Cell surface proteomics identifies glucose transporter type 1 and prion protein as candidate biomarkers for colorectal adenoma-to-carcinoma progression

Meike de Wit,1,2 Connie R Jimenez,2 Beatriz Carvalho,1 Jeroen A M Belien,1 Pien M Delis-van Diemen,1 Sandra Mongera,1 Sander R Piersma,2 Maindad Vikas,3 Sanjay Navani,3 Fredrik Pontén,4 Gerrit A Meijer,1 Remond J A Fijneman 1

ABSTRACT

Background and objective Early detection of colon adenomas at high risk of progression and early-stage colorectal cancer (CRC) is an effective approach to reduce CRC death rates. Current screening methods lack specificity as they detect many adenomas that will never progress to CRC. The authors aimed to identify cell surface protein biomarkers with extracellular domains that could be targeted for molecular imaging and discriminate low-risk adenomas and normal colon from high-risk adenomas and CRC.

Design Cell surface proteins of five CRC cell lines were biotinylated, isolated and analysed by in-depth proteomics using gel electrophoresis and nanoliquid chromatography coupled to tandem mass spectrometry. Differential expression in adenomas and CRCs was based on mRNA expression and verified by immunohistochemical staining of tissue microarrays.

Results In total, 2609 proteins were identified in the cell surface fractions. Of these, 44 proteins were selected as promising cell surface candidate biomarkers for adenoma-to-carcinoma progression based on the following criteria: protein identification in at least four out of five cell lines, a predicted (trans)membrane location and increased mRNA expression in CRCs compared to adenomas. Increased protein expression in high-risk adenomas and CRCs compared to low-risk adenomas was confirmed by immunohistochemistry for glucose transporter type 1 (gene symbol SLC2A1; p<0.00001) and prion protein (gene symbol PRNP; p<0.005).

Conclusion This study revealed glucose transporter type 1, prion protein and 42 other cell surface candidate biomarkers for adenoma-to-carcinoma progression that could potentially serve as targets for emerging molecular imaging modalities like optical imaging, 19F-MRI and positron emission tomography.

INTRODUCTION

Colorectal cancer (CRC) is the second most common cancer in the industrialised world, accounting for about 1 million new cases and 0.5 million cancer-related deaths each year.1 The stage at which CRC is determined determines patient outcome, with 5-year survival rates of more than 90% for stage I disease and less than 10% for stage IV disease.2 Therefore, early detection of CRC is a realistic approach to reduce the currently high CRC death rates. CRC develops in a multi-step manner from normal colon epithelium upon accumulation of (epi)genetic changes that stimulate proliferation and formation of non-malignant precursor lesions, that is, adenomas, which may further progress into invasive tumours with metastatic potential. One of the key biological features of adenoma-to-carcinoma progression is the introduction of genomic instability, occurring as chromosomal instability in about 85% of CRC cases or as microsatellite instability (MSI) in about 15% of

Significance of this study

What is already known on this subject?
- Early detection of high-risk premalignant colon adenomas and early-stage colorectal cancer (CRC) is an effective approach to reduce CRC death rates.
- Colon adenoma-to-carcinoma progression is characterised by non-random molecular alterations, a number of which cause overexpression of cell surface proteins.
- Cell surface proteins can be specifically targeted by antibodies and are therefore potential biomarkers suitable for molecular imaging or drug targeting.

What are the new findings?
- Identification of 2609 proteins from cell surface fractions of five CRC cell lines, one of the largest CRC protein data sets obtained to date.
- Identification of 44 candidate cell surface protein biomarkers associated with adenoma-to-carcinoma progression.
- Verification of glucose transporter type 1 and prion protein as cell surface biomarkers with increased protein expression in high-risk colon adenomas and carcinomas compared to low-risk adenomas.

How might it impact on clinical practice in the foreseeable future?
- Cell surface protein biomarkers for adenomas at high risk of progression and carcinomas are candidate targets for molecular imaging as well as candidate drug targets.
- These biomarkers may facilitate optical image-guided surgery, positron emission tomography imaging and MRI-based molecular virtual colonoscopy.
Specific chromosomal aberrations have been associated with adenoma-to-carcinoma progression, particularly gains of chromosomal regions 20q, 13q and 8q. Importantly, only approximately 5% of all adenomas will progress into carcinomas, a subset to which we will further refer to as ‘high-risk adenomas’. Therefore, biomarkers that discriminate normal colon and low-risk adenomas on the one hand from high-risk adenomas and early-stage CRC on the other hand would be of particular clinical relevance.

