ORIGINAL ARTICLE

Genomic profiling identifies *GPC5* amplification in association with sarcomatous transformation in a subset of uterine carcinosarcomas

M Herman Chui[1]* ⓘD, Cherry Have[1], Lien N Hoang[2], Patricia Shaw[1], Cheng-Han Lee[2]
and Blaise A Clarke[1]

[1] Department of Pathology, University Health Network, University of Toronto, Toronto, Canada
[2] Department of Pathology, BC Cancer Agency, University of British Columbia, Vancouver, Canada

Correspondence to: Dr M Herman Chui, Department of Pathology, University Health Network, 200 Elizabeth Street, 11th Floor, Toronto, Ontario, Canada. E-mail: michaelherman.chui@utoronto.ca

Abstract

Uterine carcinosarcoma, also known as Malignant Mixed Müllerian Tumour, is a high-grade biphasic neoplasm composed of sarcomatous elements thought to originate via transdifferentiation from high-grade endometrial carcinoma. To identify molecular factors contributing to the histogenesis of this tumour, we analyzed DNA extracted from matched carcinoma and sarcoma components from 12 cases of carcinosarcoma by a molecular inversion probe microarray to assess genomic copy number alterations (CNAs) and allelic imbalances. Widespread CNAs were identified in tumours with serous histology in the carcinoma component (9/12), while the remaining three cases with endometrioid carcinoma were near-diploid. Quantification of the extent of genomic aberrations revealed a significant increase in sarcoma relative to carcinoma in tumours with well-delineated histologic components. Focal amplification of 13q31.3 was identified in 6/12 profiled tumours, of which four harboured the aberration exclusively in the sarcoma component. This result was verified by fluorescence *in situ* hybridization against *GPC5*, the only gene situated within the minimal region of amplification. In a validation cohort composed of 97 carcinosarcomas and other uterine sarcomas, amplification of *GPC5* (*GPC5/CEP13* ratio ≥ 2.2) was identified in 11/97 (11.3%) cases (9/64 carcinosarcoma, 1/3 rhabdomyosarcoma, 1/21 leiomyosarcoma, 0/8 adenosarcoma, 0/1 undifferentiated endometrial sarcoma) and an additional 4 (2.8%) cases had low level gains (*GPC5/CEP13* ratio ≥1.5 but <2.2). The functional relevance of Glypican-5, the gene product of *GPC5*, in regulating differentiation and lineage commitment was demonstrated in an endometrial carcinoma cell line *in vitro*. In conclusion, we identified *GPC5* amplification as a molecular event mediating epithelial-mesenchymal transdifferentiation in a subset of uterine carcinosarcomas.

Keywords: carcinosarcoma; Malignant Mixed Müllerian Tumour; copy number alterations; gene amplification; transdifferentiation; epithelial-mesenchymal transition

Introduction

Uterine carcinosarcoma, also known as Malignant Mixed Müllerian Tumour (MMMT), is an aggressive biphasic neoplasm composed of high-grade epithelial and mesenchymal elements. Known histotypes of endometrial carcinoma, serous in particular, comprise the carcinoma component, while the sarcoma component may exhibit homologous (e.g. endometrial stromal sarcoma, leiomyosarcoma) or heterologous (e.g. osteosarcoma, chondrosarcoma) differentiation [1].

We were therefore interested in studying uterine carcinosarcoma as a model system to understand the molecular basis for sarcomatous transformation and to gain insight into the role of epithelial-mesenchymal transition (EMT) in tumour progression.

Three theories have been proposed to explain the histogenesis of uterine carcinosarcoma [1]. The 'collision' model stipulates that the components represent distinct clonal populations that arise independently from different cells-of-origin. The more widely accepted 'combination' and 'conversion' models imply

monoclonality; according to the former, malignant transformation of a bipotential progenitor cell results in a tumour with a 'combination' of phenotypes, while in the latter scenario, tumour progression in a pre-existing endometrial carcinoma causes 'conversion' to sarcomatous differentiation.

By applying next-generation targeted sequencing of known cancer genes to uterine carcinosarcoma, we recently demonstrated shared mutational profiles between matched carcinoma and sarcoma components, consistent with their monoclonal relationship [2]. These included *TP53* mutations, and recurrent alterations in *PTEN* and *PIK3CA* which activate the PI3K pathway. While confirming a common origin for the carcinoma and sarcoma components, we were unable to establish the directionality of tumour progression, as neither histologic component consistently harboured a higher mutational burden. In contrast, a similar approach used to study another mixed uterine tumour, namely, de-differentiated endometrial carcinoma, identified additional mutations in the undifferentiated component relative to the low-grade endometrioid carcinoma component [3].

Given that the majority of uterine carcinosarcomas contain a serous carcinoma component and that the genomic landscape of serous carcinoma is dominated by frequent copy number gains and losses [4,5], we hypothesized that an unbiased approach interrogating genome-wide copy number alterations (CNAs) may yield further insight into the underlying molecular basis of sarcomatous differentiation. In this study, using a molecular inversion probe microarray profiling strategy to assess genomic CNAs and allelic imbalances, we provide evidence supporting the conversion theory of histogenesis and identify candidate genetic alterations that may be responsible for sarcomatous transformation of endometrial carcinoma.

Methods

Cases and sample preparation

Work with human tissues has been approved by the institutional Research Ethics Board (UHN REB 10-0006-T, 5 June 2014). Twelve cases of uterine carcinosarcoma, comprising the index cohort, were selected from the archives of the Department of Pathology, University Health Network. The validation cohort consisted of 97 cases of uterine sarcomas and mixed tumours represented on a tissue microarray obtained from the BC Cancer Agency and Vancouver General Hospital.

