JOURNAL OF INTEGRATED OMICS
A METHODOLOGICAL JOURNAL
HTTP://WWW.JIOMICS.COM

ORIGINAL ARTICLE | DOI: 10.5584/jiomics.v2i2.102

Analysis of the rat primary hepatocyte nuclear proteome through sub-cellular fractionation.

Cliff Rowe, Roz E. Jenkins, Neil R. Kitteringham, B. Kevin Park, Christopher E.P. Goldring*

MRC Centre for Drug Safety Science, Department of Pharmacology & Therapeutics, Sherrington Building, Ashton Street, The University of Liverpool, L69 3GE.

Received: 09 August 2012 **Accepted:** 17 October 2012 **Available Online**: 02 November 2012

Abstract

Characterising primary hepatocytes and their de-differentiation in culture is vital for the refinement of current culture techniques and for the development of new and improved in vitro hepatocyte models. We have performed multiplexed iTRAQ proteomics on whole cell preparations and further employed nuclear fractionation to expand the coverage of this important organelle. We identify many proteins that change in abundance during culture of rat hepatocytes for 48h and map their molecular functions. 431 proteins were identified and quantified in whole cell homogenates, mapping to 69 molecular functions using the PANTHER (Protein ANalysis THrough Evolutionary Relationships) classification system. In whole cell homogenates liver-associated functions, such as oxidoreductase activity, were enriched compared with the reference rat proteome dataset but some functions, such as transcriptional activity, were under-represented. Nuclear fractionation resulted in the identification of an additional 156 proteins which mapped to 31 molecular functions. These proteins included some associated with hepatic differentiation, such as HNF4alpha and CCAAT/enhancer-binding protein beta and others with less well-defined roles.

Hierarchical clustering of samples within each experiment showed segregation of fresh and cultured sample types and stringent statistical analysis demonstrated significant changes in 36% of proteins from the whole cell homogenates and 21% of proteins from the nuclear dataset (adjusted $p < 0.05$). The molecular functions of the changed proteins in each dataset are mapped. These datasets broaden our understanding of hepatocyte de-differentiation and will aid the identification of target pathways to attenuate de-differentiation in culture and maintain hepatocytes with a more relevant physiological phenotype.

Keywords: Differentiation; Hepatocyte; Nucleus; Proteomics.

Abbreviations:

ACN, acetonitrile; **HBSS**, Hanks balanced salt solution; **iTRAQ**, isobaric tags for relative and absolute quantification; **LC-MS/MS**, Liquid chromatography mass spectrometry/mass spectrometry; **MMTS**, Methyl methanethiosulfonate; **PBS**, Phosphate buffered saline; **SDS**, Sodium Dodecyl sulphate; **TFA**, trifluoroacetic acid; **TEAB**, triethylammonium bicarbonate; **WEM**, Williams' E Medium.

1. Introduction

Mammalian hepatocyte cultures represent an established and essential tool in drug development [1-5]. However, the relatively poor availability of human cells and their limited lifespan in culture means that animal cells, particularly rat hepatocytes, are often employed. Although these cells have utility, there is an urgent need to develop new, robust, metabolically-competent human hepatocyte models with predictable availability for routine hepatotoxicity and

***Corresponding author:** Christopher EP Goldring, E -mail Adress: C.E.P.Goldring@liverpool.ac.uk; Postal Adress: MRC Centre for Drug Safety Science, Department of Pharmacology & Therapeutics, Sherrington Building, Ashton Street, The University of Liverpool, United Kingdom, L69 3GE.

metabolism studies. Despite the continuous refinement of hepatocyte culture techniques, these cells undergo profound changes during culture with loss of liver function [6]. The underlying mechanisms are partially understood [7] but our knowledge concerning which proteins, and cellular functions change during culture remains incomplete. In a recent study we profiled the proteome of freshly-isolated hepatocytes and mapped temporal changes in primary culture that were concomitant with large declines in the abundance of liver specific mRNAs [6]. We postulated that significant changes in the nuclear proteome may be driving de-differentiation of hepatocytes in culture although there was limited coverage of this organelle in our dataset.

The nucleus of a cell orchestrates its phenotype through transcriptional control of specific gene programs. 'Master' transcription factors that determine cell specificity during embryonic development have been identified for many cell types [8] including hepatocytes [9], but alone these represent an incomplete picture of the complex networks of transcription factors, co-factors, chromatin remodelling proteins, RNA processing factors, and myriad accessory proteins that combine to direct cellular phenotype. Studies have been conducted previously to identify the proteins found in the nuclei of various cell types [10-14]. In the context of hepatocytes, a wider knowledge of the nucleus, and how its proteome changes in culture, should help develop better models of primary culture and assist in defining pathways amenable to manipulation. In this study, therefore, our aim was to further our understanding of hepatocyte de-differentiation through analysis of freshly-isolated and cultured hepatocyte whole cells and purified nuclei using iTRAQ proteomics and analysis of the molecular functions associated with the proteins that significantly change in abundance.

2. Material and Methods

2.1 *Hepatocyte isolation*

All experiments were undertaken in accordance with criteria outlined in a license granted under the Animals (Scientific Procedures) Act 1986, and approved by the Animal Ethics Committees of the University of Liverpool. Hepatocytes were isolated from male Wistar rats (Charles River, Manston UK) weighing approximately 250g . Rats were housed in controlled environmental conditions of 12 h light/dark cycle with free access to food and water. Rats were anesthetized with 200 μL of pentoject by intraperitoneal injection and, a laparotomy performed. The hepatic portal vein was cannulated with a 20 g × 48 mm vialon catheter (Becton Dickinson, U.K.), which was tied in place. Hanks balanced salt solution (HBSS) (Sigma, Poole, U.K.) without calcium and magnesium was perfused through the liver at 37 °C for 10 minutes during liver excision. Once excised, 100 mL of complete HBSS containing 50 mg of collagenase A (Roche, Welwyn Garden City, U.K.) was perfused through the liver with recirculation until it was digested. The cells were dispersed into 100 mL Williams' E medium (WEM) (Sigma, Poole, U.K.) and filtered through a 125 μm gauze into a sterile container. The cells were washed three times by centrifugation (100*g* for 2 minutes at 4 °C) and an aliquot was taken for counting. Cell suspensions (0.5×10^6/mL) were prepared in WEM supplemented with 100 U/mL penicillin and 100 *μ*g/mL streptomycin, 2 mM/ L glutamine (Sigma, Poole, U.K.), 10% fetal calf serum (Lonza biologics, Slough, U.K.), and 1 μg/mL bovine insulin (Sigma, Poole, U.K.) and seeded at 70 000 cells/cm^2. The seeded plates were incubated at 37 °C in a humidified atmosphere containing 5% CO_2 for 3 h before the medium was exchanged for serum - and insulin-free WEM. Hepatocytes were cultured for 48 h as a monolayer on 100mm collagen I-coated culture dishes (Becton Dickinson, Oxford, UK). Medium was changed 3 h after seeding and each 24 h thereafter.