Anatomical imaging techniques like CT and MRI are commonly used to visualise tumour mass for diagnosis, prognosis and monitoring of disease. However, tumour masses that are phenotypically similar may have evolved through activation of distinct molecular pathways, with significant consequences for disease prognosis and treatment selection. Visualisation of molecular information by molecular imaging is an emerging field of research, with already a number of established clinical applications and with many more under study. Cell surface proteins with extracellular domains are attractive as biomarkers because they can be targeted by monoclonal antibodies with high specificity. Such antibodies can be labelled with radioactive tracers for imaging by positron emission tomography (PET), with particles for MRI, or with fluorescent labels for image-guided surgery. In addition, some of these biomarkers may function as novel targets for (antibody-based) therapeutic cancer drugs. Large-scale identification of several thousands of proteins by in-depth proteomics analysis has only become feasible in recent years due to technological advancements in mass spectrometry and bioinformatics. Nevertheless, protein identification by shotgun proteomics remains a stochastic process in which few high-abundant proteins can easily overshadow and hamper identification of relevant but intermittently or lowly expressed proteins. Enriching samples for the subproteome of interest is a powerful approach to overcome this problem. The methodology to biotinylate and isolate the membrane proteome, that is, the collection of proteins that potentially can be targeted by antibodies, has been developed and successfully applied to identify biomarkers for primary cancer and metastatic disease. In the present study, we aimed to identify cell surface protein biomarkers for colorectal adenoma-to-carcinoma progression that can distinguish clinically harmless low-risk adenomas from clinically relevant high-risk adenomas and CRC.

**MATERIALS AND METHODS**

**Cell culture**

The CRC cell lines HT-29 and Caco-2 were obtained from the American Type Culture Collection (LGC Standards GmbH, Wesel, Germany), and the COLO 205, HCT 116 and RKO cell lines were kindly provided by Dr G J Peters, Department of Oncology, VU University Medical Center. HT-29, COLO 205, HCT 116 and RKO were cultured in Dulbecco’s modified Eagle’s medium (Lonza Biowhittaker, Verviers, Belgium) containing 10% fetal bovine serum (Hyclone, Perbio, UK). Caco-2 was cultured in RPMI 1640 (Lonza Biowhittaker, Verviers, Belgium) containing 2% fetal bovine serum. Both cell culture media were supplemented with 10% FBS, 100 IU/ml penicillin, 100 μg/ml streptomycin (Fisipharma, Palomonta (SA), Italy).

**Cell surface protein isolation**

The Pierce cell surface protein isolation kit (Pierce, Rockford, Illinois, USA) was used for the isolation of cell surface protein and intracellular protein fractions, according to the manufacturer’s protocol with some modifications. Per cell line, five 75-cm² tissue culture flasks (Greiner Bio-One, Alphen a/d Rijn, The Netherlands) at 70% to 80% confluency (approximately 14×10⁶ cells per flask) were incubated with 412 μM Sulfo-NHS-SS-Biotin (Pierce) dissolved in phosphate-buffered saline (PBS) at 4°C for 30 min, washed with PBS, collected by gentle scraping and pelleted by centrifugation (3 min at 500 g). Cells were lysed in 500 μl of lysis buffer (Pierce) supplemented with 1× complete protease inhibitor cocktail (Roche, Mannheim, Germany) and incubated on ice for 30 min with intermediate vortexing every 5 min. Cell lysates were clarified by centrifugation at 4°C (10 min at 16 100 g) upon which supernatants were incubated with NeutrAvidin Protein beads (Pierce) at 4°C for 2 h. The beads were collected by centrifugation and washed according to the protocol described by Scheuer et al. Beads were washed three times in PBS containing 1% w/v Nonidet P-40, 0.1% w/v sodium dodecyl sulfate (SDS) and 1× protease inhibitor cocktail and again three times in PBS containing 0.1% w/v Nonidet P-40, 0.5 M NaCl and 1× protease inhibitor cocktail. Proteins were eluted from the beads by incubation in 330 μl of PBS containing 50 mM diethiothreitol and 62.5 mM Tris-HCl at room temperature for 1 h. The eluted proteins represented the cell surface fractions. The remaining unbound cell lysates were also collected and represented the intracellular fractions (figure 1A). Protein concentrations for the cell surface protein fractions and the intracellular protein fractions were 0.25 μg/μl and 5 μg/μl, respectively. The cell surface protein fractions were concentrated with a 10-kDa cut-off filter (Millipore Corporation, Billerica, Massachusetts, USA) prior to gel electrophoresis.

**Western blotting**

Five micromgms of protein samples was separated by gel electrophoresis using 8% SDS-polyacrylamide gel electrophoresis gels and transferred to polyvinylidene fluoride membranes (Amersham Biosciences, Little Chafont, UK). Membranes were blocked with PBS containing 0.05% Tween-20 (Sigma-Aldrich, St Louis, Missouri, USA) and 3% w/v dry milk powder, incubated with PBS containing 0.05% Tween-20, incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies at room temperature for 1 h and developed using ECL-Plus and hyperfilm (Amersham Biosciences).