All available diagnostic H&E slides and immunohistochemical stains were reviewed by the study pathologists (MHC, BAC) to confirm the diagnosis of uterine carcinosarcoma in the index cohort. The following parameters were recorded: histologic subtyping of the carcinoma component, homologous versus heterologous differentiation in the sarcoma component (with specific histotype, if possible), whether the components are admixed or well-delineated, and pathologic stage. Representative areas from the epithelial and mesenchymal components with >90% tumour cell purity were selected for molecular analysis. Circled areas on slides were matched to the corresponding formalin-fixed paraffin-embedded (FFPE) tissue blocks, from which duplicate 0.6 mm diameter tissue cores were taken.

DNA extraction and quantification

DNA was extracted from tissue cores using the QIAamp DNA FFPE Tissue Kit and quantitation of DNA was performed using the Quant-iT^{TM} Pico-Green dsDNA Assay Kit (Invitrogen, OR, USA), as per the manufacturers' instructions.

Genomic copy number profiling

DNA from FFPE samples was analysed using the OncoScan Express 2.0 service from Affymetrix, which employs a molecular inversion probe microarray platform containing 334 000 copy number and single nucleotide polymorphism (SNP) oligonucleotide probes. Median probe spacing is 4 200 kilobases (kb), with much denser coverage within known cancer genes. OncoScan copy number data were processed and normalized in reference to an Affymetrix panel of normal reference samples. Copy number calls were determined from log2 ratios and B-allele frequencies using the Affymetrix TuScan algorithm with default settings, as previously described [6]. Results were visualized using Nexus Copy Number 5.1 software (BioDiscovery). All array data were manually reviewed for subtle alterations not automatically called by the software. The complete OncoScan array dataset is presented in supplementary material, Table S1.

Immunohistochemistry

Immunohistochemistry for desmin was performed on tissue microarrays or whole sections using the Ventana Discovery XT and the Ventana Benchmark XT automated systems (Ventana Medical Systems, Tucson, AZ, USA).

Fluorescence *in situ* hybridization

Interphase fluorescence *in situ* hybridization (FISH) was performed using a bacterial artificial chromosome (BAC) clone corresponding to the *GPC5* locus (RP11–306O1; BACPAC Resources Center, Oakland, CA). CEP13 (13q12.11) was visualized using a commercial probe (SureFISH, Agilent-DAKO, Glostrup, Denmark).

The *GPC5* BAC was labeled with red (02N34–050; Abbott Molecular) dUTPs using a nick translation kit (07J00–001; Abbott Molecular) according to the manufacturer's instructions. After nick translation and before ethanol precipitation, 10 μg of human COT-1 DNA (06J31-001, 1 μg/μl; Abbott Molecular) was added per microgram of labeled BAC. Precipitated DNA was resuspended in hybridization buffer (06J67-001; Abbott Molecular) diluted in deionized water (7:10) to a final concentration of 1:10 (w/v). Probe specificity was confirmed to map to the appropriate region by hybridization to normal metaphase chromosomes before use.

Following deparaffinization in xylene, and hydration in graded alcohols, slides were incubated in sodium citrate 10 mM (pH 6.0) at 80 °C for 120 min. After rinsing in deionized water and 2X SSC at RT for 5 min each, slides were then digested in pepsin (P7012-1G, 75,000 U/ml, 1/100 dilution in 0.01 N HCl; Sigma-Aldrich, Oakville, Canada) at 37 °C for 10–15 min. Pepsin was inactivated by placing the slides in deionized water at room temperature for 5 min. The slides were further dehydrated in graded alcohols. In a thermobrite apparatus (Abbott Molecular), samples were co-denatured with the probe mix (each probe, 50 ng; human COT-1 DNA, 2 μg; hybridization buffer, 6 μl; deionized water, 1 μl) at 75 °C for 10 min. The hybridization was carried at 37 °C for 18 h in the humidified chamber. Slides were washed in 0.4× SSC/0.3% IGEPAL at 69 °C for 2 min and in 2× SSC/0.1% IGEPAL at room temperature for 1 min. The slides were counterstained with DAPI and mounted in aqueous mounting medium (1:6 dilution; H-1000; Vector Laboratories).

Signals were detected using the AxioImager Z1 (Zeiss). Red and green signals were enumerated per cell nucleus. For whole-section tissue slides, at least 20 nuclei were scored in each of the carcinoma and sarcoma components, using the H&E-stained section as a guide. For tissue microarray sections, at least 10 nuclei were scored per core (each case represented by at least 2 cores), regardless of histology, due to limited tissue sampling. Nuclei with no signals or showing signals of only one colour were not scored. For signal clusters, the score was determined by estimating the number of signals that could fit within the cluster. Scoring was performed by a single observer (MHC). For each of the cases with whole tissue sections, scoring was performed twice on two separate occasions to ensure intra-observer consistency. The average score was taken from repeated scoring of whole tissue sections and from replicate cores on the tissue microarray and used to classify *GPC5* copy number status. The cut-off values for scoring *HER2* FISH in breast carcinoma, as per the 2007 ASCO/CAP Guidelines [7], were used for classification: *GPC5*/*CEP13* ratio ≥ 2.2 for high-copy gain, and ≥ 1.5 but < 2.2 for low-copy gain.

Cell culture studies

The NOU-1 cell line was obtained from the Coriell Institute for Medical Research (NJ, USA). Cells were maintained in DMEM supplemented with 10% fetal bovine serum. The human GPC5 expression construct or LacZ control (both in the pLX304 vector, in which transgene expression is driven by the CMV promoter and includes a terminal V5 tag) was transfected into NOU-1 cells using Effectene transfection reagent (Qiagen), as per the manufacturer's instructions. At 96 h post-transfection, 2000 cells were seeded per well of a 96-well plate (Corning, NY, USA). Growth was recorded every 12 h, by automated measurement cell confluence using the Incucyte ZOOM System (EssenBio, Michigan, USA).

