# Evaluation of Serum Glycoprotein Biomarker Candidates for Detection of Esophageal Adenocarcinoma and Surveillance of Barrett's Esophagus


## Authors

Alok K. Shah, Gunter Hartel, Ian Brown, Clay Winterford, Renhua Na, Kim-Anh Lê Cao, Bradley A. Spicer, Michelle A. Dunstone, Wayne A. Phillips, Reginald V. Lord, Andrew P. Barbour, David I. Watson, Virendra Joshi, David C. Whiteman, and Michelle M. Hill

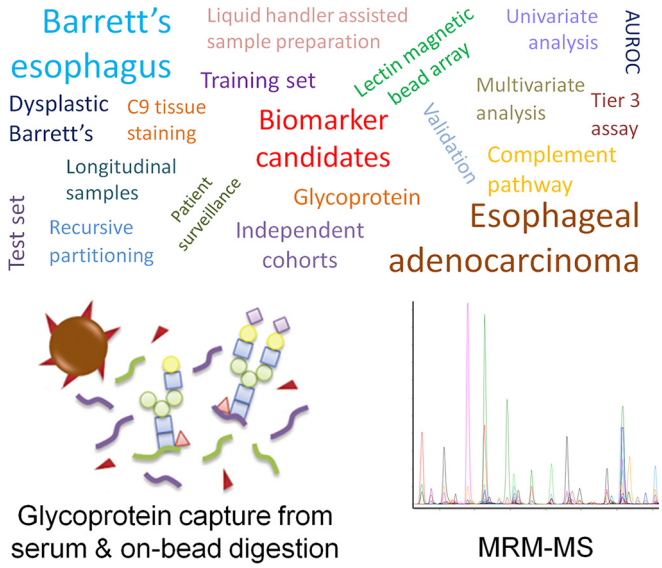
## Correspondence

Michelle.Hill@qimrberghofer.edu.au


## In Brief


This paper reports independent cohort validation of previously discovered esophageal adenocarcinoma (EAC) serum diagnostic glycoprotein biomarker candidates. To demonstrate their potential clinical application in evaluating patients during endoscopy-biopsy surveillance, we developed a panel of 10 biomarker candidates (AUC 0.93) to discriminate Barrett's esophagus (BE) patients not requiring intervention [BE+/− low-grade dysplasia] from those requiring intervention [BE with high-grade dysplasia or EAC]. Complement pathway proteins appear to be dysregulated during EAC, with C9 detected in BE and EAC tissue.


## Graphical Abstract




## Highlights

- C9, GSN, PON1, and PON3 validated as serum biomarker candidates of EAC.
- Multimarker panel with AUROC of 0.93 to aid current endoscopy surveillance of BE.
- Induction of tissue C9 in BE, dysplastic BE and EAC.
- Alteration of complement pathway glycoproteins during BE-EAC pathogenesis.


**Shah et al., 2018, Molecular & Cellular Proteomics *17*, 2324–2334**



**https://doi.org/10.1074/mcp.RA118.000734**

# Evaluation of Serum Glycoprotein Biomarker Candidates for Detection of Esophageal Adenocarcinoma and Surveillance of Barrett's Esophagus*[S]

Alok K. Shah‡§, Gunter Hartel‡, Ian Brown¶, Clay Winterford‡, Renhua Na‡, Kim-Anh Lê Cao§‖, Bradley A. Spicer**, Michelle A. Dunstone**, Wayne A. Phillips‡‡, Reginald V. Lord§§, Andrew P. Barbour§, David I. Watson¶¶, Virendra Joshi‖‖, David C. Whiteman‡, and Michelle M. Hill‡§‡‡‡

**Esophageal adenocarcinoma (EAC) is thought to develop from asymptomatic Barrett's esophagus (BE) with a low annual rate of conversion. Current endoscopy surveillance of BE patients is probably not cost-effective. Previously, we discovered serum glycoprotein biomarker candidates which could discriminate BE patients from EAC. Here, we aimed to validate candidate serum glycoprotein biomarkers in independent cohorts, and to develop a biomarker candidate panel for BE surveillance. Serum glycoprotein biomarker candidates were measured in 301 serum samples collected from Australia (4 states) and the United States (1 clinic) using previously established lectin magnetic bead array (LeMBA) coupled multiple reaction monitoring mass spectrometry (MRM-MS) tier 3 assay. The area under receiver operating characteristic curve (AUROC) was calculated as a measure of discrimination, and multivariate recursive partitioning was used to formulate a multi-marker panel for BE surveillance. Complement C9 (C9), gelsolin (GSN), serum paraoxonase/arylesterase 1 (PON1) and serum paraoxonase/lactonase 3 (PON3) were validated as diagnostic glycoprotein biomarkers in lectin pull-down samples for EAC across both cohorts. A panel of 10 serum glycoprotein biomarker candidates discriminated BE patients not requiring intervention (BE± low grade dysplasia) from those requiring intervention (BE with high grade dysplasia (BE-HGD) or EAC) with an AUROC value of 0.93. Tissue expression of C9 was found to be induced in BE, dysplastic BE and EAC. In longitudinal samples from subjects that have progressed toward EAC, levels of serum C9 were significantly ($p < 0.05$) increased with disease progression in EPHA (erythroagglutinin from *Phaseolus vulgaris*) and NPL (*Narcissus pseudonarcissus* lectin) pull-down samples. The results confirm alteration of complement pathway glycoproteins during BE-EAC pathogenesis. Further prospective clinical validation of the confirmed biomarker candidates in a large cohort is warranted, prior to development of a first-line BE surveillance blood test.** ***Molecular & Cellular Proteomics 17: 2324–2334, 2018. DOI: 10.1074/mcp.RA118.000734.***

Esophageal cancer is the sixth most common cause of cancer related mortality in men, with 3-fold higher rates in men than women (1, 2). Of the two main histological subtypes (adenocarcinoma and squamous cell carcinoma), the incidence of esophageal adenocarcinoma (EAC)[1] has been rising continuously in Western countries, and accounts for the majority of cases (3–5). Despite aggressive treatment, EAC has a 5-year survival of less than 20% (6). EAC is thought to develop from the metaplastic condition Barrett's esophagus (BE) as a consequence of chronic gastroesophageal reflux disease (GERD) through a metaplasia-dysplasia-adenocarcinoma sequence (Fig. 1*A*) (7–9).

Currently, BE patients usually undergo endoscopy-biopsy surveillance with the degree of dysplasia assessed by histopathology as a biomarker to monitor risk of neoplastic progression (10). Patients diagnosed with high-grade dysplasia (BE-HGD) are treated with endoscopic mucosal resection, radiofrequency ablation or surgery, to halt further disease

From the ‡QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia; §The University of Queensland Diamantina Institute, Faculty of Medicine, The University of Queensland, Translational Research Institute, Brisbane, Queensland, Australia; ¶Envoi Pathology, Brisbane, Queensland, Australia; ‖Melbourne Integrative Genomics and School of Mathematics and Statistics, The University of Melbourne, Victoria, Australia; **Department of Biochemistry and Molecular Biology, Monash University, Melbourne, Victoria, Australia; ‡‡Peter MacCallum Cancer Centre, and Sir Peter MacCallum Department of Oncology, The University of Melbourne, Parkville, Victoria, Australia; §§St Vincent's Centre for Applied Medical Research and University of Notre Dame School of Medicine, Sydney, Australia; ¶¶Discipline of Surgery, Flinders University, Adelaide, South Australia, Australia; ‖‖Ochsner Health System, Gastroenterology, New Orleans, LA

Received March 22, 2018, and in revised form, July 3, 2018
Published, MCP Papers in Press, August 10, 2018, DOI 10.1074/mcp.RA118.000734

progression (10–12). The significant cost of endoscopy plus the low annual progression rate to HGD or EAC means that the cost-effectiveness of endoscopic surveillance is questioned at the population level (13–16). Further, the evaluation of dysplasia in tissue biopsies by histopathology is prone to inter-observer variability and sampling error (17). A less costly and minimally invasive diagnostic procedure is needed for cost-effective screening and surveillance of at-risk populations (18, 19). We envisage that a blood test which can reliably discriminate BE-HGD and EAC from BE, GERD and healthy patients could be used to pre-select the patients for endoscopic examination, while reducing unnecessary endoscopy for most patients in a surveillance program.

As the first step to developing blood-based EAC diagnostic test, we focused on differential lectin binding of serum glycoproteins during EAC pathogenesis. We established a new glycoprotein biomarker pipeline that couples lectin-based glycoprotein isolation with state-of-the-art discovery and targeted proteomics (20–22). We then applied it to identify and verify changes in lectin binding profile of serum glycoproteins between healthy, BE and EAC patients (22). Here, we report results from validation in independent cohorts, and evaluation of biomarker panels for surveillance of BE patients.

## EXPERIMENTAL PROCEDURES

*Clinical Cohorts*—Ethical approval was obtained from all participating institutions, and all patients provided informed consent to participate in the studies. We investigated two independent cohorts recruited in Australia and the United States, respectively. The Australian samples were selected from participants recruited into The Progression of Barrett's Esophagus to Cancer Network (PROBE-NET) study across 4 states (New South Wales, Queensland, South Australia and Victoria). A total of 252 serum samples collected from 242 patients were analyzed (Normal - 43, BE - 65, BE-LGD - 39, BE-HGD - 35, and EAC - 60, at baseline). Of the 252 serum samples, 10 samples were from patients who progressed to the subsequent stage of the disease whereas 242 were baseline serum. Independently, a cross-sectional cohort of 49 serum samples collected at Ochsner Health System, New Orleans, United States were also analyzed. Nine of the Ochsner patients in the Normal group had a history of Barrett's related pathology and received endoscopic treatments. These patients were confirmed to have no Barrett's mucosa by histology at the time of serum collection, however they may have submucosal disease, hence the normal groups will be now onwards called non-specialized epithelium (NSE, epithelium without Barrett's mucosa). The Ochsner cohort comprised of 14 NSE, 13 BE, 3 BE-ID (Barrett's mucosa which is indefinite for dysplasia), 2 BE-LGD, 7 BE-HGD, and 10 EAC patients.