**Gel electrophoresis and sample preparation for proteomics analysis**

Thirty micrograms of cell surface protein fractions and two pools of intracellular protein fractions were separated by gel electrophoresis using a precast one-dimensional (1D) 4—12% gradient SDS-polyacrylamide gel electrophoresis gel (Invitrogen, Carlsbad, California, USA). The gel was fixed in 50% ethanol containing 5% phosphoric acid and stained with Coomassie R-250. Gels were washed in Milli-Q water and stored at 4°C until processing for in-gel digestion. Each lane was cut in 10 gel bands and further processed into tryptic peptides as described by Piersma et al. The volume of the desalted peptide fractions was reduced to 50 μl in a vacuum centrifuge.

**NanoLC—MS/MS proteomics analysis**

We previously demonstrated the feasibility of label-free protein profiling by 1D gel electrophoresis coupled to nanoliquid chromatography—tandem mass spectrometry (nanoLC—MS/MS), with generally 70—80% of protein reproducibly identified in three technical replicates with 17—25% coefficient of variation of quantification. Peptides were separated by an Ultimate
Figure 1  Schematic representation of the cell surface protein isolation and sample preparation for mass spectrometry. (A) Cells were incubated with Sulfo-NHS-SS-Biotin to biotinylate extracellular domains of cell surface proteins. Cell lysates were incubated with avidin-coated beads to capture biotin-labelled proteins. Subsequently, these proteins were eluted from the beads and further referred to as the cell surface protein fractions (CS).

Non-biotinylated proteins were also collected and referred to as an intracellular protein fractions (IF). (B) Detection of E-cadherin in the CS and IF of five CRC cell lines by western blotting. E-cadherin is enriched in CSs compared to the IFs of four cell lines, thereby verifying the cell surface protein isolation procedure. (C) 1-D gel electrophoresis and Coomassie staining of five CSs and two pools of IFs, one containing the chromosomal instability cell lines COLO 205, Caco-2 and HT-29 and the other containing the MSI cell lines RKO and HCT 116. The gel lanes were cut into 10 gel bands per sample and further processed for nanoLC–MS/MS.

3000 nanoLC system (Dionex LC Packings, Amsterdam, The Netherlands) equipped with a fused silica column (20 cm×75 μm internal diameter) custom packed with 3 μm 120 Å ReproSil Pur C18 aqua (Dr Maisch GmbH, Ammerbuch-Entringen, Germany) as previously described. Brieﬂy, peptides were trapped on a Pepmap C18 cartridge (5 mm×300 μm internal diameter) (Dionex) and separated at 300 nl/min in a 60-min gradient. Intact peptide mass spectra and fragmentation spectra (top ﬁve) were acquired on an LTQ-FT hybrid mass spectrometer (Thermo Fisher, Bremen, Germany). Dynamic exclusion was applied with a repeat count of 1 and an exclusion time of 50 s.

Database searching
MS/MS spectra were searched against the human International Protein Index database 3.51 (67 511 entries) using Sequest (version 27, rev 12). Scaffold version 2.01 (Proteome Software, Portland, Oregon, USA) was used to organise the data and to validate peptide identiﬁcations using PeptideProphet (probability >95%) and ProteinProphet (probability of >99% with two peptides or more in one sample). Proteins were quantified by spectral counting. For each protein, the number of assigned MS/MS spectra was summed across the 10 fractions and exported to Excel (see supplementary table 1). Normalised counts were calculated by dividing the counts per protein by the sum of all counts per sample and multiplying by the average sum across all samples.

Glucose transporter type 1 (GLUT1) and prion protein (PrP<sup>C</sup>) immunohistochemistry on tissue microarrays (TMAs)
TMAs were constructed using a series of 82 colon adenomas and 82 CRCs that were collected retrospectively from 2001 to 2008 at the VU University Medical Center, Amsterdam, The Netherlands. All tissues were used in compliance with the institutional ethical regulations for use of patient material. Two to six cores from each parafﬁn block were included in the TMAs, depending on the amount of tissue available.

The immunohistochemical stainings, scanning and scoring procedures were done employing instruments and in accordance to strategies used within the Human Protein Atlas. In brief, sections were deparafﬁnised and rehydrated, and endogenous peroxidases were blocked with 0.3% hydrogen peroxide in methanol for 30 min. For the GLUT1 staining, antigens were retrieved by autoclaving in citrate buffer solution (pH 6.0). Incubation of GLUT1 antibody (AB1341, Chemicon/Millipore) at a 1:50 dilution in antibody diluent (ThermoScienﬁc, Waltham, Massachusetts, USA) was performed for 30 min at room temperature. GLUT1 staining was detected by an HRP-coupled polymer and visualised by diaminobenzidine plus (DAB Plus) (both from ThermoScienﬁc). For the PrP<sup>C</sup> staining, antigens were retrieved by autoclaving in 1.5 mM HCl. Primary anti-PrP<sup>C</sup> antibody (clone BH4, Sigma-Aldrich) was incubated at a 1:200 dilution in PBS containing 1% bovine serum albumin and 0.1% Tween-20 (Sigma-Aldrich) at 4°C overnight and subsequently detected by an HRP-coupled antihouse polymer (PowerVision plus, Immunologic, Duiven, The Netherlands) and incubation with diaminobenzidine (DAKO, Heverlee, Belgium). Sections were counterstained with Mayer’s haematoxylin.