For protein extraction, cells were lysed in 2X sample buffer at 96 h post-transfection. Western blotting was performed as previously described, using the following primary antibodies: V5-tag (rabbit monoclonal, CST), pan-keratin (mouse monoclonal, Sigma), E-cadherin (mouse monoclonal, BD Biosciences), vimentin (mouse monoclonal, Sigma), and GAPDH (rabbit monoclonal, CST).

Results

The histopathologic features of 12 uterine carcinosarcomas subjected to molecular profiling are summarized in Table 1. Cases were selected to be representative of the heterogeneous spectrum of histologic features, including tumours with serous or endometrioid histology in the carcinoma component, and sarcoma components showing homologous or heterologous differentiation.

For each case, genome-wide copy number profiling of matched carcinoma and sarcoma components was performed using the Oncoscan Express 2.0 platform. Copy number profiles of the epithelial component

Table 1. Characteristics of the index cohort profiled by SNP array

| Case # | Histologic type | | Spatial separation of components | Pathologic stage |
	Carcinoma component	Sarcoma component		
1	Serous	Homologous	Well-Delineated	T2 NX
2	Serous	Homologous	Admixed	T1b N1
3	Serous	LMS, Undifferentiated	Well-Delineated	T1a N0
4	Serous	RMS	Well-Delineated	T3a N2
5	Serous	RMS	Well-Delineated	T1a N0
6	Serous	Osteosarcoma, Chondrosarcoma, RMS, Undifferentiated	Well-Delineated	T1a N0
7	Serous	Chondrosarcoma, Undifferentiated	Admixed	T1a NX
8	Serous	RMS, Chondrosarcoma	Admixed	T3c NX
9	Serous	Undifferentiated	Admixed	T1b N0
10	Endometrioid	Homologous	Well-delineated	T1a N2
11	Endometrioid	Homologous	Admixed	T1a N0
12	Endometrioid	Homologous	Admixed	T1b M0

LMS, leiomyosarcoma; RMS, rhabdomyosarcoma.

from each case were consistent with the subtype-associated profiles reported in The Cancer Genome Atlas endometrial carcinoma study: high frequency of CNAs in serous and low frequency in endometrioid [8] (Figure 1A,B). Considering only cases where the histologic components are spatially well-delineated, the sarcoma component typically had more CNAs and allelic imbalances than the corresponding carcinoma component, in keeping with the conversion theory of histogenesis (five out of six cases, Figure 2A,B). This could be expressed quantitatively as percentage of the genome altered, which showed a statistically significant increase when comparing matched sarcoma with carcinoma components (mean $\pm$ sem of differences in % of the genome altered $= 35 \pm 13\%$, $p = 0.048$, paired t-test, two-tailed). However, in tumours composed of an admixture of epithelial and mesenchymal elements, the molecular profiles were homogeneous throughout the tumour (5/6 cases), with no significant difference observed in the percentage of the genome altered ($p = 0.51$, paired t-test, two-tailed). Notably, in all three uterine carcinosarcomas with an endometrioid carcinoma component, copy number profiles were similar between matched samples. In two tumours (Cases #9 and #10), there were fewer CNAs in the sarcoma component.

To identify potential genetic drivers of sarcomatous transformation, matched pair analysis was used to identify recurrent CNAs remaining in the sarcoma component copy number profile after subtracting out the copy number profile of the corresponding carcinoma component. Two amplification peaks, positioned at chromosomes 8q21.2–21.3 and 13q31.3 were statistically significant after applying the GISTIC algorithm (Figure 3A). Of particular interest, while the 8q21 amplicon encompassed 28 genes,

only a single gene, namely, GPC5 (encoding Glypican-5), resides within the 13q31.3 amplification peak (Figure 3B).

We therefore chose to validate the GPC5 amplicon and designed a FISH probe against this gene locus (Figure 3C), which was used to directly interrogate GPC5 copy number changes in whole sections of the 12 cases in the discovery cohort. Adapting the conservative cut-off values established for scoring HER2 FISH in breast carcinoma, as per the 2007 ASCO/CAP Guidelines [7] (GPC5/CEP13 ratio ≥ 2.2 for high-copy gain, and ≥ 1.5 but < 2.2 for low-copy gain), amplification of GPC5 was detected exclusively in the sarcoma component in five cases (Cases #1,2, 3, 6, and 9; Table 2 and Figure 3D,F). In addition, one case harboured high-copy amplification throughout the tumour (Case #5). Results obtained by SNP array and FISH were generally concordant (Figure 3D,E, and Table 2). Due to technical differences between the assays, cut-off values used for SNP array and FISH were not identical. As a result, discordance between SNP array and FISH were due to borderline calls near threshold values (cut-off between diploid and low-gain and between low-gain and high-gain). Notably, Case #8, with high-copy gain by SNP array and GPC5/CEP13 ratio $= 1.0$, was in fact polysomic for chromosome 13 (Figure 3G). Comparison of the copy number profiles of the four GPC5-amplified cases, as determined by both SNP array and FISH (Cases #2, 5, 6, and 8), with the copy number profiles of the rest of the cases did not reveal any significant co-occurrence of other CNAs with GPC5 amplification. However, the sample size is too small to make any definitive conclusions.

