ARTICLE
Journal of Cellular Biochemistry 114:2577–2587 (2013)

Journal of Cellular Biochemistry

The Secretome Signature of Colon Cancer Cell Lines

Esther Imperlini,[1] Irene Colavita,[2] Marianna Caterino,[1] Peppino Mirabelli,[1] Daniela Pagnozzi,[3] Luigi Del Vecchio,[2,4] Rosa Di Noto,[2,4] Margherita Ruoppolo,[2,4] and Stefania Orrù[1,5]*

[1]*Fondazione IRCCS SDN, Naples, Italy*
[2]*CEINGE Biotecnologie Avanzate Scarl, Naples, Italy*
[3]*Porto Conte Ricerche, Tramariglio, Alghero, Italy*
[4]*DMMBM, University of Naples Federico II, Naples, Italy*
[5]*DISIST, University of Naples Parthenope, Naples, Italy*

ABSTRACT

The definition of the secretome signature of a cancer cell line can be considered a potential tool to investigate tumor aggressiveness and a preclinical exploratory study required to optimize the search of cancer biomarkers. Dealing with a cell-specific secretome limits the contamination by the major components of the human serum and reduces the range of dynamic concentrations among the secreted proteins, thus favouring under-represented tissue-specific species. The aim of the present study is to characterize the secretome of two human colon carcinoma cell lines, CaCo-2 and HCT-GEO, in order to evaluate differences and similarities of two colorectal cancer model systems. In this study, we identified more than 170 protein species, 64 more expressed in the secretome of CaCo-2 cells and 54 more expressed in the secretome of HCT-GEO cells; 58 proteins were shared by the two systems. Among them, more than 50% were deemed to be secretory according to their Gene Ontology annotation and/or to their SignalP or SecretomeP scores. Such a characterization allowed corroborating the potential of a cell culture-based model in order to describe the cell-specific invasive properties and to provide a list of putative cancer biomarkers. J. Cell. Biochem. 114: 2577–2587, 2013.

KEY WORDS: SECRETOME; PROTEOMICS; CACO-2; HCT-GEO; COLON CANCER

In the biomedical research field many efforts are directed towards the discovery of reliable biomarkers for early detection, tumor aggressiveness and surveillance of cancer. In the recent years, the focus has moved from intracellular components to secreted factors. In fact, all currently known biomarkers are secreted or shed proteins [Kulasingam and Diamandis, 2008]. These latter contribute to tumor growth, angiogenesis, and metastasis through autocrine, paracrine, juxtacrine, and/or intracrine pathways [Normanno et al., 1996]. The protein set secreted in the conditioned media (CM) of a serum-free cellular culture (termed secretome) represents the signature of a specific cell line and is also peculiar of the organ/tissue of origin. The definition of a cell-specific secretome can be considered a potential tool to investigate tumor invasion properties and a preclinical exploratory study required to optimize the search of cancer biomarkers. This cell-specific secretome approach is more advantageous compared to the direct analysis of the serum because: i) the major components of the human serum, such as albumins and immunoglobulins, are less abundant, if not absent, thus simplifying the search for tissue-specific biomarkers; and ii) the range of dynamic concentrations among the secreted proteins is reduced favoring under-represented tissue-specific species [Kulasingam and Diamandis, 2008]. Despite some criticisms have been raised against this procedure due to its reductionist nature, several recent reports are based on: a) secretome analysis to study cancer aggressiveness and b) cell culture models of cancer to identify novel candidate markers in different kind of carcinomas [Makridakis and Vlahou, 2010]. As for

Abbreviations: CM, conditioned media; CRC, colorectal cancer.
Conflicts of interest: None.
Esther Imperlini and Irene Colavita authors contributed equally to this work.
Grant sponsor: Fondazione SDN-IRCCS; Grant sponsor: AIRC; Grant number: 10737; Grant sponsor: Progetto FIRB Italian Human Proteomenet; Grant number: RBRN07BMCT.
*Correspondence to: Prof. Stefania Orrù, Dipartimento di Studi delle Istituzioni e dei Sistemi Territoriali (DISIST), University of Naples Parthenope, Via Medina, 40 80133 Naples, Italy. E-mail: orru@uniparthenope.it
Manuscript Received: 30 April 2013; Manuscript Accepted: 29 May 2013
Accepted manuscript online in Wiley Online Library (wileyonlinelibrary.com): 6 June 2013
DOI 10.1002/jcb.24600 • © 2013 Wiley Periodicals, Inc.

colorectal cancer (CRC), one of leading cause of death worldwide [Jemal et al., 2008], few studies have been published based on characterization of cell-specific secretome using different methodological approaches, most of them investigating the primary colon cancer cell line SW480 and its lymph node metastatic counterpart, SW620 cell line [Xue et al., 2010; Wu et al., 2008, 2010]. Another study, based on an immunoscreening procedure, involved five colon cancer cell lines, including CaCo-2 and HT-29 [Klein-Scory et al., 2010].

The aim of the present study is to characterize the secretome of two human colon carcinoma cell lines, CaCo-2 and HCT-GEO, by proteomic procedures in order to evaluate differences and similarities of two CRC model systems. CaCo-2 consists of human colonocyte-derived adenocarcinoma cells that express spontaneously the differentiation characteristic of mature enterocytes [Chantret et al., 1988]. HCT-GEO is a colon cancer cell line whose differentiation can be modulate [Chantret et al., 1988].

In this study, we identified more than 170 protein species, 64 more expressed in the secretome of CaCo-2 cells and 54 more expressed in the secretome of HCT-GEO cells; 58 proteins were shared by the two systems. Protein data were analyzed by means of different bioinformatic tools to define functional clusters and recognize secreted species according to both the classical and the non-classical secreted pathways. The characterization of the secretome signature of CaCo-2 and HCT-GEO cells allowed speculating that HCT-GEO secretome contains several pro-invasion species, whereas different adhesion proteins are identified in CaCo-2 CM. As for biomarker search, we find putative candidates both in a cell specific secretome and in the subset of secreted species shared by two cell lines. In particular, we confirm already known putative CRC biomarkers, but also we indicate new secreted proteins that could be tested on the sera from CRC patients.

MATERIALS AND METHODS

CELL CULTURES AND SERUM FREE MEDIA COLLECTION

CaCo-2 and HCT-GEO human colon cancer cell lines were obtained from the CEINGE–Biotecnologie Avanzate (Naples, Italy) Cell Culture Facility. Three independent replicates of CaCo-2 cells were seeded at 20,000–40,000 cells/cm^2 in a T-75 culture flask (Falcon, Becton Dickinson, San Jose, CA) incubated at 37°C and 5% CO_2 in EMEM (Sigma–Aldrich, St Louis, MO) supplemented with 10% of fetal bovine serum (FBS) (Lonza, Basel, Switzerland), 1% non-essential amino acids (Sigma–Aldrich) and 1% Ultraglutamine (Lonza). Three independent replicates of HCT-GEO cells were seeded at 40,000 cells/cm^2 in a T-75 culture flask in DMEM high glucose (Sigma–Aldrich) supplemented with 10% FBS and 2% Ultraglutamine. Cellular viability was assessed by trypan blue (Sigma–Aldrich) exclusion and cell count was performed using a Burker chamber. Cells were incubated at 37°C in a humidified atmosphere of 5% CO_2 until they reached about 60–70% confluence. The culture media were then aspirated and cells were washed three times with 1X Dulbecco's Phosphate Buffered Saline (DPBS; Sigma–Aldrich) w/o calcium and magnesium and once with serum-free culture media. Cells were then grown in 10 ml of complete culture medium without serum for 24 h. Thereafter, for each independent experiment, the conditioned media from 15 T-75 flasks were collected for each cellular model system, supplemented with Complete Mini protease inhibitor cocktail (Roche, Indianapolis, IN) and centrifuged at 400*g* for 15 min at 4°C. Following collection of the conditioned media, adherent cells were detached by Trypsin-EDTA (Sigma–Aldrich) and counted using a Burker chamber. We estimated that more than 98% cells were viable.

CELL LYSIS AND PROTEIN EXTRACTION

Cells were resuspended with lysis buffer containing 50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1% (v/v) Triton X-100, 10% (v/v) glycerol, 0.5 mM PMSF and Complete Mini protease inhibitor cocktail. After 30 min incubation on ice, the insoluble material was removed by centrifugation at 16,000*g* at 4°C for 30 min. Protein concentrations were determined using Bradford's reagent (Biorad).

PROTEIN PRECIPITATION FROM CONDITIONED MEDIUM AND SDS–PAGE

Proteins were precipitated from the pooled conditioned media of each independent experiment using 15% TCA. Following 2 h incubation on ice, the samples were centrifuged at 15,000*g* at 4°C for 15 min and washed twice with ice-cold acetone. The precipitated proteins were then dissolved in de-ionized water and their concentrations were determined using Bradford's reagent.