2.2 *Sample protein preparation*

For whole-cell preparations, 5 x 10^6 freshly-isolated hepatocytes or hepatocytes scraped into ice-cold phosphate-buffered saline (PBS) from two 100mm culture dishes, after 48h in culture, were washed twice with 5 mL PBS, by low speed centrifugation, 100*g* for 5 minutes at 4 °C. Cell pellets were sonicated directly in 200 μL of 0.5M triethylammonium bicarbonate, 0.1% SDS (TEAB/SDS), with ten 1s pulses of a probe and the preparation clarified by centrifugation (10,000g, 2 minutes).

Nuclei were purified from freshly-isolated hepatocytes and cells after 48h culture. Nuclei were prepared from 5 x 10^7 freshly-isolated cells, or cells from at least ten 100mm culture plates. Cells were washed with ice-cold PBS, pellets were re-suspended in 5 mL ice-cold hypotonic solution (0.2x PBS) and placed on ice for 10 minutes. Hepatocytes were disrupted using a dounce homogeniser and were centrifuged at 1000*g* for 10 minutes at 4°C. The pellet was re-homogenised in 5 mL 2M sucrose, 3.3 mM $CaCl_2$. 100 μL of 2 M sucrose, 3.3mM $CaCl_2$ was pipetted into multiple micro-tubes and carefully overlaid with 400 μL of sample homogenate and centrifuged at 18,000*g*, 4°C for 1 h. The resulting pellets of nuclei were each washed with 0.5 mL PBS per tube, and centrifuged at 1000*g* for 10 minutes to form a pellet. The fraction was checked for the absence of whole cells and large debris by microscopic examination of a 5 μL aliquot diluted 20-fold with PBS and a final concentration of 0.1% trypan blue. The pellets were combined and the proteins extracted by sonication in 30 μL TEAB/SDS). The preparation was clarified by centrifugation (10,000*g*, 2 minutes). Protein concentration was determined in the supernatant using Bradford assay reagents (Sigma, Poole, UK).

2.3 *iTRAQ labelling Cation exchange and Mass spectrometry*

Labelling was carried out according to the manufacturer's instructions for 8-plex labelling (ABSciex, Warrington, UK)

using 75 µg cellular or nuclear protein per tag. For LC-MS/MS analysis of iTRAQ labelled samples, each cation exchange fraction was resuspended in 120 ìL 5% ACN/0.05% trifluoroacetic acid (TFA) and 60 ìL loaded on to the column. Samples were analysed on a QSTAR® Pulsar i hybrid mass spectrometer (ABSciex) and were delivered into the instrument by automated in-line liquid chromatography (integrated LC Packings System, 5 mm C18 nano-precolumn and 75 ìm × 15 cm C18 PepMap column; Dionex, California, USA) via a nano-electrospray source head and 10 ìm inner diameter PicoTip (New Objective, Massachusetts, USA). The pre-column was washed for 30 minutes at 30 ìL/min with 5% ACN/0.05% TFA, prior to initiation of the solvent gradient in order to reduce the level of salt in the sample. A gradient from 5% ACN/0.05% TFA (v/v) to 60% ACN/0.05% TFA (v/v) in 70 min was applied at a flow rate of 300 nL/minute. The MS was operated in positive ion mode with survey scans of 1 second, and with an MS/MS accumulation time of 1 second for the three most intense ions. Collision energies were calculated on the fly based on the m/z of the target ion and the formula, collision energy = (slope × m/z) + intercept. The intercepts were increased by 3–5 V compared to standard data acquisition in order to improve the reporter ion intensities/quantitative reproducibility.

Data analysis was performed using ProteinPilot software (Version 3, ABSciex, Warrington, UK). The data were analysed with a fixed modification of MMTS-labelled cysteine, biological modifications allowed and with the confidence set to 10% to enable the false discovery rate to be calculated from screening the reversed SwissProt database 30-11-2009. Ratios for each iTRAQ label were obtained, using one sample of freshly-isolated cell or nuclear protein as the denominator in respective experiments. Raw iTRAQ files are available at the TRANCHE hash: yBAoOh3Q3eZCL8Dc10EJjzm5m0vcxS27Lb4EUjJYQB88/DGbPpYkVVQxJvlLmB74NomeFJR2CMvJX/jk6+gHne6lhaQAAAAAAAAi7A==

2.4 *Data processing and analysis*

Proteins above the 1% false discovery rate, identified with ≥95% confidence (for those identified with more two or more peptides) or 99% confidence (for those identified by a single peptide) were included in the final data analysis. Only proteins meeting these criteria and identified in at least three samples were used in the statistical analysis in the R computational environment, version 2.14.1 [15]. The packages *marray* (16) and *multtest* [17] were used for comparison between freshly-isolated cells and cultured cells. R scripts are available at the TRANCHE hash yBAoOh3Q3eZCL8Dc10EJjzm5m0vcxS27Lb4EUjJYQB88/DGbPpYkVVQxJvlLmB74NomeFJR2CMvJX/jk6+gHne6lhaQAAAAAAAAi7A==. Molecular functions were assigned using the PANTHER (**P**rotein **AN**alysis **TH**rough **E**volutionary **R**elationships) classification system, utilising the gene expression analysis tool 'compare gene lists' (www.pantherdb.org/) [18, 19].

2.5 *Western Blotting*

Proteins were denatured and reduced by heating to 80 °C for 10 minutes in Laemmli sample buffer (Sigma, Poole, U.K.) and 10 μg of each sample was resolved through a 10% polyacrylamide gel (Biorad, Hemel Hempstead, U.K.) and transferred to a nitrocellulose membrane (GE Healthcare, Slough, U.K.). After transfer, the nitrocellulose membrane was blocked overnight in Tris buffered saline, containing 0.1% Tween 20, (Sigma, Poole, U.K.) and 10% milk protein at 4 °C. Antibodies used were rabbit anti-Lamin A (Sigma, Poole, U.K.) and goat anti-HNF4 alpha (Insight biotechnology, Wembley, U.K.). Peroxidise-conjugated anti-rabbit, or anti-goat IgG antibodies (Sigma, Poole, U.K.) were used for chemiluminescence detection using ECL detection reagents and ECL hyperfilm (GE Healthcare, Slough,U.K.).

3. Results

3.1 *The proteome of whole cell homogenates is enriched for molecular functions associated with liver.*

iTRAQ analysis of freshly-isolated and cultured whole cell homogenates identified and quantified 431 proteins. The supporting data file, sheet 'WC iTRAQ results" details all proteins identified in whole cell extracts, their relative quantities, log2 fold-change after culture and their associated p values (T test, comparing freshly-isolated and cultured samples). These proteins were compared against the *rattus norvegicus* reference list in PANTHER using the gene expression analysis tool 'compare gene lists'. PANTHER's background reference list of 27758 proteins mapped to 120 molecular functions and the experimental dataset mapped to 69 molecular functions (supporting data file, sheet "WC MF"). 41 molecular functions were represented by significantly more or fewer proteins (p <0.05) than expected, based on the proportion of proteins from the reference list that mapped to each molecular function. Of these functions, 24 were significantly over-represented in the experimental dataset (Table 1) and 17 were under-represented. Many of the functions that are over-represented are functions associated with liver parenchyma, such as oxidoreductase-, peroxidase- and antioxidant activities. Under-represented molecular functions included those that are carried out by low abundance proteins, such as transcription regulation- or kinase activities.

