2019

Genomic Dysregulation by Overexpression of Transcription Factors

MacKenzie Scroggins

Genomic Dysregulation by Overexpression of Transcription Factors

BY

MacKenzie Scroggins

UNDERGRADUATE THESIS

Submitted in partial fulfillment of the requirement for obtaining

UNDERGRADUATE DEPARTMENTAL HONORS

Department of Biological Sciences along with the Honors College at
EASTERN ILLINOIS UNIVERSITY
Charleston, Illinois

2019

I hereby recommend this thesis to be accepted as fulfilling the thesis requirement for obtaining Undergraduate Departmental Honors

4/16/19
Date

THESIS ADVISOR

April 16, 2019
Date

HONORS COORDINATOR

4/16/19
Date

DEPARTMENT CHAIR

TABLE OF CONTENTS

ABSTRACT

As one of the major organs, the liver plays vital roles in the homeostasis of an individual. Being able to identify master regulator genes, genes whose product can affect the activation or deactivation of other genes, will enrich our understanding of hepatic function and liver disease. Previously, our lab used genome-wide microarray data to identify several transcription factors that may play a key role in liver expression. Here we ask whether overexpression of these transcription factors in a non-liver cell would result in liver phenotype activation or general disruption of gene regulation. To this end, we stably introduced seven transcription factors that we identified as being liver-specific into a non-liver cell line, followed by whole-genome expression analysis. As controls, we introduced a neo-plasmid to monitor general plasmid effects, as well as the HNF1α gene, previously shown to rescue liver-specific gene expression. Results show that introduction of the neo-plasmid alone resulted in 9 genes activated and 58 genes repressed by ≥2.5 fold. Overexpression of transcription factors resulted in between 320 and 664 genes activated, and 158 and 348 genes repressed by ≥2.5 fold even after controlling for the neo-plasmid data. In some cases, we observed much overlap. For example, 41 genes were activated by 4 of the 4 transcription factors (HHEX, CREG, CREB, and HNF6), with CREG and CREB sharing activation of an additional 94 genes. Focusing on only hepatic genes, each of the transcription factors activated between 13 and 35 liver-specific genes. However, there was little overlap between which genes that were activated in each case. These results suggest that while over-expression of transcription factors may activate tissue-specific genes, there is also a general dysregulation of gene expression that must be considered when interpreting data.

1. INTRODUCTION

Hepatocytes primarily regulate mRNA production of liver-specific genes through transcription factors (1-2). Several mechanisms of gene regulation have been proposed in the literature (3). HNF1α and HNF4α were demonstrated to regulate each other by binding to the other's promoter region (4). Both of these factors have been shown to partially rescue the hepatic phenotype, and HNF4 may prevent HNF1 from being deactivated (4-6). HNF6 has also been shown to promote HNF4α (5). Another transcription factor, XBP1, was also found to regulate lipid synthesis in the liver (6).

Throughout the literature, variant cells derived from rat hepatomas have been used to better understand liver metabolism and gene regulation (7-11). Some of these studies test for the activation of alpha-1-antritrypisin (SERPINA1), which is highly expressed in hepatic tissues.

Several transcription factors were previously identified as candidates of master regulation in rat liver tumor cells: HNF1α , HNF6, IGFBP1, HHEX, CREB, WNT4, and CREG (12). Our lab uses a system of comparing expression levels of liver-specific genes from rat liver tumor cells (Fg14), "variant" cells that have lost the hepatic phenotype (H11), and transfected variant cells (Figure 1).

Our lab demonstrated partial rescue of the hepatic phenotype in HNF1α transfected cells. HNF4α demonstrated a regulatory pathway with HNF1α via mutual binding to the other's promoter site. However, the use of RT-qPCR is limited to a small gene pool due to the impracticalities of applying this method to wide-scale analysis. Therefore, in order to understand how these transcription factors affect the cell as a whole, our lab sent RNA samples to the University of Illinois at their Urbana-Champaign location to perform a microarray analysis (12). By testing the effects of the five other candidate master regulators (CREG, CREB, IGFBP1,

HHEX, and HNF6) on the variant cell line and comparing to the hepatoma, non-transfected variant as well as the HNF1α transfected variant, we can determine whether the transcription factors are targeting tissue-specific genes or producing general disruption of the genome.

The WNT4 gene, another potential master regulator, assists in sexual development. Loss of function in this gene can cause SERKAL syndrome, which results in sex reversal (13). Current literature suggests that WNT4 has a significant role in the development of other tissues such as lung (14) and kidney (15). If WNT4 is lowly expressed in muscle tissue, there tends to be a lack of acetylcholine receptors, while overexpression causes an increase (16). Since WNT4 seems to influence various tissues, it is an interesting candidate for study on hepatic gene regulation.

Our lab's main objectives were to determine whether overexpression of transcription factors produces a liver-specific effect or a general disruption by comparing microarray data of transfected variant cells to hepatoma, non-transfected variants, and the HNF1α transfected variant line that demonstrated partial rescue of hepatic function. To do this, we compared the numbers and overlap of activated genes from each transfected cell line in the entire genome and in only genes upregulated by the hepatoma cell line. In addition, we also determined the tissue specificity of affected genes by each of the transcription factors and compared them to HNF1α.

