay (Thermo Fisher Scientific) following the manufacturer's protocol. 300 μg of proteins from the cell lysates and media were aliquoted out into a new 1.5 mL tube (Eppendorf). The sample volume was brought up to 100 μL using 50 mM ammonium bicarbonate (Sigma Aldrich) followed by vortexing for 15 s. Samples were incubated for 1 h at 65 °C with 10 mM Tris(2-carboxyethyl) phosphine hydrochloride (TCEP, Sigma Aldrich) to reduce the cysteine disulfide bonds. Then, thiolate groups on cysteines were alkylated with 15 mM of iodoacetamide (Across Organics) in the dark for 45 min at room temperature. Proteins were precipitated with 1 mL of cold acetone and stored at -20 °C overnight. Precipitated proteins were pelleted down via centrifugation at 14,000g for 10 min at 4 °C and acetone was removed. Pelleted proteins were dried at room temperature and 50 μL of 50 mM ammonium bicarbonate. Proteins were digested with sequencing grade trypsin (Thermo Fisher Scientific) at a 1:30 enzyme to protein ration and incubated at 37 °C overnight without shaking.

Glycopeptides were enriched using the strong anion exchange and electrostatic repulsion hydrophilic interaction chromatography (SAX-ERLIC)^{70,71}. SOLA SAX SPE (Thermo Fisher Scientific) columns were washed with 3 mL of acetonitrile with a flow rate 1 mL/min followed by 3 mL of 100 mM triethylammonium acetate in water, and 3 mL of 1% TFA in water. The column was equilibrated with 3 mL of 95% acetonitrile with 1% TFA in water. The organic solvent in the tryptic peptide solution was adjusted by adding 3 mL of 95% acetonitrile with 1% TFA in water. Samples were loaded on the SOLA-SAX SPE column and passed through with a flow rate of 0.5 mL/min. The SOLA-SAX SPE column was then washed with 6 mL of 95% acetonitrile 1% formic acid in water. Glycopeptides were eluted in a 2.0 mL Eppendorf vials by adding 850 μL aliquots of 50% acetonitrile 0.1% TFA in water twice, followed by another two 850 μL aliquots of 5% acetonitrile 0.1% TFA in water. The two glycopeptide fractions were concentrated, combined, and dried down using a speed vacuum (LabConco) prior to LC-MSMS analysis.

3 μL of the glycopeptides were loaded into a 20 μL sample loop using a Dionex Ultimate Rapid Separation liquid chromatography system (Thermo Fisher Scientific) and subsequently loading glycopeptides onto a C18 PepMap trap column (Thermo Fisher Scientific) at a flow rate of 5 $\mu\text{L}/\text{min}$ for 10 min. A reversed phase liquid chromatography gradient consisting of mobile phase A (0.1% formic acid in water) and mobile phase B (0.1% formic acid in acetonitrile) was used to separate glycopeptides on a 25 cm long analytical column packed with BEH C18, 130 Å, and 1.7 μm particle size (Waters) encapsulated in a column heater (MSWILL) with a temperature set at 60 °C. Eluted peptides were subject to MS/MS analysis on an Orbitrap Eclipse Tribrid mass spectrometer (Thermo Fisher Scientific). The gradient program was set to hold mobile phase B at 2% for the first 6 min, steadily ramped to 35% B over the next 80 min followed by an increase to 85% B over 5 min with a 5 min hold at a flow rate of 0.3 $\mu\text{L}/\text{min}$ and a 15 min equilibration time. Each sample was analysed in triplicate. The MS program consisted of setting the cycle time at Top-speed at 1 s with an MS mass scan range of 450–2000 m/z and mass resolution of 50,000. The most abundant precursor ions were fragmented with Higher Energy Collisional Dissociation (HCD) with a collisional energy set to 38%. Dynamic exclusion was enabled for 60 s with a mass tolerance of 10 ppm and the normalized AGC target set to 100%. MS2 fragments were detected in the Orbitrap detector with mass scan range set to auto and a mass resolution of 30,000. The injection time set to custom with the maximum injection time of 54 ms. The raw data files were searched using Byonic v3.9.6 software against a focus protein database identified from the shotgun analysis from the same sample set (2020; 4419 entries) and a glycan database of 9 common O-glycans provided by Byonic software. The glycopeptide search included parameters of trypsin digestion with maximum two missed cleavages, MS1 precursor mass tolerance of 10 ppm, and fragment mass tolerance 0.5 Da with fixed cysteine carbamidomethylation, variable methionine oxidation and asparagine deamination. To further validate glycopeptide identifications, the following process was applied. Only glycopeptides with > 6 amino acids were retained. A theoretical MS/MS spectrum was generated for each of the remaining glycopeptides. Spectral matching was then performed so that peptides with 6–7, 8–16, and 17–20 amino acids needed to have 80, 70, and 50% sequence coverage in the measured MS/MS fragmentation spectra respectively to be considered for further analysis.