The secreted proteins from each independent experiment were fractionated by 10% SDS–PAGE. Molecular masses of protein bands were estimated by using Precision Plus All Blue protein standards (Bio-Rad). Protein electrophoretic patterns were then visualized using GelCode Blue Stain Reagent (Pierce). The resulting gel image was scanned by PDquest 7.1 software (Biorad).

IN GEL DIGESTION

The whole gel lanes were cut to create 2-mm gel slices. Each slice was crushed and washed first with acetonitrile (ACN) and then with 50 mM ammonium bicarbonate (AMBIC). Enzymatic digestions were carried out by slightly modifying a previously described protocol [Orrù et al., 2002]. Briefly, protein samples were reduced by incubation in 10 mM DTT for 45 min at 56°C and alkylated with 55 mM iodoacetamide in 50 mM AMBIC for 30 min at room temperature in the dark. The gel particles were then washed with 50 mM AMBIC and ACN. Enzymatic digestions were carried out with modified trypsin (Sigma) (10 ng/μl) in 50 mM AMBIC, pH 8.5, at 4°C for 45 min. The enzymatic solution was then removed. A new aliquot of the buffer solution was added to the gel particles and incubated at 37°C for 18 h. A minimum reaction volume sufficient for complete rehydration of the gel was used. Peptides were extracted by washing the gel particles in ACN at 37°C for 15 min, and then vacuum-dried. Before mass spectrometry analysis, peptides were resuspended in 10 μl 0.5% TFA.

μLC–MS/MS ANALYSIS

Mass spectrometry analyses were performed by using the LC/MSD Trap XCT Ultra (Agilent Technologies, Palo Alto, CA) equipped with an 1100 HPLC system and a chip cube (Agilent Technologies), as reported previously [Imperlini et al., 2010]. Briefly, after an enrichment step, each peptide sample was fractionated on a C18

reverse-phase capillary column (75 μm × 43 mm) at a flow rate of 200 nl/min with a linear gradient of eluent B (0.1% FA in acetonitrile) in A (0.1% FA) from 5 to 60% in 50 min. Elution was monitored on the mass spectrometers without any splitting device. Peptide analysis was performed using data-dependent acquisition mode on the three most abundant ions in each MS scan. Dynamic exclusion was used to acquire a more complete survey of the peptides. A permanent exclusion list of the most frequent peptide contaminants was included as previously reported [Corbo et al., 2012]. Each analysis was conducted in duplicate.

PROTEIN IDENTIFICATION AND QUANTITATION

Mascot format text files were analyzed by Proteome Discoverer platform (version 1.3; Thermo Scientific, Bremen, Germany), interfaced with an in-house Mascot server (version 2.3, Matrix Science, London, UK) for protein identifications. All peak lists were processed against the UniProtKB/Swiss-Prot database (release 2012_05) with the following search parameters: Homo Sapiens as taxonomy, Precursor Mass Tolerance 300 ppm, Fragment Mass Tolerance 0.6 Da, charge state +2, and +3, cysteine carbamidomethylation as static modification, N-terminal glutamine conversion to pyro-glutammic acid and methionine oxidation as Dynamic modifications, enzyme trypsin, allowing up to one missed cleavage. For label-free quantification, all peptides with FDR ≤ 0.01 and a peptide rank of 1 were included. Spectral count (SpC) values were used as a quantitative parameter for estimating protein abundance and comparing the expression of the same protein between the two cell lines. SpC log ratio and Normalized Spectral Abundance Factor (NSAF) were calculated according to Old et al. [2005] and Zybailov et al. [2006], respectively, considering for each identified protein the average SpC value among three biological replicates.

DATA ANALYSIS

The secretome datasets identified were analyzed according to Gene Ontology annotations using the Gene Ontology Enrichment Analysis Software Toolkit (GOEAST, http://omicslab.genetics.ac.cn/GOEAST/) [Zheng and Wang, 2008]. Among the different annotation categories, the analysis was focused onto the cellular compartment annotation term.

The secretion features of identified proteins were analyzed by using Signal Peptide Predictor (SignalP, http://www.cbs.dtu.dk/services/SignalP4.0) and SecretomeP (http://www.cbs.dtu.dk/services/SecretomeP2.0) softwares. SignalP uses amino acid sequences and numerous artificial neural networks to predict the existence and location of signal peptide cleavage sites. A protein is considered classically secreted if it receives a signal peptide probability above a specified threshold (D-cutoff score ≥0.45) [Petersen et al., 2011]. SecretomeP software uses a neural network that combines six protein characteristics (number of atoms, number of positively charged residues, presence of transmembrane helices, presence of low-complexity regions, presence of pro-peptides, and subcellular localization) to determine if a protein is non-classically secreted. A protein is considered non-classically secreted if it receives a neural network score (NN-score) >0.5 [Bendtsen et al., 2004].

The identified proteins were also classified according to the Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 (http://david.abcc.ncifcrf.gov/). This tool adopts the Fisher′ exact test to measure the protein enrichment in annotation terms. If the *P*-value were equal to or smaller than 0.05, a protein would be considered strongly enriched in the annotation categories.

WESTERN BLOT

Western blot analyses were performed on additional and freshly protein samples prepared from conditioned serum-free media (CM, 15 μg) of three independent CaCo-2 and HCT-GEO cultures. The protein samples were resolved on 4–15% gradient SDS–PAGE gels and then transferred onto nitrocellulose membrane (GE Healthcare). As protein loading controls, the CM was stained with PonceauS (Sigma). The membranes were blocked as previously described [Caterino et al., 2013] and incubated for 1.5 h with 1% non-fat milk in PBS pH 7.6 and 0.05% Tween-20 (PBS-T), containing specific commercial primary antibodies (Santa Cruz Biotechnology Inc., Santa Cruz, CA): mouse anti-Fn1 (fibronectin, 1:200), mouse anti-SerpinE2 (serpin peptidase inhibitor, clade E, member 2, 1:200) and mouse anti-β tubulin (1:1,000). After a washing step in PBS-T, the membranes were incubated with a horseradish peroxidase (HRP)-conjugated secondary anti-mouse antibody (1:5,000) (GE Healthcare). Immunoblots were detected using the ECL-Advance Western Blotting Detection kit (GE Healthcare) by chemiluminescence. The resulting Western blot images were scanned by PDquest 7.1 software. Each experiment was repeated at least three times.

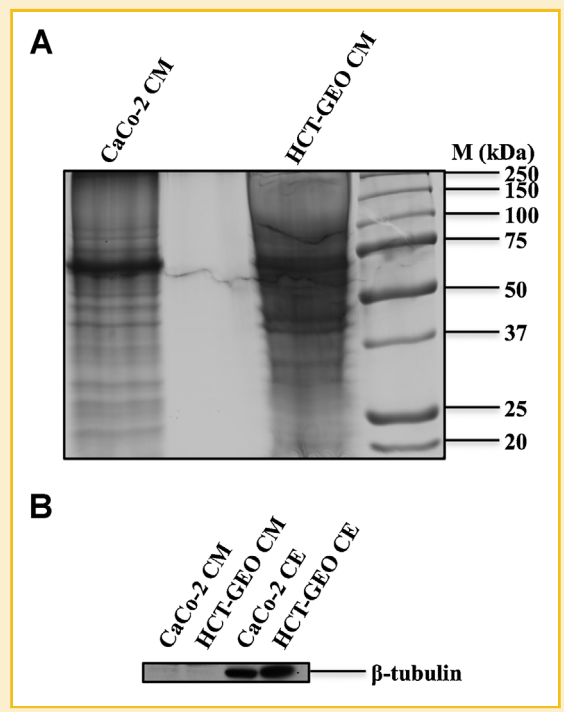

Fig. 1. SDS–PAGE analysis of secreted proteins. *A*: Coomassie-stained 10% SDS–PAGE analysis of the proteins secreted in the conditioned media by CaCo-2 (CaCo-2 CM) and HCT-GEO (HCT-GEO CM) cell cultures. *B*: Secreted proteins in the conditioned media by CaCo-2 (CaCo-2 CM) and HCT-GEO (HCT-GEO CM) cell cultures and cell extracts from both cell lines (CaCo-2 CE and HCT-GEO CE, respectively) were analyzed by Western blot with anti-β tubulin antibody.

RESULTS

IDENTIFICATION OF SECRETED PROTEINS

We analyzed proteins secreted from two human colon cancer cell lines, CaCo-2 and HCT-GEO, in order to evaluate cancer cell secretome as suitable source for evaluating cell-specific properties and searching diagnostic biomarkers. Both cell lines, at the proper confluence, were grown in serum-free medium for 24 h, and the resulting conditioned media (CM) were harvested. After precipitation, proteins were resolved by SDS–PAGE (Fig. 1A). To investigate the occurrence of intracellular protein species within the secretome of CaCo-2 and HCT-GEO cells, we compared the distribution of β tubulin, an abundant cytosolic protein, in the total cell extracts and in the CM of the two cell lines by western blot analysis (Fig. 1B). β-tubulin signal was only detected in the total cell extracts of CaCo-2 and HCT-GEO, whereas its signal was absent in the lanes of the proteins secreted by the two systems. Such a finding implies that the putative contamination of CM by intracellular species is a very limited phenomenon [Wu et al., 2008].