Similar procedure of serum collection was followed across all sites. Briefly, 10 ml of blood was collected by trained phlebotomists into Vacutainer red top collection tubes (Becton Dickinson (BD), New Jersey). Once in the lab, the blood was allowed to coagulate at room temperature for a minimum of 20 min. The tubes were centrifuged at 3000 RPM (Beckman CS-6R) for 10 min at 22 °C (brake off) making sure that all blood was separated. Using sterile pipette, the serum was removed, aliquots were prepared and stored at −80 °C. Serum samples were shipped to the Translational Research Institute, Brisbane on dry ice for this study. The clinical diagnosis linked to the sample was based on histological examination of biopsies taken at the same endoscopy. The diagnoses were provided to the researcher performing the biomarker candidate measurements to allow batch randomization design for the assay. Formalin-fixed, paraffin-embedded (FFPE) tissue sections from the Ochsner Health System were selected and shipped to Brisbane, Australia for immunohistochemistry.

Both PROBE-NET and the Ochsner cohort provided information on patients' age, sex and body mass index (BMI, calculated as weight (kg)/(height (m))$^2$); whereas ethnicity was provided by the Ochsner cohort only and education, alcohol drinking and tobacco smoking was only available in the PROBE-NET cohort. Data on demographics and lifestyle factors were compared among different clinical and histological groups using Pearson's chi-square test or Fisher's exact test as appropriate. $p < 0.05$ was statistically significant. Analyses were performed using SAS 9.4 software.

*Lectin Magnetic Bead Array (LeMBA) and Multiple Reaction Monitoring (MRM)-Mass Spectrometry*—AAL (*Aleuria aurantia* lectin), EPHA (erythroagglutinin from *Phaseolus vulgaris*), JAC (jacalin from *Artocarpus integrifolia*), and NPL (*Narcissus pseudonarcissus* lectin) (Vector Laboratories, Burlingame, CA) were individually conjugated to MyOne tosyl-activated Dynabeads® (Thermo Fisher Scientific, Massachusetts) as previously described (20). Total serum protein concentration was measured using bicinchoninic acid (BCA) assay (Thermo Fisher Scientific) (23). 50 $\mu$g of serum protein samples were spiked with chicken ovalbumin as an internal standard and were denatured as described previously (22). LeMBA pull-downs were performed using Bravo liquid handler (Agilent Technologies, California). The Bravo protocol and device files are available in supplemental method. The pull-down proteins were digested into peptides by on-bead enzymatic treatment with protease trypsin at 37 °C for overnight (Modified sequencing grade trypsin from Promega, Madison, WI), vacuum-dried, and stored at −80 °C. The dried peptide samples were reconstituted in 0.1% formic acid, and then a mixture of ten synthetic stable isotope standard (SIS) labeled peptide was spiked-into the sample. The concentration of each of these ten SIS peptides was optimized so that their response fell within ±10-fold of natural peptide. The sequence of SIS labeled peptide along with the protein candidate they represent is as follows: VTSIQDWVQ**K** (Haptoglobin or P00738 or HP), AVEVLP**K** (Gelsolin or P06396 or GSN), NLAVSQ**V**-VHK ($\alpha$-1-antichymotrypsin or P01011 or SERPINA3), LPPNVVE**E**-SAR ($\alpha$-2-macroglobulin or P01023 or A2M), LSPIYNL**V**PVK (Complement component C9 or P02748 or C9), SPAFTD**L**HLR (Apolipoprotein B-100 or P04114 or APOB), GSFEFPVGDA**V**SK

[1] The abbreviations used are: EAC, esophageal adenocarcinoma; % CV, % co-efficient of variation; AAL, *Aleuria aurantia* lectin; AUROC, area under receiver operating characteristics curve; BE, Barrett's esophagus; BE-HGD, Barrett's esophagus with high-grade dysplasia; BE-ID, Barrett's esophagus which is indefinite for dysplasia; BE-LGD, Barrett's esophagus with low-grade dysplasia; BMI, body mass index; C1QB, complement C1q subcomponent subunit B; C2, complement C2; C3, complement C3; C4B, complement C4-B; C4BPA, C4b-binding protein alpha chain; C4BPB, C4b-binding protein beta chain; C9, complement component C9; CFB, complement factor B; CFI, complement factor I; CI, confidence interval; CP, ceruloplasmin; EPHA, erythroagglutinin from *Phaseolus vulgaris*; FFPE, formalin-fixed, paraffin-embedded; GERD, gastroesophageal reflux disease; GSN, gelsolin; JAC, Jacalin from *Artocarpus integrifolia*; LeMBA, lectin magnetic bead array; MRM-MS, multiple reaction monitoring-mass spectrometry; NPL, *Narcissus pseudonarcissus* lectin; NSE, non-specialized epithelium; OR, odds ratio; PGLYRP2, N-acetylmuramoyl-L-alanine amidase; PON1, serum paraoxonase/arylesterase 1; PON3, serum paraoxonase/lactonase 3; RBP4, retinol-binding protein 4; SERPINA4, kallistatin; SIS, stable isotope-labeled internal standard.

(Complement C4-B or P0C0L5 or C4B), LTPLYE**L**VK (Complement component C7 or P10643 or C7), VASMASE**K** (Ovalbumin or P01012 or SERPINB14), and ISQAVHAAHAEINEAG**R** (Ovalbumin or P01012 or SERPINB14). The SIS peptides ISQAVHAAHAEINEAG**R** and VASMASE**K** that represent ovalbumin internal standard generated poor quality data and were not considered for data analysis. As an inter-batch quality control (QC) for processing of the PROBE-NET cohort, serum sample from a healthy volunteer was used in 4 wells of each 96-well plate. This was not required for the Ochsner cohort because of the smaller sample size per lectin.

The MRM-MS assay was set up on Agilent Technologies 6490 triple quadrupole mass spectrometer coupled with 1290 standard-flow infinity Ultra-High Performance Liquid Chromatography (UHPLC) fitted with a standard-flow ESI (Jet Stream) source. Agilent 6490 triple quadrupole mass spectrometer was operated in positive ion mode using Agilent's MassHunter Work station software (version B.06.00). The MRM acquisition parameters were 150 V high pressure RF, 60 V low pressure RF, 4000 V capillary voltage, 300 V nozzle voltage, 11 L/min sheath gas flow at a temperature of 250 °C, 15 L/min drying gas flow at a temperature of 150 °C, 30 psi nebulizer gas flow, unit resolution (0.7 Da full width at half maximum in the first quadrupole (Q1) and the third quadrupole (Q3)), and 200 V delta EMV (+). A blank was injected after every sample during LC-MS/MS data acquisition. Periodically, human serum albumin synthetic peptide standard mix (#G2455–85001, Agilent Technologies) was injected as quality control to monitor the mass spectrometer performance over a long period of time.

The precursor ions and transitions were selected based on discovery experiments described earlier (22) and through *in silico* analysis using Skyline (24). The list of transitions containing precursor *m/z*, product *m/z*, collision energy, retention time, and delta retention time for PROBE-NET and Ochsner cohorts are reported in supplemental Table S1. The UHPLC system consisted of a reverse phase AdvanceBio Peptide Mapping column (150 × 2.1 mm i.d., 2.7 $\mu$m, #653750-902, Agilent Technologies) with a 5 mm long guard column. Mobile phase A was 0.1% formic acid, and mobile phase B was 100% acetonitrile containing 0.1% formic acid. The UHPLC system was operated at a flow rate of 0.4 ml/min at 60 °C. The gradient used for peptide separation was as follows: 3% B at 0 min; 35% B at 40 min; 95% B at 40.5 min; 95% B at 44.5 min; 3% B at 45 min; followed by conditioning of column for 4 min at 3% B.

Skyline (version 3.6.0.10493) was used for inspecting and processing MRM data (24). Poor quality peptides that showed nonreproducible results (% coefficient of variation (% CV) < 30%) for repeatedly processed QC serum sample were removed. Peak picking algorithm in Skyline was optimized for improved peak integration. Peak area for each peptide (sum of all transitions) was exported. The quality of acquired data sets (supplemental Table S2*A* for PROBE-NET cohort and supplemental Table S2*B* for Ochsner cohort) were evaluated by % CV of spiked-in stable isotope-labeled internal standard (SIS) peptides and peptides derived from the spiked-in internal standard protein chicken ovalbumin. The data were normalized using the median intensity of 8 SIS peptides. Peptide intensities were converted into protein intensity with Pearson correlation coefficient cut-off set at 0.6 (22). As recently highlighted by others (25), this step serves as quality control for peptide level measurements resulting in a robust protein level quantitative dataset for down-stream statistical analysis. The normalized protein intensities were transformed using the natural logarithm and z-scores were calculated to ensure that there were no scale differences between the two datasets prior to downstream statistical analysis.

