ORIGINAL ARTICLE

Basic Study

Parallel mRNA, proteomics and miRNA expression analysis in cell line models of the intestine

Finbarr O'Sullivan, Joanne Keenan, Sinead Aherne, Fiona O'Neill, Colin Clarke, Michael Henry, Paula Meleady, Laura Breen, Niall Barron, Martin Clynes, Karina Horgan, Padraig Doolan, Richard Murphy

Finbarr O'Sullivan, Joanne Keenan, Sinead Aherne, Fiona O'Neill, Michael Henry, Paula Meleady, Laura Breen, Niall Barron, Martin Clynes, Padraig Doolan, National Institute for Cellular Biotechnology, Dublin City University, Dublin D09 W6Y4, Ireland

Colin Clarke, National Institute for Bioprocessing Research & Training, Blackrock, Dublin A94 X099, Ireland

Karina Horgan, Richard Murphy, Alltech, Dunboyne, Meath, A86 X006, Ireland

Abstract

AIM
To identify miRNA-regulated proteins differentially expressed between Caco2 and HT-29: two principal cell line models of the intestine.

METHODS
Exponentially growing Caco-2 and HT-29 cells were harvested and prepared for mRNA, miRNA and proteomic profiling. mRNA microarray profiling analysis was carried out using the Affymetrix GeneChip Human Gene 1.0 ST array. miRNA microarray profiling analysis was carried out using the Affymetrix Genechip miRNA 3.0 array. Quantitative Label-free LC-MS/MS proteomic analysis was performed using a Dionex Ultimate 3000 RSLCnano system coupled to a hybrid linear ion

trap/Orbitrap mass spectrometer. Peptide identities were validated in Proteome Discoverer 2.1 and were subsequently imported into Progenesis QI software for further analysis. Hierarchical cluster analysis for all three parallel datasets (miRNA, proteomics, mRNA) was conducted in the R software environment using the Euclidean distance measure and Ward's clustering algorithm. The prediction of miRNA and oppositely correlated protein/mRNA interactions was performed using TargetScan 6.1. GO biological process, molecular function and cellular component enrichment analysis was carried out for the DE miRNA, protein and mRNA lists *via* the Pathway Studio 11.3 Web interface using their Mammalian database.

RESULTS

Differential expression (DE) profiling comparing the intestinal cell lines HT-29 and Caco-2 identified 1795 Genes, 168 Proteins and 160 miRNAs as DE between the two cell lines. At the gene level, 1084 genes were upregulated and 711 were downregulated in the Caco-2 cell line relative to the HT-29 cell line. At the protein level, 57 proteins were found to be upregulated and 111 downregulated in the Caco-2 cell line relative to the HT-29 cell line. Finally, at the miRNAs level, 104 were upregulated and 56 downregulated in the Caco-2 cell line relative to the HT-29 cell line. Gene ontology (GO) analysis of the DE mRNA identified cell adhesion, migration and ECM organization, cellular lipid and cholesterol metabolic processes, small molecule transport and a range of responses to external stimuli, while similar analysis of the DE protein list identified gene expression/transcription, epigenetic mechanisms, DNA replication, differentiation and translation ontology categories. The DE protein and gene lists were found to share 15 biological processes including for example epithelial cell differentiation [P value $\leqslant$ 1.81613E-08 (protein list); $P \leqslant$ 0.000434311 (gene list)] and actin filament bundle assembly [P value $\leqslant$ 0.001582797 (protein list); $P \leqslant$ 0.002733714 (gene list)]. Analysis was conducted on the three data streams acquired in parallel to identify targets undergoing potential miRNA translational repression identified 34 proteins, whose respective mRNAs were detected but no change in expression was observed. Of these 34 proteins, 27 proteins downregulated in the Caco-2 cell line relative to the HT-29 cell line and predicted to be targeted by 19 unique anti-correlated/upregulated microRNAs and 7 proteins upregulated in the Caco-2 cell line relative to the HT-29 cell line and predicted to be targeted by 15 unique anti-correlated/downregulated microRNAs.

CONCLUSION

This first study providing "tri-omics" analysis of the principal intestinal cell line models Caco-2 and HT-29 has identified 34 proteins potentially undergoing miRNA translational repression.

Key words: Caco-2; HT-29; Microarray; Proteomics; miRNA; Targetscan; Gene ontology

Core tip: Unique triomics analysis of Caco-2 and HT-29, two commonly used *in vitro* cell lines models of the intestine, was conducted. This analysis not only provided data on differentially expressed mRNAs, miRNAs and proteins but also allowed the identification of miRNA-regulated proteins differentially expressed between these two cell lines.

INTRODUCTION

In recent years both research and industry, particularly in the areas of pharmacology and nutrition, have steadily increased their use of *in vitro* cell models as an alternative to animal testing. This increase has been driven in part by the principles of the 3Rs (Replacement, Reduction and Refinement) and associated legislation, for example Article 4 of EU Directive 2010/63/EU. In addition there has been a growing use of automated high throughput techniques in the laboratory which has also increased the use of such *in vitro* cell lines. Thus there has been increased interest in the development of suitable intestinal *in vitro* models. These models are key to obtaining new information on intestinal toxicity, microbial infections, bioavailability of new food additives or new drugs as well as intestinal-related diseases[1-5].

The intestinal epithelium contains a number of specialised cell types; absorptive enterocytes (the principle intestinal cell type), goblet cells (mucin secreting), enteroendocrine cells, Paneth cells, tuft cells and M(or microfold) cells[6]. However the *in vitro* primary culture of intestinal epithelium is difficult, hence the use of intestinal cell lines allow for robust, reproducible *in vitro* culture models. Two cell lines, derived human from colonic adenocarcinomas, commonly used for the creation of such models are Caco-2 and HT-29.

The Caco-2 cell line can differentiate spontaneously to yield a polarised monolayer shown to express several characteristics and markers of enterocytes[7,8]. Thus Caco-2 cells have become a cell line of choice in the making of *in vitro* models of the intestine[9,10]. HT29 cell line is heterogeneous, consisting of a main population of undifferentiated cells and a smaller subpopulation capable of producing mucus[11-13]. Using glucose-free conditions, the HT-29 line can be differentiated to an

enterocyte phenotype. In addition various selection procedures have generated homogeneous mucin-secreting populations[11,13-15]. These two cell lines form the basis of numerous models of the intestine either as single cell models or advanced multicellular models[16-18]. However, the creation of suitable complex *in vitro* models requires that cell lines used are well characterised; this includes an understanding of the molecular controls within the cells.

There is an increasing recognition of the importance of the post-transcriptional control by miRNA of cell processes such as proliferation, differentiation, and apoptosis. The importance of miRNA in intestinal development and behaviour has been implicated in a number of studies[19,20]. In addition, the miRNAs miR-99b, miR-125a-5p and miR-1269 have been identified as having a potential role in determining Caco-2 and HT-29 behaviour[21].

miRNA exerts its control by either interfering with protein initiation and elongation or by the degradation of the target mRNA[22-25]. An additional complexity of this control mechanism is that a single mRNA can be co-operatively targeted by multiple miRNAs, while a single miRNA can target multiple mRNAs[26,27]. A number of algorithms exist to predict the potential targets of miRNAs for example miRanda, TargetScan and PicTar[28-30]. These algorithms generally function by their ability to recognise regions of sequence complementarity to miRNA seed regions (nucleotides 2-7 of the miRNA) in the 3'UTR of target mRNAs and the thermodynamic feasibility of such binding.

However, validation by experimental means is often required to determine the most significant miRNAs responsible for an observed effect within a biological system. This study uniquely uses a tri-omics approach combining the expression profiles of mRNA, miRNA and proteomics generated in parallel to identify potential miRNAs controlling translational repression and thus cell behaviour. As we have previously noted[31], the use of solitary profiling methodologies (including proteomic mass spectrometry, mRNA microarrays *etc.*) is prone to several distinct disadvantages when used in isolation. For example, gene expression analysis using microarrays is capable of identifying a wide coverage of mRNAs but post-transcriptional processes will not be captured using just this method. The availability of a combined profiling approach could also reduce the possibilities of false positive/negative rates associated with purely in-silico prediction analyses. Additionally, studies examining the role of miRNAs typically rely heavily on computational methods to predict miRNA interaction and prioritise potential direct targets.

This study uses the availability of data on multiple levels to identify potential interacting miRNA-protein networks that would not have been identified by a single dataset, allowing us to provide new information

for the characterising of these two cell lines.

MATERIALS AND METHODS

Cell culture
The human colon carcinoma cell lines Caco-2 (cat. HTB37) was obtained from the American Type Culture Collection) and HT-29 (cat. 91072201) was obtained from Public Health England Culture collection. HT-29 and Caco-2 were maintained in MEM supplemented with 1% L-glutamine and 5% or 10% FBS respectively under normal conditions (37 ℃, 5% CO_2). Caco-2 cells were allowed to reach a maximum confluency of 60%-70% before trypsinizing to ensure that spontaneous differentiation was not occurring. Both cell lines were tested and found to be mycoplasma negative. All chemicals (unless otherwise stated), glutamine and cell culture media were obtained from Sigma (Poole, United Kingdom). FBS and HBSS were obtained from Invitrogen-ThermoFisher Scientific. Exponentially growing cells were set up in 25 cm^2 flasks (Corning) for RNA extraction or 75 cm^2 flasks (Corning) for protein extraction.

Affymetrix miRNA microarray analysis of miRNA expression
Total RNA was extracted from triplicate biological replicate samples using the Qiagen miRNeasy kit according to the manufacturer's instructions. RNA quantity was assessed using the NanoDrop ND-1000 spectrophotometer and RNA quality assayed using the Agilent RNA 6000 NANO KIT and Agilent Bioanalyzer (Agilent, Santa Clara, CA, United States). The FlashTag Biotin HSR RNA Labelling Kit (Affymetrix, Santa Clara, CA, United States) was used to label a total of 500 ng of total RNA according to manufacturer's procedure. The GeneChip Scanner 3000 7G System and reagents from Affymetrix were used to hybridize, wash, stain and scan the Genechip miRNA 3.0 arrays (Affymetrix, Santa Clara, CA, United States).

Microarray profiling and associated bioinformatics analysis
Gene expression analysis was carried out on the Affymetrix GeneChip Human Gene 1.0 ST array according to the manufacturer's instructions (Affymetrix, Santa Clara, CA, United States). The methodology and criteria used for total RNA purification, ssDNA sample processing and hybridisation to human microarrays have been previously described[32]. All microarray data were pre-processed as described previously. Prior to data analysis, probesets that did not reach the detection threshold (fluorescence level $\geq$ log2 (100) for at least 1 sample) were identified and designated undetected. The remaining probesets were considered

differentially expressed between the two cell types if a fold change ≥ 1.2 in either direction along with a BH adjusted *P* value ≤ 0.05 was observed

Proteomics profiling and associated bioinformatics analysis

Exponentially growing cells were harvested and pelleted by centrifugation. Cells were washed with ice cold PBS, cell pellets were then snap frozen and stored at -80 ℃ until required. Cell lysis was carried out using a buffer containing 7 mol/L Urea, 2 mol/L Thiourea, 4% CHAPS, 30 mmol/L Tris, pH 8.5, 1x HALT protease inhibitors (Thermo Fisher Scientific) and incubated on ice for 20 min with occasional vortexing. Sample lysates were centrifuged at 14000 × for 15 min at 4 ℃. The supernatant was transferred to a new microcentrifuge tube and protein concentration was determined using the Quick StartTM Bradford Protein Assay (BioRad).

Protein samples from the cell lysates were cleaned up using the ReadyPrep 2D Cleanup Kit (BioRad). Pellets were then resuspended in a buffer containing 6 M Urea, 2 mol/L Thiourea, 30 mmol/L Tris, pH 8.5. 20 μg from each sample was reduced with dithiothreitol to a final concentration of 5 mmol/L for 20 min at 56 ℃ followed by alkylation with iodoacetamide to a final concentration of 15 mmol/L, and incubated in the dark for 20 min at room temperature. Protein samples were tryptically digested (sequence grade, Promega) at a ratio of 50:1 protein/enzyme w/v at 37 ℃ overnight. Trifluoroacetic acid (TFA) was added to a final concentration of 0.5% to each sample to inactivate the trypsin. The digested peptide samples were then concentrated with C18 spin columns (Pierce, Thermo Fisher Scientific) and eluted peptides were gently dried with a vacuum evaporator (SpeedVac).

LC-MS/MS analysis

Dried peptides were re-solubilised in 25 μL of LC-MS grade water with 0.1% FA and 2% Acetonitrile (ACN). Peptides were separated and analysed using a Dionex Ultimate 3000 RSLCnano system (Thermo Fisher Scientific) coupled to a hybrid linear ion trap/Orbitrap mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific).