To define the protein secreted by CaCo-2 and HCT-GEO cells, their corresponding gel lanes (Fig. 1A) were entirely cut in very thin slices. The protein bands were excised individually, in-gel digested with trypsin, and analyzed by μLC–MS/MS analysis. Tandem mass spectrometry outputs were analyzed by the Protein Discoverer platform and submitted to label-free quantitation analysis, allowing to compare the expression profile of the secreted proteins between the two cell systems. We identified 122 proteins in the CM of CaCo-2 cell line and 112 species in the CM of HCT-GEO cell line. As shown in Figure 2A, 64 proteins were more expressed in the CM of CaCo-2, 54 species were more expressed in the CM of HCT-GEO, whereas 58 proteins were shared between the two cellular model systems. A total of 176 different species were identified in the analysis and are listed in the Table I which includes the protein names, the corresponding gene names, the gene IDs, and the cancer cell lines where the identified proteins were more expressed. Supplemental material contains the details of label-free quantitation method based on spectral counting for protein abundance estimation.

DATA ANALYSIS

The two secretome datasets were submitted to the Gene Ontology analysis using the GOEAST software. As for CaCo-2, the most represented cellular compartment term was extracellular region (GO: 0005576; *p*-value: 9.13×10^{-10}) containing 30 species; accordingly, in the HCT-GEO protein set, the most significant cellular compartment term was extracellular region part (GO: 0044421; *P*-value: 3.4×10^{-4}) characterized by 18 elements. The species classified as extracellular in the Gene Ontology database are pointed out in Table I.

To evaluate the occurrence of specific structural determinants indicating, or suggesting, the possibility to be secreted through classical or non-classical pathways, each protein was analyzed by

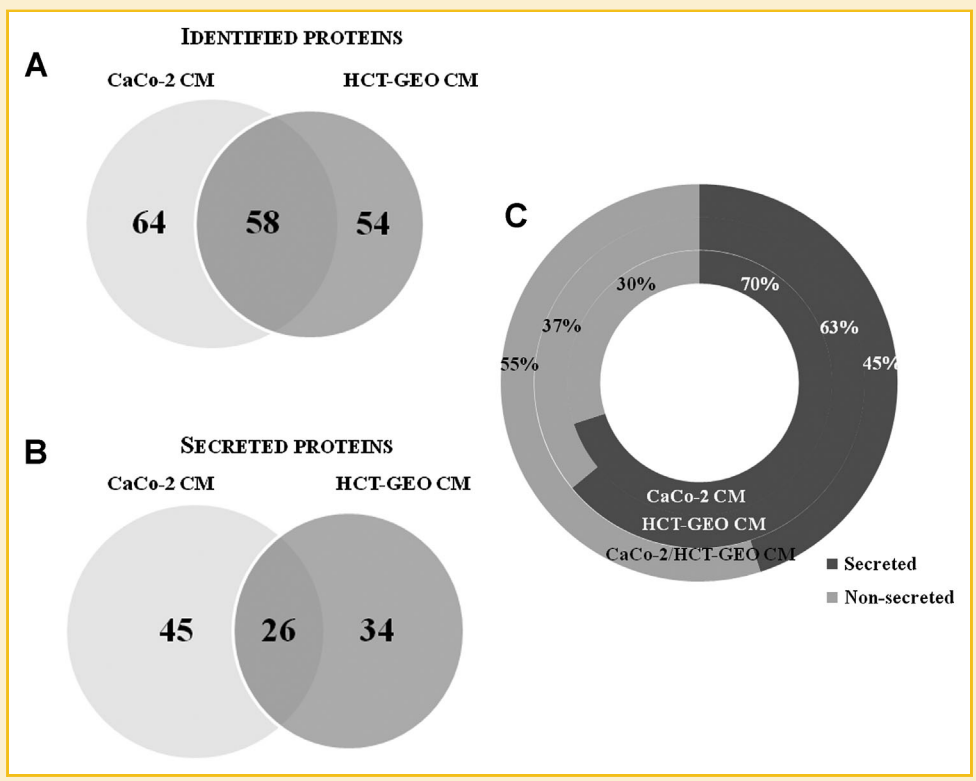

Fig. 2. Output data of the secretome signature of CaCo-2 and HCT-GEO cells. *A*: Overlaps of non-redundant proteins identified in the serum-free conditioned media of CaCo-2 (CaCo-2 CM) and HCT-GEO (HCT-GEO CM) cell lines. A total of 176 proteins were identified: 64 proteins were more expressed in CaCo-2 CM, 54 more expressed in HCT-GEO CM and 58 proteins in the conditioned media from both cell lines (CaCo-2/HCT-GEO). *B*: Overlaps of secreted proteins identified in CaCo-2 CM and HCT-GEO CM. A total of 105 secreted proteins were identified: 45 proteins were more expressed in CaCo-2 CM, 34 more expressed in HCT-GEO CM and 26 proteins in the conditioned media from both cell lines (CaCo-2/HCT-GEO). *C*: Percentage of differentially secreted and non-secreted proteins identified in the secretome of each cell lines or shared by the two cellular model systems.

TABLE I. Proteins Identified in the Conditioned Media of CaCo-2 and HCT-GEO Colon Cancer Cell Lines by LC–MS/MS Analysis

Protein	Gene	Gene ID (NCBI)	CaCo-2	HCT-GEO	GOEAST cellular component	SignalP probability[a]	SecretomeP probability[b]
Fibronectin 1	FN1	119590943	*		Extra-cell	0.598	/
Laminin subunit beta-1	LAMB1	126366	*		Extra-cell	0.846	/
Laminin B2 chain	LAMB2	186964	*		Extra-cell	0.712	/
NrCAM protein	NRCAM	2511666	*			0.547	/
Desmoglein 2	DSG2	416178	*			0.868	/
Glucosidase II	GANAB	2274968	*			0.564	/
E-cadherin	CDH1	31075	*			0.818	/
Dystroglycan	DAG1	398026	*		Extra-cell	0.824	/
LI-cadherin	CDH17	854175	*			0.812	/
HSP105 beta	HSPH1	3970829	*		Extra-cell	/	/
Fibulin-1 isoform D precursor	FBLN1	13661193	*		Extra-cell	0.869	/
Calnexin	CANX	179832	*			0.863	/
Heterogeneous nuclear Ribonucleoprotein L	HNRNPL	11527777	*			/	/
Chaperonin (HSP60)	HSPD1	306890	*		Extra-cell	/	/
Ras GTPase-activating protein-Binding protein 1	G3BP1	5031703	*			/	/
Calpastatin	CAST	951315	*			/	/
Clusterin (apolipoprotein J precursor)	CLU	178855	*		Extra-cell	0.842	/
Calreticulin precursor	CALR	4757900	*		Extra-cell	0.904	/
Alpha-1-antitrypsin	SERPINA1	177827	*		Extra-cell	0.916	/
Aaspartate aminotransferase, cytoplasmic	GOT1	4504067	*			/	/
Cyr61 protein	CYR61	2606094	*		Extra-cell	0.902	/
CKB	CKB	49457530	*			/	/
Actin-like protein 8	ACTL8	227498241	*			/	0.515
Farnesyl pyrophosphate synthetase (EC 2.5.1.1)	FDPS	182399	*			0.498	/
Agmatine ureohydrolase	AGMAT	13477245	*			/	0.876
CHORD containing protein-1	CHORDC1	6581056	*			/	0.521
Vitronectin	VTN	13477169	*		Extra-cell	0.936	/
Albumin, isoform CRA_t	ALB	119626083	*		Extra-cell	0.848	/
60S acidic ribosomal protein P0	RPLP0	4506667	*			/	/
cyclophilin	PPIB	181250	*			/	0.853
Cathepsin L2	CTSL2	3087790	*			0.834	/
Carboxyl terminal LIM domain protein	PDLIM1	1905874	*			/	/
Apolipoprotein E	APOE	178849	*		Extra-cell	0.919	/
Guanine nucleotide-binding protein subunit beta-2-like 1	GNB2L1	4503607	*			/	/
Electron transfer flavoprotein subunit alpha, mitochondrial isoform a	ETFA	2781202	*			/	0.599
Cathepsin X precursor	CTSZ	3650498	*		Extra-cell	0.845	/
Stanniocalcin-2 precursor	STC2	4507267	*		Extra-cell	0.724	/
Beta-subunit (AA 1-312)	ATP5B	28931	*			/	0.591
hnRNP JKTBP	HNRPDL	2780748	*			/	/
Inorganic pyrophosphatase	PPA1	11056044	*			/	/
F-actin-capping protein subunit alpha-1	CAPZA1	5453597	*		Extra-cell	/	/
Nucleophosmin isoform 1	NPM1	10835063	*			/	0.811
Thioredoxin-like protein 1	TXNL1	4759274	*			/	/
Voltage-dependent anion channel 1	VDAC1	4507879	*			/	/
Dickkopf-related protein 1 precursor	DKK1	7110719	*		Extra-cell	0.584	/
Proapolipoprotein	APOA1	178775	*		Extra-cell	0.847	/
Hepatoma-derived growth factor isoform a	HDGF	4758516	*		Extra-cell	/	/
Nascent polypeptide-associated complex subunit alpha-2	NACA2	40548328	*			/	/
Peroxiredoxin-6	PRDX6	4758638	*			/	/
Ribosomal protein L10a	RPL10A	531171	*			/	0.637
Isopentenyl-diphosphate delta-isomerase (EC 5.3.3.2) homolog-human	IDI1	539623	*			/	/
Ubquitin carboxy-terminal hydrolase L1	UCHL1	4185720	*			/	0.530
Transgelin variant	TAGLN	62897565	*			/	0.560
Thioredoxin-dependent peroxide reductase, mitochondrial isoform a precursor	PRDX3	5802974	*			/	0.756
Translin	TSN	4759270	*			/	/
Ribosomal protein S9	RPS9	550023	*			/	/
Ribosomal protein S5	RPS5	550021	*			/	0.780
Tissue inhibitor of metalloproteinases-2	TIMP2	1517893	*		Extra-cell	0.938	/
Ferritin light subunit	FTL	182516	*			/	/
Gamma-glutamylcyclotransferase	GGCT	13129018	*			/	0.503
Heterogeneous nuclear ribonucleoprotein A/B isoform a	HNRNPAB	55956919	*			/	/
Small nuclear ribonucleoprotein polypeptide C	SNRPC	306875	*			/	/
Histone H1b	HIST1H1B	356168	*			/	/
Endoplasmic reticulum resident protein 29 isoform 1 precursor	ERP29	5803013	*			0.826	/