*Statistical Analysis*—JMP Pro 13.2 (SAS Institute, Inc., Cary, NC, USA) was used for univariate and multivariate biomarker statistical analyses. Univariate logistic regressions were conducted on EAC *versus* NSE outcome, EAC *versus* BE outcome, and the surveillance outcome (BE-HGD or EAC *versus* BE or BE-ID or BE-LGD), against each of the glycoprotein_lectin biomarker candidates. Odds ratios (ORs) with 95% Wald confidence intervals (95% CIs), area under receiver operating characteristic curves (AUROCs) and Likelihood Ratio *p* values were calculated for both PROBE-NET and Ochsner cohorts. Recursive partitioning (also known as Classification and Regression Trees, CART) was used to identify a multivariate panel of markers that would discriminate between surveillance outcomes (BE-HGD+EAC *versus* BE+BE-ID+BE-LGD). The PROBE-NET data set was used as the training set to develop predictive models, and the Ochsner dataset was used as the validation set. The set of 217 markers that were available in both PROBE-NET and Ochsner datasets was used in the training set, as well as baseline characteristics, including age, gender, and BMI. To avoid overfitting, models were limited to 6, 8, and 10 biomarker candidates. We also implemented a permutation test to assess the performance of the model fitting approach in the training set. Specifically, we fit our model to the training set using 5-fold cross-validation and noted the AUROC. Then we randomly permuted the response variable (EAC status) and refit a new model in the same way and recorded the permutated AUROC. We repeated this procedure for 10,000 permutations to build a distribution of AUROCs against which we evaluated our observed AUROC. Thus, models that tend to overfit the data lead to large AUROCs in the permutations relative to the observed AUROC thus allowing us to tune our models to avoid overfitting. The prediction formulas derived from PROBE-NET were then applied to the Ochsner dataset to determine sensitivity and specificity in the validation set.

*Expression and Purification of Recombinant C9*—Recombinant C9 protein was produced by transient expression in HEK 293 Expi cells for 4 days. HEK cells were split to a density of 1.2 to 1.5 × $10^6$ cells/ml the day prior to transfection. Transfections mixtures were setup with a 1:4 ratio (w/w) of huC9-pSectag2a:PEI diluted in PBS to 1/10 of the volume of HEK cells containing 1 $\mu$g human IgK-C9 DNA/ml of cells. The native C9 signal peptide was replaced by an IgK secretion sequence for effective secretion into expression media. Following 4 days expression, the supernatant was obtained by centrifugation at 2000 × rcf and the conductivity of the media was adjusted with 10 mM Na phosphate pH 7.4, 20 mM NaCl for binding to a Hi Trap (GE Healthcare, Illinois) DEAE Sepharose column. All chromatography steps were eluted on ÄKTA FPLC. Protein was eluted on a linear gradient from (10 mM Na phosphate pH 7.4, 45 mM NaCl; to 10 mM Na phosphate pH 7.4, 500 mM NaCl) over 20 column volumes. Fractions containing C9 were pooled for type I CHT chromatography using a 5 ml Bio-Rad pre-packed column and bound at 10 mM Na Phosphate pH 7.0, 100 mM NaCl. CHT column was eluted with a phosphate gradient (constant pH 8.1) from 45–350 mM over 6 column volumes. The C9 containing fractions from CHT were concentrated using a 30 kDa MWCO concentrator (Amicon) and polished by size exclusion on a superdex 200 16/60 (GE Healthcare) eluted in 10 mM HEPES pH 7.2, 200 mM NaCl. Following purification, recombinant C9 was concentrated to 1.0 mg/ml and aliquots were frozen in liquid nitrogen.

*Immunohistochemical Analysis*—Immunohistochemical staining was performed on 42 tissue sections collected from 34 patients using anti-C9 primary antibody (Sigma Aldrich #SAB4503059). Sections were dewaxed in xylene, rehydrated in descending concentrations of ethanol to water and endogenous peroxidase quenched with 2% $H_2O_2$ in Tris-buffer saline for 10 min. Antigen retrieval was carried out in a Biocare Medical Decloaking Chamber using Dako Target Retrieval Solution pH 9 for 15 min at 105 °C. To eliminate nonspecific background, sections were treated at room temperature with Innovex Fc Receptor Blocker for 30 min then Biocare Medical Background Sniper for 15 min. Anti-C9 primary antibody diluted 1:300 in Background Sniper was applied for 60 min at room temperature. To

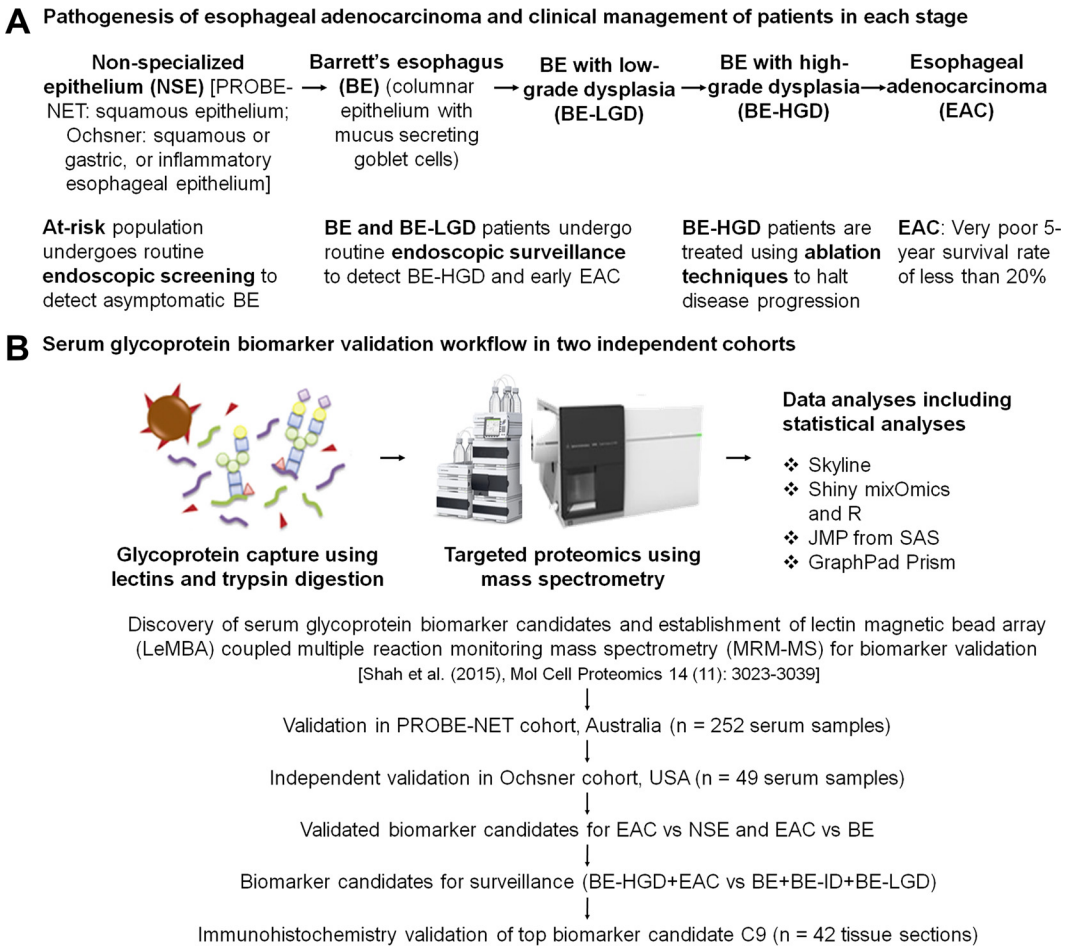



FIG. 1. *A*, Pathogenesis of esophageal adenocarcinoma (EAC) and clinical management of patients during each stage. In response to exposure to gastric and bile acid, non-specialized esophageal epithelium (NSE) converts to Barrett's esophagus (BE), and may progress through low grade dysplasia (LGD) and high grade dysplasia (HGD) stages to EAC. Patients at high risk of BE undergo endoscopic screening to detect asymptomatic metaplastic BE condition. Patients with BE or BE-LGD undergo endoscopy-biopsy surveillance to detect BE-HGD for endoscopic treatments. *B*, Workflow of the study. A total of 301 serum samples from two different patient cohorts were subjected to lectin magnetic bead array-coupled multiple reaction monitoring mass spectrometry (LeMBA-MRM-MS). Biomarker candidates for EAC and surveillance were identified by statistical analysis. Tissue expression of top glycoprotein biomarker candidate complement C9 was evaluated by immunohistochemistry.

determine antibody specificity, anti-C9 antibody was inactivated by incubation with 50 $\mu$g/ml C9 recombinant protein for 60 min at room temperature before application to the slide (supplemental Fig. S1). The slides were then incubated with Universal Polymer HRP for 30 min. A Biocare Medical MACH1 HRP kit was used for detection according to manufacturer's recommendations; Betazoid DAB, MACH1 kit, was used to visualize antibody signals. Sections were washed in water then hematoxylin counterstained, dehydrated, cleared in xylene and coverslipped using a Leica XL automated stainer and CV5030 coverslipper.

Staining intensities in squamous mucosa, columnar epithelium without intestinal metaplasia, Barrett's esophagus mucosa (with intestinal metaplasia), dysplasia (low and high grade), EAC and inflammatory infiltrate were scored by a specialist gastrointestinal pathologist. Each component, if present in the tissue, was scored separately using a 4-grade assessment of intensity (0 no staining, 1+ weak staining, 2+ moderate staining, 3+ strong staining). Where the staining was non-uniform in a component, the maximum intensity of staining was used for the score, providing at least 10% of the cells of that component stained to this intensity. Because of limited numbers, for statistical analysis, we combined dysplasia with EAC group; similarly, those with 3+ staining in Barrett's mucosa were combined with 2+ staining group. The relationship between histological features and the staining intensity was evaluated using Fisher exact test.

