

Differential gene expression analysis and interpretation

Understanding liver tumor progression using the QuantStudio™ 12K Flex Real-Time PCR System and QIAGEN® Ingenuity® iReport™

Introduction

Hepatocellular carcinoma (HCC) is the fifth most common form of cancer and is a complex and heterogeneous malignancy that arises in the context of progressive underlying liver dysfunction. Although the exact molecular mechanisms for hepatocarcinogenesis have still not been defined, WNT/ β -catenin and JAK/STAT are the two major oncogenic pathways in HCC, and angiogenesis has been found to play a major role in HCC pathogenesis.

HepG2 is an epithelial hepatocellular carcinoma cell line used for studying tumor progression and hepatocarcinogenesis. This application note presents the comparison of HepG2 RNA and normal hepatic tissue RNA, with data generated using the TaqMan® OpenArray® Human Cancer Panel and QuantStudio™ 12K Flex Real-Time PCR System. Data is analyzed using Ingenuity iReport to identify differentially expressed genes and explore the signaling



pathways and biological processes involved in HCC development.

Materials and methods

One-step qRT-PCR was performed on purified normal liver total RNA and HepG2 total RNA using the SuperScript® III Platinum® One-Step qRT-PCR Kit. The reaction mix containing samples was loaded onto 384-well plates, which were then loaded onto the TaqMan® OpenArray® Human Cancer Panel with the QuantStudio™ 12K Flex AccuFill™ System. The TaqMan® OpenArray® Human Cancer Panel targets 672 cancer-related genes and also includes 19 endogenous control genes that can be used for normalization.

The QuantStudio™ 12K Flex Real-Time PCR System combines flexible throughput capabilities with a streamlined workflow, taking you from targeted discovery through confirmation and screening, all on a single platform. Gene expression analysis of 672 cancer-related genes is accomplished using the TaqMan® OpenArray® Human Cancer Panel in combination with the QuantStudio™ 12K Flex system.

Ingenuity iReport is designed to drastically reduce time spent for gene expression data analysis. With an easy-to-use web interface that handles the complexity of data calculations from RNA-Seq, microarray, or qPCR gene expression data, Ingenuity iReport draws on the comprehensive and accurate Ingenuity® Knowledge Base to quickly highlight the key biological processes, pathways, and diseases linked to your experimental data.

The OpenArray® plate was then subjected to a one-step qRT-PCR protocol using the QuantStudio™ 12K Flex Real-Time PCR System [1]. Data was collected using the QuantStudio™ 12K Flex system analysis software, and data analysis was performed first using ExpressionSuite™ Software (provided with the system), with subsequent downstream analysis using Ingenuity iReport. Ingenuity iReport is optimized for real-time PCR and allows rapid upload and automatic mapping of several TaqMan® Gene Expression Assay formats including OpenArray® file types.

Results

We identified 608 differentially expressed genes between the normal liver and HepG2 samples, and 111 genes in the dataset have been associated with the disease state “liver tumor” in the literature. In addition, several notable interactions and pathways were also identified. A detailed review of the Ingenuity iReport analysis is presented in the following section.

Ingenuity iReport analysis overview

Ingenuity iReport can be used to filter, group, and visualize genes by function, biological process, role in pathway or disease, expression value, and other biological attributes. From a high-level summary of differentially expressed genes (DEGs) and their isoforms, it is possible to deep-dive and filter results with a range of interactive tools for clear and detailed understanding of the complex relationships of a focused set of data. All information is linked to the comprehensive bibliographies that support the report findings, providing a summary of the findings and direct links to each complete reference.

Summary Chapter

The Summary Chapter is a highly interactive tool to help direct attention towards an appropriate DEG (or set of DEGs) for detailed analysis. DEGs can be ranked and sorted by alphabetical order, fold change, number of isoforms, or their connectivity ranking using a unique score based on the number of experimentally demonstrated molecular interactions between genes or gene products, as determined by the data drawn from the Ingenuity Knowledge Base. More detailed information related to the context of a connectivity rating (including direct links to relevant publications) can be found through exploration using the Interactions Chapter.

Ingenuity iReport utilizes the Summary Chapter to provide an overview of results. DEGs are listed alongside a volcano plot showing data (with statistical significance) on fold changes between the two samples. Information on pathways, processes, and disease states that are identified to the top-scoring genes are also listed, along with related data linked to key words within the study hypothesis.