2. MATERIALS & METHODS

2.1 Cell Lines and Culture Conditions

Cell lines used for the microarray analysis and transfections descended from H4IIEC3, a rat liver tumor cell line. The cell type used as the benchmark for liver phenotype, Fg14, is adenine phosphoribosyltransferase-positive and xanthine-guanine phosphoribosyltransferase-positive. Fg14 cell types were collected from APRT- and hypoxanthine-guanine phosphoribosyltransferase-negative (HPRT-) Fado-2 cells being transfected stably with Gpt and Aprt transgenes under the influence of human SERPINA1 gene promoter (-640 to -2 bp). Cells that appear to have lost liver phenotype, H11, are descendants of the Fg14 cell type. H11 were negatively selected against the Aprt and Gpt expression using 20 mg/mL of 2,5-diaminopurine (DAP) and 30 mg/mL 6-thioxanthine, respectively. H11 regain hepatic phenotype at a 10-3 frequency. Cells were grown in 1:1 Ham's F12/Dulbecco's modified Eagle's medium (FDV) with 5% fetal bovine serum (FBS) (Gibco) and 5 µg/100 mL penicillin-streptomycin (Gibo) unless otherwise stated. Cells were incubated at 37°C in a humid 5% CO_2 chamber.

2.2 Transfection of H11 with Transcription Factors

For each well, with a total of three wells (A, B, and C) of a 6-well plate, 1,000,000 H11 cells were plated in 5 mL of growing medium. Cells were counted using a hemocytometer. The next day cells were transfected using Lipofectamine LTX and Plus Reagent from Invitrogen (Thermo Fisher Scientific, Carlsbad, CA). The cells were then incubated for 48 hours.

Slides were prepared for the GFP-transfected H11P cells and observed under UV light. Both the control and the transfected H11 cells were split each into five T-25 flasks in 5 mL of

selective G418 media. These flasks were monitored over the course of the next few weeks to allow growth of the transfected clones.

2.3 RNA Isolation & cDNA Synthesis

Once the transfected cells had reached a high density on 100 mm plates, RNA was isolated using the RNAeasy Plus Mini Kit (Qiagen, Hilden, Germany). Using the High Capacity cDNA Reverse Transcription Kit by Appliedbiosystems (Thermo Fisher Scientific), the isolated RNA was used to synthesize cDNA. The cDNA samples were diluted by a factor of ten for RT-pCR.

2.4 Real-Time quantitative Polymerase Chain Reaction

Diluted cDNA samples were used to perform Real-Time quantitative Polymerase Chain Reactions (RT-qPCR). Reactions were performed using PowerUp SYBR Green Master Mix by Appliedbiosystems (Thermo Fisher Scientific, Austin, TX). Each tube contained 10 µl SYBR Green, 2 µL of diluted cDNA (replaced by nuclease-free water in the control), 1 µL each of forward and reverse primer (2 µL total), and 6 µL of nuclease-free water. Three tubes were used for each cDNA. Depending on the primers being tested, the annealing temperature was either 60°C or 64°C. GAPDH primers were used to correct the results of liver-specific genes.

2.5 Illumina Microarray Analysis

The Illumina microarray analysis was performed by the W.M. Keck Center for Comparative and Functional Genomics in the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign. The quality of RNA was tested using the Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Using the Illumina Total Prep RNA Amplification Kit (Ambion, Inc., Austin, TX), single-stranded cDNA was synthesized via reverse transcription with an oligo(dT) primer containing a T7 promoter and 300 ng of high-

quality RNA. The single-stranded cDNA was synthesized into double-stranded cDNA and purified. The cDNA was then used as a template for a 14-hour in vitro transcription (IVT) reaction. The cRNA was purified using the same kit. After testing for quality, a 16-hour hybridization to the RatRef-12 v1 Expression BeadChips (Illumina, Inc., San Diego, CA) was performed in the 58°C Hybridization Oven using 750 ng of cRNA. The next steps of washing, staining using streptavidin-Cy3 (GE Healthcare Bio-Sciences, Piscataway, NJ), and scanning were completed via instructions in the Illumina Whole-Genome Gene Expression Direct Hybridization Assay Guide (Part#11322355, Rev. A). For each of the two BeadChips, 12 arrays were performed, which were scanned using an Illumina BeadArray Reader. All 24 arrays were screened to detect signal artifacts, scratches, or debris. Images were analyzed using the GeneExpression Module (version 1.8.0) of the Illumina GenomeStudio software.

2.5 Analysis of Data

All data were controlled for plasmid effects using the neo plasmid. Genes were sorted based on their upregulation and downregulation in the transfected cell line when compared to non-transfected H11. Genes were also sorted based on which transcription factors affected them by at least 2.5-fold. Specificity of genes were pulled from Genecards.

3. RESULTS

To address the possibility of genomic instability that could be caused by overexpression of transcription factors, I analyzed whole genome expression profiles in the hepatoma variant cell line, H11. Cells were transfected with a series of genes encoding transcription factors identified in a genomic screen comparing expression profiles of hepatoma cells (Fg14) to hepatoma variant cells (H11 cells) (12). Genes transfected included HNF1, HNF6, CREB, CREG, IGFBP1 and HHEX. Each gene was in an expression vector that included the neo gene to allow for G418 selection of transfectants.