In an independent cohort comprised of 97 uterine mixed/mesenchymal tumours represented on a tissue microarray, the frequency of GPC5 amplification was
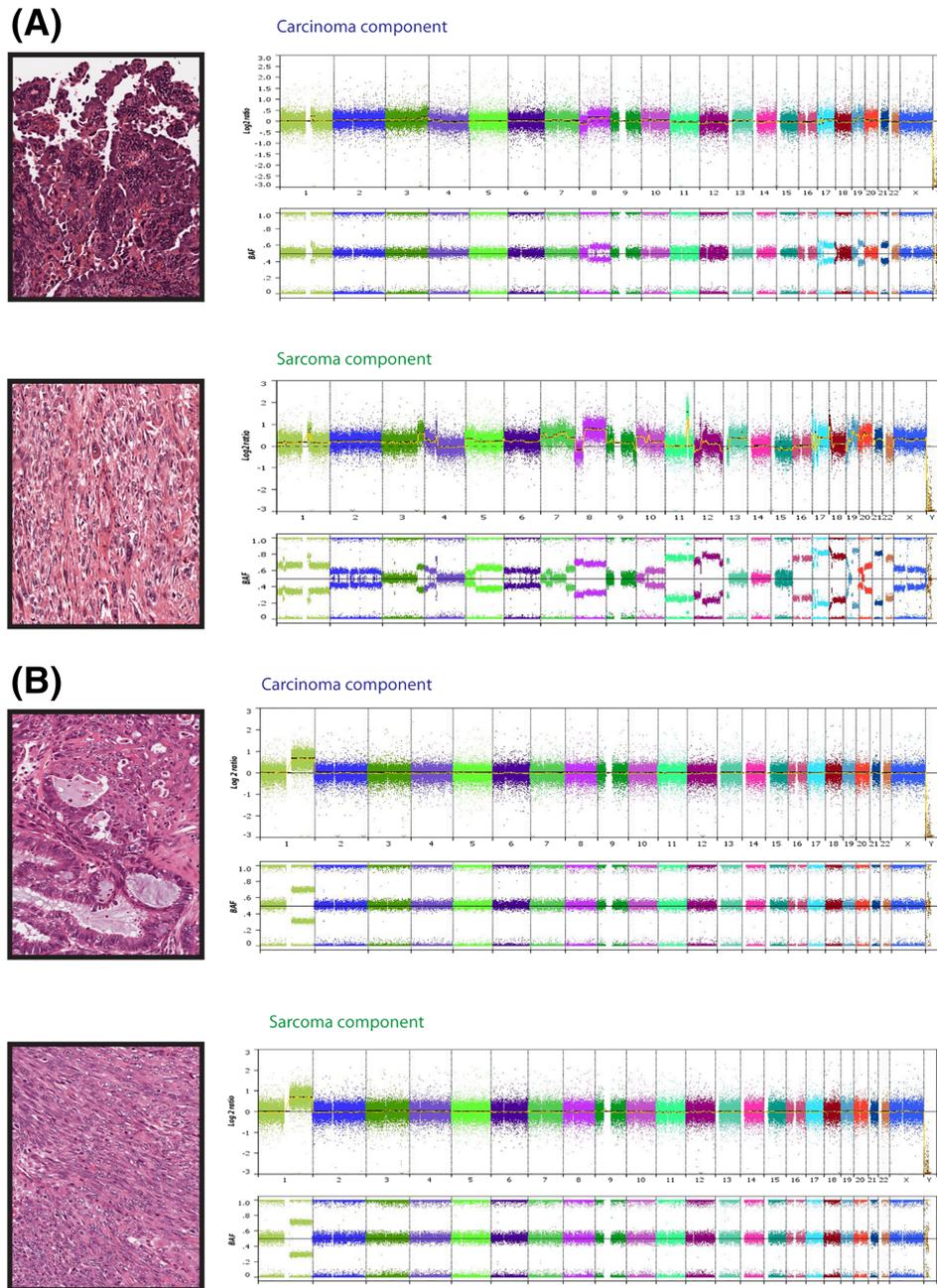
Figure 1. Genomic copy number profile of matched carcinoma and sarcoma components in carcinosarcoma. H&E stains, copy number (*right upper panel*) and allelic frequency (*right lower panel*) profiles, for (A) Case #1 – Serous carcinoma component well-delineated from the sarcoma component; and (B) Case #11 – Endometrioid carcinoma component admixed with sarcoma component. BAF – B-allele frequency.

15 [11 (11.3%) high-gain, 4 (2.8%) low-gain; Table 3]. Immunohistochemical staining for desmin was performed to determine whether there was any association between *GPC5* amplification and myogenic differentiation; no significant difference in frequency of desmin-positive cases was observed [GPC5-amplified: 11/15 (73%) versus non-amplified: 59/82 (72%)].