(*Continued*)

TABLE I. (Continued)

Protein	Gene	Gene ID (NCBI)	CaCo-2	HCT-GEO	GOEAST cellular component	SignalP probability[a]	SecretomeP probability[b]
Myosin-9	MYH9	12667788		*		0.837	/
Clathrin heavy chain 1	CLTC	4758012		*		0.707	/
Tenascin C	TNC	553348		*	Extra-cell	/	0.530
ATP-citrate (pro-S-)-lyase	ACLY	28935		*		0.864	/
Ubiquitin activating enzyme E1	UBA1	35830		*		/	/
Calsyntenin 1	CLSTN1	21706696		*		0.830	/
IQGAP1 IQ motif containing GTPase activating protein 1	IQGAP1	40674640		*		/	/
Transportin 1	TNPO1	5107636		*		0.830	/
Brain glycogen phosphorylase	PYGB	307200		*		/	0.540
Importin subunit beta-1	KPNB1	19923142		*		/	/
Sushi repeat-containing protein SRPX isoform 2	SRPX	1699163		*		0.864	/
Phosphofructokinase-P	PFKM	2344769		*		/	/
Transitional endoplasmic reticulum ATPase	VCP	6005942		*		0.785	/
hnRNP U protein	HNRNPU	32358		*		0.916	/
Elongation factor 2	EEF2	4503483		*		/	/
Transketolase	TKT	37267		*		/	/
Protein disulfide isomerase-related protein 5	PDIA5	1710248		*		0.877	/
T-complex protein 1 subunit beta	CCT2	5453603		*		/	/
Splicing factor, proline- and glutamine-rich	SFPQ	4826998		*		/	/
Galectin-3-binding protein	LGALS3BP	119609949		*	Extra-cell	0.827	/
CD44E	CD44	29801		*		/	0.588
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 precursor	PLOD3	4505891		*		/	/
Lamin A/C, isoform CRA_a	LMNA	119573381		*		/	/
Polyadenylate-binding protein 4 isoform 2	PABPC4	4504715		*		/	/
Heat shock 70 kDa protein 9	HSPA9	292059		*		/	/
T-complex protein 1 subunit eta isoform a	CCT7	5453607		*		/	/
Vimentin	VIM	340219		*		/	/
T-complex polypeptide 1	CCT3	36796		*		/	0.512
Lysosomal pepstatin insensitive protease	TPP1	2408232		*		/	/
RecName: full = beta-hexosaminidase subunit alpha	HEXA	123079		*		0.723	/
adenylyl cyclase-associated protein 1	CAP1	5453595		*		0.916	/
UDP-*N*-acetylglucosamine pyrophosphorylase	UAP1	3273316		*		/	/
ATPase, H+ transporting, lysosomal 50/57 kDa, V1 subunit H	ATP6V1H	4680661		*		/	0.526
Protein phosphatase 1F	PPM1F	47940633		*		/	/
U2 small nuclear RNA auxiliary factor 2	U2AF2	228543		*		/	0.507
Septin 7	SEPT7	560623		*		/	/
Serpin peptidase inhibitor, clade E, member 2	SERPINE2	24307907		*	Extra-cell	0.707	/
Osteonectin	SPARC	338325		*	Extra-cell	0.939	/
Gelsolin-like capping protein isoform 9	CAPG	55597035		*		/	/
Eukaryotic initiation factor 4AII	EIF4A2	485388		*		/	/
HLA-A protein	HLA-A	17512578		*		0.813	/
Pirin	PIR	4505823		*		/	/
Cathepsin D preproprotein	CTSD	4503143		*	Extra-cell	0.781	/
Heat shock protein 27	HSPB1	4504517		*		/	0.740
Annexin A1	ANXA1	4502101		*	Extra-cell	/	0.511
Scaffold protein Pbp1	SDCBP	1916850		*		/	/
Insulin-like growth factor-binding Protein 2 precursor	IGFBP2	55925576		*	Extra-cell	0.648	/
methylthioadenosine phosphorylase	MTAP	847724		*		/	/
N-acetyl-beta-glucosaminidase prepro-polypeptide	HEXB	179462		*		/	0.712
Ras suppressor protein 1 isoform 1	RSU1	6912638		*		/	0.562
Protein Dj-1	DJ-1	31543380		*		/	/
60S Ribosomal protein L12	RPL12	4506597		*		/	0.865
Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 1	PLOD1	190074		*	Extra-cell	0.910	/
Chloride intracellular channel 1	CLIC1	895845		*		/	/
Alpha-enolase	ENO1	4503571	*	*		/	0.536
Heterogeneous nuclear Ribonucleoprotein A3	HNRNPA3	34740329	*	*		/	/
Laminin-binding protein	RPSA	34234	*	*		/	0.620
Proteasome subunit alpha type-6	PSMA6	8394076	*	*		/	/
Peroxiredoxin-4	PRDX4	5453549	*	*	Extra-cell	0.783	/
Profilin-1	PFN1	4826898	*	*		/	/
14-3-3 protein epsilon	YWHAQ	5803225	*	*		/	/
Poly(rC) binding protein 1	PCBP1	444021	*	*		/	0.549
Heterogeneous nuclear Ribonucleoprotein A1 isoform a	HNRNPA1	4504445	*	*		/	/
Chromobox homolog 3	CBX3	7416937	*	*		/	0.836
Alpha-actinin-1 isoform b	ACTN1	4501891	*	*		/	/
moesin	MSN	4505257	*	*		/	0.530
S-adenosylhomocysteine hydrolase	AHCY	178277	*	*		/	/
Purine nucleoside phosphorylase	PNP	387033	*	*		/	0.509
Proteasome subunit beta type-1 precursor	PSMB1	4506193	*	*		/	/
Phospholipase C-alpha	PDIA3	303618	*	*		0.887	/
Endoplasmin precursor	HSP90B1	4507677	*	*		0.915	/
X-ray repair cross-complementing protein 5	XRCC5	10863945	*	*		/	/
Transgelin-2	TAGLN2	4507357	*	*		/	0.784

(*Continued*)

TABLE I. (Continued)