Figure 1 captures the key elements of the Summary Chapter for the HCC dataset, indicating 608 DEGs between HepG2 and normal tissue (Fold Change Cutoff = 1.5). Ingenuity iReport identifies keywords from the text provided in the Title and Hypothesis/Expectations sections of the report and highlights a subset of all significant Pathways, Biological Processes, and Diseases that are relevant to those keywords. You can search with additional keywords using the search field below.

Many of the top-scoring elements identified to these genes are clearly associated with cancer and cancer progression, as seen in the Top Results linked to Experiment Keywords, including hepatocellular carcinoma, liver tumor, apoptosis of hepatocytes, and other cancer-related or liver-related pathways. Also shown under Top Scoring Pathways, Processes, and Diseases are proliferation of cells, apoptosis, necrosis, and cell movement, as well as other processes such as cell cycle progression or diseases such as cancer and solid tumor or liver tumor. “Molecular mechanisms of cancer” canonical pathway is also

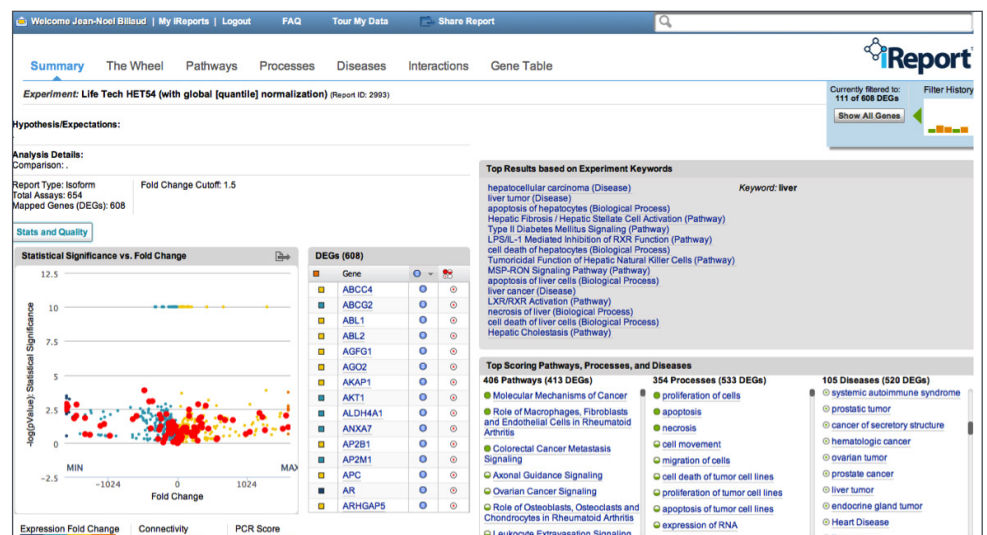


Figure 1. The Summary Chapter for analysis with Ingenuity iReport indicates 608 DEGs were found between HepG2 cells and normal cells.

indicated as well as “leukocytes extravasion signaling”.

“Liver tumor” disease was used to select further genes for drill down and detailed study. In this example, the mouse-over functionality in the Summary Chapter is highlighting the 111 genes associated with the disease state “liver tumor”, seen as red dots on the volcano plot (see Figure 1).

Diseases Chapter

Within the Diseases Chapter, experimental evidence is provided to link DEGs to specific disease states. This interface is fully interactive; mouse-over features allow for easy identification of genes involved in the list of diseases, and datasets can be subsequently filtered by specific selection of diseases or genes (Figure 2). References within the Ingenuity Knowledge Base are divided to represent genes that are known or exploratory biomarkers for the disease (Biomarker), genes that are targets for drugs that treat or are currently in clinical trials for the disease of interest (Drug), genes with altered expression in the disease state (Expression), and genes that are mutated in the disease state (Mutation).

This Diseases Chapter indicates that 111 genes in the dataset have been associated with “liver tumor”. HCC is a highly vascularized tumor, and angiogenesis is central in the initiation, growth, and metastasis of the liver carcinoma [2]. Angiogenesis in HCC is dependent on endothelial cell activation, proliferation, and migration, which occur in response to angiogenic cues (e.g., hypoxia and inflammation) and involves several molecular effectors such as growth factors, extracellular matrix proteins, and proteases. The *VEGFA* gene (encodes vascular

endothelial growth factor A) is known to be a key factor in liver tumor progression, and is highlighted in the Diseases Chapter with knowledge on biomarker, mutation, expression, and drug target information using a color code (see Figure 2 legend). *VEGFA* and its receptors are known

are represented based on their expression level as follows: inner circle blue = downregulated; inner circle yellow = upregulated.