To interrogate the functional relevance of *GPC5* in sarcomatous transformation, a plasmid construct carrying the coding sequence of *GPC5* gene, under a constitutive promoter, was transfected into NOU-1 cells, initially derived from a poorly differentiated endometrial carcinoma [9]. Expression of the construct was verified by immunoblot. There was no
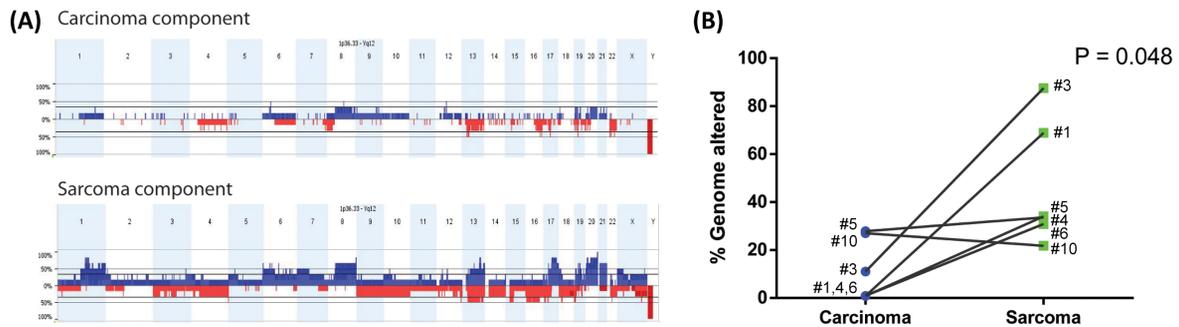
Figure 2. Increased genomic alterations detected in sarcoma relative to carcinoma components of carcinosarcoma, consistent with the 'conversion theory' of histogenesis. Analyses are limited to cases with well-delineated epithelial and mesenchymal components (n = 6). (A) Frequency of CNAs in epithelial versus mesenchymal tumour, in aggregate. (B) Comparison of percent genome altered in matched samples (paired *t*-test, two-tailed).

difference in proliferation rate between *GPC5*-transfected cells versus the empty vector-transfected control (Figure 4A). However, western blot for differentiation markers revealed that *GPC5* over-expression causes decreased expression of E-cadherin and cytokeratins, consistent with a functional role of Glypican-5 in mediating loss of epithelial differentiation in endometrial carcinoma cells (Figure 4B).

Discussion

In an earlier report, using a targeted sequencing approach to study the biphasic nature of uterine carcinosarcoma, we found identical mutations and similar mutational burdens in matched carcinoma and sarcoma components, consistent with a clonal relationship [2]. Since uterine carcinosarcomas are essentially metaplastic high-grade carcinomas, it may be reasonable to speculate that CNAs may be more prevalent than point mutations, based on extrapolation of the genomic profiling data from high-grade serous carcinoma [4,5]. With the aim of extending our earlier findings, the present study employed genomic copy number profiling. This unbiased comprehensive interrogation of genetic events in matched histologic components provides evidence supporting the sequential progression from carcinoma to sarcoma in at least a subset of tumours. As for the remainder, there are several potential explanations.

In cases consisting of an admixture of epithelial and mesenchymal tumour cells, it is likely that the tissue cores captured mixed populations, resulting in the inability to generate distinct genetic profiles for each component, which is a limitation of our study. Laser-capture microdissection (not used in the current study due to the input DNA requirement for genomic

analysis) may be considered in follow-up studies to analyze pure populations of matched carcinoma and sarcoma components.

Uterine carcinosarcomas with homogenous copy number profiles throughout could also be the result of transformation of a progenitor cell capable of epithelial and mesenchymal differentiation (i.e. 'combination' theory). Epigenetic changes affecting gene expression may be potentially involved in transdifferentiation. Since whole exomic sequencing was not performed on this cohort, we cannot exclude the possibility that an acquisition of a driver mutation was responsible for EMT.

We note that in two cases, a higher CNA burden was present in the sarcoma relative to the epithelial component. There were shared genetic alterations, consistent with a common origin. It is likely that in these cases, divergence from a common ancestral clone occurred early, followed by a more rapid progression of the sarcoma component. The derivation of uterine carcinosarcoma from Müllerian adenosarcoma has been previously proposed [10]. However, as the epithelial glands in adenosarcoma are thought to be hyperplastic in nature, independent malignant transformation of the epithelial component in a tumour with sarcomatous overgrowth would result in histologic components with completely different genomic profiles. Genomic analysis of more cases of carcinosarcoma, particularly those with adenosarcoma-like features, will be necessary to verify this hypothesis.

