RNA-Seq-based analysis of differential gene expression associated with hepatitis C virus infection in a cell culture

Anna Hojka-Osinska[1], Lucyna Budzko[1], Agnieszka Zmienko[1,2], Agnieszka Rybarczyk[1,2], Patrick Maillard[3], Agata Budkowska[3], Marek Figlerowicz[1,2] and Paulina Jackowiak[1,4✉]

[1]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland; [2]Institute of Computing Science, Poznan University of Technology, Poznań, Poland; [3]Institut Pasteur, Unité Hepacivirus et Immunité Innée, CNRS, URA 3015, Département de Virologie, Paris, France; [4]Institute of Chemical Technology and Engineering, Poznan University of Technology, Poznań, Poland

Hepatitis C virus (HCV) infection is one of the major causes of chronic liver diseases. Unfortunately, the mechanisms of HCV infection-induced liver injury and host-virus interactions are still not well recognized. To better understand these processes we determined the changes in the host gene expression that occur during HCV infection of Huh-7.5 cells. As a result, we identified genes that may contribute to the immune and metabolic cellular responses to infection. Pathway enrichment analysis indicated that HCV induced an increased expression of genes involved in mitogen-activated protein kinases signaling, adipocytokine signaling, cell cycle and nitrogen metabolism. In addition, the enrichment analyses of processes and molecular functions revealed that the up-regulated genes were mainly implicated in the negative regulation of phosphorylation. Construction of the pathway-gene-process network enabled exploration of a much more complex landscape of molecular interactions. Consequently, several essential processes altered by HCV infection were identified: negative regulation of cell cycle, response to endoplasmic reticulum stress, response to reactive oxygen species, toll-like receptor signaling and pattern recognition receptor signaling. The analyses of genes whose expression was decreased upon HCV infection showed that the latter were engaged in the metabolism of lipids and amino acids. Moreover, we observed disturbance in the cellular antiviral defense. Altogether, our results demonstrated that HCV infection elicits host response that includes a very wide range of cellular mechanisms. Our findings significantly broaden the understanding of complex processes that accompany HCV infection. Consequently, they may be used for developing new host-oriented therapeutic strategies.

INTRODUCTION

Hepatitis C virus (HCV) infection is considered to be a global healthcare problem and a major cause of chronic liver disease, with nearly 170 million people infected worldwide. Exposure to HCV typically leads to acute infection with mild symptoms. While in about 20% of cases the virus is spontaneously eradicated, in the remaining 80–85% of cases chronic hepatitis C (CHC) develops. The latter causes serious liver injury, which may progress to fibrosis, cirrhosis and hepatocellular carcinoma (Ghany *et al.*, 2009). Despite well advanced diagnostics, it is postulated that most of the infected people (50–90%) remain asymptomatic and thus undiagnosed (Galbraith *et al.*, 2015). Moreover, there is no vaccine for prevention of HCV infection. Until very recently, the main CHC treatment involved administering the pegylated interferon alpha in association with ribavirin. This therapy was effective in 50–80% of patients and the positive treatment response rates strongly depended on the HCV genotype. Currently, new combined treatment regimens have been recommended, which involve direct-acting antiviral (DAA) agents that target the viral proteins (Rehermann, 2016). Although these approaches greatly increased the treatment efficiency (with a cure rate of over 90%), some patients fail to develop a sustained response due to the presence of resistance-associated HCV variants (Larrat *et al.*, 2015). Therefore, further antiviral strategies, including host-targeting agents (HTAs), are likely to emerge (Jackowiak *et al.*, 2011; Liang & Ghany, 2013; Wendt *et al.*, 2014).

HCV is a single stranded, positive sense RNA virus. Like other RNA viruses, HCV encodes an error-prone RNA-dependent RNA polymerase that is responsible for high genetic variability of the virus. Accordingly, HCV isolates worldwide are classified into seven genotypes and more than 60 subtypes (Simmonds, 2013). What is more, each infected individual harbors a population of related but diverse HCV variants, referred to as a quasispecies, which evade the host immune response (Farci, 2011; Domingo *et al.*, 2012; Jackowiak *et al.*, 2014). The enormous genetic variability gives HCV the ability to adapt to environmental conditions and rapidly respond to their changes. Consequently, genetic diversity of HCV is considered to be one of the most important factors that underlie the HCV persistence. Moreover, it can lead to generation of well-fitted variants which become stable under certain conditions and dominate the viral population (Figlerowicz *et al.*, 2010; Jackowiak *et al.*, 2012).

A better understanding of the HCV-host interactions is indispensable to elaborate strategies of protecting against HCV infections and/or restricting its development. Thus far, many cellular factors required for the HCV entry, translation, replication, packaging and release

of viral particles have already been identified (Carnero & Fortes, 2016). However, progress in this area had been hindered for a long time by the lack of appropriate model systems that support the complete virus replication cycle in a cell culture. The latter have been developed in the past 15 years. The most exploited HCV cell culture (HCVcc) model is based on infection of the human hepatoma Huh-7.5 cell line with JFH1, an HCV strain of the 2a subtype. In this system, the virus undergoes a complete infection cycle, through virus entry, replication and production of infectious viral particles (Wakita *et al.*, 2005; Lindenbach *et al.*, 2005).

The elaboration of the HCVcc model provided a great opportunity to investigate the effects that HCV infection has on the host gene expression. Until the present day, only a few gene expression studies of infected Huh-7.5 cells have been performed (Walters *et al.*, 2009; Blackham *et al.*, 2010; Woodhouse *et al.*, 2010; Papic *et al.*, 2012). All of these studies have shown that HCV infection changes the profile of the host gene expression and causes alterations in many biological pathways, for example those involved in immune signaling, oxidative stress, apoptosis and cellular metabolism. In three of these studies, the microarray technology was applied, whereas two of them were based on RNA sequencing (RNA-seq). Recently, the latter has become a standard procedure for transcriptome analysis and is considered to be most reliable and reproducible. Nevertheless, the comparison of RNA-seq results obtained for HCV-infected Huh 7.5 cells revealed significant discrepancies (Woodhouse *et al.*, 2010; Papic *et al.*, 2012). This might be a consequence of the application of different approaches to sample preparation. Papic *et al.* performed RNA-seq analysis of 5' capped RNAs, which may be a good practice in identification of unannotated transcripts, as well as RNAs with short poly(A) tails or non-polyadenylated RNAs generated by RNA polymerase II (Papic *et al.*, 2012). However, this approach may lead to the overestimation of the expression levels of known functional transcripts due to capturing of pre-mRNAs (Zhao *et al.*, 2014). In contrast, studies by Woodhouse *et al.* were based on poly(A) RNA, which is a standard procedure for mRNA deep sequencing (Woodhouse *et al.*, 2010). However, their study was conducted when the next generation sequencing (NGS) was only emerging, and thus, the authors could not take advantage of the standardized library preparation methods and sophisticated post-sequencing analyses that are available at the present time. Both RNA-seq analyses of the HCV infected Huh-7.5 cells detected enrichment in pathways that had been also reported in previous microarray-based studies, such as apoptosis, sphingolipid metabolism, amino acid metabolism and TGF-beta signaling. Interestingly, each of the groups had also observed a distinct set of enriched pathways. For example, Woodhouse and coworkers (2010) reported that HCV infection affected the pregnane X receptor/retinoic acid receptor activation pathway and integrin-linked kinase signaling. Moreover, they also reported that HCV infection had a broad effect on cellular metabolism (Woodhouse *et al.*, 2010). Papic *et al.* identified alterations in Notch, Hedgehog and NOD-like receptor signaling pathways, that had not been reported in previous transcriptomic studies (Papic *et al.*, 2012). Considering the discordances between the earlier reports listed above, we found that further studies of the host transcriptional response to HCV infection in a cell culture are necessary.