The Interactions Chapter indicated to us that *VEGFA* and *IGF2* (insulin-like growth factor 2 or somatomedin A) are also

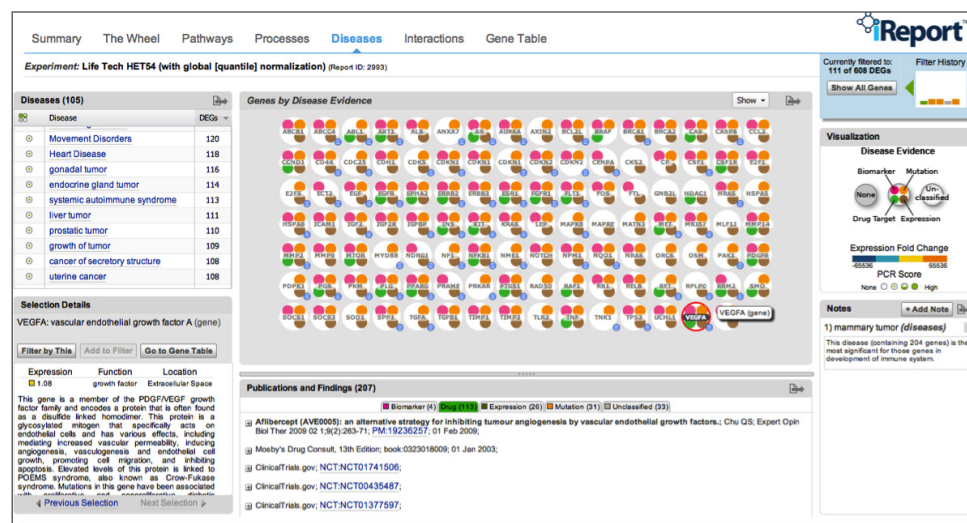


Figure 2. The Diseases Chapter displays the 111 genes that have been associated with liver tumors in the literature.

to be key mediators of HCC neovascularization [3].

Interactions Chapter

A key functionality of the Ingenuity iReport tool is demonstrated through the exploration of the Interactions Chapter. Here, DEGs for a process of interest can be displayed and then linked to other genes within the dataset based on information manually curated from the Ingenuity Knowledge Base. Connectivity includes causal interactions such as expression, activation, inhibition, or phosphorylation (where one gene product affects another), as well as binding interactions such as protein-protein binding or protein-DNA binding (physical interactions between genes and gene products). Clicking on a gene of interest will display links. Genes

connected (Figure 3). The *IGF2* gene, which is induced during hypoxia (a prevalent phenomenon in tumor establishment and progression), increases the expression of *VEGFA* and this contributes to increased angiogenesis. Many other interactions are displayed, including interactions between *VEGFA* and *ICAM1* (intercellular adhesion molecule 1, a molecule expressed on endothelial cells), and *VEGFA* and *CCND1*. Many of these interactions will influence the development of the liver carcinoma.

Looking back at the Disease Chapter (Figure 4), we observe that *VEGFA* has “Disease Evidence”, including being a known target for the drug aflibercept. Aflibercept is tested in different clinical trial phases for various oncological conditions.