3.2 *Nuclear fractionation affects the number and percentage of identified proteins mapping to specific molecular functions.*

The purity of hepatocyte nuclei isolation was checked by microscopic examination after addition of trypan blue to an aliquot of nuclei. Nuclei were seen as distinct organelles and the preparations were largely free from whole cell and cell

Table 1. Molecular functions significantly over-represented in the whole cell extract experimental dataset. The number of proteins assigned to each molecular function in the PANTHER classification system for the whole cell experimental dataset was compared to the number assigned in the reference rat proteome. Functions that were represented by significantly more proteins (p <0.05) than expected based on dataset size are reported.

Molecular Function	REFLIST (27758)	WC dataset (419)	Expected	Observed/ Expected	P-value
oxidoreductase activity	971	124	14.66	8.46	1.17E-76
catalytic activity	6555	248	98.95	2.51	1.94E-54
structural constituent of ribosome	814	54	12.29	4.39	1.97E-19
lyase activity	230	27	3.47	7.78	6.66E-16
transferase activity	2027	78	30.6	2.55	2.84E-14
isomerase activity	258	23	3.89	5.91	2.33E-11
racemase and epimerase activity	66	12	1	12.00	6.98E-10
hydro-lyase activity	71	12	1.07	11.21	1.57E-09
structural molecule activity	2389	72	36.06	2.00	1.57E-08
antioxidant activity	29	8	0.44	18.18	2.14E-08
peroxidase activity	27	7	0.41	17.07	2.48E-07
transaminase activity	27	7	0.41	17.07	2.48E-07
acyltransferase activity	209	15	3.15	4.76	1.05E-06
hydrogen ion transmembrane transporter activity	63	7	0.95	7.37	5.88E-05
transketolase activity	6	3	0.09	33.33	1.15E-04
protein disulfide isomerase activity	18	4	0.27	14.81	1.81E-04
transferase activity, transferring glycosyl groups	278	12	4.2	2.86	1.27E-03
ligase activity	652	20	9.84	2.03	2.57E-03
anion channel activity	62	5	0.94	5.32	2.73E-03
translation regulator activity	130	7	1.96	3.57	3.99E-03
translation factor activity, nucleic acid binding	133	7	2.01	3.48	4.51E-03
translation elongation factor activity	47	4	0.71	5.63	5.97E-03
transaldolase activity	1	1	0.02	50.00	1.50E-02
translation initiation factor activity	99	5	1.49	3.36	1.81E-02

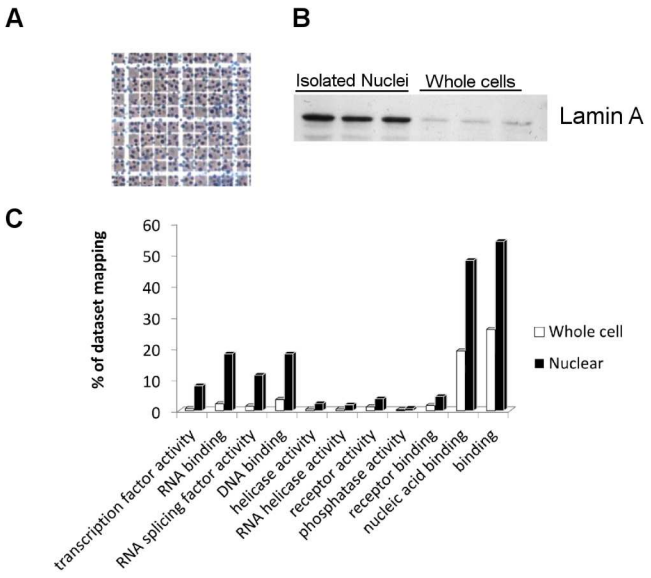

Figure 1. Enrichment of nuclei: A) Isolated nuclei were examined microscopically to verify purity and found to be largely free from whole cells and cell debris. B) Western blot for the nuclear envelope protein, Lamin A, demonstrates enrichment of nuclei. C) Bar chart of the molecular functions, identified in whole cell or nuclear preparations that were enriched 2-fold or more by nuclear fractionation. The number of proteins mapping to each function is expressed as a percentage of the total number of identified proteins in that experiment.

debris (figure 1A). As further confirmation of high levels of nuclear enrichment western blots were performed for the nuclear envelope protein, Lamin A. Immunodetection of Lamin A was much stronger in nuclear preparations compared with whole cell samples confirming the high degree of enrichment for nuclei (figure 1B). These highly pure preparations were used to analyse the hepatocyte nuclear proteome by iTRAQ proteomics. 268 distinct proteins were identified and quantified in our purified nuclear protein preparations. The supporting data file, sheet "Nuc iTRAQ results" details all proteins identified in nuclear preparations, their relative quantities, mean log2 fold change in culture and associated *p* values. 156 proteins were unique to this experiment and not represented in the whole cell experiment. This dataset was also compared against the PANTHER rat reference list of proteins for mapping to molecular function (Supporting data file, sheet "Nuc MF"). There were 31 molecular functions represented by significantly more or fewer proteins (p <0.05) than expected (23 molecular functions over-represented and 8 under represented). Table 2 details the molecular functions significantly over-represented by proteins in the nuclear dataset compared with expected values based on the reference list. Some of these molecular functions were over-or under-represented in common with the whole cell dataset but others are uniquely enriched or reduced in nuclear preparations. RNA binding, for example, was detected with the expected number of proteins in the whole cell dataset but more than 8-fold enriched in the nuclear preparations (table 2). Similarly transcription co-factor activity, which was under-represented in whole cell samples, is more than 2-fold over-represented in the nuclear preparations relative to the reference dataset.

By simultaneously comparing the whole cell and nuclear datasets against the reference list of proteins in PANTHER we showed that nuclear fractionation changed the relative abundance of proteins associated with some molecular functions. This also gave an absolute count of the numbers of proteins that map to each function for each experiment. Many additional proteins were mapped to particular molecular functions depending on whether the protein preparation was whole cell- or nuclear-derived. In the whole cell dataset there were 37 molecular functions that were proportionally enriched compared to nuclear preparations. These functions were mainly associated with hepatocyte functions such as catalytic-, oxidoreductase- and transferase activities. Nuclear-derived protein preparations were proportionally enriched for 46 molecular functions (supporting data file, sheet "MF enrichment"). The functions identified in both datasets but proportionally enriched by more than 2 fold in nuclear samples are shown in figure 1 C. Many of the functions have a strong association with the nucleus, such as nucleic acid binding and transcriptional activity. Supporting data file, sheet "MF enrichment", details, for all represented molecular functions in both datasets, the count for proteins mapping to that function for each dataset and the percentage of the each dataset mapping to that function.