Thus, the major goal of our study was to identify cellular processes and pathways deregulated by HCV infection and subsequently to better characterize the molecular mechanisms of the HCV-associated liver injury and explore the host factors engaged in the antiviral response. To this end, we determined the changes in gene expression that HCV infection induces in the Huh-7.5 cells. The comparative analysis of transcriptomes from HCV infected and non-infected cells revealed hundreds of differentially expressed genes (DEGs). Among them, we identified genes that have not been reported as differentially expressed in the HCVcc model. Moreover, we found that the products of several of them can have substantial implications for the host response to HCV infection. The pathway enrichment analysis showed alterations in the MAPK signaling, adipocytokine signaling, and cellular metabolism. Moreover, Gene Ontology (GO) analyses suggested perturbations in the regulation of phosphorylation and cell cycle, as well as induction of response to the endoplasmic reticulum (ER) stress. Finally, we constructed a pathway-gene-process network that integrates pathways with individual processes that contribute to them. This analysis revealed new factors involved in the host response to HCV infection in the HCVcc model. Taken together, here we provide a new set of information that may find application in designing new host-targeted antiviral therapeutic strategies.

MATERIALS AND METHODS

Cell culture and infection with JFH-1 HCV. Huh-7.5 human hepatoma cells were kindly provided by C. Rice and grown as previously described (Maillard *et al.*, 2011). Plasmid encoding the genome of JFH-1 HCV strain (genotype 2a), kindly provided by T. Wakita, was used to obtain high-titer stocks of infectious HCV virions, according to the published procedure (Wakita *et al.*, 2005). Huh-7.5 cells were inoculated with the viral stock of 46 400 TCID$_{50}$/ml, at a multiplicity of infection (moi) of 1 or 0.1 for 2 h at 37°C to allow infection. The inoculated cells were washed and grown at 37°C. For total RNA isolation, cells were harvested 72 hours post inoculation (hpi; for moi of 1) or 96 hpi (for moi of 0.1) when infection reached approximately 80% of cells.

Percentage of infected cells was estimated by detection of the HCV core protein in infected cells using an immunofluorescence technique described previously (Cerutti *et al.*, 2011). Mouse monoclonal anti-HCV core (ACAP27, BioRad) and Alexa Fluor 568-labeled goat anti-mouse IgG H+L (Invitrogen) were used as primary and secondary antibodies, respectively. The percentage of infected cells was expressed as an average number of infected cells per 100 cells, counted in a few random fields of view. Non-infected control cells were cultured and collected in parallel. Experiments were performed in duplicates.

RNA extraction, library preparation and sequencing. RNA was extracted from the infected and non-infected Huh-7.5 cells using the *mir*Vana™ miRNA Isolation Kit, according to the manufacturer's instructions for isolation of long RNA fraction that is depleted of small RNAs. For each of the two biological replicates, four RNA samples were obtained: from infected cells collected at 72 hpi (72I), from infected cells collected at 96 hpi (96I) and from the corresponding controls (72C and 96C, respectively). RNA quality was assessed using Nanodrop 2000 and all of the samples used for this study had an excellent purity ($A_{260}/A_{280} \geq 1.9$; $A_{260}/A_{230} \geq 2$) and showed no visible signs of degradation (RIN$\geq$9) in the 2100 Bioanalyzer RNA NANO assay (Agilent).

cDNA libraries for RNA-Seq were prepared as two technical replicates, where in each four RNA samples from one biological replicate were included: 72I, 96I, 72C and 96C. The libraries were obtained according to strand-specific protocol adapted from Parkhomchuk and coworkers (2009) and Sultan and coworkers (2012), using the dUTP method combined with Illumina TruSeq RNA Sample preparation kit v.2. Briefly, mRNA purification, fragmentation, and first strand synthesis were performed as described in the Illumina TruSeq RNA kit manual. Next, the procedure was interrupted and RNA was cleaned on the Sephadex G-50 columns. The second strand synthesis was performed using dUTP/dNTPs. After this step, the Illumina TruSeq RNA kit protocol was resumed at the step of cDNA cleanup. Before PCR enrichment, the protocol was interrupted again and degradation of the second strand was performed using USER (uracil-specific excision reagent) Enzyme mix (New England BioLabs). Next, the TruSeq RNA kit protocol was resumed at the "Enrich DNA fragment" step and continued until the end.

Sequencing was performed with Illumina Genome Analyzer (GAIIx) and single-end 100 bp reads were obtained.

Preprocessing and alignment of sequencing reads. The raw sequencing reads were checked for quality with Prinseq (Schmieder & Edwards, 2011), trimmed to remove parts of the sequences with Phred quality score (Ewing & Green, 1998; Ewing et al., 1998; Ledergerber & Dessimoz, 2011) that was below 30, and then filtered to keep sequences longer than 20 nt.

Filtered reads were aligned to *Homo sapiens* reference genome build hg38 – Ensembl GRCh38_76 (ftp://ftp.ensembl.org/pub/release-76/fasta/homo_sapiens) using TopHat v. 2.014 (Trapnell et al., 2009). Taking advantage of the strand-specific library preparation protocol (Sultan et al., 2012), TopHat was set to treat each read as the first-strand read. The aligned reads were processed with HTSeq (Anders et al., 2015) to count the abundance of genes and GTF annotation file (Gencode version 21) (Zhang, 2016) was used to guide the assembly. The generated count table (Anders et al., 2015) was used for further differential gene expression (DGE) analysis.

Differential gene expression analysis. The DGE analysis was performed using edgeR (Robinson et al., 2010). The count table of gene coverage from HTseq (Anders et al., 2015) was used as input data. First, the expression levels of the analyzed genes were presented as the number of reads per kilobase per million (RPKMs) (Mortazavi et al., 2008). Then, statistical analysis was performed and the results table (Robinson et al., 2010) was reported with values of: RPKMs, fold change (FC), $\log_2(FC)$, statistical significance (P-value), and Benjamini-Hochberg adjusted p-value (P adj. value) (Benjamini & Hochberg, 1995). FC was calculated as a ratio of expression level in infected and non-infected cells. For values less than 1, the ratio was inverted and multiplied by –1. In the next step, genes encoding proteins that showed RPKM ≥1 were selected (as genes with low coverage can produce artificially high FC values and cause more false positive results (Yuan et al., 2013)). In order to identify genes with significantly altered expression in infected cells, with respect to their controls, different stringency criteria were applied (see the Results section).

KEGG pathway, GO biological process and GO molecular function enrichment analyses. The differentially expressed genes (up- and down-regulated, separately) were subjected to functional enrichment analysis, using Enrichr (Chen et al., 2013). The KEGG (Kyoto Encyclopedia of Genes and Genomes) 2015, GO Biological Process 2015 and GO Molecular Function 2015 libraries were used and the statistical significance P-value cutoff was set at 0.05.

Construction of pathway-gene-process network. The construction of the relationship network started from the list of genes that overlapped either of the four enriched pathways (MAPK signaling, adipocytokine signaling, cell cycle and nitrogen metabolism; see Table 3). The relationship between a given pathway and a biological process was determined based on the assumption that the gene that overlapped the biological pathway also overlapped the biological process in the enrichment analysis. To create the network, we searched for any enriched GO biological process overlapped by any of the genes from the starting list. Notably, the GO biological processes have child-parental relationships, for example: positive regulation of protein serine/threonine kinase activity → regulation of protein serine/threonine kinase activity → positive regulation of protein kinase activity → positive regulation of kinase activity. Therefore, only biological processes that represent the lowest class of relationship graph for GO biological processes were visualized in the network (for example: positive regulation of protein serine/threonine kinase activity). Relationships between biological processes were then represented as an undirected graph. Biological processes presented in rectangular boxes were connected to each other via mediating circle nodes that represented genes overlapping with these processes.