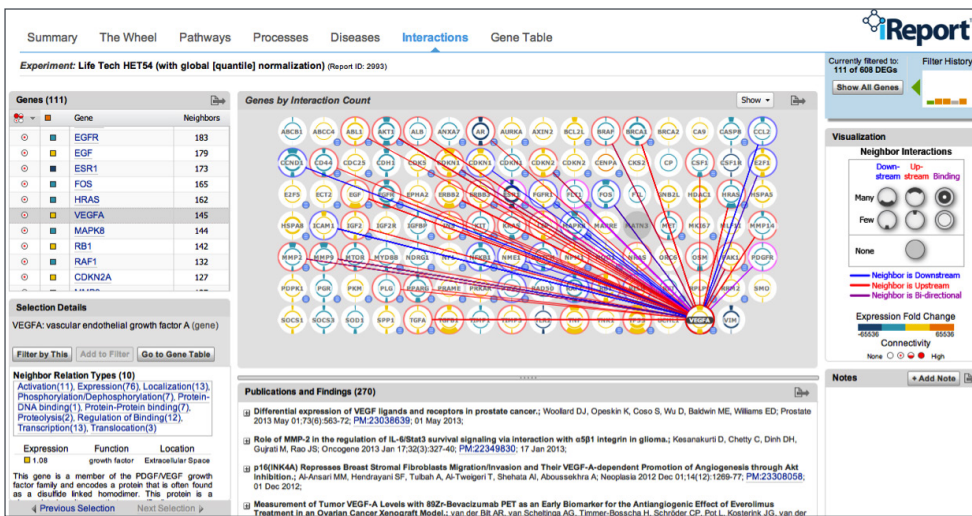


Figure 3. The Interactions Chapter shows the numerous genes that interact upstream of (red lines), downstream of (blue lines), or bind to (purple lines) *VEGFA*.

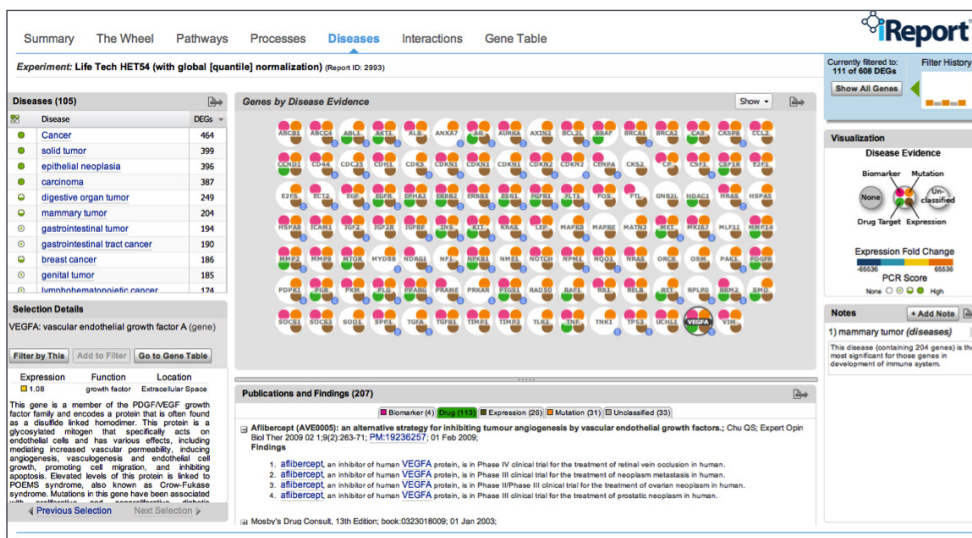


Figure 4. The Disease Chapter displays the publications and findings relevant to *VEGFA* (i.e., the known biomarker, drug, expression, or mutation findings in the Ingenuity® Knowledge Base).

Pathways Chapter

Here, pathways are identified that have a statistically significant representation by the genes in your dataset. A fully interactive, graphical view of each pathway is provided, color-coded by fold change and showing the molecular function each gene represents. As with all figures and tables in iReport, everything is fully exportable and all data can be filtered and bibliographies easily accessed.

The list of pathways shown in this chapter is displayed according

to the number of DEGs that are represented. Figure 5 displays the selection “Hepatic Fibrosis/ Hepatic Stellate Cell Activation” (39 DEGs); this pathway highlights the different molecular mechanisms that lead to fibrosis, tissue cirrhosis, and may eventually progress to hepatocellular carcinoma. Expression of many genes (e.g., *COL1A1*, *MMP2*) in this pathway are induced after engagement of the ligand to its receptor (e.g., EGF to EGFR or IGF1 to IGF1R). It is worth noting that the expression of many molecules in this pathway, such as

COL1A1 and *MMP2*, are known to participate in the fibrotic process and are downregulated in this dataset.

The Wheel Chapter

The Wheel allows you to identify, browse, or filter on any topic of interest and to rapidly visualize the key genes involved in that topic. Using the mouse-over functionality on the Wheel’s rim allows for a drill down of the biological topics in which the genes are significantly involved.

The view shown in Figure 6 is a selection of the 111 genes involved in liver tumors, clearly showing their respective expression levels (blue = downregulated; yellow = upregulated) in liver carcinoma vs. normal cells. In this view, the genes are grouped by molecular function and are also sized based on the number of diseases with which they have known association.