A 5 μL injection of sample was picked up using the autosampler and loaded onto a C18 trap column (C18 PepMap, 300 μm ID × 5 mm, 5 μm particle size, 100 Å pore size; Thermo Fisher Scientific). The sample was desalted for 3 minutes using a flow rate of 25 μL/min in 0.1% TFA containing 2% ACN. The trap column was then switched online with the analytical column [PepMap C18, 75 μm ID × 250 mm, 3 μm particle and 100 Å pore size; (Thermo Fisher Scientific)] using a column oven at 35°C and peptides were eluted with the following binary gradients of: Mobile Phase Buffer A and Mobile Phase Buffer B: 0%-25% solvent B in 240 min and 25%-50% solvent B in a further 60 min,

where solvent A consisted of 2% acetonitrile (ACN) and 0.1% formic acid in water and solvent B consisted of 80% ACN and 0.08% formic acid in water. The column flow rate was set to 300 nL/min. Data was acquired with Xcalibur software, version 2.0.7 (Thermo Fisher Scientific).

The LTQ Orbitrap XL was operated in data-dependent and externally calibrated. Survey MS scans were acquired in the Orbitrap in the 400-1800 m/z range with the resolution set to a value of 30000 at m/z 400. Up to three of the most intense ions (1+, 2+ and 3+) per scan were CID fragmented in the linear ion trap. Dynamic exclusion was enabled with a repeat count of 1, repeat duration of 30 s, exclusion list size of 500 and exclusion duration of 40 s. The minimum signal was set to 500. All tandem mass spectra were collected using a normalised collision energy of 35%, an isolation window of 2 m/z with an activation time of 30 ms.

Quantitative label-free LC-MS/MS data analysis

The raw MS data files obtained were processed using Progenesis QI for Proteomics software (version 2.0; Non-Linear Dynamics, a Waters company, Newcastle upon Tyne, United Kingdom). Peptide LC retention times from all MS data files were aligned to an assigned reference run as previously described[31]. The samples from the two experimental groups were then set up within the Progenesis software for differential analysis. Peptide features were filtered using the following parameters; (1)peptide features with ANOVA ≤ 0.05 between experimental groups; and (2)mass peaks with charge states from +1 to +3 and greater than one isotope per peptide. For the proteomic analysis a mascot generic file (mgf) was then generated from all exported MS/MS spectra which was used for peptide and protein identification *via* Proteome Discoverer 2.1 using Sequest HT (Thermo Fisher Scientific) and Percolator against the human Swissprot database containing 23053 sequences. The following search parameters were used for protein identification: (1)peptide mass tolerance set to 20 ppm; (2)MS/MS mass tolerance set to 0.6 Da; (3)up to two missed cleavages were allowed; (4)carbamidomethylation of cysteine set as a fixed modification; and (5)methionine oxidation was set as variable modifications. Only high confident peptide identifications with an FDR ≤ 0.01 [identified using a SEQUEST HT workflow coupled with Percolator validation in Proteome Discoverer 2.1 (Thermo Fisher Scientific)] were imported into Progenesis QI software for further analysis. A statistical criteria of ANOVA *P* value less than 0.05, a minimum of two peptides matched to a protein and a ≥ 1.2 fold change between the two cell lines (HT-29 and Caco-2) was used as the criteria for identification as a differentially expressed protein. Data visualisation was achieved using a heatmap which were generated using

ggplot package in R to show the distribution of all identified proteins based on statistics and fold-change across all of the samples used in the analysis.

Bioinformatics QC analysis, miRNA target prediction and GO analysis

Hierarchical cluster analysis (HCA) for all three parallel datasets (miRNA, proteomics, mRNA) was conducted in the R software environment (provided in the public domain by R Foundation for Statistical Computing, Vienna, Austria, available at http://www.r-project.org/) using the Euclidean distance measure and Ward's clustering algorithm. The prediction of miRNA and oppositely correlated protein/mRNA interactions was performed using TargetScan 6.1 (www.targetscan.org/vert_61/) and has been described previously[32]. miRNA Names (*e.g.,* hsa-miR-375) were annotated to miRNA Accession IDs (*e.g.,* MI0000783) using miRBase (http://www.mirbase.org/) for GO analysis. GO biological process, molecular function and cellular component enrichment analysis was carried out for the DE miRNA, protein and mRNA lists *via* the Pathway Studio 11.3 Web interface (https://www.elsevier.com/solutions/pathway-studio-biological-research) using their Mammalian database and has been described previously[33].

RESULTS

Data analysis methodology of combined omics/"Tri-omics" expression profiling approach

Figure 1 describes the data analysis strategy employed in this study. Differential expression (DE) profiling of HT-29 and Caco-2 cell line triplicate samples at the miRNA, protein and mRNA levels identified three separate datasets of 160 DE miRNAs, 168 DE proteins and 1795 DE mRNA transcripts (Stage 1: Supplementary Tables 1-3). Enrichment analysis against GO was conducted using the literature mining software Pathway Studio Web for the three resulting lists to determine if any biological processes were overrepresented (Stage 2: Supplementary Tables 4-6). The availability of all three data streams acquired in parallel was essential for the identification of our priority candidates - targets undergoing potential miRNA translational repression.

Prior to target prediction against the DE miRNA list, we separated miRNA, protein and mRNA targets into three groups of interest (Stage 3). "Group 1" contains those targets where a degree of post-transcriptional regulation was observed (possibly *via* miRNA mediated translational repression) which may contribute to phenotypic differences between these two cell lines and were considered the most interesting group studied. The 34 proteins in Group 1 were DE between HT-29 and Caco-2, their respective mRNAs

were expressed above the microarray detection threshold but no change in mRNA expression was observed and were predicted to be targeted by 34 DE microRNAs whose expression was correlated in the opposite direction. This group comprised 27 proteins downregulated in the Caco-2 cell line relative to the HT-29 cell line and predicted to be targeted by 19 unique anti-correlated/upregulated microRNAs (Table 1) and 7 proteins upregulated in the Caco-2 cell line relative to the HT-29 cell line and predicted to be targeted by 15 unique anti-correlated/downregulated microRNAs (Table 2).

The second group of candidates (referred to as "Group 2") were comprised of DE protein targets where the matching mRNA transcript was DE in the matching direction. These "Group 2" candidates follow the classical mRNA-protein expression control model (*i.e.,* non-miR-based expression control) and are detailed in Supplementary Table 7.

The third group of candidates ("Group 3"; detailed in Supplementary Table 8) were comprised of DE protein targets where (1) the matching mRNA transcripts were not DE in the same direction or (2) where the mRNA probeset was under the detection threshold or (3) not present on the chip and finally (4) microRNAs predicted to target these proteins (using Targetscan) were not identified as DE between the two cell lines. The role of miRNA in control of expression of these "Group 3" candidates is therefore unclear and these proteins are considered lowest-priority of the three groups studied, for the purposes of the analysis presented here.

Combined omics/"Tri-omics" expression profiling identifies three interesting DE candidate groups that characterise the differences between HT-29 and Caco-2

Tri-omics expression profiling using the three methodologies outlined identified 160 DE miRNAs, 168 DE Proteins and 1795 DE Genes from a comparison of the HT-29 and Caco-2 cell lines.

Bioinformatics analysis of the microRNA profiling data (as outlined) on the HT-29 and Caco-2 cell lines identified 1550 probesets (772 upregulated, 778 downregulated) as differentially expressed (DE). These 1550 DE probesets were species-annotated to 221 human-specific probesets (140 upregulated, 71 downregulated) and further reduced to 160 DE mature non-star human miRNAs (104 upregulated, 56 downregulated (Supplementary Table 1). Bioinformatics analysis (as outlined) of the proteomics profiling data identified a total of 168 annotated DE proteins (57 upregulated, 111 downregulated) between the HT-29 and Caco-2 cell lines (Supplementary Table 2). Bioinformatics analysis (as outlined) of the microarray profiling data identified 1795 probesets (1084 upregulated, 711 downregulated) as differentially
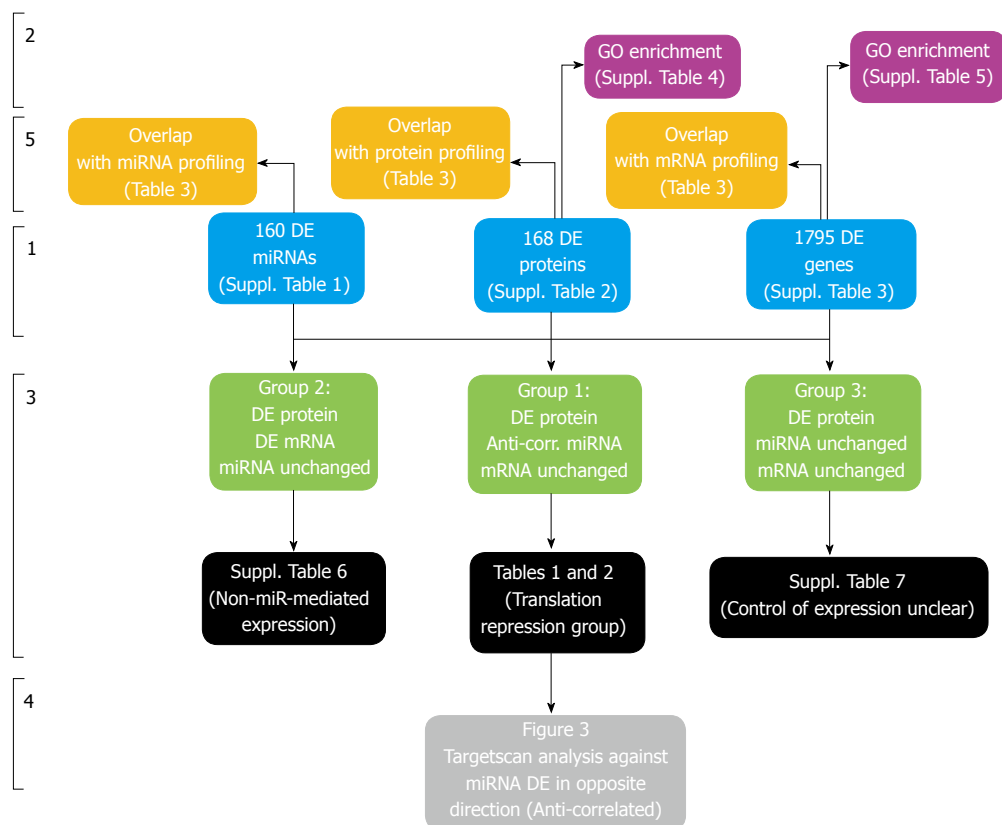
Figure 1 Overview of Tri-Omics data analysis approach. Stage 1: Differential expression analysis of "Tri-Omics" data (miRNA, protein, mRNA) between HT-29 and Caco-2 cell lines (for a full list of results, including fold-changes and associated *P* value statistics, see Supplementary Tables 1-3). Stage 2: Enrichment analysis using Pathway Studio to determine overrepresented GO biological processes in the DE miRNA, protein and mRNA lists (see Supplementary Tables 4-6). Stage 3: Integration of three parallel datasets to generate 3 groups of interesting candidates; (1) Potential targets of miRNA translational repression (protein DE without mRNA change and targeted by anti-correlated miRNA): see Table 1 (Downregulated Proteins targeted by Upregulated miRNAs) and Table 2 (Upregulated Proteins targeted by Downregulated miRNAs); (2) Non- microRNA-mediated gene-protein expression (matching mRNA transcripts and proteins were DE, no corresponding anti-correlated microRNA data or results were below detection level): see Supplementary Table 7 and (3) Targets where control of differential protein expression was unclear (protein DE without mRNA or corresponding targeting microRNA change). Stage 4: TargetScan analysis of both groups against miRNAs DE in the opposite direction. Stage 5: Overlap analysis of Supplementary Tables 1-3 with published profiling studies on same cell lines: see Table 3.

expressed (DE) between the HT-29 and Caco-2 cell lines (Supplementary Table 3).

Two-way hierarchical HCA clustering analysis as outlined on the HT-29 and Caco-2 cell line samples confirms that the DE (miRNA, proteomics, miRNA) candidates separate the samples into their respective replicate groups, indicating that they are of sufficient quality to yield high-priority candidates for phenotypic characterisation between both cell lines (Figure 2). For (A) miRNA and (C) mRNA, red indicates diminished expression and green indicates increased expression for all three replicate groups, while for (B) proteomics blue indicates diminished expression and red indicates increased expression for all three replicates in each group.

A TargetScan prediction analysis for (A) post-transcriptionally downregulated (green) proteins and anti-correlated/upregulated (red) miRNAs and (B) post-transcriptionally upregulated (red) proteins and anti-correlated/downregulated (green) miRNAs is shown in Figure 3. The prioritisation of these targets was possible

only through integration of the miRNA, protein and mRNA datasets.