Protein	Gene	Gene ID (NCBI)	CaCo-2	HCT-GEO	GOEAST cellular component	SignalP probability[a]	SecretomeP probability[b]
Alpha actinin 4	ACTN4	2804273	*	*	Extra-cell	/	/
L-lactate dehydrogenase B chain	LDHB	4557032	*	*		/	0.569
Aldolase A	ALDOA	28614	*	*	Extra-cell	/	/
Elongation factor 1-alpha 1	EEF1A2	4503471	*	*		/	/
Nucleolin	NCL	189306	*	*		/	/
Cytovillin 2	EZR	340217	*	*		/	0.563
Heat shock protein HSP 90-alpha 2	HSP90AA1	61656603	*	*		/	/
Triosephosphate isomerase isoform 1	TPI1	4507645	*	*		/	/
Glyceraldehyde-3-phosphate dehydrogenase	GAPDH	31645	*	*		/	/
Heat shock cognate 71 kDa protein isoform 1	HSPA8	5729877	*	*	Extra-cell	/	/
Phosphoglycerate kinase 1	PGK1	4505763	*	*		/	/
L-lactate dehydrogenase A chain isoform 1	LDHA	5031857	*	*		/	0.549
Beta-actin	ACTB	14250401	*	*		/	/
Glutathione S-transferase-P1c	GSTP1	2204207	*	*		/	0.546
GRP78 precursor	HSPA5	386758	*	*		0.898	/
4F2 Cell-surface antigen heavy chain	SLC3A2	177207	*	*		/	0.644
Pyruvate kinase type M2	PKM2	189998	*	*		/	0.590
Stress-induced-phosphoprotein 1	STIP1	5803181	*	*		/	/
T-complex protein 1 subunit epsilon	CCT5	24307939	*	*		/	/
T-complex protein 1 subunit theta	CCT8	48762932	*	*		/	0.508
Protein disulfide isomerase	P4HB	860986	*	*	Extra-cell	0.794	/
Polypyrimidine tract-binding protein 1 isoform a	PTBP1	4506243	*	*		/	/
Colligin	SERPINH1	30130	*	*	Extra-cell	0.890	/
6-Phosphogluconate dehydrogenase, decarboxylating	PGD	40068518	*	*		/	/
Human rab GDI	GDI1	285975	*	*		/	/
Actin, cytoplasmic 2	ACTG1	4501887	*	*		/	0.505
Proteasome subunit p42	PSMC6	1526426	*	*		/	0.619
Heterogeneous nuclear ribonucleoprotein H	HNRNPH1	5031753	*	*		/	/
Heterogeneous nuclear ribonucleoproteins A2/B1 isoform A2	HNRNPA2B1	4504447	*	*		/	/
Malate dehydrogenase precursor	MDH1	2906146	*	*		/	/
Human elongation factor-1-delta	EEF1D	38522	*	*		/	0.529
F-actin-capping protein subunit beta	CAPZB	4826659	*	*		/	0.583
Annexin A2 isoform 2	ANXA2	4757756	*	*	Extra-cell	/	0.746
Tropomyosin alpha-3 chain isoform 2	TPM3	24119203	*	*		/	/
Annexin A5	ANXA5	4502107	*	*	Extra-cell	/	0.550
14-3-3 Protein zeta/delta	YWHAZ	4507953	*	*		/	/
peroxiredoxin-1	PRDX1	4505591	*	*		/	0.528
GPI glucose-6-phosphate isomerase	GPI	6653226	*	*	Extra-cell	/	/
Heterogeneous nuclear ribonucleoprotein K	HNRNPK	55958547	*	*		/	/

[a]Signal peptides were predicted using SignalP 4.0 (proteins with D-cutoff score $\geq$0.45 are considered secreted).
[b]Non-classical secretion of proteins was evaluated by the neural network output score of SecretomeP 2.0 (proteins with NN score $\geq$0.5 are considered secreted).

using the SignalP or the SecretomeP softwares. The SignalP 4.0 software allowed defining the species that are secreted through the classical ER/Golgi pathway. All the proteins, lacking the presence of the classical signal peptide for the translocation to the ER, and having a score lower than the D-cutoff, were then tested by SecretomeP 2.0 software for the putative export through one of the so-called non-classical secretory pathways. At this regard, Table I shows the scores assigned to secreted proteins according to one of the two prediction softwares. The proteins that did not match the criteria of either SignalP (D-cutoff score $\geq$0.45) or SecretomeP softwares (NN-score $>$0.5) do not show any score in the Table I. As for CaCo-2 cells, the in silico analyses showed that 63 out of 122 identified species are secreted proteins; as for HCT-GEO, 56 out of 112 species are secreted.

By merging the information acquired by means of GO annotation and in silico analyses we defined 71 secreted proteins (58%) in CaCo-2 CM and 60 secreted species (54%) in HCT-GEO CM. Among them, 45 proteins (70%) were more expressed in CaCo-2 cells, 34 species (63%) in HCT-GEO cells and 26 elements (45%) were shared between the two cellular systems. A total of 105 secreted proteins were identified. These results are summarized in Figure 2B and C.

The two protein datasets were also analyzed in order to cluster the secreted species into functional annotation terms. To this aim, we performed functional classification of the identified proteins according to the DAVID annotation system. Table II shows the analysis of the 122 proteins identified in the CM of CaCo-2 cells. The most significative biological process is the "glucose catabolic process" (P-value 4.91×10^{-11}) defined by 11 species and the most

TABLE II. Biological Process and Molecular Function Categories Over Represented in the Secretome Signature of Caco-2 Cells According to David Software

Categories	No. of proteins	P-value
Biological process		
Glucose catabolic process	11	4.91×10^{-11}
Negative regulation of apoptosis	13	4.69×10^{-5}
Regulation of apoptosis	19	1.44×10^{-4}
RNA splicing, via transesterification reactions	9	4.74×10^{-5}
Nicotinamide nucleotide metabolic process	5	3.58×10^{-4}
Cellular carbohydrate biosynthetic process	4	2.0×10^{-2}
Hydrogen peroxide catabolic process	3	9.0×10^{-3}
Cholesterol biosynthetic process	3	2.0×10^{-2}
Blood vessel morphogenesis	6	3.4×10^{-2}
Reverse cholesterol transport	3	7.95×10^{-3}
Cholesterol metabolic process	4	4.36×10^{-2}
Negative regulation of protein complex assembly	3	3.37×10^{-2}
Molecular function		
Peroxiredoxin activity	4	3.77×10^{-5}
Protein disulfide isomerase activity	3	2.73×10^{-3}

Categories are sorted by P-value ($P \leq 0.05$).

significative molecular function is the "peroxiredoxin activity" (P-value 3.77×10^{-5}) defined by four proteins. Similarly, the analysis of the 112 proteins identified in the CM of HCT-GEO cells, reported in Table III, shows that the most significative biological process is the "glucose catabolic process" (P-value 3.46×10^{-9}) defined by 13 species and the most significative molecular function is the "protein disulfide isomerase activity" (P-value 2.24×10^{-3}) defined by three proteins. Tables II and III share several overlapping functional categories due to the presence of a common protein dataset between the two secretomes under investigation. At the same time, in CaCo-2 secretome five proteins (ApoA1, ApoE, Clu, Fdps, and Idi1), not shared by HCT-GEO system, are responsible for the functional terms "reverse cholesterol transport" and "cholesterol biosynthetic process"; the CM of HCT-GEO, on the other hand, is characterized by the presence of four proteins (Tpp1, Hexa, Hexb, and Ctsd), not shared by CaCo-2 cell line, known to be involved in the "vacuole organization". These findings indicate that cell-specific functional categories are also found, thus emphasizing the occurrence of selected biological processes in each of the two systems.

VALIDATION OF SECRETED PROTEINS BY WESTERN BLOTTING

To confirm the results obtained from LC–MS/MS analysis, Western blot analyses were used to verify the expression of selected secreted proteins in the serum-free CM of CaCo-2 and HCT-GEO. In particular, we evaluated the expression of fibronectin (Fn1), exclusively identified in CaCo-2 dataset, and serpin peptidase inhibitor, member 2 (SerpinE2), exclusively found in HCT-GEO secretome. As shown in Figure 3, Fn1 was only detected in the CM of CaCo-2 whereas SerpinE2 signal was only present in the CM of HCT-GEO, in agreement with mass spectrometry data.

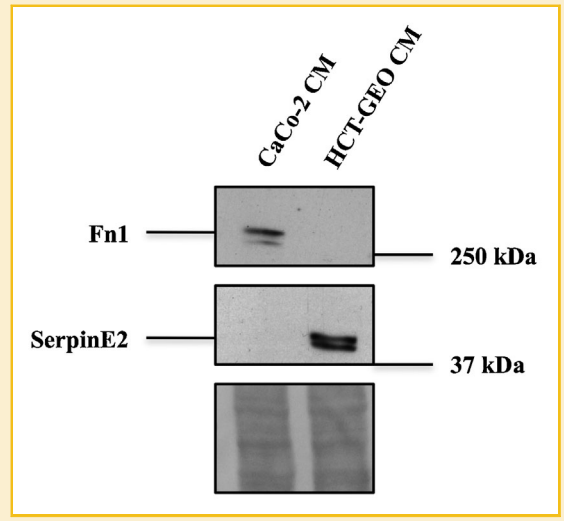

Fig. 3. Validation of proteomic data. Western blot analysis of Fn1 and SerpinE2 identified in the serum-free conditioned media of CaCo-2 (CaCo-2 CM) and HCT-GEO (HCT-GEO CM) cell lines. Protein samples (15 μg/lane) were separated in SDS–PAGE, transferred onto a nitrocellulose membrane and then probed with specific antibodies against the indicated target proteins. As protein loading controls, PonceauS-stained protein profiles from the CM of both cell lines were used.