I compared the expression profile of the Fg14 hepatoma cells to that of H11 cells to identify the number genes differentially expressed. Fg14 cells expressed 554 genes at levels ≥2.5 fold higher than the H11 cells. Surprisingly, the Fg14 cells also expressed 357 genes at ≥2.5 fold lower levels that the H11 cells.

In order to control for effects of the plasmid alone, pKOneo, a plasmid expressing only the neo gene, was used. Results showed only a small number of genes (9) were activated ≥2.5 fold, although a much larger number (58) were repressed by ≥2.5 due introduction of the neo plasmid. In contrast, H11 cells transfected with the transcription factor plasmids showed ≥2.5 fold activation of between 75 and 664 genes (Figure 2A). IGFBP1 had the highest number of activated genes (664). Excluding neo-effects, only two transcription factors had activated less than 100 genes: HNF1α (76 genes) and HHEX (68 genes). Conversely, between 76 and 215 genes were repressed by at least 2.5 fold with CREG being the highest and HNF1 being the lowest (Figure 2B). In the case of HHEX, CREG, CREB, and HNF6, 41 genes were activated by all four transcription factors out of the 1099 total genes activated by all transcription factors. (Figure 3).

To determine the effects overexpression had on genes that were highly upregulated in hepatoma, I only considered genes expressed ≥2.5 fold in the Fg14 cells compared to the H11. With this limited pool of genes, 14 to 42 genes were upregulated by at least 2.5 fold (Figure 4A). HNF1α had the highest number, while HNF6 had the lowest. CREB, CREG, IGFBP1, and HHEX upregulated 35, 23, 29, and 27 genes. Between 4 to 25 genes were repressed by ≥2.5 fold. CREG repressed the most, while CREB and HNF1α both had the lowest (Figure 4B). Very little overlap was observed between HNF1, HHEX, CREB, and CREG, especially HNF1 (Figure 5).

Similarly, I analyzed possible patterns in genes that were downregulated in hepatoma and subsequently considered only genes that were downregulated ≥2.5 fold by Fg14 compared to H11 cells. Between 1 to 10 genes were upregulated ≥2.5 fold, HNF1α being the highest and HHEX being the lowest (Figure 6A). Between 19 to 72 were repressed ≥2.5 fold, with HNF6 being the lowest and CREG being the highest (Figure 6B).

To understand the activated gene profile due to overexpression of the transcription factors, I compared the tissue-specificity of genes activated in each of the transfected cell lines when genes were expressed ≥2.5 fold relative to the non-transfected H11 cell line. The results revealed that 17% of genes activated by ≥2.5 fold in H11- HNF1α were either liver specific or selective for liver (Figure 7A). For the other five cell lines, the highest was 6% in H11-HHEX (Figure 7B) and H11-CREB (Figure 7C). When only considering genes upregulated ≥2.5 fold by Fg14 compared to H11, 41% of genes activated ≥2.5 fold in H11- HNF1α were either liver specific or selective for liver (Figure 8A). HHEX had the next highest percentage (18%) (Figure 8B), while CREG had the lowest (4%) (Figure 8C).

I next asked whether another candidate gene, Wnt4, identified in our analysis could rescue hepatoma gene expression in the H11 cells. To this end, a Wnt4 expression vector was

introduced into the H11 cells, G418 selected, and pooled clones were expanded and analyzed using RT-qPCR. The H11-WNT4 pool cells were tested for overexpression of the transfected human WNT4 gene and found to express high levels compared to the H11 cells (results not shown). I then tested for effects on a panel of other liver genes. High expression of the liver-specific genes SERPINA1 (average cycle number 23.26), KNG1 (average cycle number 21.88), and FGB1 (average cycle number 22.99) was observed. However, the non-transfected H11 cells also showed the same high expression levels (average cycle numbers 20.67, 22.12, and 22.75). Importantly, both of these cell lines had higher expression levels than the hepatoma (average cycle numbers 23.26, 28.24, and 26.51) (Figure 9), suggesting that the H11 cells were contaminated.

3. DISCUSSION

Each of the transcription factors, except for HNF1α, had high numbers of genes being either activated or repressed by ≥2.5 fold. In both cases, HNF1 had numbers below 100. Similarly, HHEX also activated less than 100 genes. Notably, HHEX deactivated 108 genes, while the next highest number was 129 from CREB. When looking only at genes upregulated ≥2.5 fold by Fg14 in comparison to H11, HNF1α was the highest at 42 genes. Most of the other transcription factors, however, were comparable to HNF1α except for HNF6. Despite high numbers in genome-wide activation, CREB had low gene repression in this category. Genes that were repressed in Fg14 compared to H11 revealed that HNF1α activated the most genes compared to other transcription factors. This suggests that HNF1α is upregulating genes that are downregulated in hepatoma despite previously being found to positively regulate liver-specific genes. Furthermore, HNF1α did not have the highest number of genes repressed in this category. Instead, CREB repressed 44 more genes than HNF1α. In general, HNF1α behaved like a liver master regulator. Although not strongly, both HHEX and CREB seemed to show that they may regulate hepatic function.