## RESULTS

Workflow of this study is depicted in Fig. 1*B*. We took a sequential approach to discover and validate serum diagnostic glycoprotein biomarker candidates. In previous study (22), we discovered candidates that differentiate between EAC from NSE and BE groups. In this study, we first validate the biomarker candidates for EAC in two independent cohorts. As BE-HGD is expected to be an intermediate phenotype progressing toward EAC, next, we evaluated ability of the biomarker candidates to aid current surveillance program *i.e.* distinguish between patients who require follow-up endos-

TABLE I
*Baseline demographics, histologic and clinical characteristics of the study cohorts (n = 291)*

| Characteristics | All[a] | Histology diagnosis | | | | | $P^b$ |
|---|---|---|---|---|---|---|---|
| | N (%) | NSE N (%) | BE N (%) | BE-LGD N (%) | BE-HGD N (%) | EAC N (%) | |
| **Ochsner Cohort (N)** | 49 | 14 | 13 | 5 | 7 | 10 | |
| Age at baseline | | | | | | | **<0.001** |
| <60 | 11 (22) | 4 (29) | 3 (23) | 2 (40) | 1 (14) | 1 (10) | |
| 60–69 | 22 (45) | 5 (36) | 4 (31) | 3 (60) | 3 (43) | 7 (70) | |
| ≥70 | 16 (33) | 5 (36) | 6 (46) | 0 | 3 (43) | 2 (20) | |
| Sex | | | | | | | 0.83 |
| Female | 11 (22) | 3 (21) | 2 (15) | 2 (40) | 2 (29) | 2 (20) | |
| Male | 38 (78) | 11 (79) | 11 (85) | 3 (60) | 5 (71) | 8 (80) | |
| Ethnicity | | | | | | | 0.57 |
| Caucasian | 45 (92) | 12 (86) | 12 (92) | 5 (100) | 6 (86) | 10 (100) | |
| Hispanic/Latino | 1 (2) | 0 | 1 (8) | 0 | 0 | 0 | |
| African-American | 3 (6) | 2 (14) | 0 | 0 | 1 (14) | 0 | |
| BMI at baseline | | | | | | | 0.19 |
| <25 | 12 (24) | 2 (14) | 7 (54) | 0 | 1 (14) | 2 (20) | |
| 25–29.9 | 15 (31) | 3 (21) | 2 (15) | 2 (40) | 3 (43) | 5 (50) | |
| ≥30 | 22 (45) | 9 (64) | 4 (31) | 3 (60) | 3 (43) | 3 (30) | |
| **PROBE-NET (N)** | 242 | 43 | 65 | 39 | 35 | 60 | |
| Age at baseline | | | | | | | **0.04** |
| <60 | 95 (39) | 24 (57) | 29 (45) | 11 (28) | 11 (31) | 20 (33) | |
| 60–69 | 74 (31) | 14 (33) | 16 (25) | 17 (43) | 11 (31) | 16 (27) | |
| ≥70 | 73 (30) | 5 (10) | 20 (31) | 11 (30) | 13 (37) | 24 (40) | |
| Sex | | | | | | | **<0.001** |
| Female | 58 (24) | 28 (64) | 10 (15) | 7 (18) | 4 (11) | 9 (15) | |
| Male | 184 (76) | 15 (36) | 55 (85) | 32 (82) | 31 (89) | 51 (85) | |
| Education | | | | | | | 0.57 |
| High School | 87 (40) | 13 (32) | 25 (40) | 17 (49) | 11 (37) | 21 (40) | |
| College equivalent | 91 (41) | 16 (39) | 26 (42) | 12 (34) | 12 (40) | 25 (48) | |
| University | 42 (19) | 12 (29) | 11 (18) | 6 (17) | 7 (23) | 6 (12) | |
| BMI at baseline | | | | | | | 0.41 |
| <25 | 45 (22) | 13 (32) | 14 (24) | 3 (10) | 6 (20) | 9 (19) | |
| 25–29.9 | 94 (45) | 13 (32) | 26 (44) | 16 (51) | 13 (43) | 26 (54) | |
| ≥30 | 69 (33) | 14 (35) | 19 (32) | 12 (39) | 11 (37) | 13 (27) | |
| Alcohol drinking | | | | | | | 0.45 |
| Never | 32 (14) | 9 (22) | 8 (13) | 5 (14) | 2 (6) | 8 (15) | |
| Ever | 190 (86) | 32 (78) | 54 (87) | 30 (86) | 30 (94) | 44 (85) | |
| Tobacco smoking | | | | | | | 0.18 |
| Never | 73 (33) | 15 (37) | 26 (42) | 12 (34) | 6 (19) | 14 (27) | |
| Ever | 149 (67) | 26 (63) | 36 (58) | 23 (66) | 26 (81) | 38 (73) | |

[a]Counts in subcategories may not add up to the total number because of missing data.
[b]$\chi^2$ test or Fisher exact test as appropriate. Bold values indicate significance (<0.05).

copy (BE-HGD and EAC) and those who do not (BE, BE-ID, and BE-LGD).

Table I details the baseline demographic and clinical characteristics of the Australian (PROBE-NET) and Ochsner cohorts. Serum glycoprotein biomarker candidates were measured using our previously reported lectin magnetic bead array (20, 21)-coupled multiple reaction monitoring (MRM) mass spectrometry assay (26). In this method, lectin binding is used to isolate glycoproteins with glycan structures. Based on our previous work for BE/EAC (22, 27), four lectins were selected for the independent validation cohorts, namely AAL, EPHA, JAC, and NPL. The targeted proteomics workflow uses a tier 3 assay with inclusion of internal standard peptides for 8 candidates. As compared with previously published MRM-MS assay that analyzed 41 glycoprotein candidates (22), ~40 more target proteins were added to the assay (supplemental Table S1). The uniqueness of target peptides was established using Skyline with human proteome as a background. The assay reproducibility was tested as described earlier (22). PROBE-NET and Ochsner cohorts were independently analyzed with block randomization design for each cohort.

Univariate analysis of the PROBE-NET cohort dataset revealed 46 and 54 biomarker candidates with $p < 0.05$ for EAC *versus* NSE and EAC *versus* BE comparisons respectively (supplemental Table S3). Data for the top 10 biomarker candidates that differentiate EAC from NSE, and EAC from BE in PROBE-NET cohort are shown in Fig. 2*A* and 2*B*, respectively. Out of these candidates, 16 candidates for EAC *versus* NSE comparison and 9 candidates for EAC *versus* BE comparison were also significantly different in the Ochsner cohort, confirming these glycoproteins as validated biomarker candidates. As illustrated in Fig. 2*C*, 8 validated biomarker candidates overlap between the two lists. These biomarker candidates are potentially most useful, being able to distinguish EAC from NSE and BE. The 8 biomarker candidates

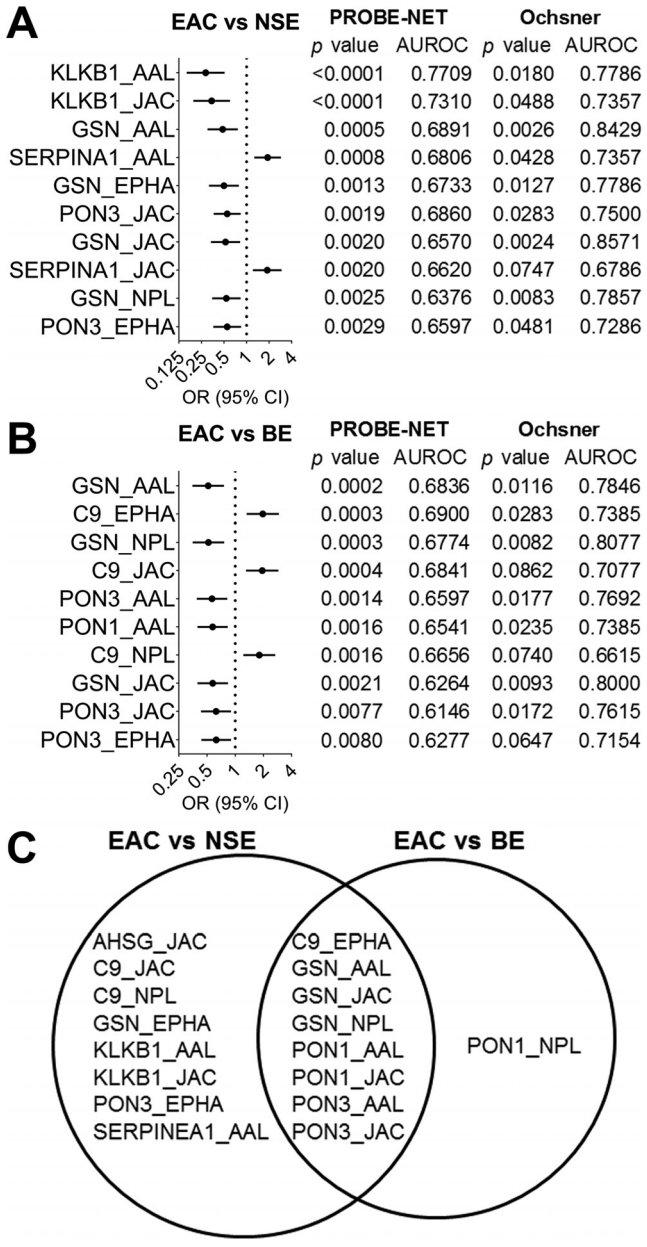



FIG. 2. **Serum glycoprotein biomarker validation in PROBE-NET and Ochsner cohorts.** Top 10 glycoprotein_lectin biomarker candidates that differentiate (*A*) esophageal adenocarcinoma (EAC) from non-specialized epithelium (NSE) and (*B*) EAC from Barrett's esophagus (BE) in PROBE-NET and Ochsner cohorts. *C*, Overlap of validated biomarker candidates ($p < 0.05$). Biomarker candidates are shown by gene name and lectin affinity, and ordered by *p* value. Horizontal bar indicates odds ratio (OR) with 95% Wald confidence intervals. Likelihood Ratio *p* values were calculated to test statistical significance in the univariate logistic regressions. Area under receiver operating characteristic curve (AUROC) indicates diagnostic ability of individual biomarker candidate. AAL = *Aleuria aurantia* lectin; C9 = complement C9; EPHA = erythroagglutinin *Phaseolus vulgaris*; GSN = gelsolin; JAC = jacalin from *Artocarpus integrifolia*; KLKB1 = plasma kallikrein; NPL = *Narcissus pseudonarcissus* lectin; PON1 = serum paraoxonase/arylesterase 1; PON3 = serum paraoxonase/lactonase 3; SERPINA1 = alpha-1-antitrypsin.

comprised of 4 glycoproteins, namely, complement C9 (C9), gelsolin (GSN), serum paraoxonase/arylesterase 1 (PON1), and serum paraoxonase/lactonase 3 (PON3), that were different between the disease groups in one or more lectin pull-down samples.