After aligning the chromatographic runs, quantitative data was extracted from the MS1 spectra of identified peptides using an R script based on the MSnbase package⁷². This approach allowed for the retrieval of precise measurements of peptide abundance. The area under the curve (AUC) of the extracted ion current (XIC) was calculated for all detected peptides, providing quantitative measurements. To analyse changes in protein abundance, the Generic Integration Algorithm was employed. This algorithm utilized the AUC values of the peptides to assess variations in protein expression between two groups: the GCNT1 overexpression group and the control DU145 group. Following a previously described approach⁷³, all quantitative data was expressed as Z-scores at the protein or peptide level. Log2 ratios were calculated by comparing the AUC of peptides in the GCNT1 overexpression group with the control DU145 group, enabling the determination of fold changes in protein expression. The WSPP model was applied at the spectrum level to calculate the corresponding statistical weight. The weight was rescaled and standardized to a normal distribution $N(0,1)$, facilitating comparisons and enabling the identification of statistically significant changes. Cumulative distribution plots were generated at each level (spectrum, peptide, and protein) to assess the validity of the null hypothesis and examine the statistical significance of observed differences. Initial analysis focused on unmodified peptides to determine variations at the scan, peptide, and protein levels, as well as changes in protein expression. Subsequently, glycan-containing peptides were included in the analysis to identify any peptides that deviated significantly from the behaviour of other non-modified peptides belonging to the same protein⁷³. The final statistical comparison was performed

using Student's t-test, with a significance level below 0.01 considered statistically significant. Only peptides meeting this criterion were selected for further analysis.

RNA sequencing analysis

RNA sequencing data can be accessed on GEO repository (GSE224036). RNA was extracted from 3 PC3 cell lines with stable *GCNT1* overexpression and 3 PC3 cell lines transduced with negative-control lentiviral particles. Samples were prepared with the Illumina TruSeq Stranded mRNA Library Prep Kit and sequenced using an Illumina NextSeq 500, giving 13 million 75 bp single reads per sample. All data analyses were performed in Galaxy version 22.01. Quality control was performed with FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and reads were trimmed with Cutadapt⁷⁴. Reads were mapped to hg38 using HISAT2⁷⁵ and quantified with featureCounts⁷⁶. Differential gene expression analysis was performed using limma-voom⁷⁷ and a volcano plot was generated with ggplot2⁷⁸. Gene ontology (GO) analysis was performed with goseq⁷⁹ applying a significance threshold of adjusted p value < 0.05 for differentially expressed genes. Gene Set Enrichment Analysis (GSEA) was performed with the package EGSEA⁸⁰. Normalised count matrix values were used to create a heatmap with gplots⁸¹.

Bioinformatic analysis of TCGA cohort

To compare *GCNT1* gene expression levels in normal prostate and prostate tumour tissue samples, data from The Cancer Genome Atlas Prostate Adenocarcinoma (TCGA PRAD) cohort was analysed using GEPIA^{31,32}. To profile gene expression patterns in the TCGA PRAD cohort, RNA-Seq data from prostate tumours (n = 493) was analysed in cBioPortal^{31,35,36}. Tumours were ranked by *GCNT1* expression levels to delineate two groups from the top and bottom quartiles. Group comparison of 124 tumours with high *GCNT1* expression vs 123 tumours with low *GCNT1* expression revealed 8110 differentially expressed genes (q-value < 0.05). Genes were pre-ranked using the t statistic and GSEA was performed using the fgsea³⁷ package in Galaxy version 22.05. Annotated gene sets were obtained from the Molecular Signatures Database (MSigDB)^{82,83} and a bar chart was generated from fgsea data in R version 4.2.1⁸⁴ with ggplot2.