Sections stained for GLUT1 were automatically scanned with an Aperio ScanScope XT Slide Scanner (Aperio Technologies, Vista, California, USA) equipped with a 20× objective with a numerical aperture of 0.75 (from Olympus, Tokyo, Japan), and stainings were scored using the ImageScope viewer (Aperio). Sections stained for PrP<sup>C</sup> were scanned with a Mirax digital slide scanner (SDHISTECH, Budapest, Hungary) also equipped with a 20× objective with a numerical aperture of 0.75 (Carl Zeiss B.
V, Sliedrecht, The Netherlands), and stainings were scored using the Mirax Viewer software (3DHISTECH). The scoring was done using a dedicated and calibrated high-resolution computer screen, and visual analogue scales of staining patterns were used for reference purposes. Any ambiguous patterns of staining were discussed at a multi-head microscope.

The scoring methods applied were guided by the staining patterns observed. For GLUT1, staining intensity (negative, weak, moderate or strong) in the epithelial cells within one tissue core was scored. In addition, for GLUT1, distinct plasma membrane staining could be observed and this was scored as either present of absent. The highest score of multiple cores from each tumour was used for further analysis. Final evaluation was performed by dichotomising the data into two groups. For GLUT1, samples with a strong intensity score of the staining had ‘high protein expression’, while all others had ‘low protein expression’. For technical reasons, in some cases, staining could not be evaluated, leaving 66 colon adenomas and 56 colon carcinomas in this study for GLUT1 (see supplementary table 2). For PrP,25 staining intensity (negative, weak, moderate or strong) and percentage of positive epithelial cells within one tissue core (0–80% or >80%) were scored. The highest score of multiple tissue cores was used and high protein expression was defined as >80% of epithelial cells showing moderate to strong staining. All other samples were considered to have low staining. For PrP,25 staining, 75 colon adenomas and 68 colon carcinomas were evaluated (see Supplementary table 2).

Analysis of DNA copy number alterations of chromosomal regions 8q, 13q and 20q

Half of all tumour samples were analysed for DNA copy number alterations by array comparative genomic hybridisation.5 The R package CGHcall was used to define copy number gains and losses.26 A gain of chromosomal region 8q, 13q or 20q was defined by a gain of at least 14 consecutive probes within these regions. In the remaining samples, DNA copy number status for 8q, 13q and 20q was analysed by multiplex ligation-dependent probe amplification, as described previously.27 A gain of 8q, 13q or 20q was defined when at least 50% of all multiplex ligation-dependent probe amplification probes in a region showed a tumour-to-normal DNA copy number ratio of 1.5 or more. Adenomas that were microsatellite stable (MSS) and showed high protein expression and samples without instability in any of these markers were considered MSI-High, whereas samples with instability in one marker and samples without instability in any of these markers were considered MSS.

MSI analysis

MSI status was determined using the MSI Analysis System version 1.2 (Promega Corporation, Madison, Wisconsin, USA). In brief, this system examines the profile of five mononucleotide repeat markers (BAT-25, BAT-26, NR-21, NR-24 and MONO-27) and two pentanucleotide repeat markers (Renta C and Penta D). Samples with instability in two or more markers were considered MSI-High, whereas samples with instability in one marker and samples without instability in any of these markers were considered MSS.

Data mining and statistical analysis

The software program Phobius (http://phobius.sbc.su.se) was used for the prediction of a transmembrane region and/or a signal peptide.28 The Ingenuity Pathway Analysis tool (IPA 7.0, Ingenuity Systems, Inc., Redwood City, CA, USA) was used to collect annotations on, for example, protein function and expected subcellular location. The χ² test or Fisher’s exact test, whichever was appropriate, was applied to evaluate associations between categorical variables. t testing was applied for investigation of associations between immunohistochemical staining results and patient age. All statistical analyses were performed using SPSS (SPSS V15, SPSS Inc). All reported p values are two sided, and p ≤ 0.05 was considered significant.

RESULTS

Identification of cell surface proteins in CRC cell lines

The chromosomal instability CRC cell lines HT-29, Caco-2 and COLO 205 and the MSI CRC cell lines HCT 116 and RKO were selected for identification of cell surface proteins. Extracellular domains of membrane proteins were biotinylated, cells were lysed and the biotinylated cell surface protein fractions and non-biotinylated intracellular protein fractions were collected (figure 1A). Western blotting for the cell adhesion molecule E-cadherin was performed to verify the cell surface protein isolation procedure and confirmed the enrichment of E-cadherin within cell surface fractions compared to the complementary intracellular fractions of four cell lines (figure 1B). E-cadherin was not detected in RKO protein fractions, consistent with literature reports.29