RESULTS

RNA-seq analysis of annotated gene transcripts

To characterize the host transcriptional response to HCV infection, RNA samples were obtained from Huh-7.5 cells 72 and 96 hpi and from the time-matched non-infected cells. Next, RNA-seq was performed and the obtained data were processed as described in Materials and Methods. As our analysis focused on the HCV infection-induced alterations in pathways and processes, we decided to restrict the obtained RNA-Seq data to protein encoding transcripts. Our bioinformatics analyses allowed us to identify more than 17 000 protein coding RNAs with at least one read and 13 847 transcripts with RPKM≥1. These findings are consistent with reports indicating that the number of expressed genes in cell lines ranges between 10 000 to 15 000 (Jongeneel et al., 2003; Marinov et al., 2014).

Identification of differentially expressed genes

Several selection criteria were tested to restrict the list of the analyzed genes to those whose expression is undoubtedly disturbed after HCV infection. The initial analysis that included only the FC threshold (FC≥2 or ≤–2), identified 1 131 and 700 DEGs at 72 and 96 hpi, respectively (Table 1). In agreement with the previous study (Walters et al., 2009), a greater number of DEGs was observed at 72 hpi, while at 96 hpi this number decreased. When the standard for RNA-seq analysis filter of P adj. value ≤0.05 was applied to genes selected with the previous threshold, the subset of DEGs was narrowed down to 710 at 72 hpi and 334 at 96 hpi. When the most stringent filter of P adj. value ≤0.001 was used, the numbers of DEGs were 494 and 272, respectively. However, the number of genes that met the applied cri-

Table 1. Assessment of differentially expressed genes.
Comparison of different stringency criteria for FC and P adj. values to determine the best cutoff for genes whose expression is constantly altered in response to an HCV infection in the host cells.

Time points post inoculation	P adj. value	Number of up-regulated genes (FC≥2)	Number of down-regulated genes (FC≤−2)
	–	805	326
	≤0.05	547	163
72 h	≤0.01	474	136
	≤0.005	451	124
	≤0.001	395	99
	–	419	281
	≤0.05	286	48
96 h	≤0.01	267	39
	≤0.005	263	35
	≤0.001	250	22
	–	291	62
	≤0.05	210	24
Shared by both time points	≤0.01	182	20
	≤0.005	175	17
	≤0.001	161	11

teria at both time points did not change so dramatically with increasing threshold stringency (378 without filter P adj. value, 234 with P adj. value ≤0.05, 172 with P adj. value ≤0.001). Because we intended to analyze the genes whose expression was constantly altered in response to HCV infection in the host cells, we focused on the subset of 172 genes for which differential expression was the most significant (2≥FC≤−2, P adj. value ≤0.001) at both time points after HCV infection (Table 2 and Supp. Table 1 at www.actabp.pl). Among them, 161 genes were up-regulated and 11 were down-regulated. Many of these genes have been previously identified as differentially expressed in the HCVcc infection model, which supports the relevance of our results. At the same time, our analysis identified 25 genes whose altered expression has not been associated with an HCV infection of the Huh-7.5 cells (Supp. Table 1 at www.actabp.pl). A great subset of

those genes' products regulate transcription. They bind nucleic acids or transcription factors. Another group of the newly discovered DEGs encodes proteins that influence the cell cycle and cell death. They accomplish their functions via chromatin, microtubule or signal adaptor binding. Two further genes whose altered expression has not been previously associated with HCV infection of the Huh-7.5 cells are related to negative regulation of phosphorylation. We were especially interested if the products of the newly identified DEGs had any association with the response to virus infections. Intriguingly, considering the available literature data, several proteins emerged as potential regulators in anti-HCV host response: DDX60 (DEAD Box Protein 60), FAM46C (family with sequence similarity 46, member C), INPP5J (Inositol Polyphosphate-5-Phosphatase J) and PIK3IP1 (Phosphoinositide-3-Kinase Interacting Protein 1).

Pathway enrichment analysis of differentially expressed genes

Next, we determined the impact of HCV infection on the cellular pathways. A pathway can be defined as a series of actions between molecules that leads to a certain product or a significant change in the cell homeostasis (Baoying, 2014). It can be assumed that if a pathway involves several gene products, and the levels of some of those products are altered, then the effect of this pathway is disturbed. Our analysis of the enriched (altered) pathways was performed on the set of DEGs identified earlier (listed in Table 2) and 179 KEGG pathways deposited in the Enrichr database. As a result, we found four significantly enriched pathways for the up-regulated genes (Table 3).

The most altered pathway was MAPK signaling with 10 overlapping genes. To determine their functions, we used annotations collected in the GeneCards database (www.genecards.org) (Safran *et al.*, 2010). All information regarding gene characteristics provided hereafter is based on the GeneCards database, unless indicated otherwise. Out of those 10 genes, five belong to a dual specificity protein phosphatase subfamily (*DUSP1*, *DUSP4*, *DUSP8*, *DUSP10*, *DUSP16*). Products of these genes dephosphorylate, and thus negatively regulate members of the mitogen-activated protein kinase (MAPK) superfamily: MAPK1 (Mitogen-Activated Protein Kinase 1), ERK1 (Extracellular Signal-Regulated Kinase 1), p38

Table 2. Complete list of differentially expressed genes.
The table assembles the Associated Gene Names for genes whose differential expression was the most significant (2≥FC≤−2, P adj. value ≤0.001) at both time points post HCV inoculation (72 and 96 hpi)

Up-regulated genes

ABLIM3, ACSL1, AHR, AJUBA, ALPK3, AMOTL2, ANKRD1, ANXA1, ANXA3, AREG, ARHGEF2, ARMCX1, ASNS, ATF3, BACH2, BAM-BI, BEST1, BEX2, BHLHE40, BHLHE41, BIRC3, BMF, BMP2, C10orf10, CDC14A, CDKN2B, CEBPB, CEBPD, CHAC1, CLIP4, CREB5, CREBRF, CSF1, CTH, CXCL2, CXCL5, DAB2, DDIT3, DDIT4, DDX60, DHX32, DLX2, DTNA, DUSP1, DUSP10, DUSP16, DUSP4, DUSP8, EIF1, ELF3, ELMSAN1, EREG, ERN1, FAM47E, FBXO32, FOSL2, FUT1, GAD1, GADD45A, GADD45B, GATA3, GCNT3, GDF15, GEM, GTPBP2, HDAC9, HERPUD1, HIVEP2, HKDC1, HOXA1, HRK, HSPB8, ICAM1, IL32, IL8, INHBE, INPP5J, IRF1, IRS2, JMY, JUN, KIF5C, KLF10, KLF11, KLF4, KLF6, KLHDC7B, KLHL24, KRT23, LARP6, LETM2, LTBP1, MAFF, MAP1LC3B, MAP1LC3B2, MBNL2, MOSPD1, MR1, MRAS, MTHFD2, MXD1, NEU1, NEURL3, NFKBIA, NIPAL2, NR1D2, NUAK2, PHLDA1, PIK3IP1, PLA1A, PLA2G4C, PMAIP1, PPARGC1A, PPM1E, PPP1R15A, PRNP, PROSER2, PRSS23, PTPRH, RASD1, RASGEF1B, RBM24, RCAN2, RCN1, RORA, RPS6KL1, S100P, SAMD4A, SESN2, SH2B3, SH2D3C, SH3BP2, SH3RF1, SKIL, SLC1A4, SLC22A15, SLC25A25, SLC34A2, SLC3A2, SLC7A11, SOCS2, SPAG1, STC2, STX3, SYT11, TCP11L2, THBS1, THSD4, TMC7, TMEM140, TNFAIP3, TNFRSF9, TP53INP1, TRIB1, TRIB3, ULBP1, VNN3, WHAMM, ZBTB21, ZEB2, ZNF669

Down-regulated genes

ACADSB, ACTN3, FAM46C, HEPACAM2, HMGCS2, IFIT1, NID2, PALM3, PCDH20, SLC38A4, TCTEX1D1

Table 3. Enriched KEGG pathways for the up-regulated genes (FC≥2, *P* adj. value ≤0.001; 72 and 96 hpi).