3.3 *Molecular function analysis of proteins changed in during culture.*

Culture of hepatocytes caused a significant change in abundance for a large proportion of proteins in whole cell and nuclear samples. Using the quantitative data for whole cell extracts to construct a heatmap with unsupervised hierarchical clustering (figure 2A) demonstrated the clear differences between freshly-isolated and cultured samples with clustering of similar sample types and clear segregation by sample type. The mean log2 fold-change for each protein was calculated and plotted as a frequency histogram (figure 2B). This showed that proteins were both up- and down-regulated with the larger changes occurring with lower frequency. Statistical analysis of these data, by T-test, showed that 206 proteins (>47% of the dataset) were significantly changed in abundance after 48h in culture (raw p <0.05). After adjustment for multiple comparisons there were still 159 proteins (>36% of the dataset) significantly changed in abundance (BH p <0.05). Log2 fold-changes and p values are listed for each protein in the supporting data file, sheet "WC iTRAQ results" and summarised graphically as a volcano plot in figure 3A, and as a frequency histogram of adjusted p values in figure 3B.

The molecular functions of the significantly changed proteins in whole cell preparations were determined through PANTHER by selecting the complete list of identified proteins as a reference list. Of the 69 molecular functions repre-

Table 2. Molecular functions significantly over-represented in the nuclear extract experimental dataset. The number of proteins assigned to each molecular function in the PANTHER classification system for the nuclear extract experimental dataset was compared to the number assigned in the reference rat proteome. Functions that were represented by significantly more proteins (p <0.05) than expected based on dataset size are reported.

Molecular Function	REFLIST (27758)	Nuclear dataset (264)	Expected	Observed / Expected	P-value
RNA binding	617	48	5.87	8.18	5.66E-29
nucleic acid binding	5086	127	48.37	2.63	3.60E-28
RNA splicing factor activity, transesterification mechanism	334	30	3.18	9.43	5.04E-20
catalytic activity	6555	120	62.34	1.92	6.65E-15
Binding	8663	143	82.39	1.74	9.89E-15
structural constituent of ribosome	814	30	7.74	3.88	3.84E-10
oxidoreductase activity	971	30	9.23	3.25	2.06E-08
antioxidant activity	29	5	0.28	17.86	1.02E-05
DNA binding	2682	48	25.51	1.88	1.53E-05
structural molecule activity	2389	43	22.72	1.89	4.03E-05
peroxidase activity	27	4	0.26	15.38	1.45E-04
translation factor activity, nucleic acid binding	133	7	1.26	5.56	3.26E-04
hydrogen ion transmembrane transporter activity	63	5	0.6	8.33	3.81E-04
translation initiation factor activity	99	6	0.94	6.38	4.18E-04
deaminase activity	38	4	0.36	11.11	5.23E-04
protein disulfide isomerase activity	18	3	0.17	17.65	7.29E-04
isomerase activity	258	9	2.45	3.67	9.33E-04
translation regulator activity	130	6	1.24	4.84	1.68E-03
RNA helicase activity	99	5	0.94	5.32	2.77E-03
transferase activity, transferring glycosyl groups	278	8	2.64	3.03	5.60E-03
helicase activity	170	6	1.62	3.70	6.15E-03
hydrolase activity, acting on ester bonds	780	13	7.42	1.75	3.75E-02
transcription cofactor activity	339	7	3.22	2.17	4.50E-02

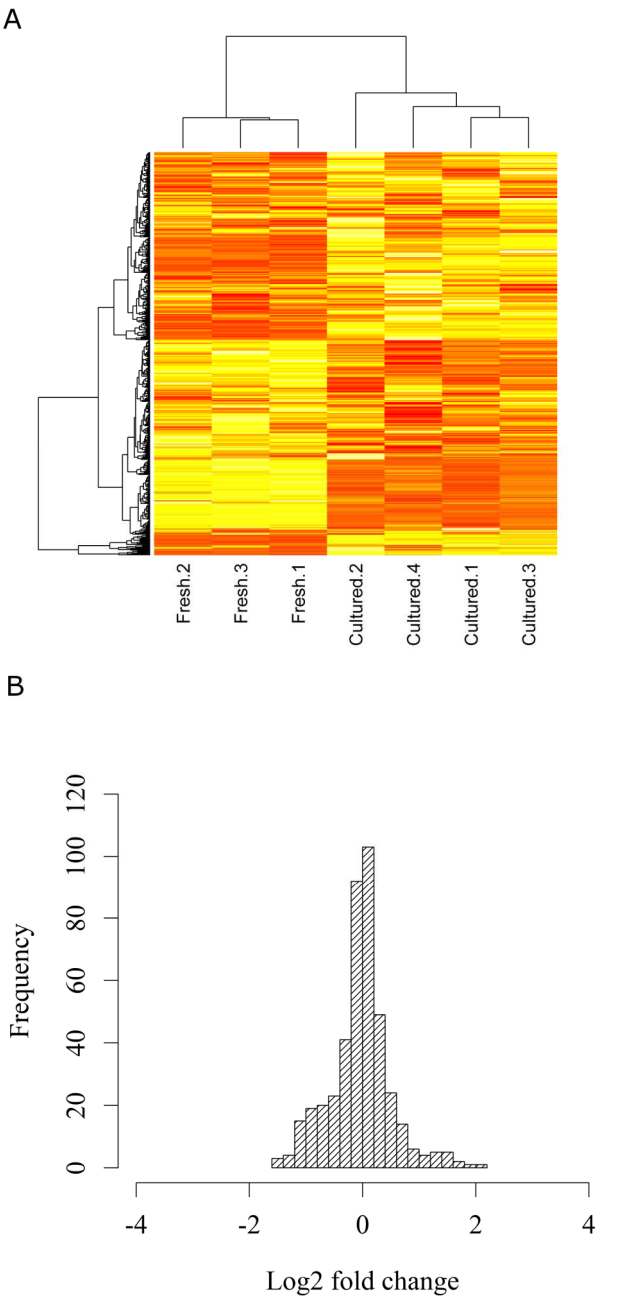

Figure 2. The changing proteome of cultured hepatocytes. Whole cell preparations were made from freshly-isolated hepatocytes or hepatocytes that were cultured for 48h. A) A heatmap with hierarchical clustering shows that samples segregate according to quantitative differences between freshly-isolated and cultured hepatocyte proteome. B) Frequency histogram for Log2-fold change during hepatocyte culture. The larger changes occur with reduced frequency.