For cell surface proteomics, the cell surface protein fractions obtained from the five CRC cell lines were separated by 1D gel electrophoresis, stained with Coomassie Brilliant Blue (figure 1C) and further processed for protein identification by nano-LC–MS/MS as described in Materials and Methods. A total of 2609 proteins were identified in the cell surface fractions with an average of 1675 proteins per cell line, ranging from 1554 proteins for Caco-2 to 1816 proteins for HT-29. Eighty-one percent of all proteins were identified in cell surface fractions of at least four of the five cell lines, indicating for common expression of these proteins among CRC cell lines (figure 2). As expected, E-cadherin was detected in all cell lines except for RKO. In addition, many other well-known plasma membrane proteins were identified, such as the epithelial cell adhesion molecules (also known as TACSTD1) and the integrins ITGA6, ITGA9, ITGB1, ITGB4 and ITGB5; receptors for signalling pathways such as the ephrin receptors EPHA1, EPHA2 and EPHB2, the epidermal growth factor...
receptor and the tumour necrosis factor receptor superfamily member FAS; transporter molecules including the ABC transporter family members ABCC1 and ABCC3, which are involved in multi-drug resistance; and immune-response-related molecules such as HLA-A and five other members of the major histocompatibility complex family and the immunoglobulin superfamily members IGSF3 and IGSF8. A complete overview of all proteins identified and their spectral counts within each of the cell surface fractions is provided by supplementary table 1.

### Cell surface proteins associated with adenoma-to-carcinoma progression

Several filters were applied to identify cell surface proteins associated with colorectal adenoma-to-carcinoma progression. To enrich for biomarker candidates that are likely to recognise a wide variety of molecularly heterogeneous colon tumours, proteins that were identified in at least four out of five cell lines were selected. Next, the Phobius program was used to select for biomarker candidates with a predicted transmembrane domain.

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>UniProt accession number</th>
<th>Protein name</th>
<th>Molecular weight (kDa)</th>
<th>Cell surface fractions (spectral counts)</th>
</tr>
</thead>
</table>
| ACSL1       | P33121                   | Long-chain fatty acid—coenzyme A ligase 1 | 78          | COLO 205  
Caco-2  
HT-29  
RKO  
HCT 116 |
| ACSL3       | Q6PM8                    | ACSL3 protein | 137         | 14 8 14 46 42 |
| AIFM1       | Q05831                   | Apoptosis-inducing factor 1, mitochondrial | 58          | 17 0 28 42 48 |
| AT2L        | Q8NH9                    | Atlastin-2    | 20          | 1 5 2 5 4 |
| AUP1        | Q0Y679                   | Apoptosis-inducing factor 1, mitochondrial | 53          | 0 4 1 1 2 |
| BCAM        | P0S0856                  | Basal cell adhesion molecule | 67          | 48 62 46 17 6 |
| BCA1        | P51572                   | B cell receptor-associated protein 31 | 36          | 18 0 8 20 1 |
| BST2        | Q10589                   | Bone marrow stromal antigen 2 | 20          | 1 3 1 8 0 |
| CNNM3       | Q8NE01                   | Metal transporter CNNM3 | 76          | 1 3 3 5 0 |
| COPS6       | Q7L5N1                   | COP9 signalosome complex subunit 6 | 36          | 9 2 4 1 2 |
| COX7B       | P36551                   | Coproporphyrinogen-III oxidase, mitochondrial | 50          | 4 4 5 8 24 |
| CYB5B       | P36551                   | Cytochrome b5 type B | 17          | 4 3 5 2 4 |
| DAG1        | Q01118                   | Dystroglycan   | 98          | 14 47 22 28 64 |
| DHCR7       | Q0UB9M7                  | 7-Dehydrocholesterol reductase | 54          | 8 3 12 1 3 |
| EPHA1       | P01709                   | Ephrin type-A receptor 1 | 108         | 8 14 14 0 4 |
| HM13        | Q8TCT9                   | Minor histocompatibility antigen H13 | 69          | 8 3 6 14 9 |
| HSD17B      | P56937                   | 3-Keto-steroid reductase | 38          | 1 2 3 1 2 |
| HTRA2       | Q04364                   | Serine protease HTRA2, mitochondrial | 28          | 14 47 22 28 64 |
| ICAM1       | P05382                   | Intercellular adhesion molecule 1 | 58          | 3 16 3 19 3 |
| IGSF8       | P06901                   | Immunoglobulin superfamily member 8 | 65          | 8 3 6 14 9 |
| ITG43       | Q9HX04                   | Protein ITG3   | 69          | 1 5 2 5 4 |
| KIAA1524    | Q8TCGO1                  | Protein CIP2A  | 102         | 1 1 1 5 0 |
| LDLR        | P01130                   | Low-density lipoprotein receptor 8 | 95          | 3 2 3 43 11 |
| LRPB        | Q014114                  | Low-density lipoprotein receptor-related protein 8 | 106         | 3 0 1 4 2 |
| MTDH        | Q5UE4                    | Protein LRYC   | 49          | 1 14 1 0 3 |
| NSDHL       | Q15739                   | Sterol-4α-carboxyline-3-dehydrogenase, dehydrogenating | 55          | 1 7 7 2 6 |
| PLX1        | Q0UVW2                   | Plexin-A1      | 211         | 14 8 14 46 42 |
| PODXL       | P00592                   | Podocalyxin-like protein 1 | 56          | 0 11 1 2 7 |
| PRNP        | P04156                   | Major prion protein | 28          | 10 6 6 15 15 |
| PRSS8       | Q16851                   | Prostasin      | 37          | 1 27 5 0 1 |
| PTGFRN      | Q9P2B2                   | Prostaglandin F2 receptor negative regulator | 99          | 17 0 28 42 48 |
| PTK7        | Q3308                    | Tyrosine-protein kinase-like 7 | 118         | 18 35 32 47 81 |
| RP2N        | P04844                   | Dolichyl-diphosphoglycerol-arachidonate—protein glycosyltransferase subunit 2 | 57          | 3 16 3 19 3 |
| RTN4        | Q5NQ23                   | Reticulin-4    | 37          | 3 2 3 43 11 |
| SCARB1      | Q8WTX0                   | Scavenger receptor class B member 1 | 57          | 5 19 5 13 24 |
| SDC2        | P31431                   | Syndecan-4     | 22          | 0 4 2 2 8 |
| SDF1        | Q908B4                   | Sideroflexin-1  | 95          | 48 62 46 17 6 |
| SLCA1       | P43003                   | Excitatory amino acid transporter 1 | 60          | 1 14 1 0 3 |
| SLCA1A      | P43007                   | Neutral amino acid transporter 1 | 56          | 1 0 1 5 1 |
| SLCA1B      | Q15758                   | Neutral amino acid transporter B(0) | 57          | 17 25 17 27 20 |
| SLCA2       | P11166                   | Solute carrier family 2, facilitated glucose transporter member 1 | 54          | 7 18 12 7 16 |
| SLC4A4      | P04920                   | Anion exchange protein 2 | 137         | 18 0 8 20 1 |
| SLC7A11     | Q5UPY5                   | Cysteine/glutamate transporter | 55          | 5 2 5 2 4 |
| SURF4       | Q15260                   | Surfeit locus protein 4 | 17          | 5 19 5 13 24 |