Figure 2 demonstrates that these transcription factors do not share many genes. For example, when considering HHEX, HNF6, CREB, and CREG, there was a large portion of overlap, but many genes were only affected by one transcription factor. CREB activated 284 genes that none of the other transcription factors significantly affected. In addition, when only looking at genes that were also upregulated by Fg14, this pattern persisted. HNF1α, the gene that partially rescues the hepatic phenotype, overlaped very little with CREB, CREG, and HHEX. If these were liver master regulators, large overlap would have likely occurred. These genes would be affecting a similar pool and therefore share many genes. However, it would be unlikely that

any one gene would affect the entire hepatic phenotype. Instead, this responsibility would likely be divided among several genes such that, while one transcription factor might regulate one part of the hepatic phenotype, a different factor would control another part.

Genome-wide analysis showed that genes activated by HNF1α had large portions dedicated to liver. The other transcription factors did not appear to have this same trend, which suggests that they may not have large roles in liver function. When looking at genes upregulated in Fg14, HNF1α again showed large portions dedicated to liver. The other transcription factors did not show this strong activation. However, HHEX and CREB still had noticeable activation.

For the strange case of the WNT4 expression profiles, it is likely that before transfection these H11 were either contaminated by Fg14 or were actually Fg14 that were mislabeled as H11. Results suggest that WNT4 overexpression had no effect on these liver-specific genes in hepatoma cells.

4. CONCLUSION

As expected, overexpression of HNF1α did seem to follow trends that were expected of hepatic gene regulators. In some areas there were unexpected results, such as having a high activation of hepatoma-downregulated genes. This suggests that one should give these transcription factors leniency when analyzing genome-wide studies. In general, none of the other transcription factors showed a strong pattern that followed HNF1α. In addition, there was little overlap between HNF1α and the other transcription factors. Out of these five, HHEX and CREB appeared to be the strongest candidates for hepatic gene regulators. When using RT-qPCR to monitor expression of tissue-specific genes, a genome-wide analysis should also be considered. Many of these transcription factors seemed to be producing general effects rather than targeting liver-specific tissues. Further studies should exercise caution when studying expression levels of genes affected by overexpression of transcription factors, as these results may be caused by cell disruption rather than targeted regulation. In order to determine WNT4's role in hepatic function, further research must be conducted.

Figure 1. Rat liver hepatoma cells underwent a selection mechanism where non-hepatoma H11 cells were identified. H11 cells were transfected with transcription factors to determine if said transcription factor could rescue the hepatic phenotype.