Visualization by disease annotation and molecular function evidence highlights several genes of interest in the promotion of tumors (Figure 7). These genes belong to the molecular group of ligand-dependent nuclear receptors and include notably: *AR* (androgen receptor), *ESR1* (estrogen receptor 1), and *PPARG* (peroxisome proliferator-activated receptor gamma). Recent studies suggest that the *AR* gene plays dual roles in promoting HCC initiation and in suppressing HCC metastasis [4], while other studies have shown that *AR* enhances HCC cell migration and invasion [5]. We observed that *AR* is underexpressed in HepG2 when compared to normal cells. *ESR1* was recently proposed as a tumor suppressor in HCC [6], and it is downregulated in the dataset. *PPARG* activation decreases liver cancer cell invasion by specific *PPARG*-dependent regulation of PAI-1 [7], and *PPARG* is also downregulated in this dataset.

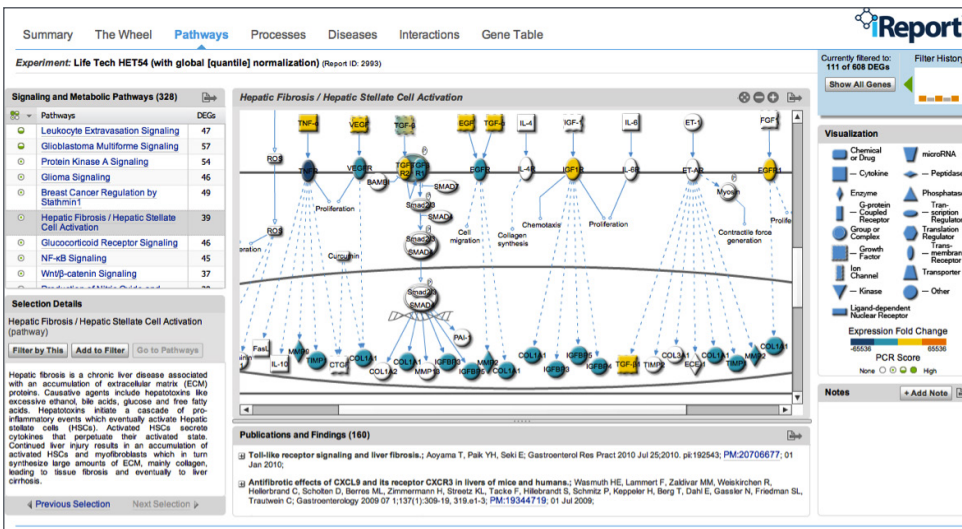


Figure 5. The Pathways Chapter highlights gene expression information using a color code (blue = downregulated; yellow = upregulated). The Hepatic Fibrosis/Hepatic Stellate Cell Activation pathway is shown here for the HepG2 vs. normal cells comparison.

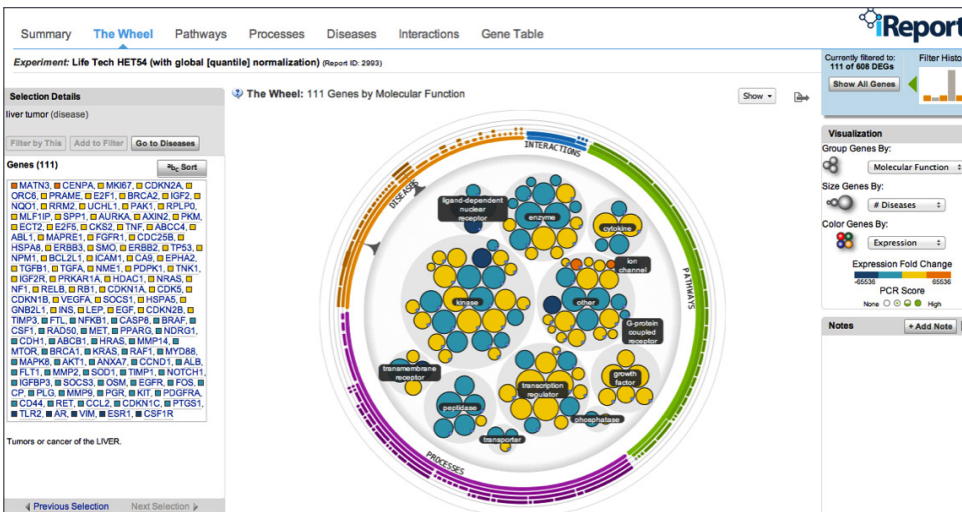


Figure 6. The Wheel Chapter displays the 111 genes with liver tumors and grouped by molecular function. The genes are sized by the number of associated diseases and color coded for expression fold change.