KEGG pathway name	Genes in pathway	DEGs	*P*-value	Up-regulated genes that overlap the pathway
mapk signaling pathway	248	10	0.00061	DUSP4; JUN; DUSP10; MRAS; GADD45B; DUSP1; GADD45A; DDIT3; DUSP8; DUSP16
adipocytokine signaling pathway	71	4	0.01087	NFKBIA; ACSL1; IRS2; PPARGC1A
cell cycle	104	4	0.03629	CDKN2B; GADD45B; GADD45A; CDC14A
nitrogen metabolism	24	2	0.03643	CTH; ASNS

MAPK, JNK (c-Jun N-terminal kinase). MAPKs are involved in cellular responses to a wide range of stimuli, such as heat shock, proinflammatory cytokines, growth factors, virus infections. In their activated state, MAPKs act as integration points of many cellular processes such as proliferation, differentiation and development. Upon activation, these kinases translocate to nucleus, where they phosphorylate their targets. Four other genes encode transcription factors (TFs), whose levels are increased during an environmental stress. Among them, Gadd45A (Growth arrest and DNA-damage-inducible protein alpha) and Gadd45B (Growth arrest and DNA-damage-inducible protein beta) respond to environmental stresses by activating the p38/JNK pathway. DDIT3 (DNA-Damage-Inducible Transcript 3 Protein) is a TF activated by the ER stress and promotes apoptosis, while Jun activates a wide range of genes in response to different types of cell stimulation. The last enriched gene in the MAPK signaling pathway was *MRAS* (Muscle RAS Oncogene Homolog) that encodes a membrane-associated protein, a signal transducer in multiple processes, including cell growth and differentiation.

Another significantly altered pathway was the adipocytokine signaling, overlapped by four up-regulated genes. Only one of these genes, *PPARGC1A* (Peroxisome Proliferator-Activated Receptor Gamma, Coactivator 1 Alpha), encodes a TF – a coactivator that regulates genes involved in the energy metabolism. The other two genes that overlapped this pathway encode proteins that function as intracellular signal transducers. The first one, *IRS2* encodes the insulin receptor substrate 2, a signaling molecule that acts as a molecular adaptor in cytokine (including insulin) signal transduction. The second one, *NFKBIA* (Nuclear Factor Of Kappa Light Polypeptide Gene Enhancer In B-Cells Inhibitor, Alpha) gene, encodes a protein that inhibits the NF-kappa-B/REL complexes, which are involved in inflammatory responses. The fourth gene, *ACSL* (Long-chain-fatty-acid – CoA ligase 1), encodes a ligase that plays a key role in lipid biosynthesis and fatty acid degradation by converting free long-chain fatty acids into fatty acyl-CoA esters.

The third significantly enriched pathway was cell cycle, which was overlapped by four up-regulated genes. Two of them (*GADD45A*, *GADD45B*), encoding TFs, were mentioned above as genes that also overlapped MAPK signaling pathway. The other two take part in cell cycle regulation by interacting with other proteins. *CDKN2B* (Cyclin-Dependent Kinase Inhibitor 2B) encodes an inhibitor which prevents activation of the CDK kinases, thus controlling progression of the G1 phase of cell cycle. *CDC14A* (Cell Division Cycle 14A) encodes a protein that is a member of the dual specificity protein tyrosine phosphatase family and is proposed to regulate the function of p53.

The last enriched pathway was nitrogen metabolism, with two overlapping genes involved in amino acids biosynthesis: (i) *CTH* encoding Cystathionine Gamma-Lyase, and (ii) *ASNS* encoding Glutamine-Dependent Asparagine Synthetase.

The pathway enrichment analysis performed for the subset of 11 down-regulated genes revealed four altered pathways, however, each of them was overlapped only by a single gene (Table 4). Among the down-regulated genes that overlapped enriched pathways were *HMGCS2* (3-Hydroxy-3-Methylglutaryl-CoA Synthase 2) and *ACADSB* (Acyl-CoA Dehydrogenase, Short/Branched Chain). Notably, *HMGCS2* overlapped three pathways: (i) synthesis and degradation of ketone bodies, (ii) valine, leucine and isoleucine degradation, and (iii) butanoate metabolism. *ACADSB* overlapped only one pathway of fatty acid metabolism. Both of these genes encode enzymes that metabolize derivatives of coenzymeA.

Table 4. Enriched KEGG pathways for the down-regulated genes (FC≤–2, *P* adj. value ≤0.001; 72 and 96 hpi).

KEGG pathway name	Genes in pathway	DEGs	*P*-value	Down-regulated genes that overlap the pathway
synthesis and degradation of ketone bodies	9	1	0.00787	HMGCS2
valine leucine and isoleucine degradation	44	1	0.03509	HMGCS2
fatty acid metabolism	45	1	0.03586	ACADSB
butanoate metabolism	45	1	0.03586	HMGCS2
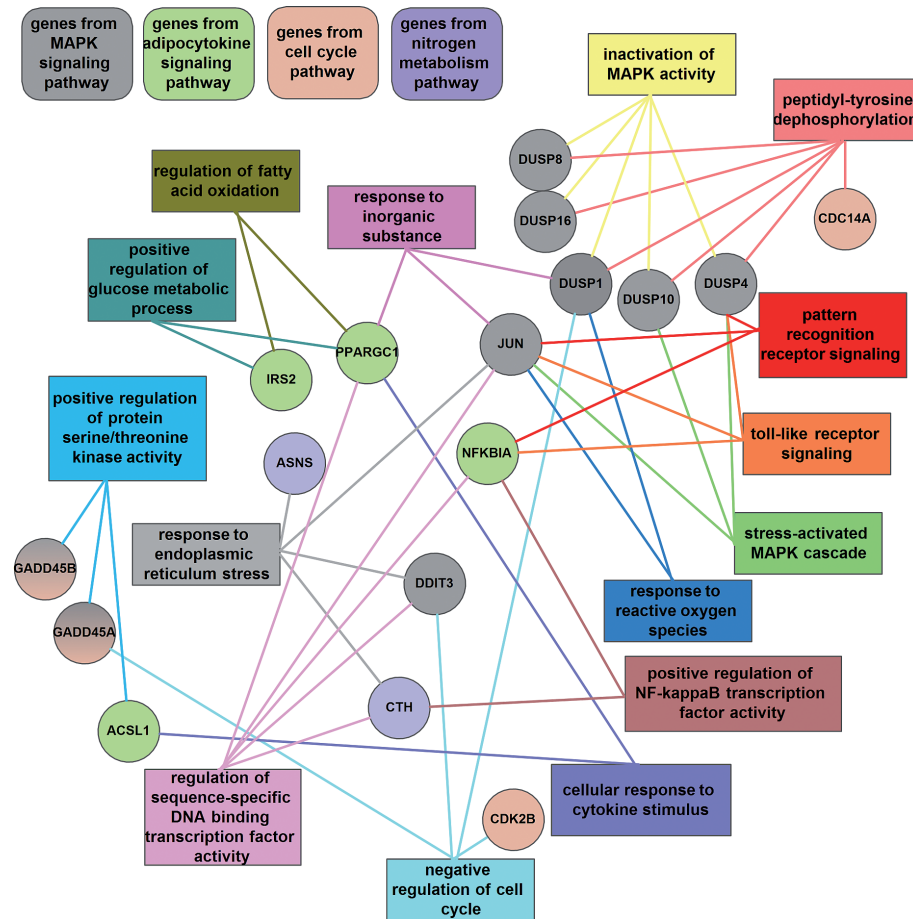
Figure 1. Pathway-gene-process network.
A graphical representation of relationships between the GO biological processes and the KEGG pathways. The circle nodes represent genes from statistically significant KEGG pathways (colored according to the overlapped pathway). The genes are connected to the GO biological process that they overlap, thus a network of interactions between biological processes themselves, as well as interactions between biological processes and pathways, is created.