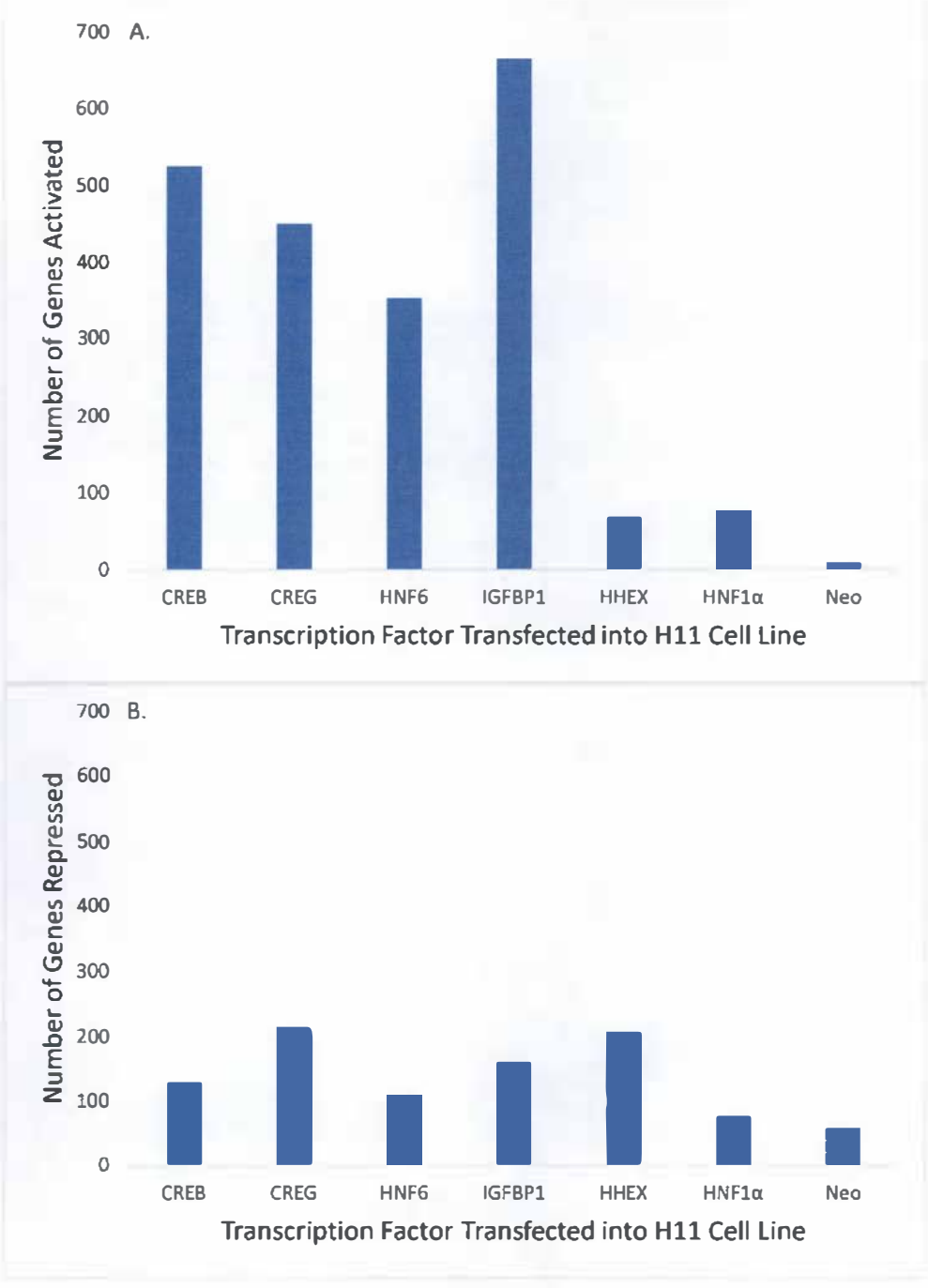

Figure 2. Genomic-wide upregulation of ≥2.5 fold (A) or repression of ≥2.5 fold (B) by each transcription factor transfected into the H11 cell line.

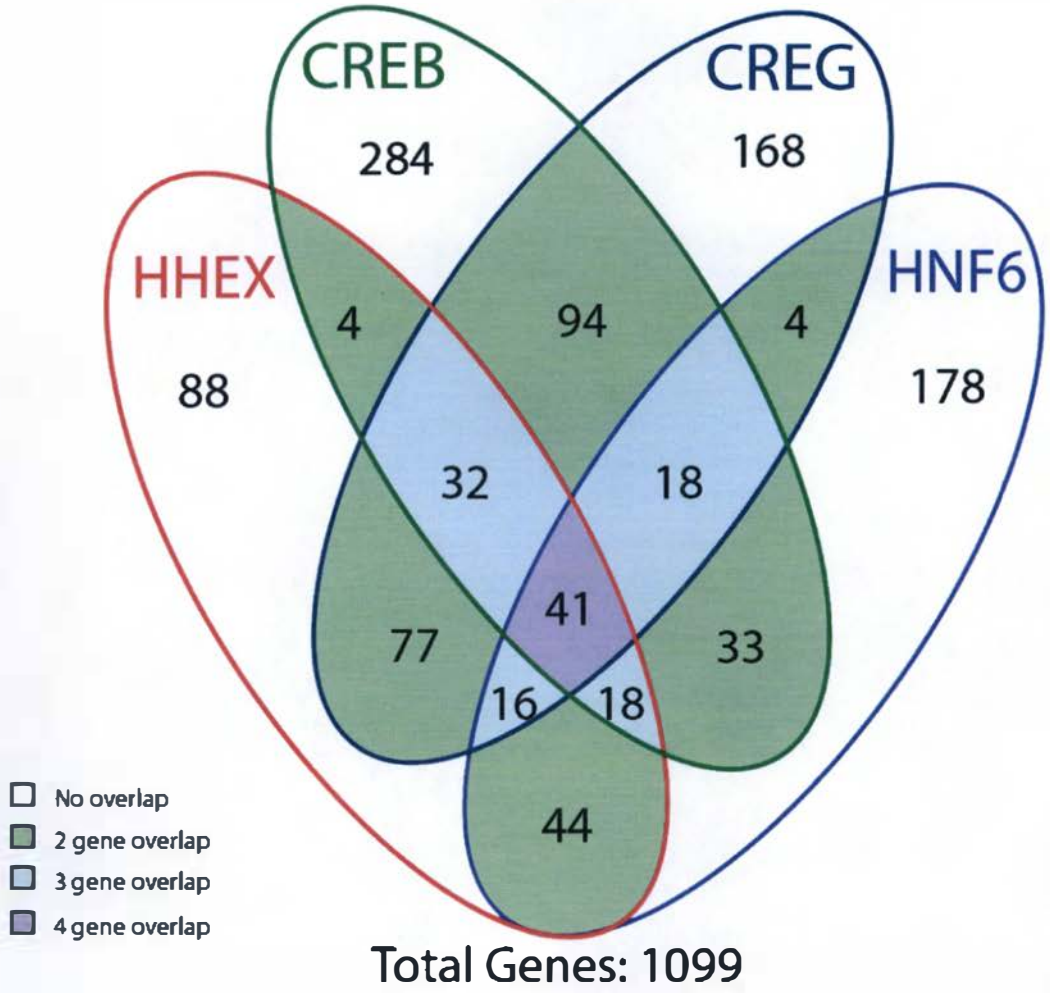

Figure 3. Venn diagram of shared genes activated ≥2.5 fold across the genome of H11 cells transfected with HHEX, CREB, CREG, or HNF6. Transfected cells were compared to non-transfected H11 to determine expression differences.