Gene Table

This view lists all DEGs and their transcripts and allows for clear visualization of the isoforms (using the RefSeq gene model) with detailed information on biological annotations found throughout the report and direct links to all relevant references and USCS Genome Browser records. However, the input data was performed here at the gene level (gene symbol) after performing a qPCR experiment and therefore the isoform view of each gene will not display the specific expression per isoform.

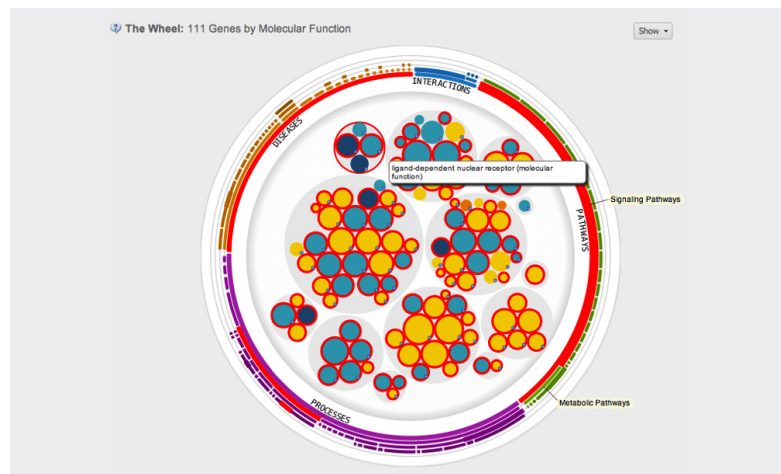


Figure 7. Partial view of the Wheel Chapter that highlights expression of the molecules in the “ligand-dependent nuclear receptor” group that are all downregulated.

The Gene Table view helps explore the parameters involved in tumor initiation or progression at the molecular level (Figure 8). For example, the Gene Table indicates that genes belonging to the WNT/beta-catenin canonical and the non-canonical pathways are differentially expressed in HepG2 vs. normal cells (e.g., *WNT1*, *WNT2*, *WNT3*, *WNT5A*, *FZD2*, *FZD5*, and *AXIN2*). This highlights the importance of understanding the role of these pathways in HCC development, as some recent studies indicate that both pathways may play a complementary role in HCC [8]. The WNT canonical pathway would be involved in tumor initiation and the WNT non-canonical pathway would be involved in tumor progression. For instance, *WNT5A* is a gene encoding a member of the WNT family that signals through the non-canonical WNT pathways. This protein is a ligand for the *FZD5* gene product (frizzled-5, a 7-transmembrane domain receptor protein) and the *ROR2* gene product (tyrosine kinase orphan receptor 2). This protein plays an essential role in regulating developmental pathways during embryogenesis. *WNT5A* has

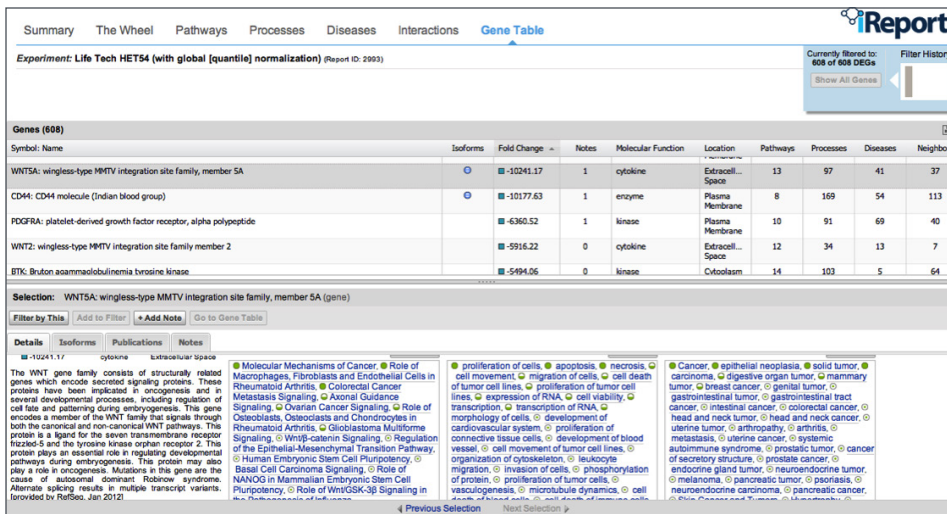


Figure 8. The Gene Table displays the list of the DEGs in the HepG2 vs. normal cells dataset by name, isoform, fold change, molecular function, and other parameters.

been shown to either promote or suppress tumors, depending on the cancer.