---

as these indicate a putative membrane localisation of the proteins identified. In addition, proteins, whose corresponding mRNA levels were previously found to be overexpressed in a series of 31 carcinomas compared to a series of 37 adenomas (supplementary table 1), were selected. When combined together, these selection criteria resulted in a list of 44 promising candidate cell surface biomarkers for adenoma-to-carcinoma progression (table 1).

**Verification of GLUT1 and PrPC as cell surface candidate biomarkers for adenoma-to-CRC progression**

Out of the 44 cell surface proteins associated with adenoma-to-carcinoma progression, two proteins were selected for further validation. GLUT1 (gene symbol SLC2A1) was selected because of its central role in cancer cell metabolism, while PrPC (gene symbol PRNP) was selected based on evidence in literature of its putative role in colon cancer. The presence of PrPC protein on the cell surface of the five CRC cell lines was verified by fluorescence-activated cell sorting analysis (supplementary figure 1). E-cadherin was taken along as a positive control, for which expression was found in all cell lines except RKO, conforming to expectations. Next, GLUT1 and PrPC protein expression were examined by immunohistochemistry using TMAs containing a series of colon adenomas and carcinomas. Within epithelial cells of adenomas and carcinomas, GLUT1 staining was observed mainly in the cytoplasm and on the plasma membrane (figure 3).

**Figure 3** GLUT1 immunohistochemical staining of colon adenomas and carcinomas. TMAs containing a series of colon adenomas and carcinomas were stained for GLUT1 by immunohistochemistry. Representative examples show that GLUT1 staining of epithelial cells in adenomas (A,B) was less intense than that in high-risk adenomas (C,D) and carcinomas (E,F). GLUT1 staining was present in the cytoplasm and at the plasma membrane of the epithelial cells. Scale bars indicate 50 μm.
High levels of GLUT1 cytoplasmic and plasma membrane staining were observed significantly more in CRCs than in adenomas (p<0.001 and p<0.00001; figure 5A,C, respectively). PrP<sup>C</sup> staining was predominantly observed in the cytoplasm within the epithelial cells (figure 4). In carcinomas, intense PrP<sup>C</sup> staining was observed at and just below the plasma membrane at the luminal side (figure 4E,D). PrP<sup>C</sup> protein expression levels were significantly higher in CRCs than in adenomas (p<0.05, figure 5E). This implies that both GLUT1 and PrP<sup>C</sup> expression are increased during adenoma-to-carcinoma progression.