Data availability

The data that support the findings of this study are openly available and can be accessed on GEO repository (GSE224036).

Received: 22 February 2023; Accepted: 18 September 2023

Published online: 09 October 2023

References

- Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J. Clin.* **72**(1), 7–33 (2022).
- Livermore, K. E., Munkley, J. & Elliott, D. J. Androgen receptor and prostate cancer. *AIMS Mol. Sci.* **3**, 280–299 (2016).
- Morote, J., Aguilar, A., Planas, J. & Trilla, E. Definition of castrate resistant prostate cancer: New insights. *Biomedicines*. **10**(3), 689 (2022).
- Sathianathan, N. J., Konety, B. R., Crook, J., Saad, F. & Lawrentschuk, N. Landmarks in prostate cancer. *Nat. Rev. Urol.* **15**(10), 627–642 (2018).
- Sandhu, S. *et al.* Prostate cancer. *Lancet* **398**(10305), 1075–1090 (2021).
- Beer, T. M. & Tombal, B. Enzalutamide in metastatic prostate cancer before chemotherapy. *N. Engl. J. Med.* **371**(18), 1755–1756 (2014).
- de Bono, J. S. *et al.* Abiraterone and increased survival in metastatic prostate cancer. *N. Engl. J. Med.* **364**(21), 1995–2005 (2011).
- Fizazi, K. *et al.* Darolutamide in nonmetastatic, castration-resistant prostate cancer. *N. Engl. J. Med.* **380**(13), 1235–1246 (2019).
- Varki, A. & Kornfeld, S. Historical background and overview. In *Essentials of Glycobiology* (eds Varki, A. *et al.*). 4th ed. 1–20 (2022).
- Munkley, J. & Elliott, D. J. Hallmarks of glycosylation in cancer. *Oncotarget* **7**(23), 35478–35489 (2016).
- Vajaria, B. N. & Patel, P. S. Glycosylation: A hallmark of cancer?. *Glycoconj J.* **34**(2), 147–156 (2017).
- Scott, E. & Munkley, J. Glycans as biomarkers in prostate cancer. *Int. J. Mol. Sci.* **20**(6), 1389 (2019).
- Munkley, J., Mills, I. G. & Elliott, D. J. The role of glycans in the development and progression of prostate cancer. *Nat. Rev. Urol.* **13**(6), 324–333 (2016).
- Pinho, S. S. & Reis, C. A. Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer*. **15**(9), 540–555 (2015).
- Munkley, J. Glycosylation is a global target for androgen control in prostate cancer cells. *Endocr. Relat. Cancer*. **24**(3), R49–R64 (2017).
- Munkley, J. *et al.* Glycosylation is an androgen-regulated process essential for prostate cancer cell viability. *EBioMedicine* **8**, 103–116 (2016).
- Bierhuizen, M. F., Mattei, M. G. & Fukuda, M. Expression of the developmental I antigen by a cloned human cDNA encoding a member of a beta-1,6-N-acetylglucosaminyltransferase gene family. *Genes Dev.* **7**(3), 468–478 (1993).
- Yeh, J. C., Ong, E. & Fukuda, M. Molecular cloning and expression of a novel beta-1, 6-N-acetylglucosaminyltransferase that forms core 2, core 4, and I branches. *J. Biol. Chem.* **274**(5), 3215–3221 (1999).
- Stone, E. L. *et al.* Glycosyltransferase function in core 2-type protein O glycosylation. *Mol. Cell Biol.* **29**(13), 3770–3782 (2009).
- Kim, J. *et al.* Tumor initiating but differentiated luminal-like breast cancer cells are highly invasive in the absence of basal-like activity. *Proc. Natl. Acad. Sci. USA* **109**(16), 6124–6129 (2012).
- Kojima, Y. *et al.* Detection of Core2 beta-1,6-N-acetylglucosaminyltransferase in post-digital rectal examination urine is a reliable indicator for extracapsular extension of prostate cancer. *PLoS ONE* **10**(9), e0138520 (2015).
- Chen, Z., Gulzar, Z. G., St Hill, C. A., Walcheck, B. & Brooks, J. D. Increased expression of *GCNT1* is associated with altered O-glycosylation of PSA, PAP, and MUC1 in human prostate cancers. *Prostate* **74**(10), 1059–1067 (2014).
- Sato, T. *et al.* Core 2 beta-1, 6-N-acetylglucosaminyltransferase-1 expression in prostate biopsy specimen is an indicator of prostate cancer aggressiveness. *Biochem. Biophys. Res. Commun.* **470**(1), 150–156 (2016).
- Chang, L. Y. *et al.* Molecular basis and role of Siglec-7 ligand expression on chronic lymphocytic leukemia B cells. *Front. Immunol.* **13**, 840388 (2022).
- Ye, Y. *et al.* Exploring prognosis-associated biomarkers of estrogen-independent uterine corpus endometrial carcinoma by bioinformatics analysis. *Int. J. Gen. Med.* **14**, 9067–9081 (2021).