Biological pathways exist due to the harmonized compilation of biological processes, which are mediated by proteins that display certain molecular functions. Assignment of a differentially expressed gene to a pathway shows potential consequences of the gene's altered expression. However, it does not provide information on the mechanisms underlying these effects. To get a better understanding of the phenomena associated with the observed changes in gene expression, we performed further enrichment analyses focused on biological processes and molecular functions.

Biological process enrichment analysis of differentially expressed genes

We analyzed the enrichment of biological processes among the DEGs (see Table 2), using data from the GO database. According to GO, biological process is defined as a series of molecular events that have a defined beginning and end. In contrast to pathways, processes do not represent molecular interactions (Gene Ontology Consortium, 2015).

The analysis was performed on the set of 5192 GO biological processes that are deposited in the Enrichr library. As a result of tests performed on the subset of up-regulated genes, 477 biological processes were found to be enriched (Supp. Table 2 at www.actabp.pl). The most significantly enriched biological process was negative regulation of phosphorylation, with 21 overlapping genes. From the top 10 enriched biological processes, two more were overlapped by the same 21 genes and those were: negative regulation of phosphate metabolic process and negative regulation of phosphorus metabolic process. Furthermore, among the top 10 enriched biological processes were also: negative regulation of transferase activity, regulation of protein serine/threonine kinase activity, negative regulation of protein phosphorylation and negative regulation of kinase activity. All of the above processes are related to cellular signal transduction. In addition, the top 10 significantly enriched biological processes also included: (i) response to ER stress, (ii) cell cycle arrest, and (iii) response to unfolded protein.

Analysis of the 11 down-regulated genes revealed 45 enriched biological processes, even though each of them was overlapped by a single gene only (Supp. Table 3 at www.actabp.pl). In the top 10 enriched biological processes, only four genes of the 11 genes analyzed were included. The most prominent observation was that the *IFIT1* gene turned up four times in the processes that can be easily connected to the response to an RNA virus infection (cellular response to exogenous dsRNA, cellular response to dsRNA, negative regulation of defense response to virus, regulation of helicase activity). The

next most frequently occurring gene was *HMGCS2*, engaged in ketone body metabolic process and isoprenoid biosynthetic process. On the list of the top 10 enriched biological processes were also: basement membrane organization and branched-chain amino acid catabolic process.

Our analysis revealed that genes overlapping one pathway were annotated to several biological processes. Thus, to visualize the relationships between pathways and biological processes, we created the pathway-gene-process network (Fig. 1). This network shows how up-regulated genes overlapping enriched KEGG pathways are connected with biological processes. By creating this network, we were able to identify relationships between biological processes that contributed to four enriched pathways: (i) MAPK signaling, (ii) adipocytokine signaling, (ii) cell cycle, and (iv) nitrogen metabolism. For example, genes that overlapped MAPK signaling pathway are assigned not only to inactivation of MAPK activity and stress-activated MAPK cascade processes, but also to negative regulation of cell cycle, response to ER stress, response to reactive oxygen species, the toll-like receptor (TLR) signaling process and the pattern recognition receptor (PRR) signaling process. Moreover, the network created also considered: (i) processes connected to the response to cellular stress (cellular response to cytokine stimulus, ER unfolded protein response), (ii) processes connected to cellular signaling (positive regulation of protein serine/threonine kinase activity, regulation of sequence-specific DNA binding transcription factor activity).

Molecular function enrichment analysis of differentially expressed genes

According to the GO definition, biological processes have a defined end, thus, they must be coherent in their action. Consequently, gene products participating in the same biological process usually have similar functions. To find out if the DEGs can be grouped on the basis of functional similarity, we performed the enrichment analysis of 1 136 GO molecular functions deposited in the Enrichr library. For 161 up-regulated genes, we found 67 enriched molecular functions (Supp. Table 4 at www.actabp.pl). On the top of the generated GO-molecular function list was a group of genes with MAP kinase tyrosine/serine/threonine phosphatase activity. Two groups of functionally related proteins were overlapped by the highest numbers of genes: sequence-specific DNA binding RNA polymerase II transcription factor activity and transcription factor binding function.

For the subset of 11 down-regulated genes, the GO molecular function enrichment analysis identified seven functions (Supp. Table 5 at www.actabp.pl). Calcium ion binding was displayed by three genes, whereas each of the remaining functions was overlapped solely by one gene. Only one of these genes, *ACADSB*, was observed earlier in the pathway enrichment analysis. Its molecular functions were acyl-CoA dehydrogenase activity and flavin adenine dinucleotide binding.

DISCUSSION

Genes involved in response to HCV infection in the Huh-7.5 cells

The HCVcc, involving Huh-7.5 hepatoma cell line infected with HCV JFH1 strain is used worldwide and is believed to be one of the best *in vitro* models for the study of the complete virus replication cycle (Wilson *et al.*, 2012). Thus, much of our understanding of HCV biology is shaped through the use of this model. To examine the host transcriptional responses to HCV infection, we performed RNA-Seq analysis of poly(A)-selected RNA isolated from the HCV-infected Huh-7.5 cells at 72 and 96 hpi. In contrast to previous analyses, our studies were focused on a very exclusive set of DEGs, selected by applying the most stringent filtering conditions. By choosing only 172 genes that were differentially expressed at both time points post inoculation, we were able to focus our analysis on genes whose altered expression can be considered to be a stable host response. It is worth to mention that a vast majority of the observed changes involved up-regulation of gene expression, which is consistent with a previous report (Walters *et al.*, 2009).