In the first in-depth interrogation of the mutational landscape of gynecologic carcinosarcomas by whole exomic sequencing, Jones *et al* confirmed the presence of carcinoma-associated mutations in *TP53* and components of the PI3K pathway, and identified novel genetic modifications in chromatin remodelling genes [11]. It should be noted that 4 of 22 cases were mismatch repair- and ARID1A-deficient, raising
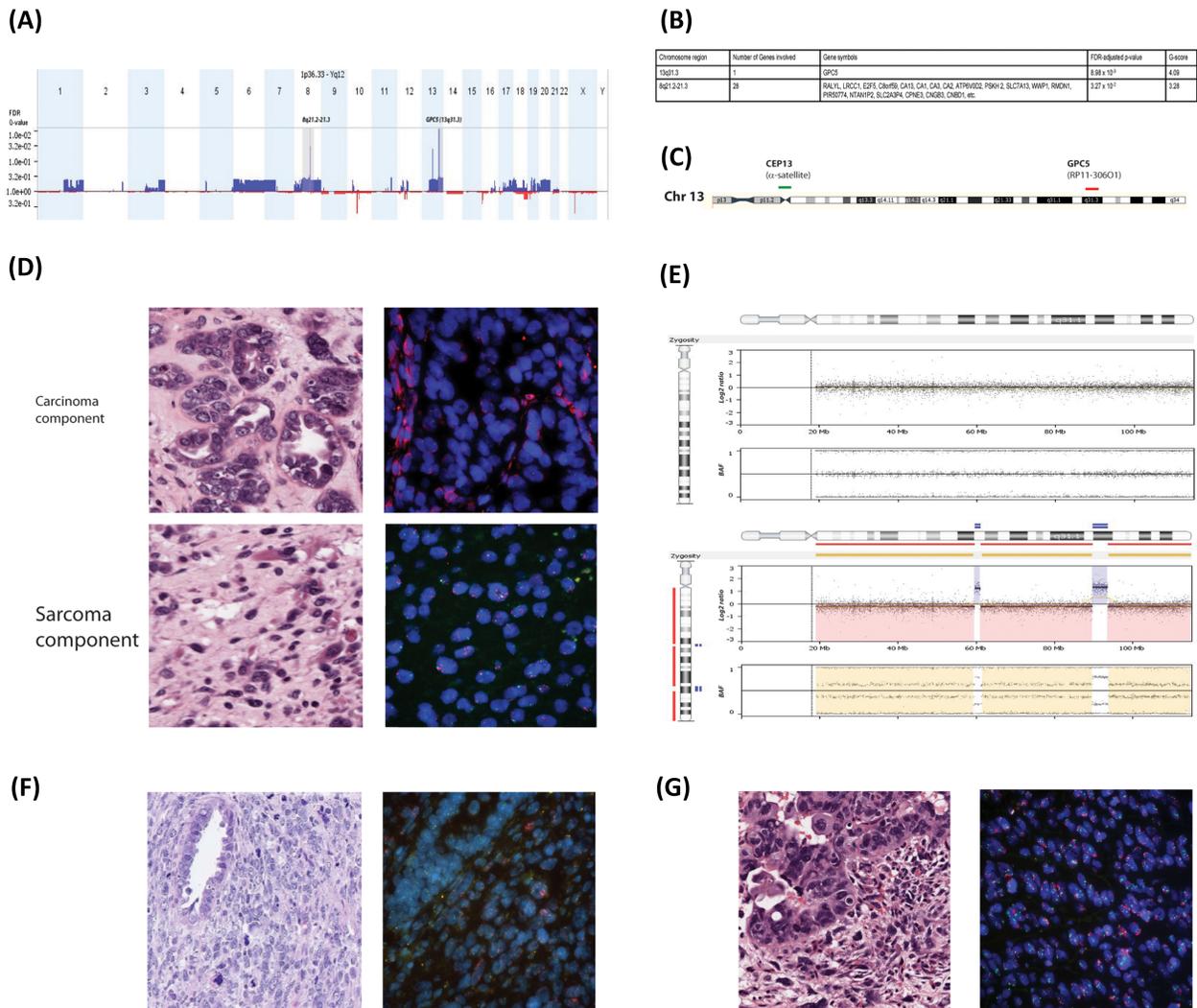
Figure 3. Recurrent amplification of 13q31–32 (encompassing *GPC5*) in the sarcoma component of carcinosarcoma. (A, B) GISTIC analysis of copy number profiles of the sarcoma component after substraction of the matched carcinoma component (n = 12). (C) FISH probes against the *GPC5* gene locus (red) and centromeric α-satellite probe for chromosome 13 (green). (D,E) Case #6, H&E and *GPC5* FISH (D), and CNAs and allelic imbalances in chromosome 13 by SNP array (E). (F) Uterine carcinosarcoma from the independent cohort, with *GPC5* amplification present only in the sarcoma component (H&E and *GPC5* FISH). (G) Case #10, showing chromosome 13 polysomy (H&E and *GPC5* FISH).BAF, B-allele frequency; for FISH, CEP13, green signal, GPC5, red signal.

concern that they may in fact represent de-differentiated endometrial carcinomas [12], rather than carcinosarcoma *sensu stricto*. In addition, with only a single tumour sample sequenced per case, carcinoma and sarcoma components were not directly compared, precluding analysis of intratumour heterogeneity and the relationship between genetic alterations and histomorphology.

These conclusions were confirmed in a subsequent whole-exome analysis of uterine and ovarian carcinosarcomas, in which novel missense mutations in genes coding for histone proteins, H2A and H2B, were identified in 21% of cases [13]. In cell cultures, enforced expression of mutant histones led to EMT

and enhanced invasiveness. In a subset of six cases, multi-region sequencing was performed on matched carcinoma and sarcoma components. Phylogenetic reconstruction of genetic events provided compelling evidence supporting a common origin of the 'present-day' carcinoma and sarcoma components, which likely started out as a carcinoma. After divergence into separate cell lineages at a variable time-point during tumour progression, mutations continue to accumulate independently in these subclonal populations.

In light of mounting evidence supporting 'sarcomatous transformation' of a pre-existing carcinoma in

Table 2. *GPC5* copy number status: SNP array and FISH validation

Case #	Carcinoma component		Sarcoma component	
	SNP array	FISH (*GPC5/CEP13* ratio)	SNP array	FISH (*GPC5/CEP13* ratio)
1	Diploid	1.3	Diploid	1.7
2	Low Copy Gain	1.2	High Copy Gain	7.2
3	Diploid	1.3	High Copy Gain	1.5
4	Diploid	1.0	Diploid	0.9
5	High Copy Gain	2.4	High Copy Gain	2.2
6	Diploid	1.0	High Copy Gain	8.4
7	Diploid	1.0	Diploid	1.0
8	High Copy Gain	1.0*	High Copy Gain	1.0*
9	Diploid	0.8	Diploid	1.6
10	Single Copy Loss	1.0	Single Copy Loss	0.7
11	Diploid	0.9	Diploid	0.9
12	Diploid	1.4	Diploid	1.0

*Polysomy with *GPC5* signals $\geq$ 4 per cell.

the histogenesis of carcinosarcoma, the major aim of the present study was to identify specific genetic events that mediate this process. We observed a novel amplification of the *GPC5* gene, located in 13q31.3, confined to the sarcoma component in a subset of uterine carcinosarcomas. It should be emphasized that only a single protein-coding region is present in this focal amplicon. Combined with the fact that this is a recurrent CNA, the evidence implicates the involvement of *GPC5* in sarcomatous transformation in some uterine carcinosarcomas. *GPC5* encodes Glypican-5, a member of the glypican family of cell surface heparin sulphate proteoglycans (which includes Glypican-3, the well-known immunohistochemical marker for hepatocellular carcinoma). These transmembrane proteins sequester extra-cellular ligands, including FGF, Wnt, and TGF-β, thereby facilitating the interaction between ligands and their cognate receptors [14]; hence, glypicans serve a co-receptor role in activating pathways known to regulate cellular differentiation.