DISCUSSION

Adenocarcinoma with moderate to well differentiated histopathology is the most common type of colon cancer in humans [Ji et al., 2011]. CaCo-2 and HCT-GEO are two well-known cellular model systems for the in vitro study of this kind of tumor because they are able to well represent differentiated colonic carcinoma in terms of specific functional features [Normanno et al., 1996; Laska et al., 2002]. Interestingly, both cell lines form glandular like tumors in nude mice, have low cloning efficiency in soft agar, and are high secretors of carcinoembryonic antigen (CEA) [Chantret et al., 1988; de Bruïne et al., 1993; Gemei et al., 2013]. On the other hand, CaCo-2 cells are able to form upon confluence a polarized monolayer with tight junctions, while HCT-GEO cells do the same only under glucose starvation [Chantret et al., 1988; Yakovich et al., 2010]. Based on these features we decided to use CaCo-2 and HCT-GEO as colon adenocarcinoma in vitro model systems for novel biomarkers discovery potentially useful for patient diagnosis and monitoring.

In colon cancer research there is a constant need to identify novel biomarkers useful for early diagnosis and prognosis. It's well known that secreted proteins might represent putative tumor biomarkers and are potential tools to define the aggressiveness of the tumor. The present study focused on the proteomic characterization of the serum-free CM of the above quoted colon cancer cell lines. We identified more than 170 protein species, 64 more expressed in the secretome of CaCo-2 cells, 54 more expressed in the secretome of HCT-GEO cells and 58 shared by the two systems, confirming that the two cell lines are characterized by common as well as cell-specific features. Among them, more than 50% were deemed to be secretory according to their Gene Ontology annotation and/or their SignalP or SecretomeP scores (58% for CaCo-2, and 54% for HCT-GEO), in agreement with other secretome studies [Kashat et al., 2010; Ji et al., 2011]. Interestingly, several identified species are not annotated as secreted within the Gene Ontology database but show a good score in one of the two used prediction softwares (Table I). On the other hand, the in silico analysis produced some false negative data, as it

TABLE III. Biological Process and Molecular Function Categories Over Represented in the Secretome Signature of HCT-GEO Cells According to DAVID Software

Categories	No. of proteins	P-value
Biological process		
Glucose catabolic process	13	3.46×10^{-9}
Actin cytoskeleton organization	10	8.66×10^{-5}
Nicotinamide nucleotide metabolic process	5	2.91×10^{-4}
RNA splicing	10	4.76×10^{-4}
Regulation of apoptosis	16	1.94×10^{-3}
Negative regulation of apoptosis	10	2.23×10^{-3}
Monosaccharide biosynthetic process	4	3.24×10^{-3}
Vacuole organization	4	4.05×10^{-3}
Proteasomal protein catabolic process	5	9.26×10^{-3}
Intracellular protein transport	9	1.06×10^{-2}
Lysosome organization	3	1.71×10^{-2}
Molecular function		
Protein disulfide isomerase activity	3	2.24×10^{-3}
Phospholipase inhibitor activity	3	4.05×10^{-3}
Ribonucleotide binding	26	5.35×10^{-3}

Categories are sorted by P-value ($P \leq 0.05$).

happened for few species already classified as extracellular. Such findings indicate that in the secretome studies more than one tool for "data analysis" must be evaluated.

The 58 proteins identified in the common subset are characterized by the presence of several components of the "glucose catabolic process"; some of them, classified as "not secreted", could be of exosomial origin according to the Exocarta database (www.exocarta.org). In fact, the protein export outside the cells can occur through various types of membrane-associated vesicles [Simpson et al., 2008]. These membranous extracellular organelles include exosomes, shedding microvesicles (SMVs) and apoptotic blebs (ABs) [Mathivanan et al., 2010]. Exosomes are reported to be secreted into the conditioned media of various mammalian cell lines and include both a common set of molecules (such as metabolic enzymes, cytoskeletal proteins, heat shock proteins and chaperones, and so on) [Simpson et al., 2008] and cell-type specific components. However, the lipid double layer membrane seems to play a key role in the exosome-cell communication; this feature adds a further level of complexity in the secretome analysis, which requires more investigations.

The label-free quantitation method showed that several secreted protein are differentially expressed between the two colorectal cancer cell lines. In particular, some of them seem to exclusively characterize the CaCo-2 or the HCT-GEO CM. As an example, both Western blot and label-free quantitation analyses confirmed that fibronectin (Fn1), involved in cell adhesion and motility, was uniquely secreted in CaCo-2 cell line; whereas serpin peptidase inhibitor, clade E (SerpinE2), implicated in impaired extracellular matrix (ECM) degradation and, consequently, in cancer invasion [Fayard et al., 2009], was exclusively detected in HCT-GEO CM. Interestingly, several proteins known to be ECM components and involved in cell adhesion were more expressed in CaCo-2 CM; moreover, several matricellular proteins previously reported as implicated in cancer invasion were more expressed in HCT-GEO CM. Besides the validated Fn1, we identified in CaCo-2 the following cell adhesion proteins, most of them already related to CRC: laminin subunit beta-1 (Lamb1) [Fujita et al., 2005], laminin B2 chain (Lamb2) [Fujita et al., 2005], dystroglycan (Dag1) [Sgambato et al., 2003], fibulin-1 isoform D precursor (Fbln1) [Kanda et al., 2011], neurone glial-related cell adhesion molecule (NrCAM) [Conacci-Sorrell et al., 2002], desmoglein 2 (Dsg2) [Kolegraff et al., 2011], E-cadherin (Cdh1) [Kroepil et al., 2012], LI-Cadherin (Cdh17) [Park et al., 2011], and guanine nucleotide-binding protein subunit beta-2-like (Gnb2l1) [Swaminathan and Cartwright, 2012]. On the other hand, the set of differentially expressed proteins secreted by HCT-GEO and involved in tumor invasion includes, besides SerpinE2, osteonectin (Sparc) [Girotti et al., 2011], galectin-3-binding protein (Lgals3bp) [Chen et al., 2011] [Wu et al., 2008], vimentin (Vim) [Cho et al., 2012], tenascin C (Tnc) [De Wever et al., 2004], myosin-9 (Myh9) [Scamuffa et al., 2011], IQGAP1 IQ motif containing GTPase activating protein 1 (Iqgap1) [Hayashi et al., 2010], protein phosphatase 1F (Ppm1f) [Susila et al., 2010], annexin A1 (Anxa1) [Sato et al., 2011] and protein DJ-1 (Dj-1) [He et al., 2012].

According to the recent secretome analyses focused on the protein markers of tumor aggressiveness, our results provide a list of candidates potentially associated with the invasive capacities of CaCo-2 and HCT-GEO cell lines. In fact, we can, speculate that HCT-GEO releases pro-invasive species into the tumor microenvironment, whereas CaCo-2 secretome has pro-adhesion features. This hypothesis is consistent with the described epithelial characteristics and tumorigenic properties of both cell lines. In particular, it's known that CaCo-2 cells, when grown on permeable supports, form a confluent monolayer with properties characteristic of differentiated absorptive epithelial cells [Chantret et al., 1988]. Contrary, HCT-GEO cells are undifferentiated when grown under standard culture conditions (i.e., in the presence of glucose) and differentiated in the absence of glucose [Chantret et al., 1988]. The tumorigenic features of the two cell lines are in agreement with our findings: in fact, HCT-GEO cell line is currently used to generate the orthotopic model of human colon cancer [Bhattacharya et al., 2001]; whereas CaCo-2 showed the lowest growth and metastatic capacities in comparison to other CRC cell lines implanted in mice [Flatmark et al., 2004].