26. Huang, M. C. *et al.* C2GnT-M is downregulated in colorectal cancer and its re-expression causes growth inhibition of colon cancer cells. *Oncogene* **25**(23), 3267–3276 (2006).
27. Hatakeyama, S. *et al.* Core 2 N-acetylglucosaminyltransferase-1 expression induces aggressive potential of testicular germ cell tumor. *Int. J. Cancer* **127**(5), 1052–1059 (2010).
28. Tsuboi, S. *et al.* A novel strategy for evasion of NK cell immunity by tumours expressing core2 O-glycans. *EMBO J.* **30**(15), 3173–3185 (2011).
29. Haggisawa, S. *et al.* Expression of core 2 beta1,6-N-acetylglucosaminyltransferase facilitates prostate cancer progression. *Glycobiology* **15**(10), 1016–1024 (2005).
30. Okamoto, T. *et al.* Core2 O-glycan-expressing prostate cancer cells are resistant to NK cell immunity. *Mol. Med. Rep.* **7**(2), 359–364 (2013).
31. Hoadley, K. A. *et al.* Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 Types of cancer. *Cell* **173**(2), 291–304.e6 (2018).
32. Tang, Z. *et al.* GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **45**(W1), W98–W102 (2017).
33. Walker, S. M. *et al.* Molecular subgroup of primary prostate cancer presenting with metastatic biology. *Eur. Urol.* **72**(4), 509–518 (2017).
34. Munkley, J. *et al.* Androgen-regulated transcription of ESRP2 drives alternative splicing patterns in prostate cancer. *Elife* **8**, e47678 (2019).
35. Cerami, E. *et al.* The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**(5), 401–404 (2012).
36. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**(269), p11 (2013).
37. Korotkevich, G. *et al.* Fast gene set enrichment analysis. *bioRxiv.* (2021).
38. Bolis, M. *et al.* Dynamic prostate cancer transcriptome analysis delineates the trajectory to disease progression. *Nat. Commun.* **12**(1), 7033 (2021).
39. Nachbar, M. S. & Oppenheim, J. D. Tomato (*Lycopersicon esculentum*) lectin. *Methods Enzymol.* **83**, 363–368 (1982).
40. Bojar, D. *et al.* A useful guide to lectin binding: Machine-learning directed annotation of 57 unique lectin specificities. *ACS Chem. Biol.* **17**(11), 2993–3012 (2022).
41. Julien, S. *et al.* Selectin ligand sialyl-Lewis x antigen drives metastasis of hormone-dependent breast cancers. *Cancer Res.* **71**(24), 7683–7693 (2011).
42. Walcheck, B. *et al.* The monoclonal antibody CHO-131 binds to a core 2 O-glycan terminated with sialyl-Lewis x, which is a functional glycan ligand for P-selectin. *Blood* **99**(11), 4063–4069 (2002).
43. Foxall, C. *et al.* The three members of the selectin receptor family recognize a common carbohydrate epitope, the sialyl Lewis(x) oligosaccharide. *J. Cell Biol.* **117**(4), 895–902 (1992).
44. Laubli, H. & Borsig, L. Selectins promote tumor metastasis. *Semin. Cancer Biol.* **20**(3), 169–177 (2010).
45. Natoni, A., Macauley, M. S. & O'Dwyer, M. E. Targeting selectins and their ligands in cancer. *Front. Oncol.* **6**, 93 (2016).
46. Bacigalupo, M. L. *et al.* Galectin-1 triggers epithelial–mesenchymal transition in human hepatocellular carcinoma cells. *J. Cell Physiol.* **230**(6), 1298–1309 (2015).
47. Zhu, J. *et al.* Galectin-1 induces metastasis and epithelial–mesenchymal transition (EMT) in human ovarian cancer cells via activation of the MAPK JNK/p38 signalling pathway. *Am. J. Transl. Res.* **11**(6), 3862–3878 (2019).
48. Zhang, P. F. *et al.* Galectin-1 induces hepatocellular carcinoma EMT and sorafenib resistance by activating FAK/PI3K/AKT signaling. *Cell Death Dis.* **7**(4), e2201 (2016).
49. Hsu, Y. L. *et al.* Galectin-1 promotes lung cancer tumor metastasis by potentiating integrin alpha6beta4 and Notch1/Jagged2 signaling pathway. *Carcinogenesis* **34**(6), 1370–1381 (2013).
50. Camby, I., Le Mercier, M., Lefranc, F. & Kiss, R. Galectin-1: A small protein with major functions. *Glycobiology* **16**(11), 137R–R157 (2006).
51. van den Brule, F. A., Waltregny, D. & Castronovo, V. Increased expression of galectin-1 in carcinoma-associated stroma predicts poor outcome in prostate carcinoma patients. *J. Pathol.* **193**(1), 80–87 (2001).
52. Laderach, D. J. *et al.* A unique galectin signature in human prostate cancer progression suggests galectin-1 as a key target for treatment of advanced disease. *Cancer Res.* **73**(1), 86–96 (2013).
53. Shih, T. C. *et al.* Targeting galectin-1 impairs castration-resistant prostate cancer progression and invasion. *Clin. Cancer Res.* **24**(17), 4319–4331 (2018).
54. Nickols, N. G. *et al.* MEK-ERK signaling is a therapeutic target in metastatic castration resistant prostate cancer. *Prostate Cancer Prostatic Dis.* **22**(4), 531–538 (2019).
55. Tome-Garcia, J., Li, D., Ghazaryan, S., Shu, L. & Wu, L. ERBB2 increases metastatic potentials specifically in androgen-insensitive prostate cancer cells. *PLoS ONE* **9**(6), e99525 (2014).
56. Miller, D. R., Ingersoll, M. A. & Lin, M. F. ErbB-2 signaling in advanced prostate cancer progression and potential therapy. *Endocr. Relat. Cancer* **26**(4), R195–R209 (2019).
57. Mustachio, L. M., Chelariu-Raicu, A., Szekevolgyi, L. & Roszik, J. Targeting KRAS in cancer: Promising therapeutic strategies. *Cancers* **13**(6), 1204 (2021).
58. Maemura, K. & Fukuda, M. Poly-N-acetylglucosaminyl O-glycans attached to leukosialin. The presence of sialyl Le(x) structures in O-glycans. *J. Biol. Chem.* **267**(34), 24379–86 (1992).
59. Munkley, J. Aberrant sialylation in cancer: Therapeutic opportunities. *Cancers* **14**(17), 4248 (2022).
60. Mittal, V. Epithelial mesenchymal transition in tumor metastasis. *Annu. Rev. Pathol.* **13**, 395–412 (2018).
61. Khan, M. I., Hamid, A., Adhami, V. M., Lall, R. K. & Mukhtar, H. Role of epithelial mesenchymal transition in prostate tumorigenesis. *Curr. Pharm. Des.* **21**(10), 1240–1248 (2015).
62. Nauseef, J. T. & Henry, M. D. Epithelial-to-mesenchymal transition in prostate cancer: Paradigm or puzzle?. *Nat. Rev. Urol.* **8**(8), 428–439 (2011).
63. Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. Epithelial–mesenchymal transitions in development and disease. *Cell* **139**(5), 871–890 (2009).
64. Stylianou, N. *et al.* A molecular portrait of epithelial–mesenchymal plasticity in prostate cancer associated with clinical outcome. *Oncogene* **38**(7), 913–934 (2019).
65. Ocana, O. H. *et al.* Metastatic colonization requires the repression of the epithelial–mesenchymal transition inducer Prrx1. *Cancer Cell* **22**(6), 709–724 (2012).
66. Roca, H. *et al.* Transcription factors OVOL1 and OVOL2 induce the mesenchymal to epithelial transition in human cancer. *PLoS ONE* **8**(10), e76773 (2013).
67. Tsai, J. H., Donaher, J. L., Murphy, D. A., Chau, S. & Yang, J. Spatiotemporal regulation of epithelial–mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* **22**(6), 725–736 (2012).
68. Celia-Terrassa, T. *et al.* Epithelial–mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. *J. Clin. Investig.* **122**(5), 1849–68 (2012).

69. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**(7), 671–675 (2012).
70. Totten, S. M., Feasley, C. L., Bermudez, A. & Pitteri, S. J. Parallel comparison of N-linked glycopeptide enrichment techniques reveals extensive glycoproteomic analysis of plasma enabled by SAX-ERLIC. *J. Proteome Res.* **16**(3), 1249–1260 (2017).
71. Bermudez, A. & Pitteri, S. J. Enrichment of intact glycopeptides using strong anion exchange and electrostatic repulsion hydrophilic interaction chromatography. *Methods Mol. Biol.* **2271**, 107–120 (2021).
72. Capuano, F. *et al.* LC-MS/MS methods for absolute quantification and identification of proteins associated with chimeric plant oil bodies. *Anal. Chem.* **83**(24), 9267–9272 (2011).
73. Navarro, P. *et al.* General statistical framework for quantitative proteomics by stable isotope labeling. *J. Proteome Res.* **13**(3), 1234–1247 (2014).
74. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**(1), 3 (2011).
75. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **12**(4), 357–360 (2015).
76. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**(7), 923–930 (2014).
77. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**(2), R29 (2014).
78. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
79. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* **11**(2), R14 (2010).
80. Alhamdoosh, M. *et al.* Combining multiple tools outperforms individual methods in gene set enrichment analyses. *Bioinformatics* **33**(3), 414–424 (2017).
81. Warnes, G. R. *et al.* editors. *gplots: Various R Programming Tools for Plotting Data* (2015).
82. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA.* **102**(43), 15545–15550 (2005).
83. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**(12), 1739–40 (2011).
84. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2020).
85. Cheng, K., Zhou, Y. & Neelamegham, S. DrawGlycan-SNFG: A robust tool to render glycans and glycopeptides with fragmentation information. *Glycobiology* **27**(3), 200–205 (2017).