Tumours were characterised for MSI status, and chromosomal regions 20q, 13q and 8q were characterised for DNA copy number status (supplementary table 2) because gains of these chromosomal regions are associated with adenoma-to-carcinoma progression. Therefore, MSS adenomas with an increased copy number in at least one of the regions (8q, 13q and 20q) were classified as high-risk adenomas, while adenomas without gains in these chromosomal regions were classified as low-risk adenomas. There was no significant correlation between MSI status and GLUT1 or PrP<sup>C</sup> expression (supplementary table 2). Classification of MSS adenomas into low-risk and high-risk adenomas indicated that high GLUT1 staining was detected more frequently in high-risk adenomas and carcinomas combined compared to the low-risk adenomas (p<0.00001; figure 5A-C, respectively).

**Figure 4** PrP<sup>C</sup> immunohistochemical staining of colon adenomas and carcinomas. TMAs containing a series of colon adenomas and carcinomas were stained for PrP<sup>C</sup> by immunohistochemistry. Representative examples show that PrP<sup>C</sup> staining of epithelial cells in adenomas (A,B) was less intense than that in high-risk adenomas (C,D) and carcinomas (E,F). Epithelial staining was mostly cytoplasmic with an increased staining intensity towards the plasma membrane at the luminal side. Scale bars indicate 50 μm.
DISCUSSION

There is an evident clinical need for biomarkers that specifically distinguish low-risk colorectal adenomas from high-risk adenomas and CRCs. Cell surface protein biomarkers are of particular interest because they can be targeted by antibodies for molecular imaging or cancer treatment. We here applied a shotgun proteomics analysis of the cell surface protein fraction of five CRC cell lines and aimed to identify proteins associated with tumour progression by integrating the proteomics data with transcriptomics data from a series of adenomas and CRCs. Combined with additional selection criteria concerning predicted membrane localisation and common expression among multiple CRC cell lines, this study identified 44 cell surface candidate biomarkers for adenoma-to-CRC progression (table 1).

The association between colon tumour progression and increased protein expression was verified by immunohistochemical analysis of two of the candidates, GLUT1 and PrPC (figure 5), thereby validating our approach.

The group of 44 candidate biomarkers comprised multiple genes that have been associated with various aspects of tumour biology. GLUT1 plays a central role in cancer metabolism. Cancer cells switch from oxidative phosphorylation to glycolysis, which results in an increase in glucose requirement well known as the Warburg effect. To compensate for this higher demand for glucose, cancer cells tend to upregulate glucose transporters, among which GLUT1, resulting in a higher glucose influx. Although GLUT1 is known to be expressed at low levels in several normal tissues including colon, consistent with our own findings (supplementary figure 2A,B), its increased expression in cancer is used in clinical practice for PET imaging with the 18FDG glucose analogue, for example, to visualise CRC liver metastases. In addition to GLUT1, several others from the group of 44 candidate biomarkers are also associated with cancer metabolism. SLC1A5 is a transporter protein for neutral amino acids such as glutamine, a high uptake of which is required for cancer cells to keep up with the proliferation rate and synthesis of new macromolecules. High levels of the SLC1A5 protein were found to be associated with a worse survival of patients with CRC. Moreover, the ACSL1, ACSL3, DHCR7, HSD17B7, LDLR, NSDHL and SCARB1 candidate biomarkers are all functionally linked to lipid metabolism (IPA 7.0, Ingenuity Systems, Inc.).

While PrPC is well known for its role in prion diseases, PrPC is expressed in multiple tissues, including normal colon and CRCs. The group of 44 candidate biomarkers comprised multiple CRC cell lines and aimed to identify proteins associated with disease recurrence in patients with stage II and stage III CRC. Several other candidate cell surface biomarkers for adenoma-to-CRC progression affect oncoprogenic pathways. KIAA1524/CIPI2A stabilises MYC in cancer cells. ALF1M1 is a central mediator of the caspase-independent apoptosis pathway. Ephrin receptors such as EPHA1 are well-established factors in cancer progression and neovascularisation. MDH1 and BST2 have been shown to promote metastasis of breast cancer. Likewise, the cellular adhesion molecules BCAM and ICAM1 have adhesive properties that may facilitate cell migration and metastasis. BCAM is overexpressed by a variety of cancers and a role in metastasis has been suggested, while for ICAM1,
a role in metastasis has been demonstrated.47 Furthermore, a soluble form of ICAM1 can be detected in serum and has been proposed as a prognostic factor for colon cancer.48 An anti-ICAM1 antibody was successfully applied to reduce tumour growth of several cancer types in an in vivo model system.49

Taken together, these data confirm that the list of cell surface biomarkers for adenoma-to-CRC progression (table 1) indeed contains many candidates that are likely to be associated with colon tumour progression.