As for the search of putative CRC biomarker, among the 176 species listed in Table I, we could distinguish between proteins already related to CRC detection and species proposed as markers for other types of carcinomas. Within the first set, we identified clusterin (Clu) [Rodríguez-Piñeiro et al., 2012], Cyr61 protein (Cyr61) [Lau, 2011], annexin A5 (AnxA5) [Xue et al., 2009], cyclophilin B (Ppib) [Kim et al., 2011], glucose-6-phosphate isomerase (Gpi) [Tsutsumi et al., 2009], besides Lgals3bp, SerpinE2, Sparc quoted above. By comparing these data with secreted proteins by three other CRC cell lines (Colo205, SW480, and SW620) [Wu et al., 2010], it is noteworthy to observe that some secreted CRC biomarkers are shared, such as Clu, AnxA5, Ppib, Gpi, Lgals3bp, and SerpinE2. On the other hand, some others (Cyr61 and Sparc) are more specific to the site/lesion of the colon cancer cell line. In fact, Cyr61, more expressed in the CaCo-2 CM, is secreted only by SW480, whereas Sparc, more expressed in HCT-GEO CM, is not secreted in the above quoted CRC cell lines. Cyr61 is a multifunctional extracellular protein whose expression is altered in cancer [Lau, 2011]; in CRC, high-levels of Cyr61 mRNA occur at an early stage of tumor development, whereas protein expression is reduced in more advanced stages [Lau, 2011]. Sparc is a matricellular protein involved in tissue remodeling, cell migration and angiogenesis, whose role in cancer seems to be tissue-specific. In CRC, high Sparc is associated with better disease outcome in stage II and may be prognostic indicators of cancer-specific survival [Chew et al., 2011]. Thus, our study points out that colon cancer cell line CM is a reservoir of two type of putative biomarkers: those shared by a wide spectrum of CRC cell lines and others that are apparently more associated to a cancer cell type.

Moreover, the CM of CaCo-2 and HCT-GEO were also characterized by putative cancer biomarkers not already related to CRC, such as vitronectin (Vtn) [Kadowaki et al., 2011], colligin (SerpinH1) [Mustafa et al., 2010], annexin A2 (AnxA2) [Wu et al., 2012] and CD44 molecule (Cd44). As for Cd44, a cell-surface glycoprotein highly expressed in primary and metastatic colon cancer [Cho et al., 2012], Van der Berg's group found that elevated levels of Cd44 and of other six secreted factors occurred in Hodgkin and Reed-Sternberg patient plasma [Ma et al., 2008]. Interestingly, CD44 is also found secreted in the Colo205 secretome [Wu et al., 2010] and further studies are needed to clarify its role in the extracellular compartment.

In conclusion, this study characterizes the secretome signature of CaCo-2 and HCT-GEO cell lines and confirms the potential of a cell

culture-based model in order to describe the cell-specific invasive properties and to provide a list of putative cancer biomarkers. In fact, we speculate that HCT-GEO secretome contains several pro-invasion species whereas different adhesion proteins are identified in CaCo-2 CM. As for biomarker search, we find putative candidates both in a cell specific secretome and in the subset of secreted species shared by two cell lines. In particular, we confirm already known putative CRC biomarkers, but also we indicate new secreted proteins that could be tested on the sera from CRC patients.

ACKNOWLEDGMENTS

This work was supported by a grant from Fondazione SDN-IRCCS to SO, grant 10737 from AIRC to LDV, and by Progetto FIRB "Italian Human Proteomenet" RBRN07BMCT. The authors thank Vittorio Lucignano, CEINGE—Biotecnologie Avanzate, for technical assistance.

REFERENCES

Bhattacharya A, Turowski SG, San Martin ID, Rajput A, Rustum YM, Hoffman RM, Seshadri M. 2001. Magnetic resonance and fluorescence-protein imaging of the anti-angiogenic and anti-tumor efficacy of selenium in an orthotopic model of human colon cancer. Anticancer Res 31:387–393.

Bendtsen JD, Jensen LJ, Blom N, von Heijne G, Brunak S. 2004. SecretomeP: Feature based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel 17:349–356.

Caterino M, Corbo C, Imperlini E, Armiraglio M, Pavesi E, Aspesi A, Loreni F, Dianzani I, Ruoppolo M. 2013. Differential proteomic analysis in human cells subjected to ribosomal stress. Proteomics 13:1220–1227.

Chantret I, Barbat A, Dussaulx E, Brattain MG, Zweibaum A. 1988. Epithelial polarity, villin expression, and enterocytic differentiation of cultured human colon carcinoma cells: A survey of twenty cell lines. Cancer Res 48:1936–1942.

Chen R, Tan Y, Wang M, Wang F, Yao Z, Dong L, Ye M, Wang H, Zou H. 2011. Development of glycoprotein capture-based label-free method for the high-throughput screening of differential glycoproteins in hepatocellular carcinoma. Mol Cell Proteomics 10:1–13.

Chew A, Salama P, Robbshaw A, Klopcic B, Zeps N, Platell C, Lawrance IC. 2011. SPARC, FOXP3, CD8, and CD45 correlation with disease recurrence and long-term disease-free survival in colorectal cancer. PLoS ONE 6:e22047.

Cho SH, Park YS, Kim HJ, Kim CH, Lim SW, Huh JW, Lee JH, Kim HR. 2012. CD44 enhances the epithelial-mesenchymal transition in association with colon cancer invasion. Int J Oncol 41:211–218.

Conacci-Sorrell ME, Ben-Yedidia T, Shtutman M, Feinstein E, Einat P, Ben-Ze′ev A. 2002. Nr-CAM is a target gene of the beta-catenin/LEF-1 pathway in melanoma and colon cancer and its expression enhances motility and confers tumorigenesis. Genes Dev 16:2058–2072.

Corbo C, Orrù S, Gemei M, Noto RD, Mirabelli P, Imperlini E, Ruoppolo M, Vecchio LD, Salvatore F. 2012. Protein cross-talk in CD133+ colon cancer cells indicates activation of the Wnt pathway and upregulation of SRp20 that is potentially involved in tumorigenicity. Proteomics 12:2045–2059.

de Bruïne AP, de Vries JE, Dinjens WN, Moerkerk PT, van der Linden EP, Pijls MM, ten Kate J, Bosman FT. 1993. Human Caco-2 cells transfected with c-Ha-Ras as a model for endocrine differentiation in the large intestine. Differentiation 53:51–60.

De Wever O, Nguyen QD, Van Hoorde L, Bracke M, Bruyneel E, Gespach C, Mareel M. 2004. Tenascin-C and SF/HGF produced by myofibroblasts in vitro provide convergent pro-invasive signals to human colon cancer cells through RhoA and Rac. FASEB J 18:1016–1018.

Fayard B, Bianchi F, Dey J, Moreno E, Djaffer S, Hynes NE, Monard D. 2009. The serine protease inhibitor protease nexin-1 controls mammary cancer metastasis through LRP-1-mediated MMP-9 expression. Cancer Res 69:5690–5698.

Flatmark K, Maelandsmo GM, Martinsen M, Rasmussen H, Fodstad Ø. 2004. Twelve colorectal cancer cell lines exhibit highly variable growth and metastatic capacities in an orthotopic model in nude mice. Eur J Cancer 40:1593–1598.

Fujita M, Khazenzon NM, Bose S, Sekiguchi K, Sasaki T, Carter WG, Ljubimov AV, Black KL, Ljubimova JY. 2005. Overexpression of beta1-chain-containing laminins in capillary basement membranes of human breast cancer and its metastases. Breast Cancer Res 7:R411–R421.

Gemei M, Mirabelli P, Di Noto R, Corbo C, Iaccarino A, Zamboli A, Troncone G, Galizia G, Lieto E, Del Vecchio L, Salvatore F. 2013. CD66c is a novel marker for colorectal cancer stem cell isolation, and its silencing halts tumor growth in vivo. Cancer 119:729–738.

Girotti MR, Fernández M, López JA, Camafeita E, Fernández EA, Albar JP, Benedetti LG, Valacco MP, Brekken RA, Podhajcer OL, Llera AS. 2011. SPARC promotes cathepsin B-mediated melanoma invasiveness through a collagen I/α2β1 integrin axis. J Invest Dermatol 131:2438–2447.

Hayashi H, Nabeshima K, Aoki M, Hamasaki M, Enatsu S, Yamauchi Y, Yamashita Y, Iwasaki H. 2010. Overexpression of IQGAP1 in advanced colorectal cancer correlates with poor prognosis-critical role in tumor invasion. Int J Cancer 126:2563–2574.

He X, Zheng Z, Li J, Ben Q, Liu J, Zhang J, Ji J, Yu B, Chen X, Su L, Zhou L, Liu B, Yuan Y. 2012. DJ-1 promotes invasion and metastasis of pancreatic cancer cells by activating SRC/ERK/uPA. Carcinogenesis 33:555–562.

Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ. 2008. Cancer statistics. CA Cancer J Clin 58:71–96.

Ji H, Goode RJ, Vaillant F, Mathivanan S, Kapp EA, Mathias RA, Lindeman GJ, Visvader JE, Simpson RJ. 2011. Proteomic profiling of secretome and adherent plasma membranes from distinct mammary epithelial cell subpopulations. Proteomics 11:4029–4039.

Kadowaki M, Sangai T, Nagashima T, Sakakibara M, Yoshitomi H, Takano S, Sogawa K, Umemura H, Fushimi K, Nakatani Y, Nomura F, Miyazaki M. 2011. Identification of vitronectin as a novel serum marker for early breast cancer detection using a new proteomic approach. J Cancer Res Clin Oncol 137:1105–1115.