Previous experimental work has implicated the involvement of glypicans in mesenchymal lineage commitment, such as Glypican-1 in promoting myogenic differentiation [15], and osteogenesis mediated by Glypican-3 [16]. While the literature concerning specifically Glypican-5 is limited, it is amplified and

over-expressed in rhabdomyosarcoma, and synergistically enhances FGF-induced mesoderm differentiation in *Xenopus laevis* embryos [17]. In keeping with these prior reports, the present study shows that introducing *GPC5* into an endometrial carcinoma cell line results in reduced expression of epithelial markers, supporting a causal role of *GPC5* amplification/over-expression in mediating EMT. More work is, however, necessary to characterize the underlying molecular mechanisms, including the interrelationships between Glypican-5 and known regulators of EMT, such as the Snail and Twist families of transcription factors.

Another open question is the role of Glypican-5 in the epigenetic regulation of cellular differentiation. Methylation profiling of MMMT in a recent study revealed aberrant methylation of the MIR200 family of microRNAs [18], which are implicated in maintaining the epithelial phenotype. Subsequent work has shown that nanoliposome-mediated delivery of miR-200c into carcinosarcoma-xenografted mice induced mesenchymal-epithelial transition and slowed tumour growth [19].

It should be emphasized that in addition to *GPC5*, gains of chromosome 8q21 in sarcoma relative to matched carcinoma components were also observed. This gene-rich region is frequently amplified in other cancers, including breast, prostate, and ovarian carcinomas and osteosarcoma [20], and thus likely contains oncogenes that generally support the malignant state.

In a broader context, our work provides insight into the role of cell lineage commitment in cancer progression. The ability of epithelial cancer cells to undergo distant metastasis has been controversially proposed to be dependent on transient EMT; however, conclusive histologic evidence of mesenchymal

Table 3. *GPC5* alterations in the independent validation cohort

Tumour type	Number of cases	*GPC5* high copy gain	*GPC5* low copy gain
MMMT	64	9	3
Rhabdomyosarcoma	3	1	0
Leiomyosarcoma	21	1	1
Adenosarcoma	8	0	0
Endometrial Stromal Sarcoma	1	0	0
Total	97	11 (11.3%)	4 (2.8%)
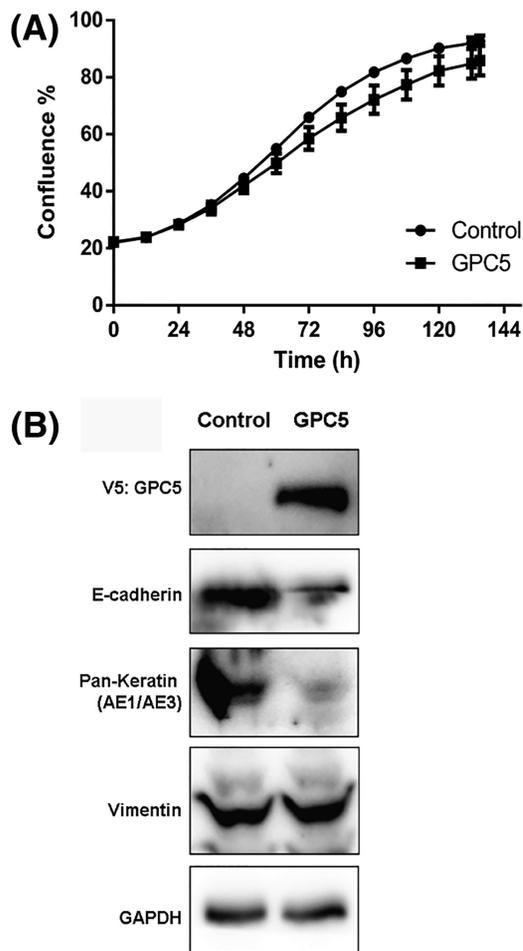
Figure 4. GPC5 mediates loss of epithelial differentiation in NOU-1 endometrial carcinoma cells. (A) No effect of *GPC5* transfection on proliferation rate. (B) Western blot of lysates from *GPC5*-transfected cells and empty vector-transfected control for differentiation markers.

transdifferentiation is lacking [21]. True EMT is uncommon and may account for the biphasic nature of rare tumour types with mixed histology, of which uterine carcinosarcoma serves as the prototypic example. However, the finding of genetic events (e.g. gains in *GPC5*, and histone gene mutations) causing a permanent switch in lineage commitment does not fit with the transient EMT model. Moreover, it is difficult to reconcile how *bona fide* EMT does not cause systemic dissemination of the sarcoma component of carcinosarcoma; it is, in fact, the carcinoma component of carcinosarcoma that travels via the lymphatic route to establish metastases. Nonetheless, the fact that these genetic alterations are recurring suggests that they are not mere passenger events. Clinicopathologic studies to evaluate the impact of EMT-inducing genetic changes on survival and

treatment outcomes will provide insight into their biologic and clinical relevance. Incorporating measures of genomic instability [22,23] in these analyses will enable determination of whether EMT-related genetic features carry independent prognostic value or simply represent surrogate markers of a highly-damaged genome.