GO enrichment analysis identifies overrepresented GO categories related to cell differentiation, actin assembly and cytoskeleton organization shared between DE protein and gene lists

GO biological processes found to be overrepresented within the DE protein list included gene expression/transcription, epigenetic mechanisms, DNA replication, differentiation and translation (Supplementary Table 4).

A similar analysis of the DE mRNAs identified cell adhesion, migration and ECM organization, cellular lipid and cholesterol metabolic processes, small molecule transport and a range of responses to external stimuli to be the most significantly-enriched GO categories from this list (Supplementary Table 5).

Fifty shared biological processes were enriched by the DE protein and microarray datasets, including epithelial cell differentiation [*P* value $\leq$ 1.81613E-08 (protein list); *P* $\leq$ 0.000434311 (gene list), actin

A

B

C
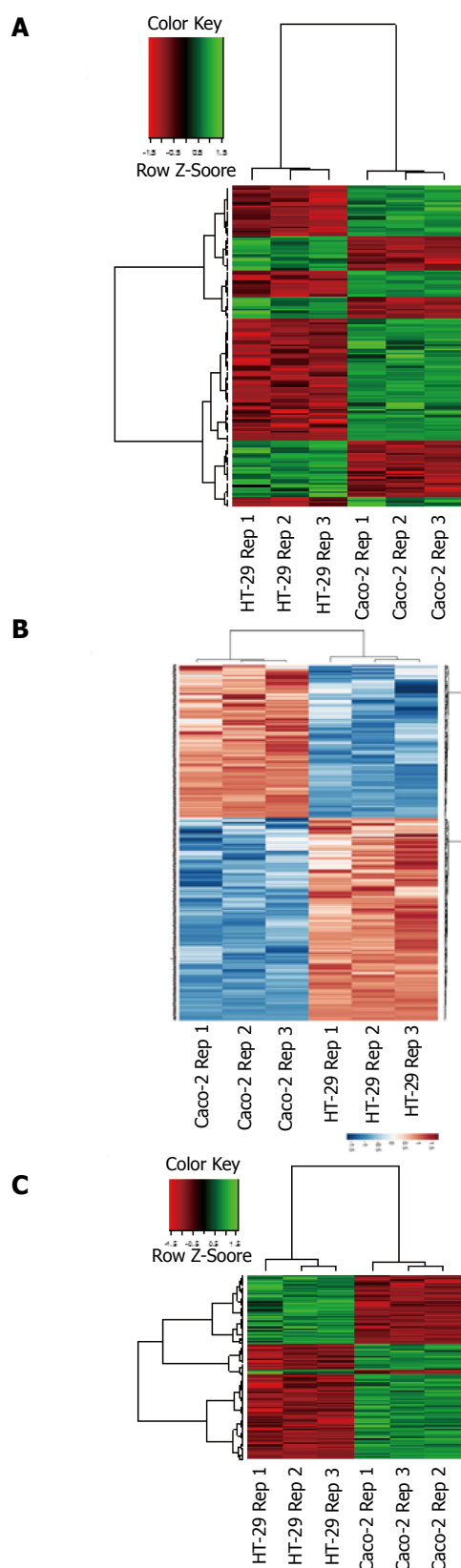
Figure 2 Bioinformatics QC analysis of HT-29 and Caco-2 replicates. 2-way hierarchical HCA clustering analysis on the HT-29 and Caco-2 cell line replicates using (A) DE miRNA probeset list (160 probesets) (B) DE proteomics list (168 proteins) and (C) DE mRNA probeset list (1795 probesets) confirms that the DE candidates separate the samples into their respective cell line replicate groups. For (A) and (C), red indicates diminished expression and green indicates increased expression for all three replicate groups; for (B) blue indicates diminished expression and red indicates increased expression for all three replicate groups.

filament bundle assembly [P value $\leqslant 0.001582797$ (protein list); $P \leqslant 0.002733714$ (gene list)], cell morphogenesis involved in differentiation [P value $\leqslant 0.003493544$ (protein list); $P \leqslant 0.005671882$ (gene list)] and cytoskeleton organization [P value $\leqslant 0.004078084$ (protein list); $P \leqslant 0.005154161$ (gene list)].

GO analysis using the 160 DE microRNA list of entities did not identify any enriched biological processes in the Pathway Studio Web database.

Integration of 3 parallel datasets identifies a priority group of 37 DE proteins and 44 associated miRs that may characterise phenotypic differences between HT-29 and Caco-2

The availability of all three level of expression data (miRNA, protein, mRNA) acquired in parallel on identical cell line samples was essential for the identification of our priority candidates for characterisation of the differences between the HT-29 and Caco-2 cell lines - protein targets that could undergo potential miRNA translational repression.

These potential targets of miRNA translational repression were identified as differentially-expressed (DE) proteins that could be targeted by a DE miRNA in the opposite direction (anti-correlated miRNA) and where the corresponding gene transcript was unchanged (protein DE without mRNA change and targeted by anti-correlated miRNA). From this 3-way overlap analysis, a total of 34 DE proteins and 34 DE microRNAs were identified to be of interest (Tables 1 and 2). Table 1 (Downregulated Proteins targeted by Upregulated miRNAs) displays the 27 proteins observed to be downregulated in the Caco-2 cell line relative to the HT-29 cell line and predicted to be targeted by 19 unique anti-correlated/upregulated microRNAs. Table 2 (Upregulated Proteins targeted by Downregulated miRNAs) contains the 7 proteins upregulated in the Caco-2 cell line relative to the HT-29 cell line and predicted to be targeted by 15 unique anti-correlated/downregulated microRNAs. Annotation information and associated statistical data for both DE proteins and microRNAs are included.

Integration of 3 parallel datasets identifies a group of 36 DE proteins and corresponding mRNA transcripts that may characterise phenotypic differences between HT-29 and Caco-2 that are not miR-mediated

The data-overlapping strategy similarly identified a group of 35 DE proteins (16 upregulated in Caco-2 relative to HT-29 and 19 downregulated in Caco-2 relative to HT-29) where the corresponding gene transcript was also DE in the direction and where the protein was not predicted to be targeted by any of the DE microRNAs (or where the DE miRNA data results were below detection levels). These proteins and genes (listed together with their attendant annotation information and associated statistical results in Supplementary Table 6) were considered to follow the standard molecular model for gene-protein expression,

A

B
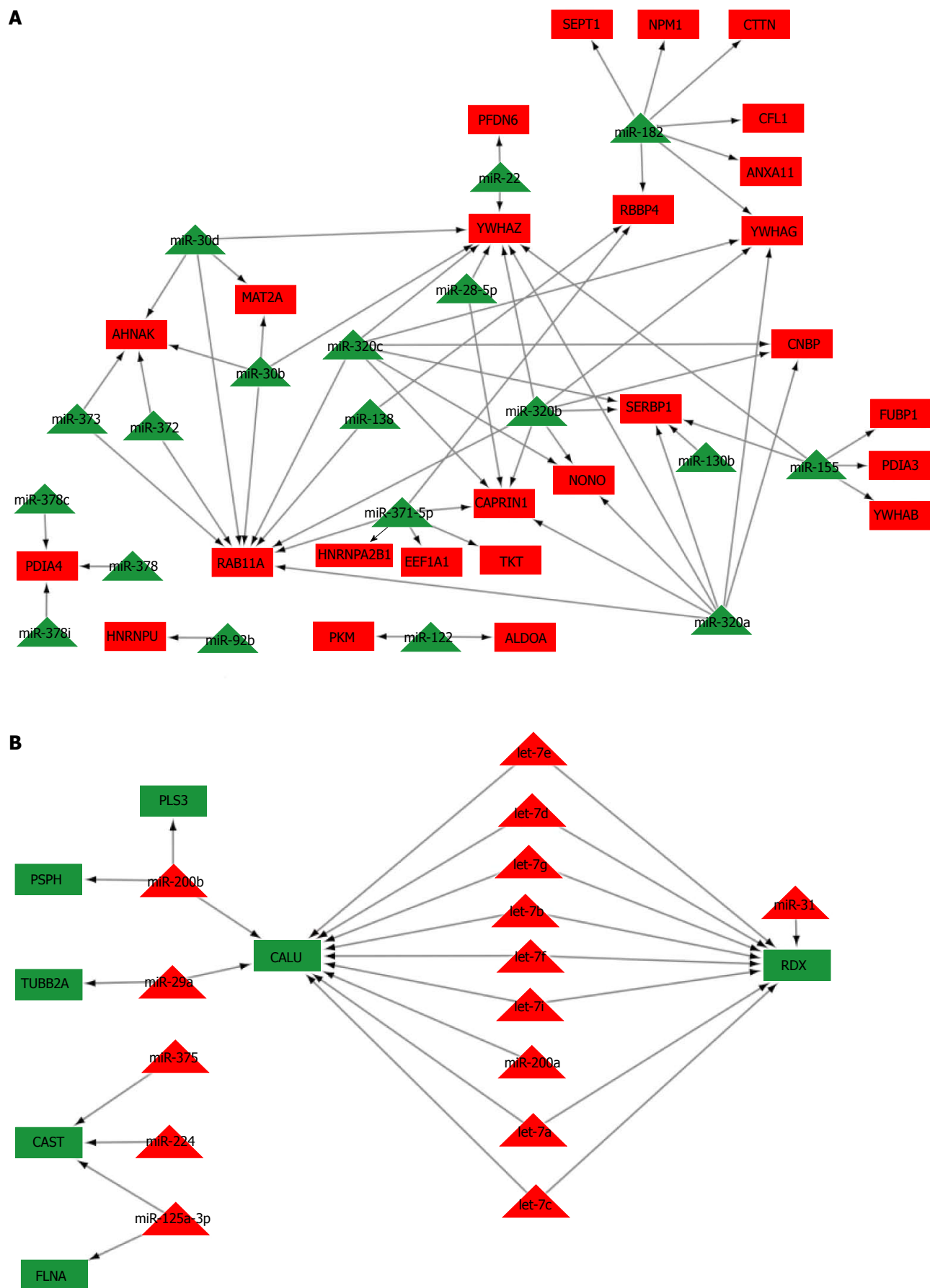
Figure 3 Summary of TargetScan predicted miRNA targets. A: TargetScan prediction analysis for post-transcriptionally downregulated (green) proteins and anti-correlated/upregulated (red) miRNAs; B: TargetScan prediction analysis for post-transcriptionally upregulated (red) proteins and anti-correlated/downregulated (green) miRNAs. The prioritisation of these targets was possible only through integration of the miRNA, protein and mRNA datasets.

Table 1 Twenty-seven downregulated proteins with no DE mRNA targeted by 19 upregulated miRNAs