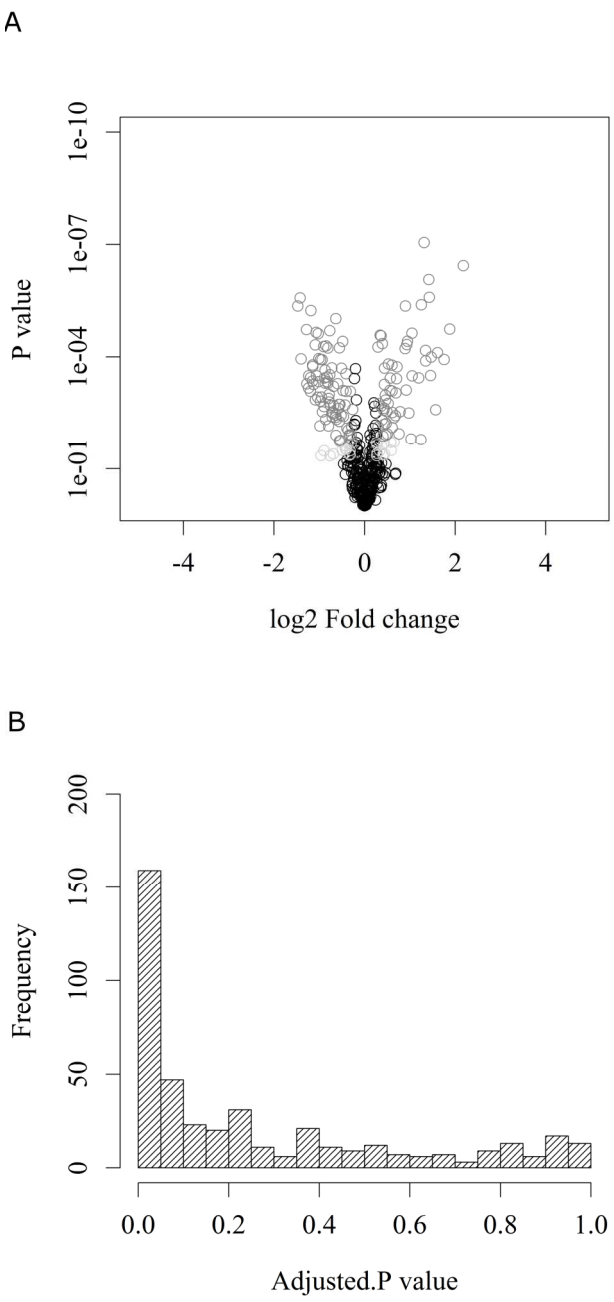

Figure 3. Statistical analysis of the changes in the cultured hepatocyte proteome. A) Volcano plot of Log2 fold-change for each protein identified in whole cell preparations versus the significance (p value). Proteins not significantly changed are indicated with a black circle, those significantly changed +/- 20% with a raw p value <0.05 are coloured light grey and those significant after adjustment for multiple comparisons (Benjamini Hochberg p value <0.05) are dark grey. B) frequency histogram depicting the number of proteins that fall into each p value range.

sented in the whole cell dataset, 65 functions were represented in the significantly changed subset when considering those with a raw p value <0.05, or 59 functions when using adjusted p values <0.05. 35 molecular functions were over-represented in this subset of proteins when compared against the complete whole cell dataset as a reference (supporting data file, sheet "WCPvalueMFs").

3.4 *Nuclear preparation extends the list of molecular functions that are identified as significantly changed in culture*

Nuclear preparations also showed significant changes in a large proportion of proteins after 48h culture. Constructing a heatmap with unsupervised hierarchical clustering again demonstrated the clear differences between fresh and cul-

tured samples (figure 4A). The mean log2 fold-change, plotted as a frequency histogram (figure 4B), shows that approximately equal numbers of proteins were up- and down-regulated. Hepatic nuclear factor 4 alpha was among the proteins identified was significantly reduced in expression in cultured cells. This was verified by western blot in 2 samples each of freshly-isolated and cultured nuclei (figure 4C). 99 proteins (>36% of the nuclear dataset) were significantly changed in abundance after 48h in culture (raw p <0.05). After adjustment for multiple comparisons 58 proteins (>21%) remained significantly changed in abundance (BH p <0.05). A total of 28 molecular functions were over-represented in the adjusted p value subset of data when compared against the complete nuclear dataset as a reference, including many functions performed in the nucleus (supporting data file, sheet "NucPvalueMFs"). Log2 fold changes and p values are listed for each protein in supporting data file, sheet "Nuc iTRAQ results" and summarised graphically as a volcano plot in figure 5A, and as a frequency histogram of adjusted p values in figure 5B.

The functions of the changed proteins were assessed using PANTHER. The complete dataset from each experiment were used as a background list and the significantly changed proteins (raw p <0.05 or BH p <0.05) were compared for the numbers of proteins mapping to each molecular function. The results are summarised in figure 6 and the full results are given in supporting data file, sheet "nucPvlaueMFs".