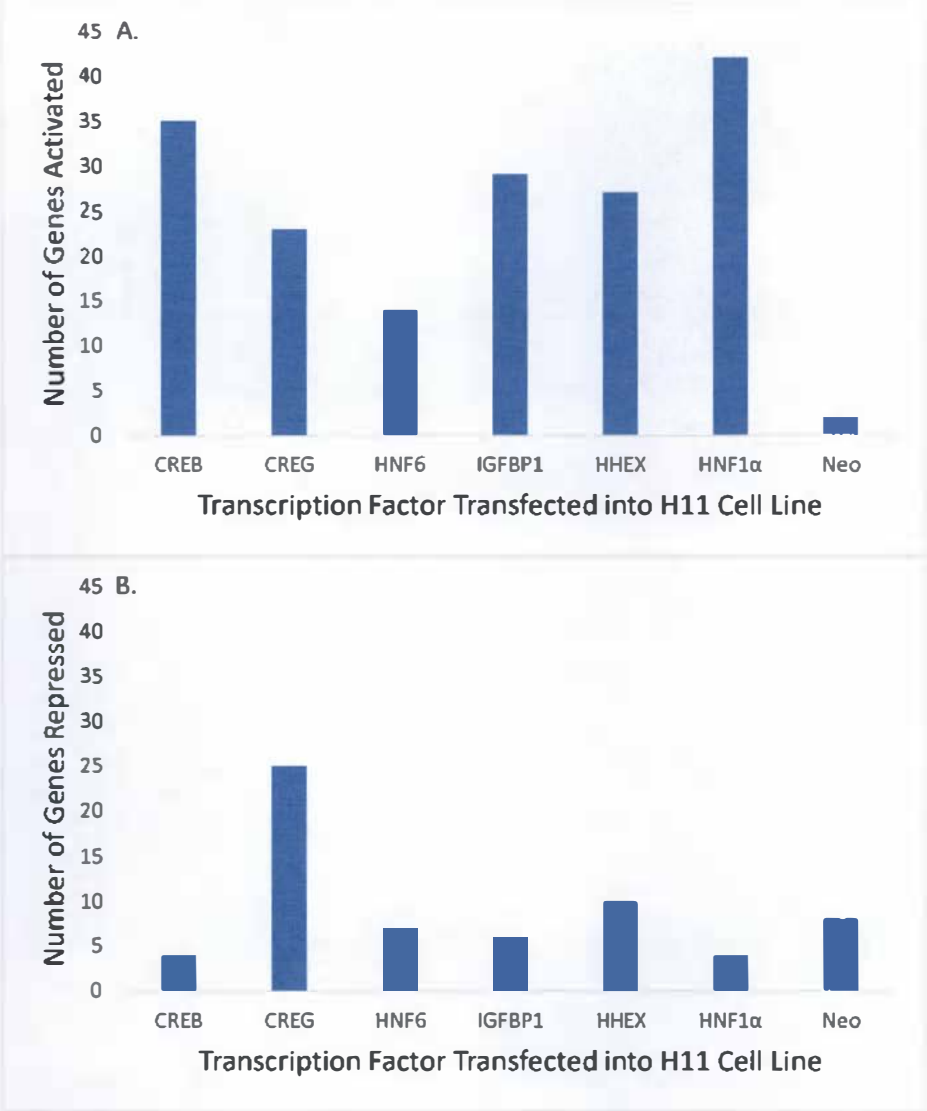

Figure 4. Number of genes that were activated (A) or repressed (B) by ≥2.5 fold with transcription factors transfected into the H11 cell line relative to non-transfected H11 cells. Only genes activated ≥2.5 fold in Fg14 compared to H11 cells were considered.