Protein					microRNA					
Uniprot	Symbol	Title	Linear FC	Adj. *P* value	miRBase ID	logFC	LinFC	AveExpr	*P* value	Adj. *P* value
Q09666	AHNAK	Neuroblast differentiation-associated protein AHNAK	1.66	0.02036	hsa-miR-30b	1.39	2.62	10.38	5.04E-04	0.011296975
					hsa-miR-30d	1.36	2.56	11.05	1.17E-04	0.003668745
					hsa-miR-372	12.21	4743.91	7.03	6.55E-11	2.27E-07
					hsa-miR-373	12.12	4451.32	7.23	2.25E-11	1.88E-07
P04075	ALDOA	Fructose-bisphosphate aldolase A	2.03	0.00001	hsa-miR-122	5.85	57.63	5.04	3.71E-05	0.001621788
P50995	ANXA11	Annexin A11	4.57	0.00012	hsa-miR-182	1.09	2.13	12.01	2.48E-04	0.006341084
Q14444	CAPRIN1	Caprin-1	1.42	0.00658	hsa-miR-28-5p	1.16	2.24	9.32	7.05E-05	0.002531521
					hsa-miR-320a	1.46	2.76	12.49	2.74E-05	0.001333363
					hsa-miR-320b	1.46	2.76	12.49	4.80E-05	0.001959043
					hsa-miR-320c	1.46	2.74	12.35	5.43E-05	0.002111768
					hsa-miR-371-5p	11.89	3799.61	7.79	1.04E-10	2.62E-07
P23528	CFL1	Cofilin-1	1.53	0.00962	hsa-miR-182	1.09	2.13	12.01	2.48E-04	0.006341084
P62633	CNBP	Cellular nucleic acid-binding protein	1.79	0.00064	hsa-miR-320a	1.46	2.76	12.49	2.74E-05	0.001333363
					hsa-miR-320b	1.46	2.76	12.49	4.80E-05	0.001959043
					hsa-miR-320c	1.46	2.74	12.35	5.43E-05	0.002111768
Q14247	CTTN	Src substrate cortactin	1.33	0.00716	hsa-miR-182	1.09	2.13	12.01	2.48E-04	0.006341084
P68104	EEF1A1	Elongation factor 1-alpha 1	1.46	0.01148	hsa-miR-371-5p	11.89	3799.61	7.79	1.04E-10	2.62E-07
P06733	ENO1	Alpha-enolase	2.06	0.00015	hsa-miR-22	1.96	3.9	10.84	1.01E-04	0.003309338
Q96AE4	FUBP1	Far upstream element-binding protein 1	2.17	0.00290	hsa-miR-155	1.85	3.61	9.51	5.54E-05	0.002127503
P22626	HNRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1	1.72	0.00053	hsa-miR-371-5p	11.89	3799.61	7.79	1.04E-10	2.62E-07
Q00839	HNRNPU	Heterogeneous nuclear ribonucleoprotein U	4.51	0.00152	hsa-miR-92b	1.22	2.33	6.95	4.51E-04	0.010275397
					hsa-miR-122	5.85	57.63	5.04	3.71E-05	0.001621788
P31153	MAT2A	S-adenosylmethionine synthase isoform type-2	1.92	0.00023	hsa-miR-30b	1.39	2.62	10.38	5.04E-04	0.011296975
					hsa-miR-30d	1.36	2.56	11.05	1.17E-04	0.003668745
Q15233	NONO	Non-POU domain-containing octamer-binding protein	2.74	0.00522	hsa-miR-320a	1.46	2.76	12.49	2.74E-05	0.001333363
					hsa-miR-320b	1.46	2.76	12.49	4.80E-05	0.001959043
					hsa-miR-320c	1.46	2.74	12.35	5.43E-05	0.002111768
P06748	NPM1	Nucleophosmin	1.51	0.01869	hsa-miR-182	1.09	2.13	12.01	2.48E-04	0.006341084
P30101	PDIA3	Protein disulfide-isomerase A3	1.84	0.00465	hsa-miR-155	1.85	3.61	9.51	5.54E-05	0.002127503
P13667	PDIA4	Protein disulfide-isomerase A4	1.74	0.00006	hsa-miR-378	1.35	2.55	11.31	7.79E-05	0.002720502
					hsa-miR-378c	1.17	2.25	10.21	5.74E-04	0.012576882
					hsa-miR-378i	0.85	1.80	9.00	7.98E-04	0.01627753
O15212	PFDN6	Prefoldin subunit 6	2.26	0.02810	hsa-miR-22	1.96	3.90	10.84	1.01E-04	0.003309338
P14618	PKM	Pyruvate kinase PKM	1.91	0.00020	hsa-miR-122	5.85	57.63	5.04	3.71E-05	0.001621788
P62491	RAB11A	Ras-related protein Rab-11A	1.74	0.00318	hsa-miR-138	5.84	57.23	4.36	6.08E-08	2.62E-05
					hsa-miR-30b	1.39	2.62	10.38	5.04E-04	0.011296975
					hsa-miR-30d	1.36	2.56	11.05	1.17E-04	0.003668745
					hsa-miR-320a	1.46	2.76	12.49	2.74E-05	0.001333363
					hsa-miR-320b	1.46	2.76	12.49	4.80E-05	0.001959043
					hsa-miR-320c	1.46	2.74	12.35	5.43E-05	0.002111768
					hsa-miR-371-5p	11.89	3799.61	7.79	1.04E-10	2.62E-07
					hsa-miR-372	12.21	4743.91	7.03	6.55E-11	2.27E-07
					hsa-miR-373	12.12	4451.32	7.23	2.25E-11	1.88E-07
Q09028	RBBP4	Histone-binding protein RBBP4	1.76	0.00339	hsa-miR-138	5.84	57.23	4.36	6.08E-08	2.62E-05
					hsa-miR-138	5.84	57.23	4.36	6.08E-08	2.62E-05
					hsa-miR-182	1.09	2.13	12.01	2.48E-04	0.006341084
					hsa-miR-371-5p	11.89	3799.61	7.79	1.04E-10	2.62E-07
Q9UHD8	9-Sep	Septin-9	1.98	0.00016	hsa-miR-182	1.09	2.13	12.01	2.48E-04	0.006341084
Q8NC51	SERBP1	Plasminogen activator inhibitor 1 RNA-binding protein	2.35	0.00473	hsa-miR-130b	1.77	3.42	11.03	2.57E-05	0.00127274
					hsa-miR-155	1.85	3.61	9.51	5.54E-05	0.002127503
					hsa-miR-320a	1.46	2.76	12.49	2.74E-05	0.001333363
					hsa-miR-320b	1.46	2.76	12.49	4.80E-05	0.001959043
					hsa-miR-320c	1.46	2.74	12.35	5.43E-05	0.002111768
P29401	TKT	Transketolase	1.56	0.00000	hsa-miR-371-5p	11.89	3799.61	7.79	1.04E-10	2.62E-07
P31946	YWHAB	14-3-3 protein beta/alpha	1.35	0.01490	hsa-miR-155	1.85	3.61	9.51	5.54E-05	0.002127503
P61981	YWHAG	14-3-3 protein gamma	2.05	0.00007	hsa-miR-182	1.09	2.13	12.01	2.48E-04	0.006341084
					hsa-miR-320a	1.46	2.76	12.49	2.74E-05	0.001333363
					hsa-miR-320b	1.46	2.76	12.49	4.80E-05	0.001959043
					hsa-miR-320c	1.46	2.74	12.35	5.43E-05	0.002111768

P63104	YWHAZ	14-3-3 protein zeta/delta	2.52	0.00005	hsa-miR-155	1.85	3.61	9.51	5.54E-05	0.002127503
					hsa-miR-22	1.96	3.90	10.84	1.01E-04	0.003309338
					hsa-miR-28-5p	1.16	2.24	9.32	7.05E-05	0.002531521
					hsa-miR-30b	1.39	2.62	10.38	5.04E-04	0.011296975
					hsa-miR-30d	1.36	2.56	11.05	1.17E-04	0.003668745
					hsa-miR-320a	1.46	2.76	12.49	2.74E-05	0.001333363
					hsa-miR-320b	1.46	2.76	12.49	4.80E-05	0.001959043
					hsa-miR-320c	1.46	2.74	12.35	5.43E-05	0.002111768

Twenty-seven proteins found downregulated in the Caco-2 cell line relative to the HT-29 cell line and predicted to be targeted by 19 unique anti-correlated/upregulated microRNAs.

without being miRNA-mediated (matching mRNA transcripts & proteins were DE (36), no corresponding anti-correlated miRNA data).

Integration of 3 parallel datasets identifies a group of 132 DE proteins where the control of differential protein expression was unclear

Finally, the data-overlapping strategy identified a list of 132 DE proteins (40 upregulated in Caco-2 relative to HT-29 and 92 downregulated in Caco-2 relative to HT-29) where the corresponding gene transcript was (1) not changing or probeset expression levels were below the detection threshold and (2) were not predicted to be targeted by any of the DE microRNAs, or where the DE miRNA data results were below detection levels. As a result, the control of expression of these proteins (listed together with their attendant annotation information and associated statistical results in Supplementary Table 7) is currently unclear (protein DE without mRNA or corresponding targeting miRNA change).

DISCUSSION

Data analysis methodology identifies unique targets involved in actin organisation

The utilisation of three expression profiling technologies (miRNA, proteomics, mRNA) in parallel as part of an integrative methodology on replicate HT-29 and Caco-2 cell lines provides a unique opportunity for characterising these two cell lines, complementing and adding to the previous literature. A total of 37 differentially-expressed proteins and targeting miRNAs were identified as a result of this unique approach, which has (to our knowledge) never before been employed on analysing these two key intestinal cell line models and which will assist in characterising miRNA/protein regulatory networks that may contribute to their respective phenotypic differences.

The key aspect of the experimental design is the combination of data from multiple expression profiling methods and in-silico prediction to study potential miRNA-protein networks that may give rise to differences between the two cell lines. We have previously shown that the use of such a combined data can address some of the disadvantages of solitary profiling methodologies, while also allowing us to identify potential interacting miRNA-protein networks that would not have been identified by a single dataset[31]. As a result, we have prioritised our discussion of these potential interacting networks (whose identification is possible only *via* the integration of the "tri-omics" (miRNA, protein, mRNA) datasets) which may undergo classical miRNA-mediated translation repression when comparing mucin-producing HT-29 cell phenotype to the more absorptive Caco-2 cell phenotype.

When differing gene- and protein-expression profiles between the cell lines were compared, the key areas of common ontological enrichment between the larger, non-overlapped DE gene and protein lists were in cytoskeleton organization and actin filament bundle assembly, possibly pointing to differences in the shape, support and movement of the cells. We have focused our discussion on a selection of these DE protein and miRNA targets that have demonstrated potential roles in cytoskeleton organization and biogenesis.

Cofilin1 (CFL1), which was downregulated at the protein level in the Caco-2 cell line relative to HT-29 and is targeted by the upregulated miR-182 (Table 1), is part of the ADF/cofilin family of actin-binding proteins which disassemble actin filaments. The role of CFL1 in positively regulating actin filament depolymerisation and severing as well as in cytoskeleton organization, function & biogenesis has been extensively reviewed[34-36]. It has been previously reported that CFL1 levels increase transiently in Caco-2 cells during early differentiation and that this may be linked to the changes required in the actin cytoskeleton[37]. Additionally, several groups have reported that CFL1 has been shown to positively regulate lamellipodium formation and cytoskeletal protein actin projections on the leading edge of highly mobile cells[38-44], although a recent study[45], indicated that it may both inhibit and promote lamellipodium formation. Loss of Cofilin1 has also been shown to hinder apical constriction in mouse endothelial cells[46].

Also targeted by the upregulated miR-182 was the downregulated (in Caco-2 relative to HT-29) protein Cortactin (CTTN)(Table 1), which is a monomeric protein located in the cytoplasm of cells that can be activated by external stimuli to promote actin filament branching and rearrangement of the actin cytoskeleton[47,48] and lamellipodium formation[49,50].

Proteins that were upregulated (in the Caco-2

Table 2 Seven upregulated proteins with no DE mRNA targeted by 15 downregulated miRNAs

Protein					microRNA					
Uniprot	Symbol	Title	Linear FC	Adj. *P* value	miRBase ID	logFC	LinFC	AveExpr	*P* Value	Adj. *P* value
O43852	CALU	Calumenin	1.86	0.0162	hsa-let-7a	-1.44	-2.72	12.58	1.30E-04	0.00396
					hsa-let-7b	-2.61	-6.11	12.84	1.71E-06	0.00017
					hsa-let-7c	-3.45	-10.94	11.26	2.39E-06	0.00021
					hsa-let-7d	-2.66	-6.34	12.12	2.20E-06	0.00020
					hsa-let-7e	-1.40	-2.65	12.39	6.41E-05	0.00236
					hsa-let-7f	-2.85	-7.20	8.75	5.04E-06	0.00034
					hsa-let-7g	-1.06	-2.08	8.90	2.99E-04	0.00737
					hsa-let-7i	-2.31	-4.95	10.22	2.35E-06	0.00021
					hsa-miR-29a	-1.63	-3.09	9.69	1.14E-04	0.00361
					hsa-miR-200a	-2.68	-6.43	9.16	4.14E-04	0.00965
					hsa-miR-200b	-2.54	-5.81	10.40	2.63E-05	0.00129
P20810	CAST	Calpastatin	2.39	0.00011	hsa-miR-125a-3p	-3.14	-8.83	5.84	3.26E-05	0.00150
					hsa-miR-224	-2.44	-5.42	7.35	8.21E-04	0.01667
					hsa-miR-375	-4.06	-16.67	8.29	1.50E-05	0.00083
P21333	FLNA	Filamin-A	3.59	0.00048	hsa-miR-125a-3p	-3.14	-8.83	5.84	3.26E-05	0.00150
P13797	PLS3	Plastin-3	5.88	0.00006	hsa-miR-200b	-2.54	-5.81	10.40	2.63E-05	0.00129
P78330	PSPH	Phosphoserine phosphatase	2.22	0.0004	hsa-miR-200b	-2.54	-5.81	10.40	2.63E-05	0.00129
P35241	RDX	Radixin	5.57	0.00012	hsa-let-7a	-1.44	-2.72	12.58	1.30E-04	0.00396
					hsa-let-7b	-2.61	-6.11	12.84	1.71E-06	0.00017
					hsa-let-7c	-3.45	-10.94	11.26	2.39E-06	0.00021
					hsa-let-7d	-2.66	-6.34	12.12	2.20E-06	0.00020
					hsa-let-7e	-1.40	-2.65	12.39	6.41E-05	0.00236
					hsa-let-7f	-2.85	-7.20	8.75	5.04E-06	0.00034
					hsa-let-7g	-1.06	-2.08	8.90	2.99E-04	0.00737
					hsa-let-7i	-2.31	-4.95	10.22	2.35E-06	0.00021
					hsa-miR-31	-1.94	-3.83	12.79	7.91E-05	0.00274
Q13885	TUBB2A	Tubulin beta-2A chain	1.46	0.00388	hsa-miR-29a	-1.63	-3.09	9.69	1.14E-04	0.00361

Seven proteins found upregulated in the Caco-2 cell line relative to the HT-29 cell line and predicted to be targeted by 15 unique anti-correlated/downregulated microRNAs.

cell line relative to HT-29) included the actin binding protein Radixin (upregulated 5.57 fold) which was targeted by several DE members of the let-7 miRNA family, as well as miR-31 (Table 2). Radixin is one of three members of the ERM family of proteins, which have a number of roles in the stabilisation and regulation of the membrane cytoskeleton interface[51]. Radixin has been found to be important in the linking of ABCC2 (MRP2) and P-glycoprotein to the actin cytoskeleton in Caco-2 and murine small intestinal cells respectively[52,53].