Processes Chapter

Using information from the range of published sources in the Ingenuity Knowledge Base, genes are linked to a list of biological processes, and clear visual representation identifying the dominant genes (those involved in the most number of processes) is provided. Dominant processes can be identified in terms of the number of genes involved as well as the p-value, which indicates the likelihood of a chance association. By clicking on a biological process of interest, it is easy to establish which genes are involved and to filter the dataset accordingly. This view can quickly highlight unfamiliar genes or gene activity that might warrant further investigation.

The Processes Chapter determined that 220 DEGs are involved in cell cycle progression, corresponding to a third of the entire DEGs list (Figure 9). This strongly indicates that the molecular machinery for

cellular division is heavily regulated in normal cells when compared to HepG2 cells.

Using the Wheel based on this filter (Figure 10), we notice a strong pool of transcription regulators (93 DEGs) that are involved in p53 signaling, cell cycle checkpoint, apoptosis, and inflammation-driven pathways such as NF- κ B signaling. All these intervening signaling pathways contribute to regulating cell processes, allowing cells to become invasive and motile later in the tumor cycle.

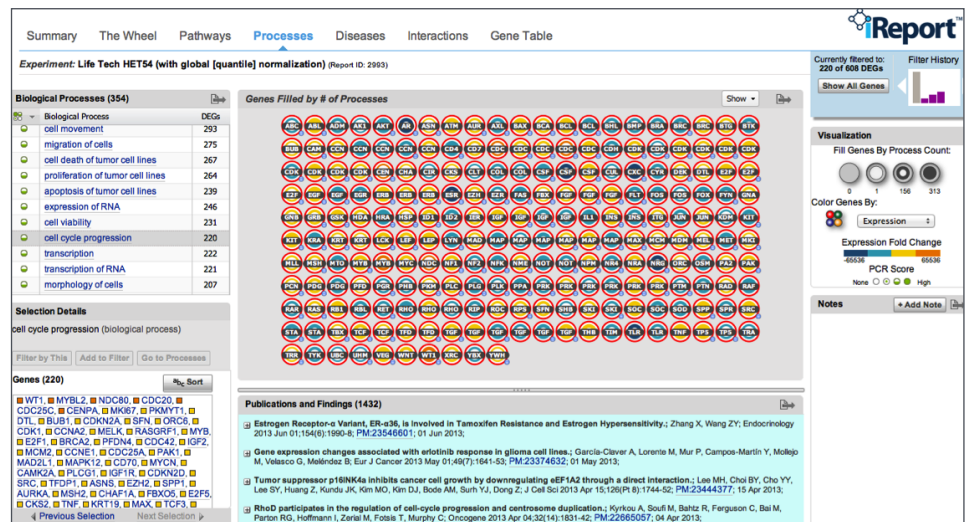


Figure 9. The Processes Chapter displays the 220 genes that participate in the biological process “cell cycle progression”.