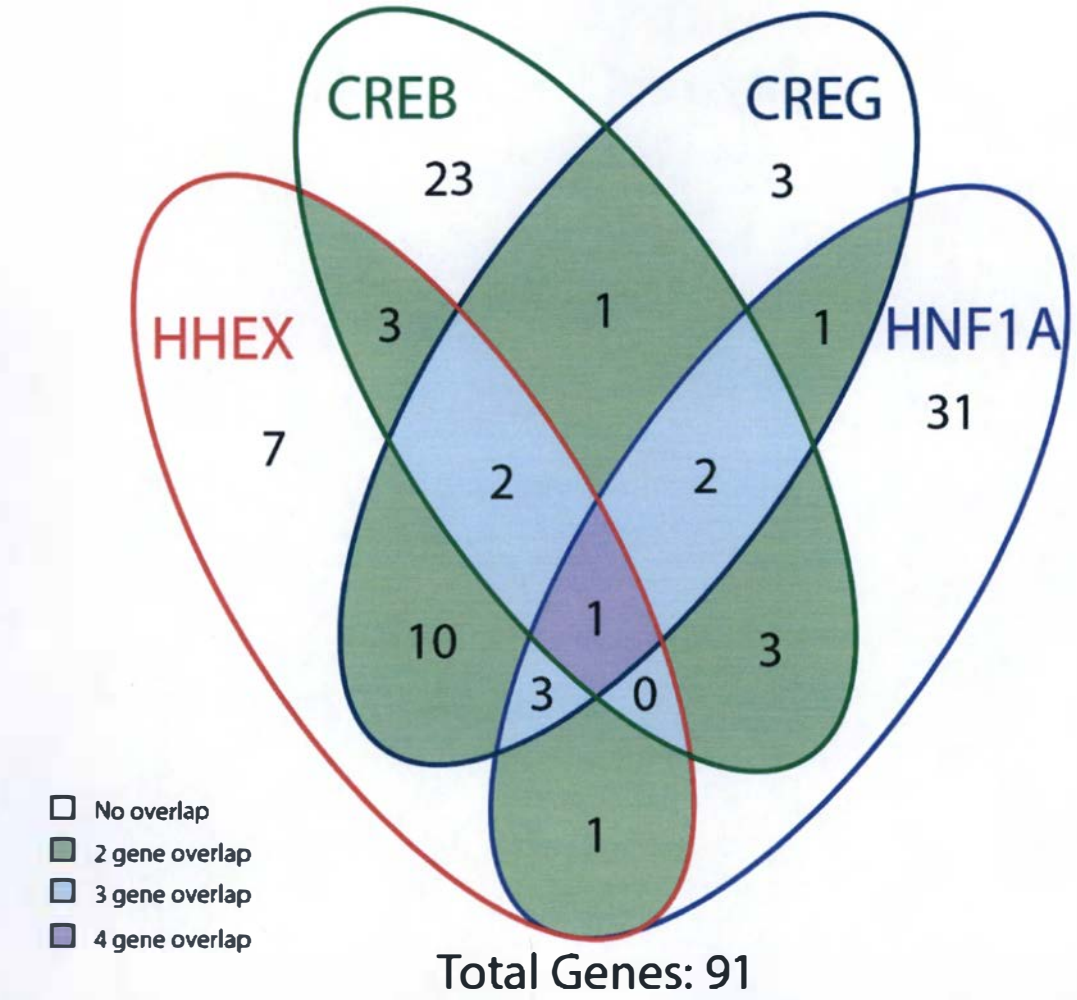

Figure 5. Figure 3. Venn diagram of shared genes activated ≥2.5 fold across the genome of H11 cells transfected with HHEX, CREB, CREG, or HNF1. Transfected cells were compared to non-transfected H11 to determine expression differences. Only genes upregulated in Fg14 cells by ≥2.5 fold compared to H11 were considered.