New differentially expressed genes implicated in an antiviral response

Many of the DEGs observed in this study have been reported earlier (Walters *et al.*, 2009; Blackham *et al.*, 2010; Papic *et al.*, 2012), however, 25 novel host genes that change expression levels in the Huh-7.5 cells in the response to HCV infection were also identified (Supp. Table 1 at www.actabp.pl). Analysis of the molecular functions of the 25 genes mentioned above and their participation in biological processes strengthened previous observations that HCV exerts its major impact on the host cell via changes in the cell cycle and transcription regulation. In the newly identified group of DEGs, several genes were of particular interest due to their putative engagement in the antiviral response. The first of them was *DDX60*. It has been shown that DDX60 mRNA level was increased in primary liver cultures upon HCV infection (Marukian *et al.*, 2011). Transient transfection experiments demonstrated that DDX60 had an anti-HCV activity (Schoggins *et al.*, 2011), and ectopic expression of this gene inhibited the HCV replication (Grünvogel *et al.*, 2015). DDX60 was shown not only to be involved in RIG-I activation, but also in RIG-I independent viral RNA degradation (Oshiumi *et al.*, 2015). Our finding of DDX60 upregulation upon HCV infection in the Huh-7.5 cells suggests the role of this gene in the mechanism governing an innate antiviral response in cells without functional RIG-I. The second gene that could be involved in the antiviral response was *FAM46C*. It belongs to type I interferon stimulated genes (ISG). However, the role of the product of this gene is unclear, as it was reported that expression of *FAM46C* enhanced replication of some viruses, including HCV (Schoggins *et al.*, 2011). After all, this observation is in agreement with our study, as we detected a decrease in FAM46C mRNA upon HCV infection. Consequently, down-regulation of this gene may be a part of immune response to virus infection.

Apart from the genes that can be involved in the host immune response, we also detected differential expression of two genes encoding proteins engaged in phosphatidylinositol metabolism: INPP5J (Inositol Polyphosphate-5-Phosphatase J) and PIK3IP1 (Phosphoinositide-3-Kinase Interacting Protein 1). Lipids that are derivatives of phosphatidylinositol are known as phosphoinositides (PIs) and represent the main fraction of cell membrane phospholipids (Balla, 2013). It has been reported that HCV targets the PI metabolism (Benedicto *et al.*, 2011; Bianco *et al.*, 2012). One of the effects of this event is the reduced membrane polarity (Mee *et al.*,

2010). Accordingly, one can hypothesize that the observed up-regulation of *INPP5J* and *PIK3IP1* may constitute the host response to HCV infection to somehow compensate the negative influence of the virus on the PIs metabolism.

The products of the newly identified genes implicated in the antiviral response might be considered as potential targets for HTAs. In particular, DDX60 appears to be the most attractive. Recently, treatment with the HTAs has been considered as a promising alternative to DAAs (Zeisel *et al.*, 2013). The main advantage of the former is the ability to overcome genetic barrier of resistance, because they target host factors that, in comparison to pathogen targets, are far less prone to variation. There are two main concepts in HTAs development: (i) to target the host elements directly engaged in virus replication cycle, or (ii) to boost the host innate immune response. However, our data suggested that there are new promising host factor-based targets that are more distantly connected to the anti-HCV response.

Molecular mechanisms of cellular response to HCV infection

The main part of our study was the enrichment analysis of DEGs. We decided to perform not only the enrichment analysis of pathways but also of processes and molecular functions. As we were interested in increasing our knowledge of biological mechanisms of the host response to HCV, the pathway analysis alone turned out to be insufficient. Because pathway is a dynamic representation of molecular interactions, the detection of some pathways' alteration per se does not give the whole image of the biological effects. Notably, for the DEGs, we found only a few enriched pathways, in comparison to more than 500 enriched biological processes.

Taken together, our enrichment analyses demonstrated that HCV infection mainly caused an increased expression of genes whose products belong to three functional groups: (i) negative regulators of phosphorylation, (ii) transcription regulators, and (iii) signal transducers. The first group included phosphatases and inhibitors of kinases from different signaling pathways. Their increased level may have potentially far-reaching consequences, as they control fundamental cellular processes, such as transcription, translation and biogenesis of small regulatory RNAs (Roux & Blenis, 2004; Malumbres *et al.*, 2009; Kurzynska-Kokorniak *et al.*, 2015). Genes that belong to the last two groups are mainly associated with the cell cycle arrest and response to the ER stress. Multiple steps of HCV replication in the liver cells are related to ER, thus the infection induces ER stress that diminishes the ER protein-folding capacity (Chan, 2014). The unfolded protein response is a cellular mechanism that restores homeostasis upon increased protein misfolding or elicits apoptosis in case of severe disturbances (Walter & Ron, 2011). Our results suggest that in Huh-7.5 cells infected with HCV this response directs the cells towards apoptosis (see Supp. Table 2 at www.actabp.pl). This may indicate that ER stress in the HCVcc model is very intense.

A small number of the identified down-regulated genes resulted in the majority of enriched pathways, processes and functions being overlapped by a single gene. This fact compromises the inference of a generalized characteristics of these genes. Nevertheless, it is evident that the down-regulated genes represented clearly different functional groups than the up-regulated ones. The genes whose expression decreased upon HCV infection encode enzymes involved in the metabolism of coenzyme A, in the processes of ketogenesis, isoprenoid biosynthesis and branched-chain amino acid degradation. It has been shown that valine, one of the branched-chain amino acids, promotes formation of viral particles but restricts the HCV RNA synthesis (Ishida *et al.*, 2013). It can be speculated that down-regulation of genes implicated in branched-chain amino acid degradation may represent a mechanism by which HCV favors production of progeny virions over RNA replication at certain stages of infection. In addition, HCV infection directly targeted the cellular defense mechanisms via down-regulation of the gene whose product governs the response to RNA viruses.

Pathway-gene-process network

To integrate the obtained results, we constructed a network connecting up-regulated genes from enriched pathways with processes that are annotated to these genes and indispensable for pathway functioning. In our analysis, the most altered pathway was MAPK signaling for which disturbances were reported previously (Papic *et al.*, 2012). Nonetheless, our approach indicated that under perturbations of that pathway there are hidden much more complex biological effects. Up-regulated genes that overlap MAPK signaling pathway, like *JUN*, *DUSP10* and *DUSP4*, also overlap the process of TLR signaling. Their altered expression was insufficient to detect disruption of the latter during pathway enrichment analysis. However, it is known that TLR signaling is affected by HCV infection (Howell *et al.*, 2013). Therefore, the applied combination of enrichment analyses at multiple levels (pathway, processes and functions) proved to efficiently extract information that would be missed otherwise. Another valuable remark from our pathway-gene-process network is the detected enrichment in the PRR signaling process. The latter is one of the mechanisms underlying the detection of HCV and activation of innate antiviral response in the host cell (Gokhale *et al.*, 2014). Three genes included in the constructed network overlapped the PRR signaling process. In addition, based on the biological process enrichment analysis, we were able to identify two more up-regulated genes related to the PRR signaling process (see Supp. Table 2 at www.actabp.pl). The latter were: *TNFAIP3* (Tumor Necrosis Factor, Alpha-Induced Protein 3) and *BIRC3* (Baculoviral IAP Repeat Containing 3). To sum up, our results illustrated that creating a pathway-gene-process network may be an effective approach to investigate the molecular mechanisms affected by the altered gene expression.

In conclusion, our data expand the understanding of acute HCV infection and may be a starting point for further studies of the mechanisms of virus-induced liver injury and HCV-host interactions. Moreover, by revealing new host factors involved in the response to HCV infection, our results open new perspectives in planning and designing the host-oriented therapeutic approaches.

awarded on the basis of a decision number DEC-2012/05/D/NZ2/02238, to PJ.

This work was partially financed by the European Union within the European Regional Development Fund through the MPD program.

This publication was also supported by the Polish Ministry of Science and Higher Education, under the KNOW program.

REFERENCES

Anders S, Pyl PT, Huber W (2015) HTSeq – a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169. http://doi.org/10.1093/bioinformatics/btu638

Balla T (2013) Phosphoinositides: tiny lipids with giant impact on cell regulation. *Physiol Rev* **93**: 1019–1137. http://doi.org/10.1152/physrev.00028.2012

Baoying W (2014) *Big Data Analytics in Bioinformatics and Healthcare*. IGI Global.