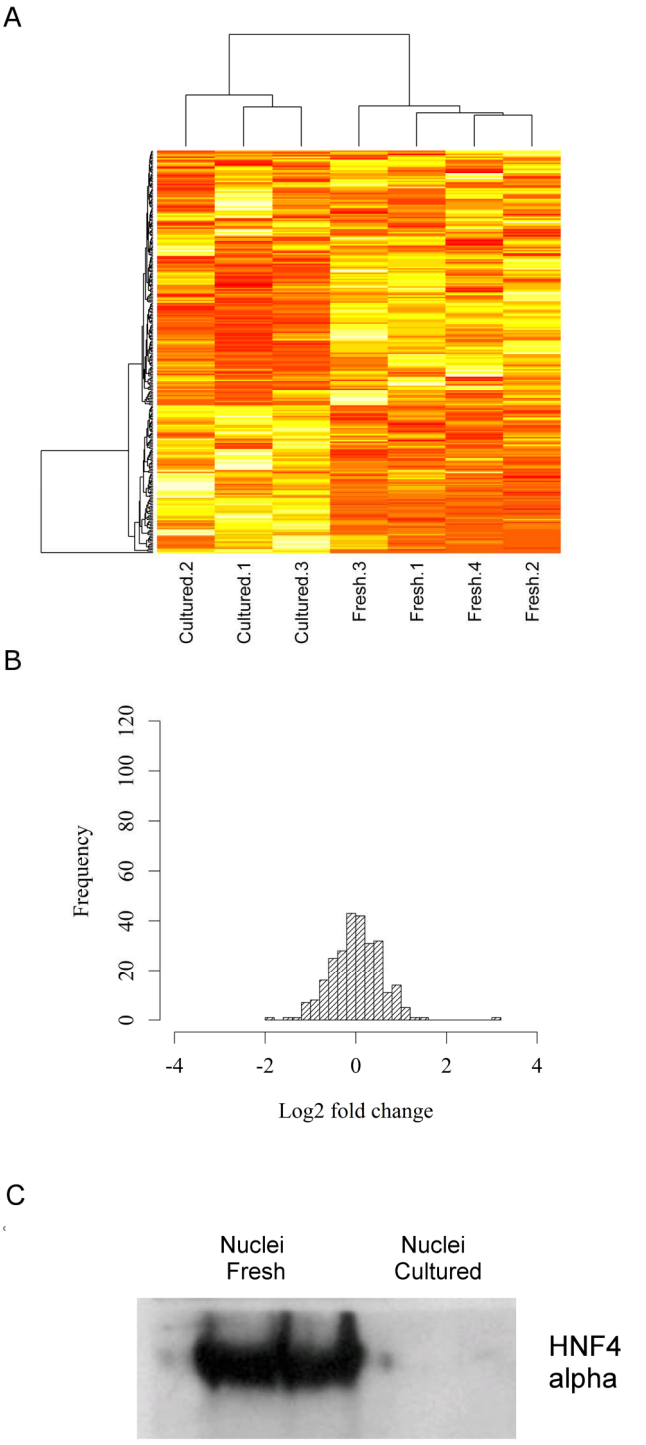

Figure 4. The changing nuclear proteome of cultured hepatocytes. A) Heatmap with hierarchical clustering shows the quantitative differences between freshly-isolated and cultured hepatocyte nuclear proteome. B) Frequency histogram for Log2 fold change demonstrating the number of proteins up-regulated or down-regulated after culture. C) Western blot verification of reduced HNF4 alpha expression in cultured hepatocytes.

4. Discussion

The short supply of primary human hepatocytes or alternative functional hepatocyte-like cells is a problem in drug development programs and is, in part, driving research into the development of stem cell-derived hepatocyte-like cells [20-24]. However, it is probable that any stem cell-derived model reaching hepatocyte-like status will suffer the same de-differentiation phenomenon as mature cells. It is therefore imperative to fully characterise the changes in primary cells in culture in order to identify possible interventions to prevent this loss of function.

Quantitative proteomic technologies such as iTRAQ [6, 25], isotope coded affinity tags [iCAT] [26] and stable isotope labelling by amino acids in cell culture (SILAC) [27] are becoming more and more accessible. The datasets typically encompass at least several hundred proteins, which represents the higher abundance or more easily detectable proteins. These multiplexed labelled proteomics technologies have the advantage of identifying and relatively quantifying proteins in multiple samples in a single experiment without gel casting and spot cutting associated with 2D gel-based proteomics. For our studies we use iTRAQ labelling as it allows relative quantification of hundreds of proteins in up to 8 samples simultaneously allowing statistical analysis of proteins from two experimental groups.

Many of the important proteins that control cellular phenotype and dynamic response are compartmentalised and/or are low abundance nuclear proteins. The nucleus consists of an estimated 14% of the entire cellular proteome [11] but many of its components are poorly represented in

datasets from un-fractionated cells. The method we employ results in clean nuclear preparations (Figure 1A). This reduced the amount of high abundance cytoplasmic/cytoskeletal proteins to allow the detection of more low abundance nuclear proteins. Others have performed experiments on purified nuclei [10-14] but to our knowledge this is the first time primary hepatocyte nuclear preparations have been prepared for shotgun proteomics analysis to decipher the changes that occur to this organelle in hepatocytes during culture. This work therefore forms a basis for the understanding of the hepatocyte nuclear proteome and expands the knowledge of the hepatocyte proteins that change during culture, providing datasets describing the proteomic signature of hepatocyte de-differentiation.

We postulated that an altered nuclear proteome in cultured cells could underpin the change in the phenotype of primary hepatocytes during culture, as significant changes in the expression of hepatocyte-specific genes are evident after hepatocyte isolation and culture [6, 28-29]. Previous quantitative analysis of nuclear proteomes in other cell types has been undertaken but the scale of identification of changed proteins was smaller than presented here. Pewsey *et al* [12] used iTRAQ technology to investigate the changing nuclear proteome of differentiating embryonic carcinoma

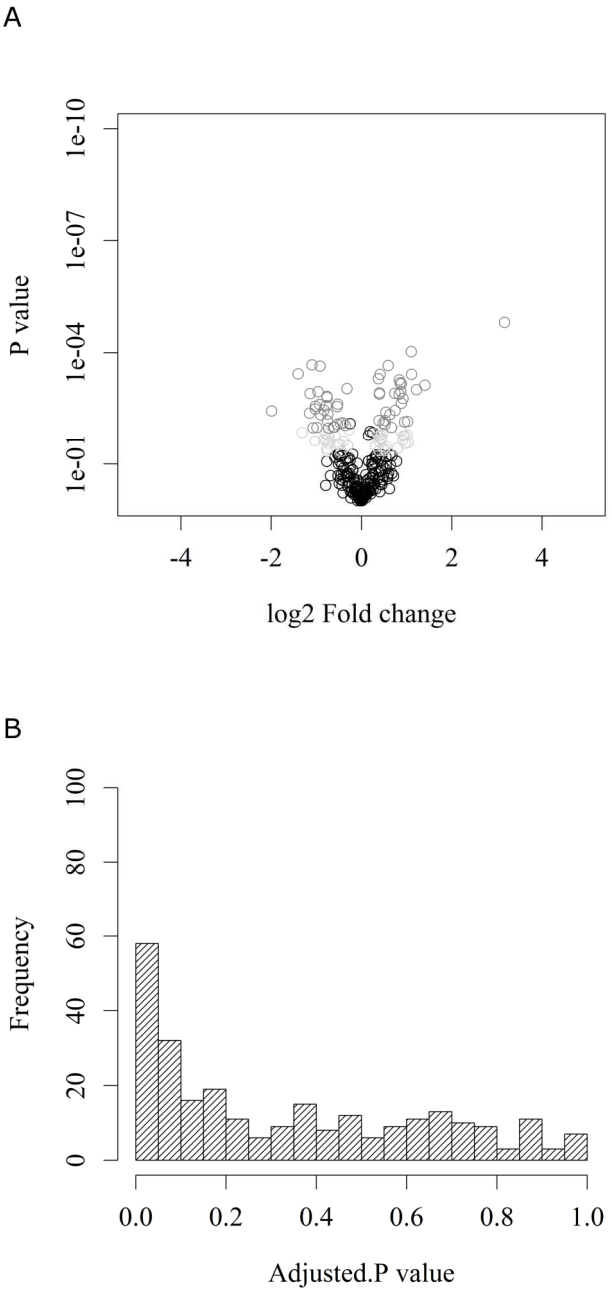

Figure 5. Statistical analysis of the changes in the cultured hepatocyte proteome: A) Volcano plot of Log2 fold change for each protein identified in nuclear preparations versus the significance (*p* value). Proteins not significantly changed are indicated with a black circle, those significantly changed +/- 20% with a raw *p* value <0.05 are coloured light grey and those significant after adjustment for multiple comparisons (Benjamini Hochberg *p* value <0.05) are dark grey. B) frequency histogram depicting the number of proteins that fall into each *p* value range.