In summary, the present work builds upon the existing literature supporting sarcomatous transformation of a pre-existing carcinoma in the histogenesis of uterine carcinosarcoma. Copy number profiling of matched carcinoma and sarcoma components identified a recurring *GPC5* gain/amplification preferentially in the sarcoma component in a subset of cases, which was validated by FISH in an independent cohort. Integrated molecular profiling of larger clinically-annotated cohorts will be needed to identify other genetic or epigenetic changes associated with EMT and to evaluate their prognostic relevance in patients with uterine carcinosarcoma.

Acknowledgement

Author contributions statement

MHC conceived and performed experiments, analysed data and drafted the manuscript. CH performed experiments. LNH, PAS, CHL, and BAC contributed and reviewed cases. All authors were involved in writing the paper and had final approval of the submitted and published versions.

References

1. McCluggage WG. A practical approach to the diagnosis of mixed epithelial and mesenchymal tumours of the uterus. *Mod Pathol* 2016; **29**(Suppl 1): S78–S91.
2. McConechy MK, Hoang LN, Chui MH, *et al*. In-depth molecular profiling of the biphasic components of uterine carcinosarcomas. *J Pathol Clin Res* 2015; **1**: 173–185.
3. Kuhn E, Ayhan A, Bahadirli-Talbott A, *et al*. Molecular characterization of undifferentiated carcinoma associated with endometrioid carcinoma. *Am J Surg Pathol* 2014; **38**: 660–665.
4. The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011; **474**: 609–615.

5. Kuhn E, Wu R-C, Guan B, *et al*. Identification of molecular pathway aberrations in uterine serous carcinoma by genome-wide analyses. *J Natl Cancer Inst* 2012; **104:** 1503–1513.

6. Foster JM, Oumie A, Togneri FS, *et al*. Cross-laboratory validation of the OncoScan(R) FFPE Assay, a multiplex tool for whole genome tumour profiling. *BMC Med Genomics* 2015; **8:** 5.

7. Wolff AC, Hammond MEH, Schwartz JN, *et al*. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 2006; **25:** 118–145.

8. Kandoth C, Schultz N, Cherniack AD, *et al*. Integrated genomic characterization of endometrial carcinoma. *Nature* 2013; **497:** 67–73.

9. Faruqi SA, Satyaswaroop PG, LiVolsi VA, *et al*. Establishment and characterization of a poorly differentiated lethal human endometrial carcinoma cell line (NOU-1) with karyotype 46, XX. *Cancer Genet Cytogenet* 2002; **138:** 44–49.

10. Seidman JD, Chauhan S. Evaluation of the relationship between adenosarcoma and carcinosarcoma and a hypothesis of the histogenesis of uterine sarcomas. *Int J Gynecol Pathol* 2003; **22:** 75–82.

11. Jones S, Stransky N, McCord CL, *et al*. Genomic analyses of gynaecologic carcinosarcomas reveal frequent mutations in chromatin remodelling genes. *Nat Commun* 2014; **5:** 5006.

12. Garg K, Soslow RA. Lynch syndrome (hereditary non-polyposis colorectal cancer) and endometrial carcinoma. *J Clin Pathol* 2009; **62:** 679–684.

13. Zhao S, Bellone S, Lopez S, *et al*. Mutational landscape of uterine and ovarian carcinosarcomas implicates histone genes in epithelial-mesenchymal transition. *Proc Natl Acad Sci USA* 2016; **113:** 12238–12243.

14. Hacker U, Nybakken K, Perrimon N. Heparan sulphate proteoglycans: the sweet side of development. *Nat Rev Mol Cell Biol* 2005; **6:** 530–541.

15. Gutierrez J, Brandan E. A novel mechanism of sequestering fibroblast growth factor 2 by glypican in lipid rafts, allowing skeletal muscle differentiation. *Mol Cell Biol* 2010; **30:** 1634–1649.

16. Haupt LM, Murali S, Mun FK, *et al*. The heparan sulfate proteoglycan (HSPG) glypican-3 mediates commitment of MC3T3-E1 cells toward osteogenesis. *J Cell Physiol* 2009; **220:** 780–791.

17. Williamson D, Selfe J, Gordon T, *et al*. Role for amplification and expression of glypican-5 in rhabdomyosarcoma. *Cancer Res* 2007; **67:** 57–65.

18. Li J, Xing X, Li D, *et al*. Whole-genome DNA methylation profiling identifies epigenetic signatures of uterine carcinosarcoma. *Neoplasia* 2017; **19:** 100–111.

19. Tseng JH, Bisogna M, Hoang LN, *et al*. miR-200c-driven mesenchymal-to-epithelial transition is a therapeutic target in uterine carcinosarcomas. *Sci Rep*. 2017; **7:** 3614.

20. Byrne JA, Chen Y, Martin La Rotta N, *et al*. Challenges in identifying candidate amplification targets in human cancers: chromosome 8q21 as a case study. *Genes Cancer* 2012; **3:** 87–101.

21. Chui MH. Insights into cancer metastasis from a clinicopathologic perspective: Epithelial-Mesenchymal Transition is not a necessary step. *Int J Cancer* 2013; **132:** 1487–1495.

22. Bayani J, Paderova J, Murphy J, *et al*. Distinct patterns of structural and numerical chromosomal instability characterize sporadic ovarian cancer. *Neoplasia* 2008; **10:** 1057–1065.

23. Cope L, Wu R-C, Shih I-M, *et al*. High level of chromosomal aberration in ovarian cancer genome correlates with poor clinical outcome. *Gynecol Oncol* 2013; **128:** 500–505.

SUPPLEMENTARY MATERIAL ONLINE

Table S1. Complete OncoScan SNP array dataset