Benedicto I, Molina-Jiménez F, Moreno-Otero R, López-Cabrera M, Majano PL (2011) Interplay among cellular polarization, lipoprotein metabolism and hepatitis C virus entry. *World J Gastroenterol WJG* **17**: 2683–2690. http://doi.org/10.3748/wjg.v17.i22.2683

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol* **57**: 289–300

Bianco A, Reghellin V, Donnici L, Fenu S, Alvarez R, Baruffa C, Peri F, Pagani M, Abrignani S, Neddermann P, De Francesco R (2012) Metabolism of Phosphatidylinositol 4-Kinase IIIα-Dependent PI4P Is Subverted by HCV and Is Targeted by a 4-Anilino Quinazoline with Antiviral Activity. *PLoS Pathog* **8**: e1002576. http://doi.org/10.1371/journal.ppat.1002576

Blackham S, Baillie A, Al-Hababi F, Remlinger K, You S, Hamatake R, McGarvey MJ (2010) Gene expression profiling indicates the roles of host oxidative stress, apoptosis, lipid metabolism, and intracellular transport genes in the replication of hepatitis C virus. *J Virol* **84**: 5404–5414. http://doi.org/10.1128/JVI.02529-09

Carnero E, Fortes P (2016) HCV infection, IFN response and the coding and non-coding host cell genome. *Virus Res* **212**: 85–102. http://doi.org/10.1016/j.virusres.2015.10.001

Cerutti A, Maillard P, Minisini R, Vidalain P-O, Roohvand F, Pecheur E-I, Pirisi M, Budkowska A (2011) Identification of a Functional, CRM-1-Dependent Nuclear Export Signal in Hepatitis C Virus Core Protein. *PLOS ONE* **6**: e25854. http://doi.org/10.1371/journal.pone.0025854

Chan SW (2014) The unfolded protein response in virus infections. *Front Microbiol* **5**: 518. http://doi.org/10.3389/fmicb.2014.00518

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**: 128. http://doi.org/10.1186/1471-2105-14-128

Domingo E, Sheldon J, Perales C (2012) Viral Quasispecies Evolution. *Microbiol Mol Biol Rev* **76**: 159–216. http://doi.org/10.1128/MMBR.05023-11

Ewing B, Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**: 186–194

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**: 175–185

Farci P (2011) New insights into the HCV quasispecies and compartmentalization. *Semin Liver Dis* **31**: 356–374. http://doi.org/10.1055/s-0031-1297925

Figlerowicz M, Jackowiak P, Formanowicz P, Kędziora P, Alejska M, Malinowska N, Błażewicz J, Figlerowicz M (2010) Hepatitis C virus quasispecies in chronically infected children subjected to interferon–ribavirin therapy. *Arch Virol* **155**: 1977–1987. http://doi.org/10.1007/s00705-010-0789-7

Galbraith JW, Franco RA, Donnelly JP, Rodgers JB, Morgan JM, Viles AF, Overton ET, Saag MS, Wang HE (2015) Unrecognized chronic hepatitis C virus infection among baby boomers in the emergency department. *Hepatology* **61**: 776–782. http://doi.org/10.1002/hep.27410

Gene Ontology Consortium (2015) Gene Ontology Consortium: going forward. *Nucleic Acids Res* **43**: D1049–1056. http://doi.org/10.1093/nar/gku1179

Ghany MG, Strader DB, Thomas DL, Seeff LB (2009) Diagnosis, management, and treatment of hepatitis C: An update. *Hepatology* **49**: 1335–1374. http://doi.org/10.1002/hep.22759

Gokhale NS, Vazquez C, Horner SM (2014) Hepatitis C virus. Strategies to evade antiviral responses. *Future Virol* **9**: 1061–1075. http://doi.org/10.2217/fvl.14.89

Grünvogel O, Esser-Nobis K, Reustle A, Schult P, Müller B, Metz P, Trippler M, Windisch MP, Frese M, Binder M, Fackler O, Barten-

schlager R, Ruggieri A, Lohmann V (2015) DDX60L is an interferon-stimulated gene product restricting hepatitis C virus replication in cell culture. *J Virol* **89**: 10548–10568. http://doi.org/10.1128/JVI.01297-15

Howell J, Angus P, Gow P, Visvanathan K (2013) Toll-like receptors in hepatitis C infection: Implications for pathogenesis and treatment. *J Gastroenterol Hepatol* **28**: 766–776. http://doi.org/10.1111/jgh.12170

Ishida H, Kato T, Takehana K, Tatsumi T, Hosui A, Nawa T, Kodama T, Shimizu S, Hikita H, Hiramatsu N, Kanto T, Hayashi N, Takehara T (2013) Valine, the branched-chain amino acid, suppresses hepatitis C virus RNA replication but promotes infectious particle formation. *Biochem Biophys Res Commun* **437**: 127–133. http://doi.org/10.1016/j.bbrc.2013.06.051

Jackowiak P, Figlerowicz M, Kurzynska-Kokorniak A, Figlerowicz M (2011) Mechanisms involved in the development of chronic hepatitis C as potential targets of antiviral therapy. *Curr Pharm Biotechnol* **12**: 1774–1780

Jackowiak P, Kowala-Piaskowska A, Figlerowicz M, Alejska M, Malinowska N, Figlerowicz M (2012) Evolution of hepatitis C virus hypervariable region 1 in chronically infected children. *Virus Res* **167**: 380–384. http://doi.org/10.1016/j.virusres.2012.05.005

Jackowiak P, Kuls K, Budzko L, Mania A, Figlerowicz M, Figlerowicz M (2014) Phylogeny and molecular evolution of the hepatitis C virus. *Infect Genet Evol* **21**: 67–82. http://doi.org/10.1016/j.meegid.2013.10.021

Jongeneel CV, Iseli C, Stevenson BJ, Riggins GJ, Lal A, Mackay A, Harris RA, O'Hare MJ, Neville AM, Simpson AJG, Strausberg RL (2003) Comprehensive sampling of gene expression in human cell lines with massively parallel signature sequencing. *Proc Natl Acad Sci U S A* **100**: 4702–4705. http://doi.org/10.1073/pnas.0831040100

Kurzynska-Kokorniak A, Koralewska N, Pokornowska M, Urbanowicz A, Tworak A, Mickiewicz A, Figlerowicz M (2015) The many faces of Dicer: the complexity of the mechanisms regulating Dicer gene expression and enzyme activities. *Nucleic Acids Res* **43**: 4365–4380. http://doi.org/10.1093/nar/gkv328

Larrat S, Vallet S, David-Tchouda S, Caporossi A, Margier J, Ramière C, Scholtes C, Haïm-Boukobza S, Roque-Afonso AM, Besse, André-Garnier, Mohamed, Halfon, Pivert A, LeGuillou-Guillemette H, Abravanel F, Guivarch M, Mackiewicz V, Lada O, Mourez T, Plantier JC, Baazia Y, Alain S, Hantz S, Thibault V, Gaudy-Graffin C, Bouvet D, Mirand A, Henquell C, Gozlan J, Lagathu G, Pronier C, Velay A, Schvoerer E, Trimoulet P, Fleury H, Bouvier-Alias M, Brochot E, Duverlie G, Maylin S, Gouriou, Pawlotsky JM, Morand P (2015) Naturally occurring resistance-associated variants of hepatitis C virus protease inhibitors in poor responders to pegylated interferon-ribavirin. *J Clin Microbiol* **53**: 2195–2202. http://doi.org/10.1128/JCM.03633-14

Ledergerber C, Dessimoz C (2011) Base-calling for next-generation sequencing platforms. *Brief Bioinform*: bbq077. http://doi.org/10.1093/bib/bbq077