Next we considered BMI as a potential confounding factor for EAC biomarker validation. Correlation analysis between BMI and biomarker candidate levels in all PROBE-NET samples revealed no substantial correlation ( $r < 0.6$) (supplemental Table S4), although three proteins namely complement C3 (C3), C4b-binding protein alpha chain (C4BPA), and complement factor I (CFI) with multiple lectin pull-downs showed positive correlations with BMI in NSE patients only ( $r$ 0.6234–0.7069). Importantly, none of the top 10 biomarker candidates showed strong correlation with BMI. This suggests that top glycoprotein biomarker candidates measured using our workflow are not affected by a common confounding covariate, BMI.

*Biomarkers for BE Surveillance*—Having confirmed univariate biomarkers for detection of EAC from NSE and BE in independent cohorts, we next evaluated the ability of serum glycoproteins to be used as a surveillance tool for BE and BE-LGD patients, *i.e.* to distinguish between patients who require treatment (BE-HGD and EAC) and those who do not (BE, BE-ID, and BE-LGD). Eight biomarker candidates that showed AUROC > 0.6 in a BE surveillance setting for both PROBE-NET and Ochsner cohort (Fig. 3, supplemental Table S3), comprised of GSN measured in AAL, NPL, and JAC pull-downs, C9 measured in JAC and EPHA pull-downs, AAL-binding PON1, AAL-binding PON3, as well as EPHA-binding Complement factor B (CFB).

Next we sought to generate a multimarker panel for BE surveillance, using the PROBE-NET cohort for modeling and Ochsner cohort for model validation. The minimal panel of six biomarker candidates showed 0.83 AUROC, 83% sensitivity and 69% specificity for the PROBE-NET cohort, and a moderate sensitivity of 61% and specificity of 65% for the Ochsner cohort (Table II). Addition of 4 more biomarker candidates to the panel increased the AUROC to 0.93, and improved the specificity and sensitivity measures for PROBE-NET, as well as the specificity for Ochsner cohort (Table II). Further evaluation of this multimarker panel of 10 biomarker candidates in a subset of Ochsner cohort comprising of only treatment naïve patients ($n = 20$) showed sensitivity of 71% and specificity of 77% to distinguish between the two groups. The performance of multimarker panel in a subset of treatment naïve Ochsner cohort is similar, if not slightly improved, as compared with results obtained using all available Ochsner cohort samples.

*Complement C9 Expression in BE and EAC Tissue*—In agreement with our previous finding of complement pathway dysregulation in EAC pathogenesis (22), 5 of the 10 glycoprotein biomarker candidates in the final surveillance biomarker panel (Table II) belonged to the complement pathway. As a

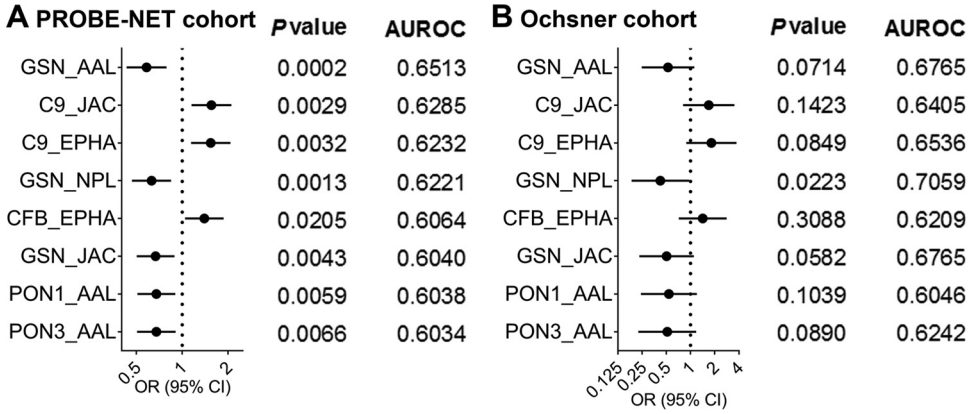



FIG. 3. **Serum glycoprotein biomarker candidates for BE surveillance in PROBE-NET and Ochsner cohorts.** Candidates that distinguish patients who require confirmatory endoscopy (BE-HGD+EAC) from those who do not (BE+BE-ID+BE-LGD) with an area under receiver operating characteristic curve (AUROC) value of $> 0.6$ in both PROBE-NET and Ochsner cohorts are shown. Odds ratio (OR) with 95% Wald confidence intervals, likelihood ratio *p* values and AUROC values for (*A*) PROBE-NET cohort and (*B*) Ochsner cohort. AAL = *Aleuria aurantia* lectin; AUROC = area under receiver operating characteristic curve; BE = Barrett's esophagus; BE-HGD = BE with high-grade dysplasia; BE-LGD = BE with low-grade dysplasia; C9 = complement C9; CFB = complement factor B; CI = confidence interval; EAC = esophageal adenocarcinoma; EPHA = erythroagglutinin *Phaseolus vulgaris*; GSN = gelsolin; JAC = jacalin from *Artocarpus integrifolia*; NPL = *Narcissus pseudonarcissus* lectin; PON1 = serum paraoxonase/arylesterase 1; PON3 = serum paraoxonase/lactonase 3.

TABLE II
*Glycoprotein biomarker panel performance for BE surveillance (BE-HGD+EAC vs BE+BE-ID+BE-LGD) in PROBE-NET and Ochsner cohorts*

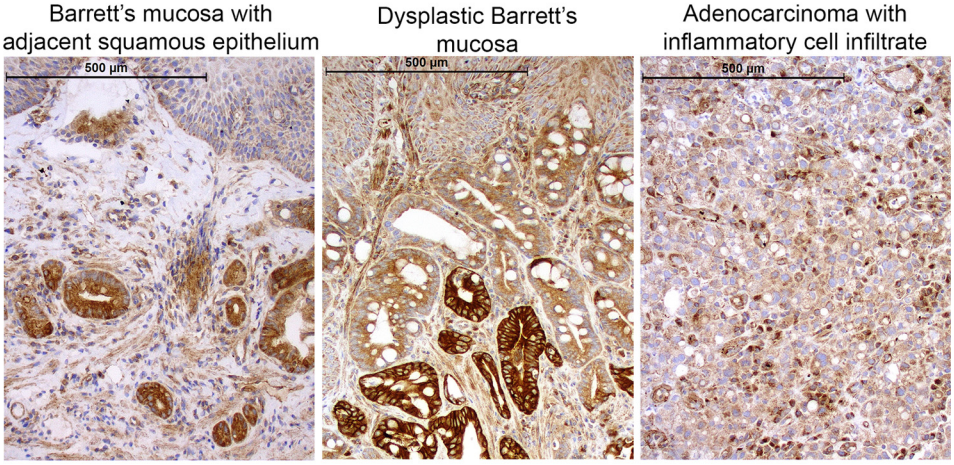
| Multimarker panel | PROBE-NET cohort | | | Ochsner cohort | |
|---|---|---|---|---|---|
| | AUROC | Sensitivity | Specificity | Sensitivity | Specificity |
| Panel 1: C9_EPHA + PON1_AAL + SERPINA4_EPHA + RBP4_AAL + C1QB_JAC + PGLYRP2_NPL | 0.83 | 83% | 69% | 61% | 65% |
| Panel 2: Panel 1 + CP_NPL + C4BPB_JAC + C2_EPHA + C4B_AAL | 0.93 | 86% | 85% | 61% | 71% |

AAL = *Aleuria aurantia* lectin; AUROC = area under receiver operating characteristic curve; BE = Barrett's esophagus; BE-HGD = BE with high-grade dysplasia; BE-LGD = BE with low-grade dysplasia; C1QB = Complement C1q subcomponent subunit B; C2 = Complement C2; C4B = Complement C4-B; C4BPB = C4b-binding protein beta chain; C9 = complement component C9; CP = Ceruloplasmin; EAC = esophageal adenocarcinoma; EPHA = erythroagglutinin *Phaseolus vulgaris*; GSN = gelsolin; JAC = jacalin from *Artocarpus integrifolia*; NPL = *Narcissus pseudonarcissus* lectin; PGLYRP2 = N-acetylmuramoyl-L-alanine amidase; PON1 = serum paraoxonase/arylesterase 1; PON3 = serum paraoxonase/lactonase 3; RBP4 = Retinol-binding protein 4; SERPINA4 = Kallistatin.