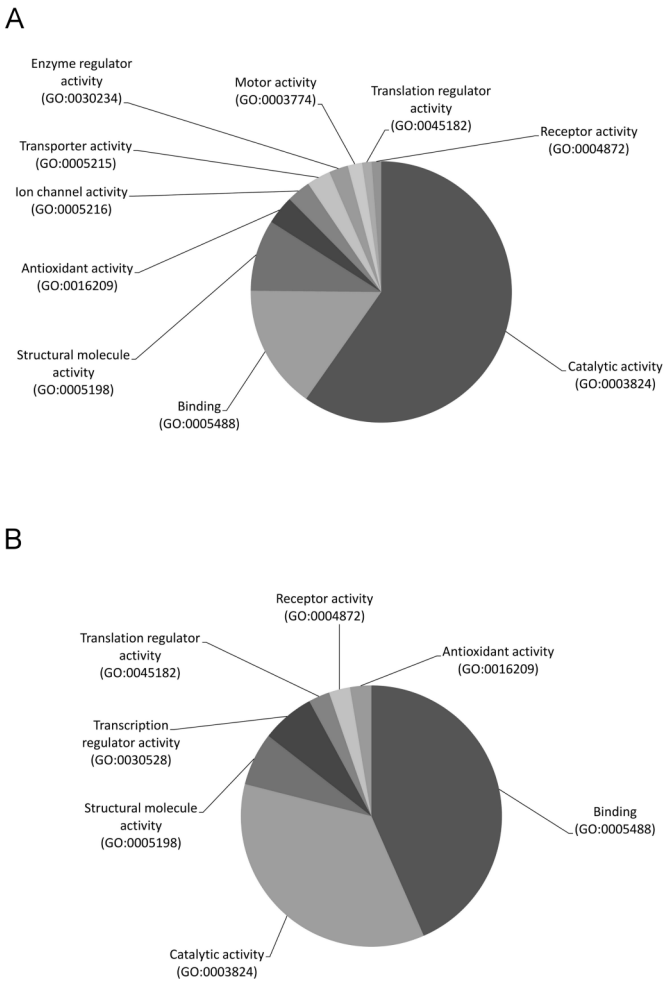

Figure 6. Molecular functions of proteins with alerted abundance in culture. Pie charts summarising the molecular functions that are performed by the proteins that were identified as significantly changed after culture (*p* value <0.05) in whole cell (A), or nuclear protein preparation (B). Significantly changed proteins were assigned to a molecular function using the Batch ID tool in PANTHER.

(NTERA-2) cells, identifying 37 proteins that were significantly altered in expression. Zhang *et al* [13] identified 22 differentially expressed nuclear proteins in cisplatin-treated HeLa cells by two-dimensional electrophoresis, using a similar approach Cao *et al* [14] identified 19 nuclear proteins changed in abundance by Epstein-Barr virus (EBV) infection. Besides our own work [6] shotgun proteomic study of changes in cultured primary hepatocytes is not extensively reported. Pan *et al* [30] performed a wide scale proteomics analysis of mouse primary hepatocytes with the purpose of comparing their phenotype with that of mouse hepatic cell lines but did not address the question of de-differentiation of the primary cells during culture.

Here we present proteomic datasets for freshly-isolated and cultured rat hepatocytes, generated using whole cells or purified nuclei. The value of this approach is clear with an expanded coverage of nuclear proteins and molecular functions represented by the 156 proteins that were uniquely detected in the nuclear fractions. The quantitative nature of these data allows a robust statistical analysis to identify quantitative changes that occur in the nucleus during cell culture. 64 proteins that were significantly changed in abundance after culture were identified in nuclei, but not detected in the whole cell experiment. These unique proteins were filtered for those annotated to the nucleus and the resulting 53 proteins are listed and their functions indicated in the supporting data file, sheet "UniqueSigNuc". These proteins include the transcription factors, established as vital to hepatic differentiation, HNF4alpha [9] and CCAAT/enhancer-binding protein beta [31], both of which were significantly reduced in cultured cells, as well as others with less well-defined roles in hepatocyte phenotype, but which are worthy of further investigation such as Nuclear factor 1 A-type and Methyl-CpG-binding protein 2 (supporting data file, sheet "Nuc iTRAQ results"). We mapped a greater number of proteins to molecular functions associated with the nucleus using the nuclear fractionation approach and by combining the protein lists for both experiments we increased the mapped molecular functions to a total of 83. However, proteins are redundant with respect to GO terms and an individual protein may appear in multiple categories. For example 'GO:0016563 : transcription factor activity' and 'GO:0030528: transcription regulator activity' were each represented by 3 proteins in the whole cell dataset and 21 proteins in the nuclear dataset, as the same proteins map to both terms. Nevertheless, this 7-fold increase in proteins detected in nuclear samples demonstrates the power of nuclear fractionation in improving the coverage of functions for a given organelle.

Sub-cellular expression proteomic profiling demands relatively large sample sizes in order to purify the required quantity of protein from organelles. This amount of protein is achievable with highly proliferative cells types, such as mammary epithelial cells [31], or proliferating undifferentiated stem cells [32, 33], or proliferative immune cells [34]. Currently, the number of differentiated stem cells typically produced in a research laboratory-scale experiment prohibits such shotgun proteomic investigations. However, the 'top-down' approach we have applied here, to more fully characterise mature, differentiated cells is essential to establish the changes that occur as a result of de-differentiation in primary culture. Figures 6A and 6B summarise the molecular functions that are performed by the proteins that were significant changed with culture (BH $p < 0.05$), identified in each dataset by performing a Batch ID tool in PANTHER.

5. Concluding Remarks

Hepatocyte de-differentiation in culture is well documented [7] but still poorly understood. Despite great progress in hepatocyte culture systems over the past 20 years or more the conditions that allow the maintenance of physiologically relevant hepatocyte phenotype in culture for prolonged periods is still elusive. The application of global quantitative technologies is making it possible to undertake much broader phenotyping of cell models and assess what is changing during cell culture. Publication of datasets such as those presented here, will help address the question of what is changing in culture and aid the development of better hepatocyte model systems.

6. Supplementary material

A microsoft excel file is provided with supporting information on 9 sheets:

Sheet "WC iTRAQ results"; All proteins identified and quantified after applying filtering cut-offs as described in methods, for whole cell extract prepared from freshly-isolated and cultured hepatocytes. The values are reported are quantities relative to 1 fresh sample.

Sheet "WC MF"; Molecular functions performed by the proteins identified and quantified in whole cell (WC) hepatocyte experiment. The number of proteins expected to fall into each category was estimated based on the proportion of proteins mapping to each function from the rattus norvegicus reference proteome in PANTHER. Over or under-representation indicated with a "+" or " - " respectively and the associated P value is reported.

Sheet "Nuc iTRAQ results"; All proteins identified and quantified after applying filtering cut-offs as described in methods, for nuclear protein extracts prepared from freshly-isolated and cultured hepatocytes. The values are reported are quantities relative to an experimental pool made from all nuclear samples.

Sheet "Nuc MF"; Molecular functions performed by the proteins identified and quantified in nuclear (Nuc) hepatocyte experiment. The number of proteins expected to fall into each category was estimated based on the proportion of

proteins mapping to each function from the rattus norvegicus reference proteome in PANTHER. Over or under-representation indicated with a "+" or " - " respectively and the associated *p* value is reported.