In addition to roles involving the actin cytoskeleton, the group of 37 proteins also includes those with metabolic functions, such as α-Enolase (Eno1) which catalyses 2-phosphoglycerate to phosphoenolpyruvate in glycolysis, is expressed at a 2.06 fold lower level in Caco-2 cells relative to HT-29 cells and is potentially targeted by miR-22 (Table 1). Down regulation of α-enolase has previously been reported as increasing and then decreasing in early differentiating Caco-2 cells and decreased in formula fed preterm pigs developing necrotizing enterocolitis[37,54]. NONO (2.74 fold lower in Caco-2 v's HT-29, Table 1) is involved in regulating androgen receptor and carbonic anhydrase activity[55,56].

Identification of differentially expressed miRNA
The profiling of miRNA expression in Caco-2 compared to HT-29 found 160 miRNAs differentially expressed, a number of which (for example miR-152, miR-99b, miR-125a-5p, miR-10a, miR-196b, miR-222) have been previously identified with intestinal differentiation, function and immune response in a number of species[19-21,57-59]. While in the top 10 upregulated miRNAs in Caco-2 compared to HT-29, an association with cell viability and proliferation has been reported previously for example miR-372 (4743.91 fold up), miR-373 (4451.32 fold up)and miR-122 (57.63 fold up)[60-62]. Similarly among the top 10 downregulated miRNAs in Caco-2 compared to HT-29, there is again an association with cell viability and proliferation for example miR-10a (46.76 fold down), miR-935 (13.38 fold down)[63,64].

Identification of differentially expressed mRNA
A total of 1795 differentially expressed mRNA transcripts were identified DE between HT-29 and Caco-2 cells, comprising of 1084 upregulated and 711 downregulated transcripts (Supplementary Table 3).

GO analysis found that biological processes linked to lipid metabolism, extracellular matrix organisation,

cell adhesion and cytoskeleton were over represented. This enrichment for processes linked to cytoskeleton can be seen in a number of targets both down and up regulated.

For example, Moesin, is one of the largest down regulated transcripts (173 fold) in Caco-2 compared to HT-29 (Supplementary Table 3). Which in addition to being down regulated at transcript level, was also downregulated at a protein level (6.48 fold) in Caco-2 compared to HT-29 (Supplementary Table 2). Moesin is a member of the ERM family of proteins and is involved in regulating actin filament depolymerisation along with apical junction assembly and focal adhesion assembly[51]. Moesin often shows different cellular distribution and expression from the other two ERM proteins, radxin and ezrin[65]. It has been previously reported that it is not expressed Caco-2 or in normal colorectal mucosa epithelial cells[52,66].

Similarly, Galectin 4 (LGALS4), whose transcript was down regulated in Caco-2 compared to HT-29 (109 fold, Supplementary Table 3), has as one of its principal functions the improved stabilisation of the apical membrane rafts and the trafficking of apical cell membrane[67,68]. Variable staining for Galectin 4 has been reported in gut villi, with the upper 1/4 of the villus showing negative staining[69]. Galectin 4 is expressed after confluence in Caco-2 cells along with the formation of brush boarder[68]. Galectin 4 linked to the segregation of glycolproteins for apical secretion in HT-29 cells[68].

The Ca^{2+} binding protein Calbindin 2 (CALB2), also known as Calretinin was down regulated 102 fold in Caco-2 compared to HT-29 cells (Supplementary Table 3). The exact role of calretinin in intestinal epithelia is unknown but it is believed to have roles in the stabilisation of cytokeratins and microtubules, differentiation status and butyrate response[70-72]. The expression profile obtained is in agreement with previous studies which report expression in undifferentiated colon adenocarcinoma cell lines (*e.g.,* HT-29) and no expression in for normal differentiated colon epithelial cells and more differentiated cell lines (*e.g.,* Caco-2) studies[71].

In case of transcripts that were up regulated in Caco-2 relative to HT-29, again a number of targets were linked to cytoskeleton function, for example Fibronectin1 (FN1), whose, mRNA was upregulated 441 fold in the Caco-2 cell line relative to HT-29 (Supplementary Table 3). In addition to having an elevated transcript, fibronectin protein was upregulated 5.15 fold up at a protein in Caco-2 relative to HT-29. Fibronectin is a glycoprotein of the extracellular matrix that binds to membrane-spanning receptor integrins and plays a major role in cell adhesion and actin cytoskeleton organisation[73]. Previously the differentiation of Caco-2 cells has been associated with

a down regulation of fibronectin[74,75], but our studies were on proliferating cells.

Another transcript with associated cytoskeleton function that was upregulated was BicC 1 (up regulated 206.75 fold in the Caco-2 cell line relative to HT-29, Supplementary Table 3). BicC1 is believed to bind and regulate mRNA translation *via* a number of mechanisms such as clustering of target mRNAs or as a chaperone which recruits specific miRNA precursors and Dicer to target mRNAs and subsequently transfers these to AGO for silencing[76,77]. Its expression has been linked positively to the formation of E-cadherin adherens junctions and cortical actin distribution[78].

Also upregulated (192.74 fold) at a transcript level was LRP2 also know Megalin, a member of the low density lipoprotein family (Supplementary Table 3). LRP2/Megalin has been found expressed on the apical surface of a number of epithelial cells, including the brush border membrane of Caco-2 cells and to have upregulated expression in the ileum of sucking rats[79-81].

Identification of differentially expressed protein

A total of 168 differentially expressed proteins were in expressed between Caco-2 and HT-29 (57 were upregulated in Caco-2 and 111 were upregulated in HT-29). Using Pathway Studio to generate GO biological processes for the DE proteins (Supplemental Table 1), it was found that there was an over representation of processes linked to Translation, Carbohydrate metabolism and Actin cytoskeleton organisation. For example Insulin-like growth factor 2 mRNA binding protein (IGF2BP1, IMP1) was upregulated 51.58 fold in Caco-2 relative to HT-29. The expression of IGF2BP1 has previously been linked to morphogenesis of the small intestine[82]

Interestingly, the expression of Insulin-like growth factor 2 mRNA binding protein has been linked to SERPINH1 a 47 KDa stress protein, also called Heat Shock Protein 47 (HSP47). HSP47 is upregulated 9.11 fold in this study in Caco-2 relative to HT-29 (Supplementary Table 2). HSP47 is localised to the endoplasmic reticulum where it is a principle chaperone in the collagen biosynthesis pathway. In addition to HSP47's role as a collagen chaperone it has been suggested that it plays an important role in reorganising of the actin cytoskeleton in Caco-2, increasing permeability[83].

The protein showing the biggest down regulation (down regulated 91.89 fold) in Caco-2 relative to HT-29 was the secreted glycoprotein Galectin-3-binding protein (also known as Mac2-binding protein) (Supplementary Table 2). Galectin-3-Binding protein has been linked to the innate immune response[84,85].

While the second most down regulated protein in Caco-2 relative to HT-29 was Annexin A13 (down regulated in 42.38 fold, Supplementary Table 2).

Annexin A13 is found expressed in the small intestine, involved in the localisation of protein to the apical membrane of cells[86,87].

Overlap comparison of individual datasets with already-published studies identifies several common DE gene, protein and miRNA candidates

In order to determine the comparability of our analysis of these cell line models to other published studies on those same cell lines, we overlapped our three individual DE datasets (gene, protein and miRNA) with a selection of two studies that each combined both cell lines (HT-29 and Caco-2) in the one analysis.

For the miRNA lists, we utilised two studies examining the role of miRNAs in colorectal cancer metastasis in the HT-29 and Caco-2 cell lines; Ma et al[88] (2011) which reported a DE list of just 8 microRNAs and Zhang et al[57] (2012) who disclosed a DE list of 16 microRNAs. Despite the shortness of these previously published lists, we identified the hsa-miR-196a microRNA as commonly DE in both our study and that of Ma et al[88] (2011). Additionally, overlapping our 160 microRNAs with the 16 microRNAs identified by Zhang et al[57]. (2012) identified a further 3 microRNAs (hsa-let-7a, hsa-miR-10a and hsa-miR-98) which were commonly DE between HT-29 and Caco-2 in both studies (see Table 3)[88].

At the protein level, we incorporated two proteomics profiling studies combining HT-29 and Caco-2; one utilising two-dimensional gel electrophoresis and a newer study incorporating LC-MS/MS[89,90]. Overlap analysis of our DE gene dataset with these two studies identified only one protein, LGALS3 (Galectin-3) as DE between all three studies (Table 3). Individual overlaps with each of the two studies separately identified an additional four proteins [ANXA3 (Annexin A3); NCL (Nucleolin); NPM1 (Nucleophosmin) and SORD (Sorbitol dehydrogenase)] commonly DE between our study and that of Uzozie et al[90]., 2014, with a further 12 proteins (ATP5B (ATP synthase subunit beta, mitochondrial); CALR (Calreticulin); CFL1 (Cofilin-1); CTSB (Cathepsin B); ENO1 (Alpha-enolase); HNRNPA2B1 (Heterogeneous nuclear ribonucleoproteins A2/B1); HNRNPL (Heterogeneous nuclear ribonucleoprotein L); HNRNPK (Heterogeneous nuclear ribonucleoprotein K); KRT8 (Keratin, type II cytoskeletal 8); P4HB (Protein disulfide-isomerase); PDIA3 (Protein disulfide-isomerase A3) and PGAM1 (Phosphoglycerate mutase 1) commonly DE between Lenaerts et al[89] (2007) and our study (Table 3)[90].

To overlap our DE genes, we utilised a study examining the expression of 377 xenobiotics metabolism genes (incl. xenobiotic-metabolizing enzymes, transporters, and nuclear receptors and transcription factors)[91], as well as a newer study examining Copper homeostasis and distribution, which utilised the same microarray chip that was used in this study[92] but only disclosed data on 33 DE genes. Overlap analysis of our DE gene dataset with these two studies identified only one gene, the ATF7B (ATPase, Cu^{2+} transporting, beta polypeptide) transporter as DE between all three studies (Table 3). Overlapping our dataset with only that of Bourgine et al[91]. (2012) identified an additional 25 genes (ABCA3, ABCB1, AKR1E2, ALDH5A1, ALDH6A1, CDCA2, CYP20A1, CYP27A1, EPHX2, GSTA4, NCOA1, NR5A2, RBP2, SLC15A1, SLC19A3, SLC1A3, SLC22A5, SLC6A4, SLC7A6, SLC7A8, SLCO2B1, SLCO4C1, SULT1C4, THRB and VDAC3) commonly DE between both cell lines[91] . The greater degree of overlap identified was most likely due to the roughly 10-fold greater number of DE genes disclosed (377) by Bourgine et al[91] (2012), compared to the 33 genes disclosed by Barresi et al[92] (2016).

DE proteins targeted by microRNAs in other studies

In this study, we propose that several of the proteins described here as DE between the HT-29 and Caco-2 cell lines may be targeted by anti-correlated microRNAs (also DE here between these cell lines) and several of these protein-miRNA regulatory interactions have been demonstrated in other cell systems. For example, microRNA-148a has been shown to inhibit the proliferation and promote the paclitaxel-induced apoptosis of ovarian cancer cells by targeting PDIA3[93]. Additionally, YWHAZ (14-3-3 zeta) was confirmed to be a target of miR-302d by proteomic comparison of Human embryonic stem (ES) cells and their differentiated T3DF fibroblasts[94], while LASP1 has been shown to be directly regulated by miR-145 in bladder cancer cell lines[95-97].