first step to evaluate alterations of the complement pathway in EAC at a tissue level, we optimized immunohistochemistry staining for the top candidate, complement C9. Staining specificity of the method was confirmed by neutralization of the antibody with recombinant C9 protein prior to staining. In supplemental Fig. S1, the disappearance of dark brown tissue staining after pre-incubation of anti-C9 antibody with recombinant C9 indicates the antibody specificity. We then evaluated expression of C9 in esophageal tissue sections from a subset of the Ochsner cohort. As shown in Fig. 4*A*, the adjacent normal squamous epithelium tissue showed no or minimal C9 staining as compared with light to dark brown staining that was detected in BE and EAC. Dysplastic BE showed particularly strong staining in the plasma membrane and/or cytoplasm (Fig. 4*A*). Strong staining in immune infiltrates served as an expected positive control. Quantitation of staining intensity score against the tissue phenotype (Fig. 4*B*) showed statistically significant association between C9 expression levels and histology groups of squamous epithelium, columnar epithelium, Barrett's mucosa, dysplasia/EAC ($p <$ 0.001). Specifically, 87% of the squamous epithelium showed no C9 staining as compared with all the BE, dysplastic, or EAC tissue that showed some degree of C9 presence. In addition, 87% of BE and 67% dysplastic/EAC observations were positive for C9 with scoring intensity $\geq 2$. In contrast, none of the squamous epithelia were stained strongly (intensity $\geq 2$) for C9. Columnar epithelium with no evidence of intestinal metaplasia showed intermediate C9 staining. These novel data demonstrating elevated C9 protein in Barrett's and EAC cells provides some biological basis, to the use of elevated serum C9 as a biomarker candidate. The exact correlation between tissue and serum C9 remains to be established.

*Serum Complement C9 in Progressor Samples*—As an additional evaluation, we examined C9 lectin pull-down levels in samples from PROBE-NET participants who had progressed from BE to BE-LGD ($n = 4$), BE-LGD to BE-HGD ($n = 3$), BE to BE-HGD ($n = 1$), BE to EAC ($n = 1$), and NSE to gastric type mucosa ($n = 1$) phenotype during follow-up. The progression time varied from 110 days up to 5 years. Significant elevation

**A** Complement C9 staining under different conditions




**B** Complement C9 staining score

| Score | Histology groups | | | | | *P* |
|---|---|---|---|---|---|---|
| | Squamous epithelium N (%) | Columnar epithelium (no IM) N (%) | Barrett's mucosa N (%) | Dysplasia or EAC N (%) | Total N (%) | |
| **Total (N)** | 15 | 12 | 23 | 6 | 56 | |
| | | | | | | **<0.001** |
| 0 | 13 (87) | 2 (17) | 0 | 0 | 15 (27) | |
| 1+ | 2 (13) | 3 (25) | 3 (13) | 2 (33) | 10 (18) | |
| ≥2+ | 0 | 7 (58) | 20 (87) | 4 (67) | 31 (55) | |

FIG. 4. **Immunohistochemical analysis of complement C9 (C9) in lower esophageal biopsies.** *A*, A panel of representative images of C9 staining in squamous epithelium, Barrett's mucosa, dysplastic epithelium, and esophageal adenocarcinoma conditions. The left image shows 2+ staining in the non-dysplastic Barrett's epithelium. The middle image shows 1+ and 3+ staining in the BE. The right image shows 2+ staining in cancer cells and 3+ staining in inflammatory cells. *B*, Summary of C9 staining intensity according to histology groups. *p* value from Fisher exact test. EAC = esophageal adenocarcinoma; IM = intestinal metaplasia.

of C9_EPHA and C9_NPL was observed following progression in this small patient cohort (Fig. 5).

## DISCUSSION

This study progresses our previous serum glycoprotein research (22, 27) toward developing cost-effective EAC surveillance by validating serum biomarker candidates in two independent cohorts. Over the years, several studies have been carried out to identify circulatory biomarker candidates to diagnose BE-dysplasia-EAC disease spectrum (19, 28). These studies have explored genetic alterations in cell free circulating DNA (29), serum miRNA changes (30), circulating tumor cells (31), glycan profile alteration in serum (32, 33), circulatory autoantibodies (against cancer antigens) (34), volatile organic compounds found in breath analysis (35), metabolic changes in urine (36), and a panel of serum proteins (37) as promising diagnostic biomarker candidates for BE and/or EAC. However, none of these biomarker candidates have progressed from bench to bedside, likely because of the lack of subsequent validation studies in large independent cohort of patients. Here, we have addressed this gap by validating eight serum glycoprotein biomarkers for EAC in two independent patient cohorts including dysplastic samples.

In addition to demonstrating the robustness of our mass spectrometry-based glycoprotein-centric proteomics workflow for biomarker validation, the current study confirmed our previous finding of complement activation in EAC (22). The complement system consists of a cascade of circulating proteases that are locally activated leading to the formation of the membrane attack complex on the immunogen and recruitment of phagocytes. Complement components are predominantly expressed and secreted into the plasma by the liver but are also found to be expressed in other tissues (38), including tumor cells (39). Although complement components are primarily involved in mediating innate immune response, recent studies have revealed an apparently paradoxical tumor-promoting role of the complement system (39, 40). Complement C3 has been reported to play an autocrine role in ovarian and lung cancer tumor growth (41). C5a is elevated in serum of lung cancer patients (42) and increases the invasiveness of C5aR+ tumors (43). Recently, Franc and colleagues found 15 co-occurring serum C9 proteoforms based on glycan expression profiling (44). Specifically, fucosylated C9 was reported to be elevated in the serum of lung cancer patients (45).

Two previous publications reported on complement component changes in BE and EAC. Bobryshev *et al.* reported reduced C1q expression in dendritic cells and macrophages in the epithelium of BE and EAC and suggest this to be an immune-escape mechanism (46). Song *et al.* identified com-

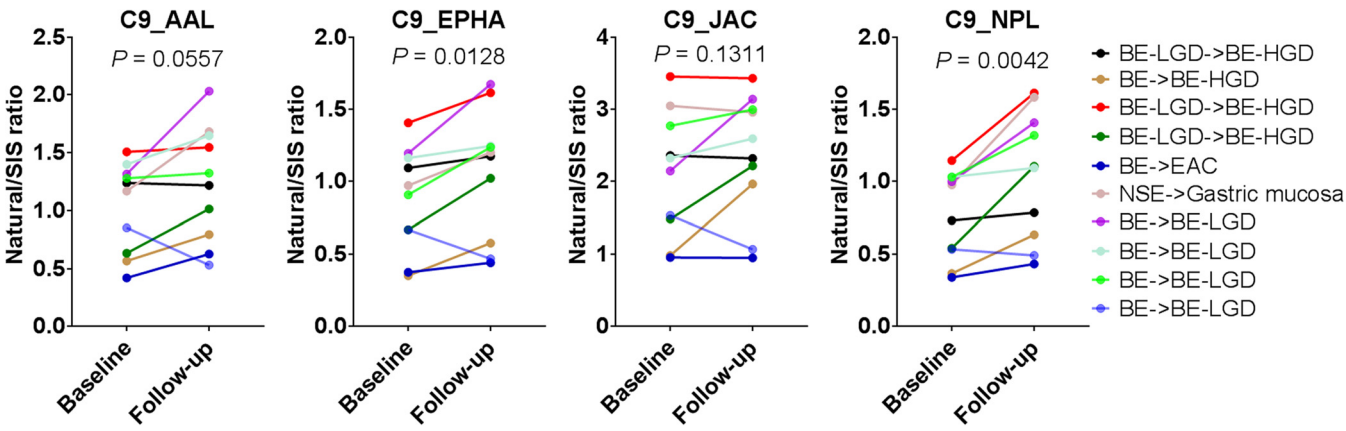



FIG. 5. **Serum complement C9 (C9) levels in specific lectin-pulldowns in PROBE-NET progressor serum samples.** C9 levels in lectin pull-downs are expressed as ratio of intensity of natural peptide LSPIYNLVPVK to stable isotope labeled spiked-in peptide response. The changes in C9 in lectin pull-down samples of participants who progressed from one stage (baseline) to another stage of the disease at follow-up. *p* value from paired *t* test. AAL = *Aleuria aurantia* lectin; BE = Barrett's esophagus; BE-HGD = BE with high-grade dysplasia; BE-LGD = BE with low-grade dysplasia; C9 = complement C9; EAC = esophageal adenocarcinoma; EPHA = erythroagglutinin *Phaseolus vulgaris*; JAC = jacalin from *Artocarpus integrifolia*; NPL = *Narcissus pseudonarcissus* lectin; SIS = stable isotope standard.

plement pathway proteins, complement C3 and complement C1r subcomponent to be increased in serum collected from BE-HGD and EAC patients respectively as compared with disease free individuals (47). The current study is the first to report C9 protein expression in BE and EAC cells, as demonstrated on tissue sections. In addition to validating our previous finding of elevated circulating C9 in LeMBA pull-downs in EAC, we also evaluated the utility of serum C9_JAC and C9_EPHA for BE surveillance in independent cohorts. Together, our results suggest a pathological change in the secretion, glycosylation, or expression of C9 during the progression of BE to BE-HGD/EAC. Modulation of C9 expression is unlikely to be the mechanism, because we detected strong C9 staining in both BE and BE-HGD tissue. Interestingly, bile (deoxycholic acid) treatment was reported to alter glycosylation and Golgi structure in esophageal epithelial and BE cells, resulting in impaired protein secretion via the classical pathway (48). Hence, C9 may be sequentially regulated by expression level and glycosylation/secretion during BE-EAC progression. Further studies are required to determine the precise molecular changes in serum C9 with respect to normal-BE-dysplasia-EAC spectrum.