Sheet "MF enrichment"; The absolute count and percentage of identified proteins performing each Molecular function mapped by PANTHER.

Sheet "WCPvlaueMFs"; Molecular functions performed by the proteins identified as significantly changed in abundance after culture (T test raw *p* value or adjusted (BH) *p* value <0.05) in the whole cell proteomics experiment. The number of proteins expected to fall into each category was estimated based on the proportion of proteins mapping to each function in the whole experimental dataset. Over or under-representation in the significantly changed subpopulation is indicated with a "+" or " - " respectively and the associated *p* value is reported.

Sheet "NucPvalueMFs"; Molecular functions performed by the proteins identified as significantly changed in abundance after culture (T test raw *p* value or adjusted (BH) *p* value <0.05) in the nuclear proteomics experiment. The number of proteins expected to fall into each category was estimated based on the proportion of proteins mapping to each function in the whole experimental dataset. Over or under-representation in the significantly changed subpopulation is indicated with a "+" or " - " respectively and the associated *p* value is reported.

Sheet "UniqueSigNuc"; Proteins identified as significantly changed in abundance during culture were filtered for those annotated to the nucleus in the swissprot database. The change in abundance is given (log2), along with raw *p* value (t test), their annotated cellular locations and protein function, where known.

Acknowledgements

This work was funded by the MRC Centre for Drug Safety Science (grant number G0700654).

References

1. Soars MG, McGinnity DF, Grime K, Riley RJ. Chem Biol Interact 2007;168:2-15.
2. Mingoia RT, Nabb DL, Yang CH, Han X. Toxicol In Vitro 2007;21:165-173.
3. Rogiers V, Vercruysse A. Toxicology 1993;82:193-208.
4. LeCluyse EL. Eur J Pharm Sci 2001;13:343-368.
5. Chesne C, Guyomard C, Fautrel A, Poullain MG, Fremond B, De Jong H, Guillouzo A. Hepatology 1993;18:406-414.
6. Rowe C, Goldring CE, Kitteringham NR, Jenkins RE, Lane BS, Sanderson C, Elliott V, Platt, V. Metcalfe, P. Park, B. K. Journal of proteome research 2010;9:2658-2668.
7. Elaut G, Henkens T, Papeleu P, Snykers S, Vinken M, Vanhaecke T, Rogiers V. Current drug metabolism 2006;7:629-660.
8. Barthelery M, Salli U, Vrana KE. Stem Cells Dev 2007;16:905-919.
9. Lemaigre F, Zaret KS. Curr Opin Genet Dev 2004;14:582-590.
10. Dellaire G, Farrall R, Bickmore WA. Nucleic Acids Res 2003;31:328-330.
11. Fink JL, Karunaratne S, Mittal A, Gardiner DM, Hamilton N, Mahony D, Kai C, Suzuki, H. Hayashizaki, Y. Teasdale, R. D. Towards defining the nuclear proteome. Genome Biol 2008;9:R15.
12. Pewsey E, Bruce C, Tonge P, Evans C, Ow SY, Georgiou AS, Wright PC, Andrews, P. W. Fazeli, A. Journal of Proteome Research 2010;9:3412-3426.
13. Zhang G, Sun L, Lu X, Chen Z, Duerksen-Hughes PJ, Hu H, Zhu X, Yang J. Mutation research 2012;746:66-77.
14. Cao JY, Mansouri S, Frappier L. Journal of virology 2012;86:382-394.
15. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini A J, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J. Genome biology 2004;5:R80.
16. Yang YH, Paquet A and Dudoit S. 2009. R package version 1.32.0;.
17. Pollard KS, Gilbert HN. and Ge Y, Taylor S, and Dudoit S. R package version 2.10.0.
18. Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N, Ladunga I, Ulitsky-Lazareva B, Muruganujan A, Rabkin S, Vandergriff J A, Doremieux O. Nucleic acids research 2003;31:334-341.
19. Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, Muruganujan A, Lazareva-Ulitsky B. Nucleic acids research 2006;34:W645-650.
20. Cerec V, Glaise D, Garnier D, Morosan S, Turlin B, Drenou B, Gripon P, Kremsdorf D, Guguen-Guillouzo C, Corlu A. Hepatology 2007;45:957-967.
21. Dalgetty DM, Medine CN, Iredale JP, Hay DC. Am J Physiol Gastrointest Liver Physiol 2009;297:G241-248.
22. Hay DC, Fletcher J, Payne C, Terrace JD, Gallagher RC, Snoeys J, Black JR, Wojtacha D. Samuel K, Hannoun Z, Pryde A, Filippi C, Currie IS, Forbes SJ, Ross JA, Newsome PN, Iredale JP. Proc Natl Acad Sci U S A 2008;105:12301-12306.
23. Basma H, Soto-Gutierrez A, Yannam GR, Liu L, Ito R, Yamamoto T, Ellis E, Carson SD, Sato S, Chen Y, Muirhead D, Navarro-Alvarez N, Wong RJ, Roy-Chowdhury J, Platt JL, Mercer DF, Miller JD, Strom SC, Kobayashi N, Fox IJ. Gastroenterology 2009;136:990-999.
24. Baxter MA, Rowe C, Alder J, Harrison S, Hanley KP, Park BK, Kitteringham NR, Goldring CE, Hanley NA. Stem Cell Res 2010;5:4-22.
25. Kitteringham NR, Abdullah A, Walsh J, Randle L, Jenkins RE, Sison R, Goldring CEP, Sanderson C, Williams S, Higgins L, Yamamoto M, Hayes J, Park BK. J Proteomics 2010.
26. Jenkins RE, Kitteringham NR, Hunter CL, Webb S, Hunt TJ, Elsby R, Watson RB, Williams D, Pennington SR, Park BK. Proteomics 2006;6:1934-1947.
27. Geiger T, Wehner A, Schaab C, Cox J, Mann M. Molecular and Cellular Proteomics 2012; 11(3):M111.014050. Epub 2012 Jan 25.
28. Boess F, Kamber M, Romer S, Gasser R, Muller D, Albertini S, Suter L. Toxicol Sci 2003;73:386-402.
29. Kienhuis AS, Wortelboer HM, Maas WJ, van Herwijnen M, Kleinjans JC, van Delft JH, Stierum RH. Toxicol In Vitro

2007;21:892-901.
30. Pan C, Kumar C, Bohl S, Klingmueller U, Mann, M. Mol Cell Proteomics 2009. **8**(3): 443-450.
31. Williams SC, Cantwell CA, Johnson PF. Genes Dev 1991;5:1553-1567.
32. Jaishankar A, Barthelery M, Freeman WM, Salli U, Ritty TM, Vrana KE. Stem Cells Dev 2009;18:793-802.
33. Sarkar P, Collier TS, Randall SM, Muddiman DC, Rao BM. Proteomics 2012;12:421-430.
34. Jarboui MA, Wynne K, Elia G, Hall WW, Gautier VW. Molecular immunology 2011;49:441-452.