In conclusion the intestinal cell lines Caco-2 and HT-29 are commonly used for the creation of in vitro models for the intestine. However, the use of such cell lines to model the intestine requires that they are fully characterised which will require a better understanding on the molecular controls that these cells exhibit. One such molecular control is the miRNA control mechanism of translation. This study, by performing a tri-omics analysis (genes, proteomics, miRNAs) on parallel data sets generated from these two cell lines, potential miRNA-protein networks were identified that would not have been immediately apparent otherwise. An advantage of such an approach was that the potential biological noise of the study was reduced by running identical samples in parallel to generate each of the profiles. This has the effect of reducing the potential false negative false positive rates of the in silico prediction and allows for high priority list of candidates for functional validation. It is interesting to note that a large number of targets were associated with the actin cytoskeleton and its associated cell processes such as motility, polarisation, endocytosis and cell division. Most of the identified proteins in table I and table II have important roles in the establishment of apical-basal membrane polarity and the formation of the apical microvilli brush border

Table 3 Overlap of DE genes, proteins, microRNAs with other HT-29/Caco-2 profiling studies

Identifier	Symbol	Description/Family	LinFC	*P* value	Overlapping study	Dataset
hsa-miR-196a_st	hsa-miR-196a	miR-196abc	-5.06	8.53E-05	88	microRNA
hsa-let-7a_st	hsa-let-7a	let-7/98/4458/4500	-2.72	6.34E-04	57	
hsa-miR-10a_st	hsa-miR-10a	miR-10abc/10a-5p	-46.76	7.40E-06	57	
hsa-miR-98_st	hsa-miR-98_st	miR-98	-11.90	4.50E-06	57	
P17931	LGALS3	Galectin-3	2.09	1.10E-03	89; 90	Protein
P06576	ATP5B	ATP synthase subunit beta, mitochondrial	2.07	1.20E-03	89	
P27797	CALR	Calreticulin	1.62	3.35E-02	89	
P23528	CFL1	Cofilin-1	1.53	5.94E-03	89	
P07858	CTSB	Cathepsin B	5.15	5.36E-05	89	
P06733	ENO1	Alpha-enolase	2.06	4.65E-03	89	
P22626	HNRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1	1.72	9.62E-03	89	
P61978	HNRNPK	Heterogeneous nuclear ribonucleoprotein K	1.38	4.38E-04	89	
P14866	HNRNPL	Heterogeneous nuclear ribonucleoprotein L	1.69	4.02E-03	89	
P05787	KRT8	Keratin, type Ⅱ cytoskeletal 8	2.21	1.81E-03	89	
P07237	P4HB	Protein disulfide-isomerase	1.41	3.49E-03	89	
P30101	PDIA3	Protein disulfide-isomerase A3	1.84	7.49E-05	89	
P18669	PGAM1	Phosphoglycerate mutase 1	1.93	1.42E-04	89	
P12429	ANXA3	Annexin A3	2.53	1.55E-06	90	
P19338	NCL	Nucleolin	2.74	4.80E-03	90	
P06748	NPM1	Nucleophosmin	1.51	1.15E-02	90	
Q00796	SORD	Sorbitol dehydrogenase	4.21	7.73E-05	90	
16779311	ATP7B	ATPase, Cu++ transporting, beta polypeptide	3.60	1.57E-05	91; 92	Gene
16823109	ABCA3	ATP-binding cassette, sub-family A (ABC1), member 3	10.07	1.75E-05	91	
17059491	ABCB1	ATP-binding cassette, sub-family B (MDR/TAP), member 1	36.64	2.14E-06		
16701957	AKR1E2	aldo-keto reductase family 1, member E2	4.56	1.09E-05		
17005368	ALDH5A1	aldehyde dehydrogenase 5 family, member A1	3.38	4.08E-06		
16794632	ALDH6A1	aldehyde dehydrogenase 6 family, member A1	2.27	7.54E-05		
17067102	CDCA2	carboxylesterase 1	2.32	2.28E-03		
16889762	CYP20A1	cytochrome P450, family 20, subfamily A, polypeptide 1	2.78	8.71E-04		
16891082	CYP27A1	cytochrome P450, family 27, subfamily A, polypeptide 1	20.68	2.05E-06		
17067284	EPHX2	epoxide hydrolase 2, cytoplasmic	6.51	3.62E-06		
17020103	GSTA4	glutathione S-transferase alpha 4	2.25	4.04E-04		
16877728	NCOA1	nuclear receptor coactivator 1	2.99	7.90E-04		
16675638	NR5A2	nuclear receptor subfamily 5, group A, member 2	2.33	1.73E-03		
16959797	RBP2	retinol binding protein 2, cellular	27.20	2.06E-06		
16780481	SLC15A1	solute carrier family 15 (oligopeptide transporter), member 1	20.05	1.95E-06		
16909257	SLC19A3	solute carrier family 19 (thiamine transporter), member 3	35.13	6.97E-08		
16984056	SLC1A3	solute carrier family 1 (glial high affinity glutamate transporter), member 3	50.97	1.62E-07		
16989018	SLC22A5	solute carrier family 22 (organic cation/carnitine transporter), member 5	2.21	5.49E-04		
16843078	SLC6A4	solute carrier family 6 (neurotransmitter transporter), member 4	9.33	1.33E-06		
16820398	SLC7A6	solute carrier family 7 (amino acid transporter light chain, y+L system), member 6	4.71	5.00E-06		
16790744	SLC7A8	solute carrier family 7 (amino acid transporter light chain, L system), member 8	11.99	4.99E-06		
16729064	SLCO2B1	solute carrier organic anion transporter family, member 2B1	10.15	8.38E-06		
16998551	SLCO4C1	solute carrier organic anion transporter family, member 4C1	11.95	2.59E-03		
16884050	SULT1C4	sulfotransferase family, cytosolic, 1C, member 4	15.89	2.49E-07		
16951567	THRB	thyroid hormone receptor, beta	4.38	1.90E-04		
17068541	VDAC3	voltage-dependent anion channel 3	2.28	3.11E-04		

List of 49 gene, protein and miRNA annotated candidates (6 microRNAs, 17 proteins, 26 genes) which are differentially expressed between the HT-29 and Caco-2 cell lines in this and other published profiling studies that combine both cell lines.

in enterocytes. The data in this study substantially expands the lists of differentially expressed miRNAs, Proteins and mRNAs between Caco-2 and HT-29 cell lines. Furthermore it is to our knowledge the first to provide "tri-omics" analysis (genes, proteomics, miRNAs) on the Caco-2 and HT29 cell lines in combination. It uses the availability of data on multiple levels to identify potential interactions that would not have been identified by a single dataset. Allowing us to provide new information for the characterising of these two cell lines, which are important in intestinal modelling, and help focus on the most likely miRNA candidates for further analysis.

COMMENTS

Background

The development of suitable *in vitro* models of the intestine is of great interest

to the food and pharmaceutical industries. Two commonly used cell lines for the generation of such *in vitro* models are Caco-2 and HT-29. The use of "-omics" studies (transcriptomics, proteomics and metabolomics) has provided insights into how *in vitro* models work in comparison to in vivo scenarios and how to improve them. However, the complexity involved identifying of miRNA targets still remains a significant challenge to researchers. This study addresses that complexity by using a "tri-omics" approach to identify targets.

Research frontiers

The development and maturation of "-omics" technologies has revolutionised our understanding of biological processes. However, while there have been previous studies investigating the differences between theses cell lines, these have focused on only one aspect at a time.

Innovations and breakthroughs

This study is the first to use an integrated "tri-omics" analysis (mRNA, proteomics, miRNAs) on Caco-2 and HT-29 to identify potential miRNA-protein networks were identified that would not have been immediately apparent otherwise. These networks contain a large number of targets associated with the actin cytoskeleton and its associated cell processes such as motility, polarisation, and endocytosis.

Applications

The use of an integrated "tri-omics" analysis will support a better understanding of the molecular events underpinning the different biological behaviours of these two cell lines, which are so important to pharmacological and nutritional research.

Terminology

Ward's method is a robust clustering and is defined as the proximity between two clusters as the increase in squared error that results when two clusters are merged.

Peer-review

This study is well elaborated, uses innovative methodology and supports organoid studies aiming at greater knowledge of enteric epithelial cells.

REFERENCES

1 **Vázquez M**, Calatayud M, Vélez D, Devesa V. Intestinal transport of methylmercury and inorganic mercury in various models of Caco-2 and HT29-MTX cells. *Toxicology* 2013; **311**: 147-153 [PMID: 23793072 DOI: 10.1016/j.tox.2013.06.002]

2 **Pullakhandam R**, Nair KM, Pamini H, Punjal R. Bioavailability of iron and zinc from multiple micronutrient fortified beverage premixes in Caco-2 cell model. *J Food Sci* 2011; **76**: H38-H42 [PMID: 21535765 DOI: 10.1111/j.1750-3841.2010.01993.x]

3 **Gagnon M**, Zihler Berner A, Chervet N, Chassard C, Lacroix C. Comparison of the Caco-2, HT-29 and the mucus-secreting HT29-MTX intestinal cell models to investigate Salmonella adhesion and invasion. *J Microbiol Methods* 2013; **94**: 274-279 [PMID: 23835135 DOI: 10.1016/j.mimet.2013.06.027]

4 **Calatayud M**, Devesa V, Vélez D. Differential toxicity and gene expression in Caco-2 cells exposed to arsenic species. *Toxicol Lett* 2013; **218**: 70-80 [PMID: 23353816 DOI: 10.1016/j.toxlet.2013.01.013]

5 **Nakamura Y**, Yogosawa S, Izutani Y, Watanabe H, Otsuji E, Sakai T. A combination of indol-3-carbinol and genistein synergistically induces apoptosis in human colon cancer HT-29 cells by inhibiting Akt phosphorylation and progression of autophagy. *Mol Cancer* 2009; **8**: 100 [PMID: 19909554 DOI: 10.1186/1476-4598-8-100]

6 **Barker N**. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nat Rev Mol Cell Biol* 2014; **15**: 19-33 [PMID: 24326621 DOI: 10.1038/nrm3721]

7 **Fogh J**, Wright WC, Loveless JD. Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *J Natl Cancer Inst* 1977; **58**: 209-214 [PMID: 833871]

8 **Rousset M**, Laburthe M, Pinto M, Chevalier G, Rouyer-Fessard C, Dussaulx E, Trugnan G, Boige N, Brun JL, Zweibaum A. Enterocytic differentiation and glucose utilization in the human colon tumor cell line Caco-2: modulation by forskolin. *J Cell Physiol* 1985; **123**: 377-385 [PMID: 2985631 DOI: 10.1002/jcp.1041230313]

9 **Sambuy Y**, De Angelis I, Ranaldi G, Scarino ML, Stammati A, Zucco F. The Caco-2 cell line as a model of the intestinal barrier: influence of cell and culture-related factors on Caco-2 cell functional characteristics. *Cell Biol Toxicol* 2005; **21**: 1-26 [PMID: 15868485 DOI: 10.1007/s10565-005-0085-6]

10 **Shah P**, Jogani V, Bagchi T, Misra A. Role of Caco-2 cell monolayers in prediction of intestinal drug absorption. *Biotechnol Prog* 2006; **22**: 186-198 [PMID: 16454510 DOI: 10.1021/bp050208u]

11 **Huet C**, Sahuquillo-Merino C, Coudrier E, Louvard D. Absorptive and mucus-secreting subclones isolated from a multipotent intestinal cell line (HT-29) provide new models for cell polarity and terminal differentiation. *J Cell Biol* 1987; **105**: 345-357 [PMID: 3611191 DOI: 10.1083/jcb.105.1.345]

12 **Maoret JJ**, Font J, Augeron C, Codogno P, Bauvy C, Aubery M, Laboisse CL. A mucus-secreting human colonic cancer cell line. Purification and partial characterization of the secreted mucins. *Biochem J* 1989; **258**: 793-799 [PMID: 2658973]

13 **Augeron C**, Laboisse CL. Emergence of permanently differentiated cell clones in a human colonic cancer cell line in culture after treatment with sodium butyrate. *Cancer Res* 1984; **44**: 3961-3969 [PMID: 6744312]

14 **Lesuffleur T**, Barbat A, Dussaulx E, Zweibaum A. Growth adaptation to methotrexate of HT-29 human colon carcinoma cells is associated with their ability to differentiate into columnar absorptive and mucus-secreting cells. *Cancer Res* 1990; **50**: 6334-6343 [PMID: 2205381]

15 **Chastre E**, Emami S, Rosselin G, Gespach C. Vasoactive intestinal peptide receptor activity and specificity during enterocyte-like differentiation and retrodifferentiation of the human colonic cancerous subclone HT29-18. *FEBS Lett* 1985; **188**: 197-204 [PMID: 2993022 DOI: 10.1016/0014-5793(85)80371-9]

16 **Mahler GJ**, Shuler ML, Glahn RP. Characterization of Caco-2 and HT29-MTX cocultures in an in vitro digestion/cell culture model used to predict iron bioavailability. *J Nutr Biochem* 2009; **20**: 494-502 [PMID: 18715773 DOI: 10.1016/j.jnutbio.2008.05.006]

17 **Antunes F**, Andrade F, Araújo F, Ferreira D, Sarmento B. Establishment of a triple co-culture in vitro cell models to study intestinal absorption of peptide drugs. *Eur J Pharm Biopharm* 2013; **83**: 427-435 [PMID: 23159710 DOI: 10.1016/j.ejpb.2012.10.003]