Liang TJ, Ghany MG (2013) Current and future therapies for hepatitis C virus infection. *N Engl J Med* **368**: 1907–1917. http://doi.org/10.1056/NEJMra1213651

Lindenbach BD, Evans MJ, Syder AJ, Wölk B, Tellinghuisen TL, Liu CC, Maruyama T, Hynes RO, Burton DR, McKeating JA, Rice CM (2005) Complete replication of hepatitis C virus in cell culture. *Science* **309**: 623–626. http://doi.org/10.1126/science.1114016

Maillard P, Walic M, Meuleman P, Roohvand F, Huby T, Goff WL, Leroux-Roels G, Pécheur E-I, Budkowska A (2011) Lipoprotein lipase inhibits hepatitis C virus (HCV) infection by blocking virus cell entry. *PLOS ONE* **6**: e26637. http://doi.org/10.1371/journal.pone.0026637

Malumbres M, Harlow E, Hunt T, Hunter T, Lahti JM, Manning G, Morgan DO, Tsai L-H, Wolgemuth DJ (2009) Cyclin-dependent kinases: a family portrait. *Nat Cell Biol* **11**: 1275–1276. http://doi.org/10.1038/ncb1109-1275

Marinov GK, Williams BA, McCue K, Schroth GP, Gertz J, Myers RM, Wold BJ (2014) From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing. *Genome Res* **24**: 496–510. http://doi.org/10.1101/gr.161034.113

Marukian S, Andrus L, Sheahan TP, Jones CT, Charles ED, Ploss A, Rice CM, Dustin LB (2011) Hepatitis C virus induces interferon-λ and interferon-stimulated genes in primary liver cultures. *Hepatology* **54**: 1913–1923. http://doi.org/10.1002/hep.24580

Mee CJ, Farquhar MJ, Harris HJ, Ramma W, Ahmed A, Maurel P, Bicknell R, Balfe P, McKeating JA (2010) Hepatitis C virus infection reduces hepatocellular polarity in a vascular endothelial growth factor dependent manner. *Gastroenterology* **138**: 1134–1142. http://doi.org/10.1053/j.gastro.2009.11.047

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5**: 621–628. http://doi.org/10.1038/nmeth.1226

Oshiumi H, Miyashita M, Okamoto M, Morioka Y, Okabe M, Matsumoto M, Seya T (2015) DDX60 Is involved in RIG-I-dependent and independent antiviral responses, and its function is attenuated

by virus-induced EGFR activation. *Cell Rep* **11**: 1193–1207. http://doi.org/10.1016/j.celrep.2015.04.047

Papic N, Maxwell CI, Delker DA, Liu S, Heale BSE, Hagedorn CH (2012) RNA-sequencing analysis of 5' capped RNAs identifies many new differentially expressed genes in acute hepatitis C virus infection. *Viruses* **4**: 581–612. http://doi.org/10.3390/v4040581

Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A (2009) Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res* **37**: e123. http://doi.org/10.1093/nar/gkp596

Rehermann B (2016) HCV in 2015: Advances in hepatitis C research and treatment. *Nat Rev Gastroenterol Hepatol* **13**: 70–72. http://doi.org/10.1038/nrgastro.2015.227

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–140. http://doi.org/10.1093/bioinformatics/btp616

Roux PP, Blenis J (2004) ERK and p38 MAPK-activated protein kinases: a family of protein kinases with diverse biological functions. *Microbiol Mol Biol Rev* **68**: 320–344. http://doi.org/10.1128/MMBR.68.2.320-344.2004

Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, Nativ N, Bahir I, Doniger T, Krug H, Sirota-Madi A, Olender T, Golan Y, Stelzer G, Harel A, Lancet D (2010) GeneCards Version 3: the human gene integrator. *Database J Biol Databases Curation* **2010**: baq020. http://doi.org/10.1093/database/baq020

Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. *Bioinformatics*: btr026. http://doi.org/10.1093/bioinformatics/btr026

Schoggins JW, Wilson SJ, Panis M, Murphy MY, Jones CT, Bieniasz P, Rice CM (2011) A diverse array of gene products are effectors of the type I interferon antiviral response. *Nature* **472**: 481–485. http://doi.org/10.1038/nature09907

Simmonds P (2013) The origin of hepatitis C virus. *Curr Top Microbiol Immunol* **369**: 1–15. http://doi.org/10.1007/978-3-642-27340-7_1

Sultan M, Dökel S, Amstislavskiy V, Wuttig D, Sültmann H, Lehrach H, Yaspo M-L (2012) A simple strand-specific RNA-Seq library preparation protocol combining the Illumina TruSeq RNA and the dUTP methods. *Biochem Biophys Res Commun* **422**: 643–646. http://doi.org/10.1016/j.bbrc.2012.05.043

Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105–1111. http://doi.org/10.1093/bioinformatics/btp120

Wakita T, Pietschmann T, Kato T, Date T, Miyamoto M, Zhao Z, Murthy K, Habermann A, Kräusslich H-G, Mizokami M, Bartenschlager R, Liang TJ (2005) Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat Med* **11**: 791–796. http://doi.org/10.1038/nm1268

Walter P, Ron D (2011) The unfolded protein response: from stress pathway to homeostatic regulation. *Science* **334**: 1081–1086. http://doi.org/10.1126/science.1209038

Walters KA, Syder AJ, Lederer SL, Diamond DL, Paeper B, Rice CM, Katze MG (2009) Genomic analysis reveals a potential role for cell cycle perturbation in HCV-mediated apoptosis of cultured hepatocytes. *PLoS Pathog* **5**: e1000269. http://doi.org/10.1371/journal.ppat.1000269

Wendt A, Adhoute X, Castellani P, Oules V, Ansaldi C, Benali S, Bourlière M (2014) Chronic hepatitis C: future treatment. *Clin Pharmacol Adv Appl* **6**: 1–17. http://doi.org/10.2147/CPAA.S30338

Wilson GK, Stamataki Z, Wilson GK, Stamataki Z (2012) *In Vitro* systems for the study of hepatitis C virus infection. *Int J Hepatol* **2012**: e292591. http://doi.org/10.1155/2012/292591, 10.1155/2012/292591

Woodhouse SD, Narayan R, Latham S, Lee S, Antrobus R, Gangadharan B, Luo S, Schroth GP, Klenerman P, Zitzmann N (2010) Transcriptome sequencing, microarray, and proteomic analyses reveal cellular and metabolic impact of hepatitis C virus infection *in vitro*. *Hepatol Baltim Md* **52**: 443–453. http://doi.org/10.1002/hep.23733

Yuan Y-C, Wu X, Warden C (2013) Optimal Calculation of RNA-Seq Fold-Change Values. *Int J Comput Bioinforma Silico Model* **2**: 285–292. Retrieved from https://www.researchgate.net/publication/258629398_Optimal_Calculation_of_RNA-Seq_Fold-Change_Values

Zeisel MB, Lupberger J, Fofana I, Baumert TF (2013) Host-targeting agents for prevention and treatment of chronic hepatitis C – Perspectives and challenges. *J Hepatol* **58**: 375–384. http://doi.org/10.1016/j.jhep.2012.09.022

Zhang H (2016) Overview of Sequence Data Formats. *Methods Mol Biol Clifton NJ* **1418**: 3–17. http://doi.org/10.1007/978-1-4939-3578-9_1

Zhao W, He X, Hoadley KA, Parker JS, Hayes DN, Perou CM (2014) Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics* **15**: 419. http://doi.org/10.1186/1471-2164-15-419