The strengths of this study include the validation of biomarker candidates in independent cohorts, evaluation of biomarker panel for BE surveillance, and the use of immunohistochemistry to determine the cellular origin of the top serum glycoprotein biomarker candidate. There are several limitations to our study. Although our glycoprotein biomarker pipeline allows high-throughput glycoprotein biomarker discovery and validation, changes in the glycosylation site occupancy or glycan structural changes for the biomarker candidate are not determined. Furthermore, the functional consequences, if any, of these glycosylation changes in cancer progression also remain to be evaluated. The tier 3 mass spectrometry assay employed in our work may suit early stage validation studies, but the final biomarker panel should be evaluated with tier 1 MRM-MS assay (or similar other assay platform) that allows absolute quantification. The small number of progressor samples available is a limitation, not only for the current study, but for evaluation of BE/EAC biomarkers in general. International collaborations and longitudinal cohort sampling are therefore required to advance the search for EAC biomarkers. Finally, the tissue expression of C9 and other candidates also need to be confirmed in a larger number of samples.

In summary, we have validated several serum glycoprotein biomarkers that warrant further clinical testing in large independent cohorts including longitudinal patient samples. Further evaluation and development of these biomarkers to a blood test may aid the current endoscopy-biopsy surveillance program of BE patients for early detection of dysplastic progression.

*Acknowledgments*—We thank Dr Thomas Hennessy (Agilent Technologies) and Mr Elliot McElroy (Agilent Technologies) for technical assistance, and Mr Cris Molina (Ochsner Health System) for clinical research management support.

## DATA AVAILABILITY

All the raw data and analysis results have been deposited to the ProteomeXchange Consortium (49) via PASSEL (50) with the dataset identifier PASS01166 (Server name: ftp.peptideatlas.org; Full URL: ftp://PASS01166:KZ8528wq@ftp.peptideatlas.org/) and PanoramaWeb (https://panoramaweb.org/rng6uH.url and https://panoramaweb.org/q7vWBk.url) (51, 52) (uploaded to PanoramaWeb using Skyline version 4.1.0.11714).

* Grant support: This work was supported by The University of Queensland-Ochsner Seed Fund for Collaborative Research Grant

2014, The University of Queensland Faculty of Medicine and Biomedical Science Cancer Bequest Grant 2014, and internal funding support from The University of Queensland Diamantina Institute 2015. PROBE-NET was supported by a National Health and Medical Research Council (NHMRC) Centre of Research Excellence grant (APP1040947). KALC is a recipient of NHMRC Career Development Fellowship (APP1087415). DCW is supported by an NHMRC Research Fellowship (APP1058522). The sponsors had no influence on the study design, collection, analysis, and interpretation of data.

[S] This article contains supplemental Tables. The authors have no conflicts of interest to disclose.

‡‡‡ To whom correspondence should be addressed: QIMR Berghofer Medical Research Institute, Locked Bag 2000, Royal Brisbane Hospital, 300 Herston Road, Herston QLD 4029 Australia. Tel.: +617 3845 3020; E-mail: Michelle.Hill@qimrberghofer.edu.au.

Author contributions: A.K.S., M.A.D., and M.M.H. designed research; A.K.S., C.W., and B.A.S. performed research; A.K.S., G.H., and R.N. analyzed data; A.K.S., G.H., D.C.W., and M.M.H. wrote the paper; G.H., K.-A.L.C., B.A.S., and M.A.D. contributed new reagents/analytic tools; I.B. evaluated and scored immunohistochemistry slides; R.N. patient sample selection for the study; W.A.P., R.V.L., A.P.B., and D.I.W. contributed to sample collection; V.J. contributed to sample collection and selection; D.C.W. contributed to sample collection and selection.

REFERENCES

1. Siegel, R. L., Miller, K. D., and Jemal, A. (2017) Cancer Statistics, 2017. *CA Cancer J. Clin.* **67,** 7–30
2. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., and Bray, F. (2015) Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **136,** E359–E386
3. Chen, Z., Ren, Y., Du, X. L., Yang, J., Shen, Y., Li, S., Wu, Y., Lv, M., Dong, D., Li, E., Li, W., Liu, P., Yang, J., and Yi, M. (2017) Incidence and survival differences in esophageal cancer among ethnic groups in the United States. *Oncotarget* **8,** 47037–47051
4. Thrift, A. P. (2016) The epidemic of oesophageal carcinoma: Where are we now? *Cancer Epidemiol.* **41,** 88–95
5. Hur, C., Miller, M., Kong, C. Y., Dowling, E. C., Nattinger, K. J., Dunn, M., and Feuer, E. J. (2013) Trends in esophageal adenocarcinoma incidence and mortality. *Cancer* **119,** 1149–1158
6. Malhotra, G. K., Yanala, U., Ravipati, A., Follet, M., Vijayakumar, M., and Are, C. (2017) Global trends in esophageal cancer. *J. Surg. Oncol.* **115,** 564–579
7. Reid, B. J., Li, X., Galipeau, P. C., and Vaughan, T. L. (2010) Barrett's oesophagus and oesophageal adenocarcinoma: time for a new synthesis. *Nat. Rev. Cancer* **10,** 87–101
8. Anaparthy, R., and Sharma, P. (2014) Progression of Barrett oesophagus: role of endoscopic and histological predictors. *Nat. Rev. Gastroenterol. Hepatol.* **11,** 525–534
9. Rubenstein, J. H., and Shaheen, N. J. (2015) Epidemiology, diagnosis, and management of esophageal adenocarcinoma. *Gastroenterology* **149**, 302–317 e301
10. Fitzgerald, R. C., di Pietro, M., Ragunath, K., Ang, Y., Kang, J. Y., Watson, P., Trudgill, N., Patel, P., Kaye, P. V., Sanders, S., O'Donovan, M., Bird-Lieberman, E., Bhandari, P., Jankowski, J. A., Attwood, S., Parsons, S. L., Loft, D., Lagergren, J., Moayyedi, P., Lyratzopoulos, G., de Caestecker, J., and British Society of, G. (2014) British Society of Gastroenterology guidelines on the diagnosis and management of Barrett's oesophagus. *Gut* **63**, 7–42
11. Shaheen, N. J., Falk, G. W., Iyer, P. G., Gerson, L. B., and American College of, G. (2016) ACG Clinical Guideline: Diagnosis and Management of Barrett's Esophagus. *Am. J. Gastroenterol.* **111**, 30–50; quiz 51
12. Whiteman, D. C., Appleyard, M., Bahin, F. F., Bobryshev, Y. V., Bourke, M. J., Brown, I., Chung, A., Clouston, A., Dickins, E., Emery, J., Eslick, G. D., Gordon, L. G., Grimpen, F., Hebbard, G., Holliday, L., Hourigan, L. F., Kendall, B. J., Lee, E. Y., Levert-Mignon, A., Lord, R. V., Lord, S. J., Maule, D., Moss, A., Norton, I., Olver, I., Pavey, D., Raftopoulos, S., Rajendra, S., Schoeman, M., Singh, R., Sitas, F., Smithers, B. M., Taylor, A. C., Thomas, M. L., Thomson, I., To, H., von Dincklage, J., Vuletich, C., Watson, D. I., and Yusoff, I. F. (2015) Australian clinical practice guidelines for the diagnosis and management of Barrett's esophagus and early esophageal adenocarcinoma. *J. Gastroenterol. Hepatol.* **30,** 804–820
13. Desai, T. K., Krishnan, K., Samala, N., Singh, J., Cluley, J., Perla, S., and Howden, C. W. (2012) The incidence of oesophageal adenocarcinoma in non-dysplastic Barrett's oesophagus: a meta-analysis. *Gut* **61,** 970–976
14. Barbiere, J. M., and Lyratzopoulos, G. (2009) Cost-effectiveness of endoscopic screening followed by surveillance for Barrett's esophagus: a review. *Gastroenterology* **137,** 1869–1876
15. Gerson, L. B., Groeneveld, P. W., and Triadafilopoulos, G. (2004) Cost-effectiveness model of endoscopic screening and surveillance in patients with gastroesophageal reflux disease. *Clin. Gastroenterol. Hepatol.* **2,** 868–879
16. Yang, Y., Chen, H. N., Wang, R., Tang, Y. J., and Chen, X. Z. (2015) Cost-Effectiveness Analysis on Endoscopic Surveillance Among Western Patients With Barrett's Esophagus for Esophageal Adenocarcinoma Screening. *Medicine* **94,** e1105
17. Goldblum, J. R. (2003) Barrett's esophagus and Barrett's-related dysplasia. *Mod. Pathol.* **16,** 316–324
18. Vaughan, T. L., and Fitzgerald, R. C. (2015) Precision prevention of oesophageal adenocarcinoma. *Nat. Rev. Gastroenterol. Hepatol.* **12,** 243–248
19. Shah, A. K., Saunders, N. A., Barbour, A. P., and Hill, M. M. (2013) Early diagnostic biomarkers for esophageal adenocarcinoma–the current state of play. *Cancer Epidemiol. Biomarkers Prev.* **22,** 1185–1209
20. Choi, E., Loo, D., Dennis, J. W., O'Leary, C. A., and Hill, M. M. (2011) High-throughput lectin magnetic bead array-coupled tandem mass spectrometry for glycoprotein biomarker discovery. *Electrophoresis* **32,** 3564–3575
21. Loo, D., Jones, A., and Hill, M. M. (2010) Lectin magnetic bead array for biomarker discovery. *J. Proteome Res.* **9,** 5496–5500
22. Shah, A. K., Le Cao, K. A., Choi, E., Chen, D., Gautier, B., Nancarrow, D., Whiteman, D. C., Saunders, N. A., Barbour, A. P., Joshi, V., and Hill, M. M. (2015) Serum glycoprotein biomarker discovery and qualification pipeline reveals novel diagnostic biomarker candidates for esophageal adenocarcinoma. *Mol. Cell. Proteomics* **14,** 3023–3039
23. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J., and Klenk, D. C. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150,** 76–85
24. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26,** 966–968
25. Zhang, B., Pirmoradian, M., Zubarev, R., and Kall, L. (2017) Covariation of peptide abundances accurately reflects protein concentration differences. *Mol. Cell. Proteomics* **16,** 936–948
26. (2013) Method of the Year 2012. *Nat. Methods* **10,** 1
27. Shah, A. K., Le Cao, K. A., Choi, E., Chen, D., Gautier, B., Nancarrow, D., Whiteman, D. C., Baker, P. R., Clauser, K. R., Chalkley, R. J., Saunders, N. A., Barbour, A. P., Joshi, V., and Hill, M. M. (2016) Glyco-centric lectin magnetic bead array (LeMBA) - proteomics dataset of human serum samples from healthy, Barretts esophagus and esophageal adenocarcinoma individuals. *Data Brief* **7,** 1058–1062
28. Yazbeck, R., Jaenisch, S. E., and Watson, D. I. (2016) From blood to breath: New horizons for esophageal cancer biomarkers. *World J. Gastroenterol.* **22,** 10077–10083
29. Openshaw, M. R., Richards, C. J., Guttery, D. S., Shaw, J. A., and Thomas, A. L. (2017) The genetics of gastroesophageal adenocarcinoma and the use of circulating cell free DNA for disease detection and monitoring. *Expert Rev. Mol. Diagn.* **17,** 459–470
30. Chiam, K., Wang, T., Watson, D. I., Mayne, G. C., Irvine, T. S., Bright, T., Smith, L., White, I. A., Bowen, J. M., Keefe, D., Thompson, S. K., Jones, M. E., and Hussey, D. J. (2015) Circulating serum exosomal miRNAs as potential biomarkers for esophageal adenocarcinoma. *J Gastrointest. Surg.* **19,** 1208–1215
31. Gallerani, G., and Fabbri, F. (2016) Circulating Tumor Cells in the Adenocarcinoma of the Esophagus. *Int. J. Mol. Sci.* **17**
32. Gaye, M. M., Valentine, S. J., Hu, Y., Mirjankar, N., Hammoud, Z. T., Mechref, Y., Lavine, B. K., and Clemmer, D. E. (2012) Ion mobility-mass