Kanda M, Nomoto S, Okamura Y, Hayashi M, Hishida M, Fujii T, Nishikawa Y, Sugimoto H, Takeda S, Nakao A. 2011. Promoter hypermethylation of fibulin 1 gene is associated with tumor progression in hepatocellular carcinoma. Mol Carcinog 50:571–579.

Kashat L, So AK, Masui O, Wang XS, Cao J, Meng X, Macmillan C, Ailles LE, Siu KW, Ralhan R, Walfish PG. 2010. Secretome-based identification and characterization of potential biomarkers in thyroid cancer. J Proteome Res 9:5757–5769.

Kim Y, Jang M, Lim S, Won H, Yoon KS, Park JH, Kim HJ, Kim BH, Park WS, Ha J, Kim SS. 2011. Role of cyclophilin B in tumorigenesis and cisplatin resistance in hepatocellular carcinoma in humans. Hepatology 54:1661–1678.

Klein-Scory S, Kubler S, Diehl H, Eilert-Micus C, Reinacher-Schick A, Stuhler K, Warscheid B, Meyer HE, Schmiegel W, Schwarte-Waldhoff I. 2010. Immunoscreening of the extracellular proteome of colorectal cancer cells. BMC Cancer 10:70–88.

Kolegraff K, Nava P, Helms MN, Parkos CA, Nusrat A. 2011. Loss of desmocollin-2 confers a tumorigenic phenotype to colonic epithelial cells through activation of Akt/β-catenin signaling. Mol Biol Cell 22:1121–1134.

Kroepil F, Fluegen G, Totikov Z, Baldus SE, Vay C, Schauer M, Topp SA, Esch JS, Knoefel WT, Stoecklein NH. 2012. Down-regulation of CDH1 Is associated with expression of snai1 in colorectal adenomas. PLoS ONE 7:e46665.

Kulasingam V, Diamandis EP. 2008. Tissue culture-based breast cancer biomarker discovery platform. Int J Cancer 123:2007–2012.

Imperlini E, Mancini A, Spaziani S, Martone D, Alfieri A, Gemei M, Del Vecchio L, Buono P, Orrù S. 2010. Androgen receptor signaling induced by supraphysiological doses of dihydrotestosterone in human peripheral blood lymphocytes. Proteomics 10:3165–3175.

Laska DL, Houchins JO, Pratt SE, Horn J, Xia X, Hanssen BR, Williams DC, Dantzig AH, Lindstrom AE. 2002. Characterization and application of a vinblastine-selected caco-2 cell line for evaluation of *p*-glycoprotein in vitro cell. Dev Biol Animal 38:401–410.

Lau LF. 2011. CCN1/CY R61: The very model of a modern matricellular protein. Cell Mol Life Sci 68:3149–3163.

Ma Y, Visser L, Roelofsen H, de Vries M, Diepstra A, van Imhoff G, van der Wal T, Luinge M, Alvarez-Llamas G, Vos H, Poppema S, Vonk R, van den Berg A. 2008. Proteomics analysis of Hodgkin lymphoma: Identification of new players involved in the cross-talk between HRS cells and infiltrating lymphocytes. Blood 111:2339–2346.

Makridakis M, Vlahou A. 2010. Secretome proteomics for discovery of cancer biomarkers. J Proteomics 73:2291–2305.

Mathivanan S, Ji H, Simpson RJ. 2010. Exosomes: Extracellular organelles importantin intercellular communication. J Proteomics 73:1907–1920.

Mustafa D, van der Weiden M, Zheng P, Nigg A, Luider TM, Kros JM. 2010. Expression sites of colligin 2 in glioma blood vessels. Brain Pathol 20:50–65.

Normanno N, Bianco C, Damiano V, de Angelis E, Selvam MP, Grassi M, Magliulo G, Tortora G, Bianco AR, Mendelsohn J, Salomon DS, Ciardiello F. 1996. Growth Inhibition of Human Colon Carcinoma Cells by Combinations of Anti-Epidermal Growth Factor-related Growth Factor Antisense Oligonucleotides. Clin Cancer Res 2:601–609.

Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG. 2005. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 4:1487–1502.

Orrù S, Ruoppolo M, Francese S, Vitagliano L, Marino G, Esposito C. 2002. Identification of tissue transglutaminase-reactive lysine residues in glyceraldehyde-3-phosphate dehydrogenase. Protein Sci 11:137–146.

Park JH, Seol JA, Choi HJ, Roh YH, Choi PJ, Lee KE, Roh MS. 2011. Comparison of cadherin-17 expression between primary colorectal adenocarcinomas and their corresponding metastases: The possibility of a diagnostic marker for detecting the primary site of metastatic tumour. Histopathology 58:315–318.

Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. Nature Methods 8:785–786.

Rodríguez-Piñeiro AM, García-Lorenzo A, Blanco-Prieto S, Alvarez-Chaver P, Rodríguez-Berrocal FJ, Cadena MP, Martínez-Zorzano VS. 2012. Secreted clusterin in colon tumor cell models and its potential as diagnostic marker for colorectal cancer. Cancer Invest 30:72–78.

Sato Y, Kumamoto K, Saito K, Okayama H, Hayase S, Kofunato Y, Miyamoto K, Nakamura I, Ohki S, Koyama Y, Takenoshita S. 2011. Up-regulated Annexin A1 expression in gastrointestinal cancer is associated with cancer invasion and lymph node metastasis. Exp Ther Med 2:239–243.

Scamuffa N, Metrakos P, Calvo F, Khatib AM. 2011. Identification of the myosin heavy polypeptide 9 as a downstream effector of the proprotein convertases in the human colon carcinoma HT-29 cells. Methods Mol Biol 768:207–215.

Sgambato A, Migaldi M, Montanari M, Camerini A, Brancaccio A, Rossi G, Cangiano R, Losasso C, Capelli G, Trentini GP, Cittadini A. 2003. Dystroglycan expression is frequently reduced in human breast and colon cancers and is associated with tumor progression. Am J Pathol 162:849–860.

Simpson RJ, Jensen SS, Lim JW. 2008. Proteomic profiling of exosomes: Current perspectives. Proteomics 8:4083–4099.

Susila A, Chan H, Loh AX, Phang HQ, Wong ET, Tergaonkar V, Koh CG. 2010. The POPX2 phosphatase regulates cancer cell motility and invasiveness. Cell Cycle 9:179–187.

Swaminathan G, Cartwright CA. 2012. Rack1 promotes epithelial cell–cell adhesion by regulating E-cadherin endocytosis. Oncogene 31:376–389.

Tsutsumi S, Fukasawa T, Yamauchi H, Kato T, Kigure W, Morita H, Asao T, Kuwano H. 2009. Phosphoglucose isomerase enhances colorectal cancer metastasis. Int J Oncol 35:1117–1121.

Wu B, Zhang F, Yu M, Zhao P, Ji W, Zhang H, Han J, Niu R. 2012. Up-regulation of Anxa2 gene promotes proliferation and invasion of breast cancer MCF-7 cells. Cell Prolif 45:189–198.

Wu CC, Chen HC, Chen SJ, Liu HP, Hsieh YY, Yu CJ, Tang R, Hsieh LL, Yu JS, Chang YS. 2008. Identification of collapsin response mediator protein-2 as a potential marker of colorectal carcinoma by comparative analysis of cancer cell secretomes. Proteomics 8:316–332.

Wu CC, Hsu CW, Chen CD, Yu CJ, Chang KP, Tai DI, Liu HP, Su WH, Chang YS, Yu JS. 2010. Candidate serological biomarkers for cancer identified from the secretomes of 23 cancer cell lines and the human protein atlas. Mol Cell Proteomics 9:1100–1117.

Xue G, Hao LQ, Ding FX, Mei Q, Huang JJ, Fu CG, Yan HL, Sun SH. 2009. Expression of annexin a5 is associated with higher tumor stage and poor prognosis in colorectal adenocarcinomas. J Clin Gastroenterol 43:831–837.

Xue H, Lu B, Zhang J, Wu M, Huang Q, Wu Q, Sheng H, Wu D, Hu J, Lai M. 2010. Identification of serum biomarkers for colorectal cancer metastasis using a differential secretome approach. J Proteome Res 9:545–555.

Yakovich AJ, Huang Q, Du J, Jiang B, Barnard JA. Vectorial TGFbeta signaling in polarized intestinal epithelial cells. J Cell Physiol 2010. 224:398–404.

Zheng Q, Wang XJ. 2008. GOEAST: A web-based software toolkit for Gene Ontology enrichment analysis. Nucleic Acids Res 36:W358–W363.

Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP. 2006. Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. J Proteome Res 5:2339–2347.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Supplemental Material: Table S1. Spectral counting and protein ratios for all identified proteins.