18 **Pereira C**, Araújo F, Barrias CC, Granja PL, Sarmento B. Dissecting stromal-epithelial interactions in a 3D in vitro cellularized intestinal model for permeability studies. *Biomaterials* 2015; **56**: 36-45 [PMID: 25934277 DOI: 10.1016/j.biomaterials.2015.03.054]

19 **Tao X**, Xu Z. MicroRNA transcriptome in swine small intestine during weaning stress. *PLoS One* 2013; **8**: e79343 [PMID: 24260202 DOI: 10.1371/journal.pone.0079343]

20 **Liang G**, Malmuthuge N, McFadden TB, Bao H, Griebel PJ, Stothard P, Guan le L. Potential regulatory role of microRNAs in the development of bovine gastrointestinal tract during early life. *PLoS One* 2014; **9**: e92592 [PMID: 24682221 DOI: 10.1371/journal.pone.0092592]

21 **Dalmasso G**, Nguyen HT, Yan Y, Laroui H, Srinivasan S, Sitaraman SV, Merlin D. MicroRNAs determine human intestinal epithelial cell fate. *Differentiation* 2010; **80**: 147-154 [PMID: 20638171 DOI: 10.1016/j.diff.2010.06.005]

22 **Bagga S**, Bracht J, Hunter S, Massirer K, Holtz J, Eachus R, Pasquinelli AE. Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* 2005; **122**: 553-563 [PMID: 16122423 DOI: 10.1016/j.cell.2005.07.031]

23 **Wu L**, Belasco JG. Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. *Mol Cell Biol* 2005; **25**: 9198-9208 [PMID: 16227573 DOI: 10.1128/MCB.25.21.9198-9208.2005]

24   **Pillai RS**, Bhattacharyya SN, Artus CG, Zoller T, Cougot N, Basyuk E, Bertrand E, Filipowicz W. Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* 2005; **309**: 1573-1576 [PMID: 16081698 DOI: 10.1126/science.1115079]

25   **Mathonnet G**, Fabian MR, Svitkin YV, Parsyan A, Huck L, Murata T, Biffo S, Merrick WC, Darzynkiewicz E, Pillai RS, Filipowicz W, Duchaine TF, Sonenberg N. MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. *Science* 2007; **317**: 1764-1767 [PMID: 17656684 DOI: 10.1126/science.1146067]

26   **Fu L**, Shi Z, Luo G, Tu W, Wang X, Fang Z, Li X. Multiple microRNAs regulate human FOXP2 gene expression by targeting sequences in its 3' untranslated region. *Mol Brain* 2014; **7**: 71 [PMID: 25269856 DOI: 10.1186/s13041-014-0071-0]

27   **Lim LP**, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 2005; **433**: 769-773 [PMID: 15685193 DOI: 10.1038/nature03315]

28   **John B**, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human MicroRNA targets. *PLoS Biol* 2004; **2**: e363 [PMID: 15502875 DOI: 10.1371/journal.pbio.0020363]

29   **Lewis BP**, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell* 2003; **115**: 787-798 [PMID: 14697198 DOI: 10.1016/S0092-8674(03)01018-3]

30   **Krek A**, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N. Combinatorial microRNA target predictions. *Nat Genet* 2005; **37**: 495-500 [PMID: 15806104 DOI: 10.1038/ng1536]

31   **Clarke C**, Henry M, Doolan P, Kelly S, Aherne S, Sanchez N, Kelly P, Kinsella P, Breen L, Madden SF, Zhang L, Leonard M, Clynes M, Meleady P, Barron N. Integrated miRNA, mRNA and protein expression analysis reveals the role of post-transcriptional regulation in controlling CHO cell growth rate. *BMC Genomics* 2012; **13**: 656 [PMID: 23170974 DOI: 10.1186/1471-2164-13-656]

32   **Gallagher C**, Clarke C, Aherne ST, Katikireddy KR, Doolan P, Lynch V, Shaw S, Bobart-Hone A, Murphy C, Clynes M, Power W, O'Sullivan F. Comparative transcriptomic analysis of cultivated limbal epithelium and donor corneal tissue reveals altered wound healing gene expression. *Invest Ophthalmol Vis Sci* 2014; **55**: 5795-5805 [PMID: 25125605 DOI: 10.1167/iovs.14-14664]

33   **Nikitin A**, Egorov S, Daraselia N, Mazo I. Pathway studio--the analysis and navigation of molecular networks. *Bioinformatics* 2003; **19**: 2155-2157 [PMID: 14594725]

34   **Bravo-Cordero JJ**, Magalhaes MA, Eddy RJ, Hodgson L, Condeelis J. Functions of cofilin in cell locomotion and invasion. *Nat Rev Mol Cell Biol* 2013; **14**: 405-415 [PMID: 23778968 DOI: 10.1038/nrm3609]

35   **Ono S**. Mechanism of depolymerization and severing of actin filaments and its significance in cytoskeletal dynamics. *Int Rev Cytol* 2007; **258**: 1-82 [PMID: 17338919 DOI: 10.1016/S0074-7696(07)58001-0]

36   **De La Cruz EM**, Gardel ML. Actin Mechanics and Fragmentation. *J Biol Chem* 2015; **290**: 17137-17144 [PMID: 25957404 DOI: 10.1074/jbc.R115.636472]

37   **Stierum R**, Gaspari M, Dommels Y, Ouatas T, Pluk H, Jespersen S, Vogels J, Verhoeckx K, Groten J, van Ommen B. Proteome analysis reveals novel proteins associated with proliferation and differentiation of the colorectal cancer cell line Caco-2. *Biochim Biophys Acta* 2003; **1650**: 73-91 [PMID: 12922171 DOI: 10.1016/S1570-9639(03)00204-8]

38   **Nagata-Ohashi K**, Ohta Y, Goto K, Chiba S, Mori R, Nishita M, Ohashi K, Kousaka K, Iwamatsu A, Niwa R, Uemura T, Mizuno K. A pathway of neuregulin-induced activation of cofilin-phosphatase Slingshot and cofilin in lamellipodia. *J Cell Biol* 2004; **165**: 465-471 [PMID: 15159416 DOI: 10.1083/jcb.200401136]

39   **Kiuchi T**, Ohashi K, Kurita S, Mizuno K. Cofilin promotes stimulus-induced lamellipodium formation by generating an abundant supply of actin monomers. *J Cell Biol* 2007; **177**:

465-476 [PMID: 17470633 DOI: 10.1083/jcb.200610005]

40   **van Rheenen J**, Condeelis J, Glogauer M. A common cofilin activity cycle in invasive tumor cells and inflammatory cells. *J Cell Sci* 2009; **122**: 305-311 [PMID: 19158339 DOI: 10.1242/jcs.031146]

41   **Staser K**, Shew MA, Michels EG, Mwanthi MM, Yang FC, Clapp DW, Park SJ. A Pak1-PP2A-ERM signaling axis mediates F-actin rearrangement and degranulation in mast cells. *Exp Hematol* 2013; **41**: 56-66.e2 [PMID: 23063725 DOI: 10.1016/j.exphem.2012.10.001]

42   **Mizuno K**. Signaling mechanisms and functional roles of cofilin phosphorylation and dephosphorylation. *Cell Signal* 2013; **25**: 457-469 [PMID: 23153585 DOI: 10.1016/j.cellsig.2012.11.001]

43   **Maizels Y**, Oberman F, Miloslavski R, Ginzach N, Berman M, Yisraeli JK. Localization of cofilin mRNA to the leading edge of migrating cells promotes directed cell migration. *J Cell Sci* 2015; **128**: 1922-1933 [PMID: 25908858 DOI: 10.1242/jcs.163972]

44   **Zhang L**, Luo J, Wan P, Wu J, Laski F, Chen J. Regulation of cofilin phosphorylation and asymmetry in collective cell migration during morphogenesis. *Development* 2011; **138**: 455-464 [PMID: 21205790 DOI: 10.1242/dev.046870]

45   **Wu Q**, Jiang Y, Cui S, Wang Y, Wu X. The role of cofilin-l in vulvar squamous cell carcinoma: A marker of carcinogenesis, progression and targeted therapy. *Oncol Rep* 2016; **35**: 2743-2754 [PMID: 26936386 DOI: 10.3892/or.2016.4625]

46   **Jodoin JN**, Coravos JS, Chanet S, Vasquez CG, Tworoger M, Kingston ER, Perkins LA, Perrimon N, Martin AC. Stable Force Balance between Epithelial Cells Arises from F-Actin Turnover. *Dev Cell* 2015; **35**: 685-697 [PMID: 26688336 DOI: 10.1016/j.devcel.2015.11.018]

47   **Ammer AG**, Weed SA. Cortactin branches out: roles in regulating protrusive actin dynamics. *Cell Motil Cytoskeleton* 2008; **65**: 687-707 [PMID: 18615630 DOI: 10.1002/cm.20296]

48   **Gallo G**. More than one ring to bind them all: recent insights into the structure of the axon. *Dev Neurobiol* 2013; **73**: 799-805 [PMID: 23784998 DOI: 10.1002/dneu.22100]

49   **Weed SA**, Karginov AV, Schafer DA, Weaver AM, Kinley AW, Cooper JA, Parsons JT. Cortactin localization to sites of actin assembly in lamellipodia requires interactions with F-actin and the Arp2/3 complex. *J Cell Biol* 2000; **151**: 29-40 [PMID: 11018051]

50   **Kinley AW**, Weed SA, Weaver AM, Karginov AV, Bissonette E, Cooper JA, Parsons JT. Cortactin interacts with WIP in regulating Arp2/3 activation and membrane protrusion. *Curr Biol* 2003; **13**: 384-393 [PMID: 12620186]

51   **Arpin M,** Chirivino D, Naba A, Zwaenepoel I. Emerging role for ERM proteins in cell adhesion and migration. *Cell Adh Migr* 2011; **5**: 199-206 [PMID: 21343695]

52   **Yang Q**, Onuki R, Nakai C, Sugiyama Y. Ezrin and radixin both regulate the apical membrane localization of ABCC2 (MRP2) in human intestinal epithelial Caco-2 cells. *Exp Cell Res* 2007; **313**: 3517-3525 [PMID: 17825285 DOI: 10.1016/j.yexcr.2007.07.033]

53   **Kobori T**, Harada S, Nakamoto K, Tokuyama S. Radixin influences the changes in the small intestinal P-glycoprotein by etoposide treatment. *Biol Pharm Bull* 2013; **36**: 1822-1828 [PMID: 24189426 DOI: 10.1248/bpb.b13-00511]

54   **Jiang P**, Siggers JL, Ngai HH, Sit WH, Sangild PT, Wan JM. The small intestine proteome is changed in preterm pigs developing necrotizing enterocolitis in response to formula feeding. *J Nutr* 2008; **138**: 1895-1901 [PMID: 18806098]

55   **Kivelä AJ**, Kivelä J, Saarnio J, Parkkila S. Carbonic anhydrases in normal gastrointestinal tract and gastrointestinal tumours. *World J Gastroenterol* 2005; **11**: 155-163 [PMID: 15633208 DOI: 10.3748/wjg.v11.i2.155]

56   **Kuwahara S**, Ikei A, Taguchi Y, Tabuchi Y, Fujimoto N, Obinata M, Uesugi S, Kurihara Y. PSPC1, NONO, and SFPQ are expressed in mouse Sertoli cells and may function as coregulators of androgen receptor-mediated transcription. *Biol Reprod* 2006; **75**: 352-359 [PMID: 16641145 DOI: 10.1095/biolreprod.106.051136]

57   **Zhang P**, Ma Y, Wang F, Yang J, Liu Z, Peng J, Qin H. Comprehensive gene and microRNA expression profiling reveals the crucial role of hsa-let-7i and its target genes in colorectal cancer

metastasis. *Mol Biol Rep* 2012; **39**: 1471-1478 [PMID: 21625861 DOI: 10.1007/s11033-011-0884-1]

58   **Liang G**, Malmuthuge N, Guan le L, Griebel P. Model systems to analyze the role of miRNAs and commensal microflora in bovine mucosal immune system development. *Mol Immunol* 2015; **66**: 57-67 [PMID: 25467799 DOI: 10.1016/j.molimm.2014.10.014]

59   **McKenna LB**, Schug J, Vourekas A, McKenna JB, Bramswig NC, Friedman JR, Kaestner KH. MicroRNAs control intestinal epithelial differentiation, architecture, and barrier function. *Gastroenterology* 2010; **139**: 1654-1664, 1664.e1 [PMID: 20659473 DOI: 10.1053/j.gastro.2010.07.040]

60   **Nakano H**, Miyazawa T, Kinoshita K, Yamada Y, Yoshida T. Functional screening identifies a microRNA, miR-491 that induces apoptosis by targeting Bcl-X(L) in colorectal cancer cells. *Int J Cancer* 2010; **127**: 1072-1080 [PMID: 20039318 DOI: 10.1002/ijc.25143]