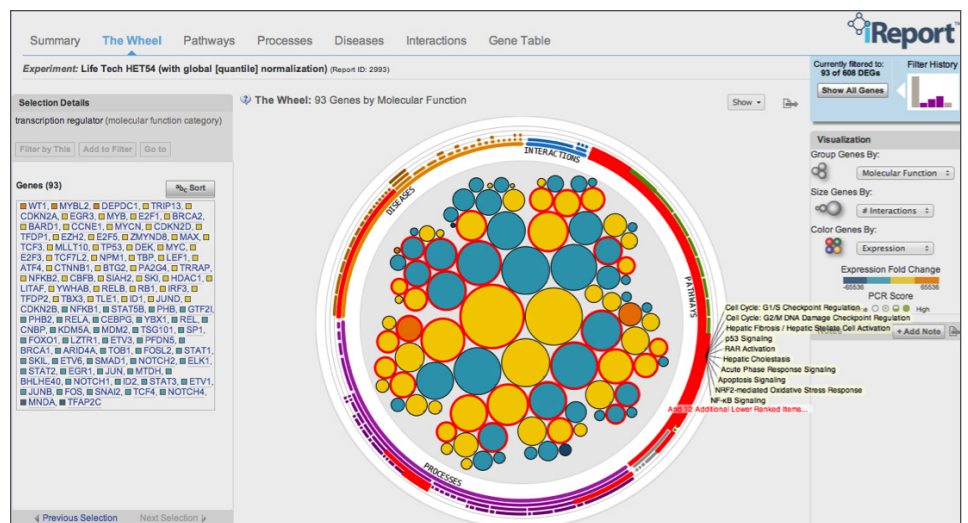


Figure 10. Partial view of the Wheel Chapter that highlights the 93 transcription regulators that are differentially expressed in the dataset. The relative expression level and the canonical pathway represented by some of these transcription regulators (red outline circles) are also shown.

Conclusions

Differential gene expression analysis is an important first step in understanding disease processes such as tumor initiation and progression. Biological interpretation of this data is necessary to ascertain the relevance of differentially expressed genes in a dataset. The powerful combination of the QuantStudio™ 12K Flex Real-Time PCR System and OpenArray® technology provides a fast, flexible platform for gene expression analysis, and Ingenuity iReport allows results to be generated and interpreted quickly and comprehensively. These advanced capabilities enable usable and relevant information to be discovered from screening and profiling applications.

References

1. Streamline high-throughput gene expression studies with SuperScript® III Platinum® One-Step Quantitative RT-PCR and OpenArray® Technology on the QuantStudio™ 12K Flex System. Application note available at lifetechnologies.com.
2. Finn RS. (2013) Emerging targeted strategies in advanced hepatocellular carcinoma. *Semin Liver Dis* 33:S11–S19.
3. Sanz-Cameno P, Trapero-Marugan M, Chaparro M et al. (2010) Angiogenesis: from chronic liver inflammation to hepatocellular carcinoma. *J Oncol* 2010:272170.
4. Ma WL, Hsu CL, Wu MH et al. (2008) Androgen receptor is a new potential therapeutic target for the treatment of hepatocellular carcinoma. *Gastroenterology* 135:947–955.
5. Ao J, Meng J, Zhu L et al. (2012) Activation of androgen receptor induces ID1 and promotes hepatocellular carcinoma cell migration and invasion. *Mol Oncol* 6:507–515.
6. Hishida M, Nomoto S, Inokawa Y et al. (2013) Estrogen receptor 1 gene as a tumor suppressor gene in hepatocellular carcinoma detected by triple-combination array analysis. *Int J Oncol* 43:88–94.
7. Pang X, Wei Y, Zhang Y et al. (2013) Peroxisome proliferator-activated receptor-γ activation inhibits hepatocellular carcinoma cell invasion by upregulating plasminogen activator inhibitor-1. *Cancer Sci* 104:672–680.
8. Yuzugullu H, Benhaj K, Ozturk N et al. (2009) Canonical Wnt signaling is antagonized by noncanonical Wnt5a in hepatocellular carcinoma cells. *Mol Cancer* 8:90.

Ordering Information

Product	Cat. No.
SuperScript® III Platinum® One-Step qRT-PCR Kit	11732020
TaqMan® OpenArray® Real-Time PCR Master Mix	4462164
Human Liver Total RNA	AM7960
TaqMan® OpenArray® Human Cancer Panel, QuantStudio™ 12K Flex	4475391
QuantStudio™ 12K Flex Real-Time PCR System with OpenArray® Block (with AccuFill™ System)	4471090
QuantStudio™ 12K Flex OpenArray® AccuFill™ System	4471021
OpenArray® AccuFill™ System Tips	4458107
QuantStudio™ 12K Flex OpenArray® Accessories Kit	4469576
QIAGEN® Ingenuity® iReport™	ingenuity.com/ireport

Find out more at lifetechnologies.com/ireport



A Thermo Fisher Scientific Brand

For Research Use Only. Not for use in diagnostic procedures. © 2014 Thermo Fisher Scientific Inc. All rights reserved. All trademarks are the property of Thermo Fisher Scientific and its subsidiaries unless otherwise specified. TaqMan is a registered trademark of Roche Molecular Systems, Inc., used under permission and license. QIAGEN and Ingenuity are registered trademarks, and iReport is a trademark of the QIAGEN Group. All other trademarks are the property of Thermo Fisher Scientific and its subsidiaries. C034951 0414