Acknowledgements

This work was funded by Prostate Cancer UK and The Bob Willis Fund through a Research Innovation Award [RIA16-ST2-011], a Prostate Cancer Research and the Mark Foundation for Cancer Research grant (reference 6961), a Prostate Cancer UK Travelling Prize Fellowship [TLD-PF19-002] and the J.G.W. Patterson Foundation.

Author contributions

K.H., M.O.M., E.S., R.G., K.L., K.B., G.H., and E.A.G. performed in vitro experiments. K.H. and M.O.M. performed bio-informatic analyses. A.B. and F.J.G.M. carried out the glycoproteomics study. E.S., H.T., Y.Z., J.H., and N.W. performed the in vivo studies. K.H., E.S., R.H., S.P., N.W. and D.J.E. supervised the study and and/or provided critical review. JM wrote the manuscript and made the figures. K.H., M.O.M., E.S., R.H., D.J.E. and N.W. contributed to critical review and paper writing. K.H., M.O.M., E.S. and J.M. designed, analysed and interpreted the study. J.M. conceived the study and is senior author and corresponding author. All authors read the manuscript, agree with the content, and were given the opportunity to provide input.

Competing interests

The authors would like to declare the following potential competing interests. JM, ES, and GH are shareholders of GlycoScoreDx Ltd. All other authors declare no potential conflict of interest.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-43019-8>.

Correspondence and requests for materials should be addressed to J.M.

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

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023