61   **Tanaka T**, Arai M, Wu S, Kanda T, Miyauchi H, Imazeki F, Matsubara H, Yokosuka O. Epigenetic silencing of microRNA-373 plays an important role in regulating cell proliferation in colon cancer. *Oncol Rep* 2011; **26**: 1329-1335 [PMID: 21785829 DOI: 10.3892/or.2011.1401]

62   **Kanaan Z**, Rai SN, Eichenberger MR, Barnes C, Dworkin AM, Weller C, Cohen E, Roberts H, Keskey B, Petras RE, Crawford NP, Galandiuk S. Differential microRNA expression tracks neoplastic progression in inflammatory bowel disease-associated colorectal cancer. *Hum Mutat* 2012; **33**: 551-560 [PMID: 22241525 DOI: 10.1002/humu.22021]

63   **Stadthagen G**, Tehler D, Høyland-Kroghsbo NM, Wen J, Krogh A, Jensen KT, Santoni-Rugiu E, Engelholm LH, Lund AH. Loss of miR-10a activates lpo and collaborates with activated Wnt signaling in inducing intestinal neoplasia in female mice. *PLoS Genet* 2013; **9**: e1003913 [PMID: 24204315 DOI: 10.1371/journal.pgen.1003913]

64   **Yang M**, Cui G, Ding M, Yang W, Liu Y, Dai D, Chen L. miR-935 promotes gastric cancer cell proliferation by targeting SOX7. *Biomed Pharmacother* 2016; **79**: 153-158 [PMID: 27044823 DOI: 10.1016/j.biopha.2016.01.011]

65   **Berryman M**, Franck Z, Bretscher A. Ezrin is concentrated in the apical microvilli of a wide variety of epithelial cells whereas moesin is found primarily in endothelial cells. *J Cell Sci* 1993; **105 (Pt 4)**: 1025-1043 [PMID: 8227193]

66   66 **Kim CY**, Jung WY, Lee HJ, Kim HK, Kim A, Shin BK. Proteomic analysis reveals overexpression of moesin and cytokeratin 17 proteins in colorectal carcinoma. *Oncol Rep* 2012; **27**: 608-620 [PMID: 22076435 DOI: 10.3892/or.2011.1545]

67   **Danielsen EM**, Hansen GH. Lipid raft organization and function in the small intestinal brush border. *J Physiol Biochem* 2008; **64**: 377-382 [PMID: 19391463 DOI: 10.1007/BF03174093]

68   **Stechly L**, Morelle W, Dessein AF, André S, Grard G, Trinel D, Dejonghe MJ, Leteurtre E, Drobecq H, Trugnan G, Gabius HJ, Huet G. Galectin-4-regulated delivery of glycoproteins to the brush border membrane of enterocyte-like cells. *Traffic* 2009; **10**: 438-450 [PMID: 19192249 DOI: 10.1111/j.1600-0854.2009.00882.x]

69   **Nio-Kobayashi J**, Takahashi-Iwanaga H, Iwanaga T. Immunohistochemical localization of six galectin subtypes in the mouse digestive tract. *J Histochem Cytochem* 2009; **57**: 41-50 [PMID: 18796404 DOI: 10.1369/jhc.2008.952317]

70   **Cargnello R**, Celio MR, Schwaller B, Gotzos V. Change of calretinin expression in the human colon adenocarcinoma cell line HT29 after differentiation. *Biochim Biophys Acta* 1996; **1313**: 201-208 [PMID: 8898855]

71   **Häner K**, Henzi T, Pfefferli M, Künzli E, Salicio V, Schwaller B. A bipartite butyrate-responsive element in the human calretinin (CALB2) promoter acts as a repressor in colon carcinoma cells but not in mesothelioma cells. *J Cell Biochem* 2010; **109**: 519-531 [PMID: 19998412 DOI: 10.1002/jcb.22429]

72   **Marilley D**, Schwaller B. Association between the calcium-binding protein calretinin and cytoskeletal components in the human colon adenocarcinoma cell line WiDr. *Exp Cell Res* 2000; **259**: 12-22 [PMID: 10942575 DOI: 10.1006/excr.2000.4942]

73   **Akiyama SK**, Yamada KM. Fibronectin. *Adv Enzymol Relat Areas Mol Biol* 1987; **59**: 1-57 [PMID: 2949539]

74   **Levy P**, Loreal O, Munier A, Yamada Y, Picard J, Cherqui G, Clement B, Capeau J. Enterocytic differentiation of the human Caco-2 cell line is correlated with down-regulation of fibronectin and laminin. *FEBS Lett* 1994; **338**: 272-276 [PMID: 8307193 DOI: 10.1016/0014-5793(94)80282-3]

75   **Vachon PH**, Simoneau A, Herring-Gillam FE, Beaulieu JF. Cellular fibronectin expression is down-regulated at the mRNA level in differentiating human intestinal epithelial cells. *Exp Cell Res* 1995; **216**: 30-34 [PMID: 7813630 DOI: 10.1006/excr.1995.1004]

76   **Rothé B**, Leal-Esteban L, Bernet F, Urfer S, Doerr N, Weimbs T, Iwaszkiewicz J, Constam DB. Bicc1 Polymerization Regulates the Localization and Silencing of Bound mRNA. *Mol Cell Biol* 2015; **35**: 3339-3353 [PMID: 26217012 DOI: 10.1128/MCB.00341-15]

77   **Piazzon N**, Maisonneuve C, Guilleret I, Rotman S, Constam DB. Bicc1 links the regulation of cAMP signaling in polycystic kidneys to microRNA-induced gene silencing. *J Mol Cell Biol* 2012; **4**: 398-408 [PMID: 22641646 DOI: 10.1093/jmcb/mjs027]

78   **Fu Y**, Kim I, Lian P, Li A, Zhou L, Li C, Liang D, Coffey RJ, Ma J, Zhao P, Zhan Q, Wu G. Loss of Bicc1 impairs tubulomorphogenesis of cultured IMCD cells by disrupting E-cadherin-based cell-cell adhesion. *Eur J Cell Biol* 2010; **89**: 428-436 [PMID: 20219263 DOI: 10.1016/j.ejcb.2010.01.002]

79   **Marzolo MP**, Farfán P. New insights into the roles of megalin/LRP2 and the regulation of its functional expression. *Biol Res* 2011; **44**: 89-105 [PMID: 21720686 DOI: 10.4067/S0716-97602011000100012]

80   **Bose S**, Kalra S, Yammani RR, Ahuja R, Seetharam B. Plasma membrane delivery, endocytosis and turnover of transcobalamin receptor in polarized human intestinal epithelial cells. *J Physiol* 2007; **581**: 457-466 [PMID: 17347267 DOI: 10.1113/jphysiol.2007.129171]

81   **Vázquez-Carretero MD**, Palomo M, García-Miranda P, Sánchez-Aguayo I, Peral MJ, Calonge ML, Ilundain AA. Dab2, megalin, cubilin and amnionless receptor complex might mediate intestinal endocytosis in the suckling rat. *J Cell Biochem* 2014; **115**: 510-522 [PMID: 24122887 DOI: 10.1002/jcb.24685]

82   **Hansen TV**, Hammer NA, Nielsen J, Madsen M, Dalbaeck C, Wewer UM, Christiansen J, Nielsen FC. Dwarfism and impaired gut development in insulin-like growth factor II mRNA-binding protein 1-deficient mice. *Mol Cell Biol* 2004; **24**: 4448-4464 [PMID: 15121863 DOI: 10.1128/MCB.24.10.4448-4464.2004]

83   **Han J**, Isoda H, Maekawa T. Analysis of the mechanism of the tight-junctional permeability increase by capsaicin treatment on the intestinal Caco-2 cells. *Cytotechnology* 2002; **40**: 93-98 [PMID: 19003109 DOI: 10.1023/A:1023922306968]

84   **Chen C**, Chi H, Sun BG, Sun L. The galectin-3-binding protein of Cynoglossus semilaevis is a secreted protein of the innate immune system that binds a wide range of bacteria and is involved in host phagocytosis. *Dev Comp Immunol* 2013; **39**: 399-408 [PMID: 23178402 DOI: 10.1016/j.dci.2012.10.008]

85   **Yilmaz Z**, Eralp Inan O, Kocaturk M, Baykal AT, Hacariz O, Hatipoglu I, Tvarijonaviciute A, Cansev M, Ceron J, Ulus IH. Changes in serum proteins after endotoxin administration in healthy and choline-treated calves. *BMC Vet Res* 2016; **12**: 210 [PMID: 27646125 DOI: 10.1186/s12917-016-0837-y]

86   **Iglesias JM**, Morgan RO, Jenkins NA, Copeland NG, Gilbert DJ, Fernandez MP. Comparative genetics and evolution of annexin A13 as the founder gene of vertebrate annexins. *Mol Biol Evol* 2002; **19**: 608-618 [PMID: 11961095 DOI: 10.1016/j.matbio.2007.04.007]

87   **Babusiak M**, Man P, Petrak J, Vyoral D. Native proteomic analysis of protein complexes in murine intestinal brush border membranes. *Proteomics* 2007; **7**: 121-129 [PMID: 17205597 DOI: 10.1002/pmic.200600382]

88   **Ma YL**, Zhang P, Wang F, Moyer MP, Yang JJ, Liu ZH, Peng JY, Chen HQ, Zhou YK, Liu WJ, Qin HL. Human embryonic stem cells and metastatic colorectal cancer cells shared the common endogenous human microRNA-26b. *J Cell Mol*

Med 2011; **15**: 1941-1954 [PMID: 20831567 DOI: 10.1111/j.1582-4934.2010.01170.x]

89 **Lenaerts K**, Bouwman FG, Lamers WH, Renes J, Mariman EC. Comparative proteomic analysis of cell lines and scrapings of the human intestinal epithelium. *BMC Genomics* 2007; **8**: 91 [PMID: 17407598 DOI: 10.1186/1471-2164-8-91]

90 **Uzozie A**, Nanni P, Staiano T, Grossmann J, Barkow-Oesterreicher S, Shay JW, Tiwari A, Buffoli F, Laczko E, Marra G. Sorbitol dehydrogenase overexpression and other aspects of dysregulated protein expression in human precancerous colorectal neoplasms: a quantitative proteomics study. *Mol Cell Proteomics* 2014; **13**: 1198-1218 [PMID: 24567419 DOI: 10.1074/mcp.M113.035105]

91 **Bourgine J**, Billaut-Laden I, Happillon M, Lo-Guidice JM, Maunoury V, Imbenotte M, Broly F. Gene expression profiling of systems involved in the metabolism and the disposition of xenobiotics: comparison between human intestinal biopsy samples and colon cell lines. *Drug Metab Dispos* 2012; **40**: 694-705 [PMID: 22217464 DOI: 10.1124/dmd.111.042465]

92 **Barresi V**, Trovato-Salinaro A, Spampinato G, Musso N, Castorina S, Rizzarelli E, Condorelli DF. Transcriptome analysis of copper homeostasis genes reveals coordinated upregulation of SLC31A1,SCO1, and COX11 in colorectal cancer. *FEBS Open Bio* 2016; **6**: 794-806 [PMID: 27516958 DOI: 10.1002/2211-5463.12060]

93 **Zhao S**, Wen Z, Liu S, Liu Y, Li X, Ge Y, Li S. MicroRNA-148a inhibits the proliferation and promotes the paclitaxel-induced apoptosis of ovarian cancer cells by targeting PDIA3. *Mol Med Rep* 2015; **12**: 3923-3929 [PMID: 26004124 DOI: 10.3892/mmr.2015.3826]

94 **Tsai ZY**, Chou CH, Lu CY, Singh S, Yu SL, Li SS. Proteomic comparison of human embryonic stem cells with their differentiated fibroblasts: Identification of 206 genes targeted by hES cell-specific microRNAs. *Kaohsiung J Med Sci* 2011; **27**: 299-306 [PMID: 21802640 DOI: 10.1016/j.kjms.2011.03.010]

95 **Chiyomaru T,** Enokida H, Kawakami K, Tatarano S, Uchida Y, Kawahara K, Nishiyama K, Seki N, Nakagawa M. Functional role of LASP1 in cell viability and its regulation by microRNAs in bladder cancer. *Urol Oncol* 2012; **30**: 434-443 [PMID: 20843712 DOI: 10.1016/j.urolonc.2010.05.008]

96 **Delanote V**, Vandekerckhove J, Gettemans J. Plastins: versatile modulators of actin organization in (patho)physiological cellular processes. *Acta Pharmacol Sin* 2005; **26**: 769-779 [PMID: 15960882 DOI: 10.1111/j.1745-7254.2005.00145.x]

97 **Chafel MM**, Shen W, Matsudaira P. Sequential expression and differential localization of I-, L-, and T-fimbrin during differentiation of the mouse intestine and yolk sac. *Dev Dyn* 1995; **203**: 141-151 [PMID: 7655078 DOI: 10.1002/aja.1002030203]

Baishideng®