spectrometry analysis of serum N-linked glycans from esophageal adenocarcinoma phenotypes. *J. Proteome Res.* **11,** 6102–6110
33. Mitra, I., Zhuang, Z., Zhang, Y., Yu, C. Y., Hammoud, Z. T., Tang, H., Mechref, Y., and Jacobson, S. C. (2012) N-glycan profiling by microchip electrophoresis to differentiate disease states related to esophageal adenocarcinoma. *Anal. Chem.* **84,** 3621–3627
34. Kilic, A., Schuchert, M. J., Luketich, J. D., Landreneau, R. J., Lokshin, A. E., Bigbee, W. L., and El-Hefnawy, T. (2008) Use of novel autoantibody and cancer-related protein arrays for the detection of esophageal adenocarcinoma in serum. *J. Thorac. Cardiovasc. Surg.* **136,** 199–204
35. Kumar, S., Huang, J., Abbassi-Ghadi, N., Mackenzie, H. A., Veselkov, K. A., Hoare, J. M., Lovat, L. B., Spanel, P., Smith, D., and Hanna, G. B. (2015) Mass spectrometric analysis of exhaled breath for the identification of volatile organic compound biomarkers in esophageal and gastric adenocarcinoma. *Ann. Surg.* **262,** 981–990
36. Davis, V. W., Schiller, D. E., Eurich, D., and Sawyer, M. B. (2012) Urinary metabolomic signature of esophageal cancer and Barrett's esophagus. *World J. Surg. Oncol.* **10,** 271
37. Zaidi, A. H., Gopalakrishnan, V., Kasi, P. M., Zeng, X., Malhotra, U., Balasubramanian, J., Visweswaran, S., Sun, M., Flint, M. S., Davison, J. M., Hood, B. L., Conrads, T. P., Bergman, J. J., Bigbee, W. L., and Jobe, B. A. (2014) Evaluation of a 4-protein serum biomarker panel-biglycan, annexin-A6, myeloperoxidase, and protein S100-A9 (B-AMP)-for the detection of esophageal adenocarcinoma. *Cancer* **120,** 3902–3913
38. Lubbers, R., van Essen, M. F., van Kooten, C., and Trouw, L. A. (2017) Production of complement components by cells of the immune system. *Clin. Exp. Immunol.* **188,** 183–194
39. Pio, R., Corrales, L., and Lambris, J. D. (2014) The role of complement in tumor growth. *Adv. Exp. Med. Biol.* **772,** 229–262
40. Mamidi, S., Hone, S., Teufel, C., Sellner, L., Zenz, T., and Kirschfink, M. (2015) Neutralization of membrane complement regulators improves complement-dependent effector functions of therapeutic anticancer antibodies targeting leukemic cells. *Oncoimmunology* **4,** e979688
41. Cho, M. S., Vasquez, H. G., Rupaimoole, R., Pradeep, S., Wu, S., Zand, B., Han, H. D., Rodriguez-Aguayo, C., Bottsford-Miller, J., Huang, J., Miyake, T., Choi, H. J., Dalton, H. J., Ivan, C., Baggerly, K., Lopez-Berestein, G., Sood, A. K., and Afshar-Kharghan, V. (2014) Autocrine effects of tumor-derived complement. *Cell Rep.* **6,** 1085–1095
42. Corrales, L., Ajona, D., Rafail, S., Lasarte, J. J., Riezu-Boj, J. I., Lambris, J. D., Rouzaut, A., Pajares, M. J., Montuenga, L. M., and Pio, R. (2012) Anaphylatoxin C5a creates a favorable microenvironment for lung cancer progression. *J. Immunol.* **189,** 4674–4683
43. Nitta, H., Wada, Y., Kawano, Y., Murakami, Y., Irie, A., Taniguchi, K., Kikuchi, K., Yamada, G., Suzuki, K., Honda, J., Wilson-Morifuji, M., Araki, N., Eto, M., Baba, H., and Imamura, T. (2013) Enhancement of human cancer cell motility and invasiveness by anaphylatoxin C5a via aberrantly expressed C5a receptor (CD88). *Clin. Cancer Res.* **19,** 2004–2013
44. Franc, V., Yang, Y., and Heck, A. J. (2017) proteoform profile mapping of the human serum complement component C9 revealing unexpected new features of N-, O-, and C-glycosylation. *Anal. Chem.* **89,** 3483–3491
45. Narayanasamy, A., Ahn, J. M., Sung, H. J., Kong, D. H., Ha, K. S., Lee, S. Y., and Cho, J. Y. (2011) Fucosylated glycoproteomic approach to identify a complement component 9 associated with squamous cell lung cancer (SQLC). *J. Proteomics* **74,** 2948–2958
46. Bobryshev, Y. V., Lu, J., and Lord, R. V. (2010) Expression of C1q complement component in Barrett's esophagus and esophageal adenocarcinoma. *J. Gastrointest. Surg.* **14,** 1207–1213
47. Song, E., Zhu, R., Hammoud, Z. T., and Mechref, Y. (2014) LC-MS/MS quantitation of esophagus disease blood serum glycoproteins by enrichment with hydrazide chemistry and lectin affinity chromatography. *J. Proteome Res.* **13,** 4808–4820
48. Byrne, A. M., Sharma, R., Duggan, G., Kelleher, D., and Long, A. (2012) Deoxycholic acid impairs glycosylation and fucosylation processes in esophageal epithelial cells. *Glycobiology* **22,** 638–648
49. Vizcaino, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P. A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H. J., Albar, J. P., Martinez-Bartolome, S., Apweiler, R., Omenn, G. S., Martens, L., Jones, A. R., and Hermjakob, H. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32,** 223–226
50. Farrah, T., Deutsch, E. W., Kreisberg, R., Sun, Z., Campbell, D. S., Mendoza, L., Kusebauch, U., Brusniak, M. Y., Huttenhain, R., Schiess, R., Selevsek, N., Aebersold, R., and Moritz, R. L. (2012) PASSEL: the PeptideAtlas SRMexperiment library. *Proteomics* **12,** 1170–1175
51. Sharma, V., Eckels, J., Schilling, B., Ludwig, C., Jaffe, J. D., MacCoss, M. J., and MacLean, B. (2018) Panorama public: A public repository for quantitative data sets processed in Skyline. *Mol. Cell. Proteomics* **17,** 1239–1244
52. Sharma, V., Eckels, J., Taylor, G. K., Shulman, N. J., Stergachis, A. B., Joyner, S. A., Yan, P., Whiteaker, J. R., Halusa, G. N., Schilling, B., Gibson, B. W., Colangelo, C. M., Paulovich, A. G., Carr, S. A., Jaffe, J. D., MacCoss, M. J., and MacLean, B. (2014) Panorama: a targeted proteomics knowledge base. *J. Proteome Res.* **13,** 4205–4210