In this study, we focused on identification of cell surface biomarkers for adenoma-to-CRC progression. To select for the protein biomarkers of interest, we made use of a list of genes that exhibited increased mRNA expression in a series of CRCs compared to adenomas. The complete list of 2609 proteins that were identified in the cell surface fractions (supplementary table 1), one of the largest CRC protein data sets obtained to date, also comprises known drug targets. This includes epidermal growth factor receptor against which the monoclonal antibodies cetuximab and panitumumab have been registered for use in CRC as well as the cell surface proteins IGF1R, CD40 and EPCAM that can be targeted by cixutumab, SGN-40 and tucutuzumab cemoluxin, respectively. Besides EPCAM, other putative stem cell markers were also identified in this study, like CD44, ALCAM, ALDH1 and FROML1 (also known as CD135).50

Almost half (48%) of the cell surface proteins identified in this study were expressed by four or five out of the five cell lines we examined (figure 2). Moreover, the present findings are consistent with the identification of nearly 500 cell surface proteins expressed by the non-metastatic and metastatic CRC cell lines KM12C and KM12SM, respectively, as more than 80% of these proteins were expressed by four or five of the CRC cell lines. The complete list of 2609 proteins that were identified in the cell surface fractions (supplementary table 1), one of the largest CRC protein data sets obtained to date, also comprises known drug targets. This includes epidermal growth factor receptor against which the monoclonal antibodies cetuximab and panitumumab have been registered for use in CRC as well as the cell surface proteins IGF1R, CD40 and EPCAM that can be targeted by cixutumab, SGN-40 and tucutuzumab cemoluxin, respectively. Besides EPCAM, other putative stem cell markers were also identified in this study, like CD44, ALCAM, ALDH1 and FROML1 (also known as CD135).50

Almost half (48%) of the cell surface proteins identified in this study were expressed by four or five out of the five cell lines we examined (figure 2). Moreover, the present findings are consistent with the identification of nearly 500 cell surface proteins expressed by the non-metastatic and metastatic CRC cell lines KM12C and KM12SM, respectively, as more than 80% of these proteins were also detected in the present study.10 These data indicate that many cell surface proteins are commonly expressed by a variety of CRC cell lines, despite their biological and molecular heterogeneity. Nevertheless, biomarkers that specifically detect the far majority of high-risk colorectal adenomas and carcinomas have not been identified so far, indicating that a combination of markers may be needed for this purpose.

It is of interest that GLUT1 plasma membrane staining was observed in only 57% of the high-risk adenoma and carcinoma tissue samples (figure 5D and supplementary table 2), while in clinical practice, 13FDG PET imaging has a sensitivity for carcinomas of 89%.51 52 Sensitivity, specificity and ultimately predictive value of candidate biomarkers for use as imaging target depend on multiple variables, including expression of the target molecule in the diseased tissues of interest compared to background, mode of administration (oral vs systemic), accessibility of the molecule for the probe or tracer, analytical sensitivity of the assay/imaging modality and clinical and program sensitivity. Therefore, while immunohistochemical evaluation demonstrates increased expression levels of GLUT1 and PrPC in high-risk colon adenomas and carcinomas (figure 5), studies using imaging modalities will be needed to conclusively determine suitability of these markers for molecular imaging of CRC.

In conclusion, we here present high-resolution cell surface proteomics data derived from five CRC cell lines. For the purpose of identification of biomarkers for adenoma-to-carcinoma progression, data were integrated with previously obtained transcriptomics information, a strategy that was validated by verification of GLUT1 and PrPC as cell surface candidate biomarkers for high-risk adenomas and CRCs. Moreover, the complete set of CRC cell surface protein data may support a variety of translational research questions for which clinical applications are emerging, particularly related to current trends in molecular imaging and targeted (personalised) therapies.

Acknowledgements We thank M Tijssen and A Bolijn for technical assistance and Dr C Postma for expert feedback.

Funding This work was supported by Philips Healthcare (MW) and the VUMc—Cancer Center Amsterdam (CRJ and proteomics infrastructure). This study was performed within the framework of CTMM, the Center for Translational Molecular Medicine. DeCoDe project (grant 03O-101).

Competing interests None.

Ethics approval This study was conducted with the approval of the Medical Ethics Review Board of the VU University Medical Center Amsterdam.

Contributors 1. Study concept and design. 2. Acquisition of data. 3. Analysis and interpretation of data. 4. Drafting of the manuscript. 5. Critical revision of the manuscript for important intellectual content. 6. Statistical analysis. 7. Obtained funding. 8. Technical or material support. 9. Study supervision.

References


Cell surface proteomics identifies glucose transporter type 1 and prion protein as candidate biomarkers for colorectal adenoma-to-carcinoma progression

Meike de Wit, Connie R Jimenez, Beatriz Carvalho, et al.

*Gut* published online September 2, 2011
doi: 10.1136/gutjnl-2011-300511

Updated information and services can be found at:
http://gut.bmj.com/content/early/2011/08/31/gutjnl-2011-300511.full.html

These include:

**References**
This article cites 52 articles, 10 of which can be accessed free at:
http://gut.bmj.com/content/early/2011/08/31/gutjnl-2011-300511.full.html#ref-list-1

**P<sub>P</sub>**
Published online September 2, 2011 in advance of the print journal.

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

Colon cancer (1022 articles)

Advance online articles have been peer reviewed, accepted for publication, edited and typeset, but have not yet appeared in the paper journal. Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/