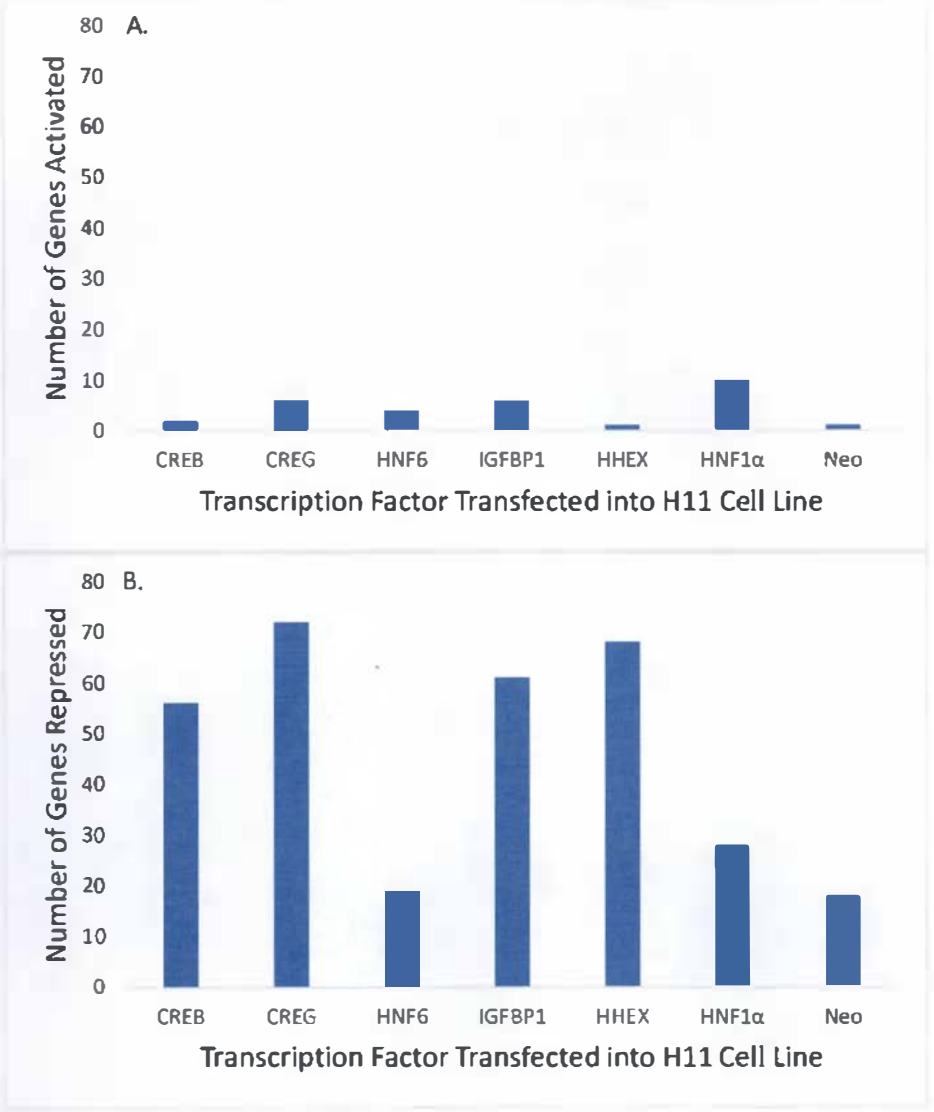

Figure 5. Number of genes that were activated (A) or repressed (B) by ≥2.5 fold with transcription factors transfected into the H11 cell line relative to non-transfected H11 cells. Only genes repressed 2.5 fold in Fg14 compared to H11 cells were considered.

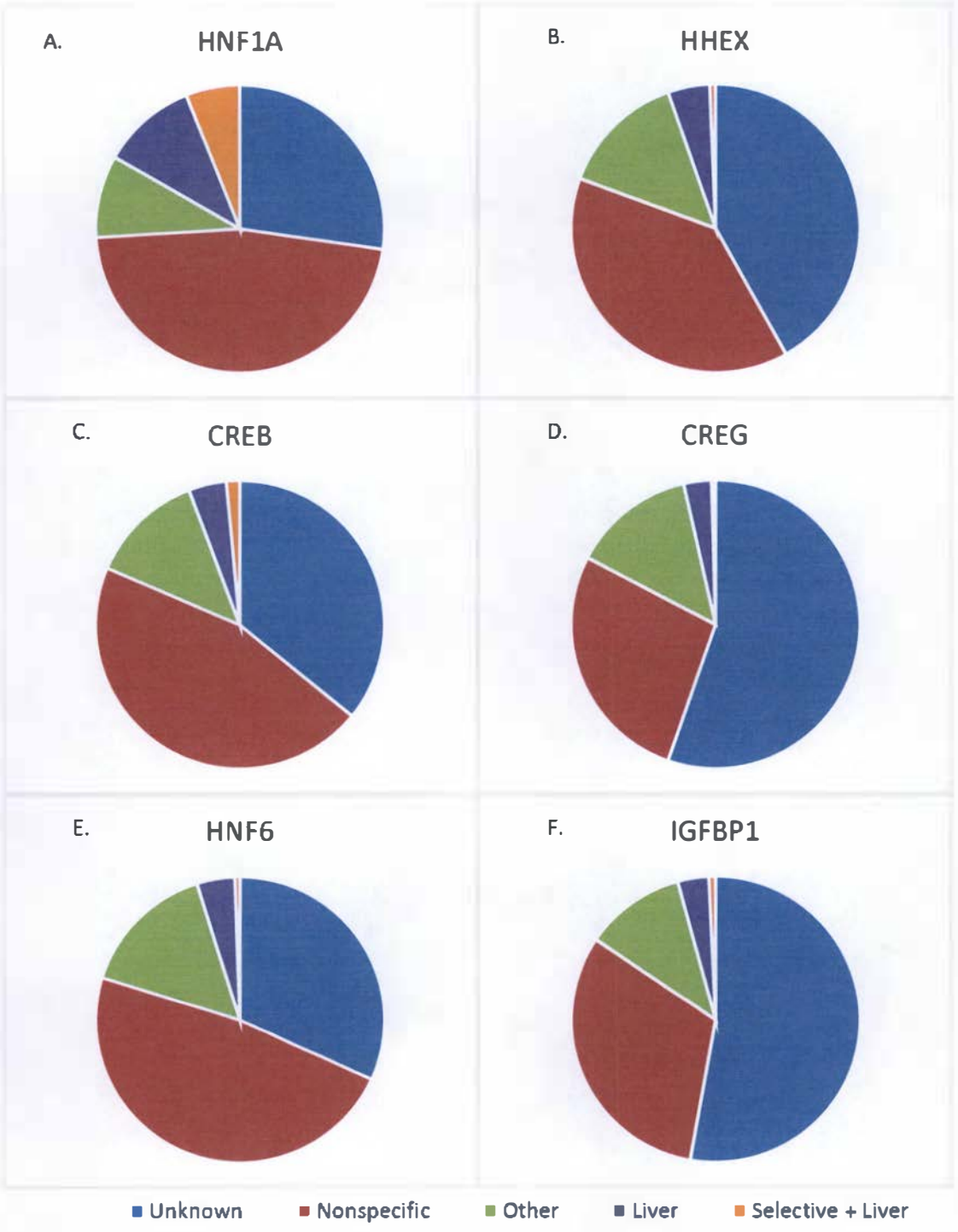

Figure 7. Genome-wide expression of genes activated ≥2.5 fold by HNF1α (A), HHEX (B), CREB (C), CREG (D), IGFBP1 (E), and HNF6 (F) when transfected into H11 cells compared to non-transfected H11 cells.

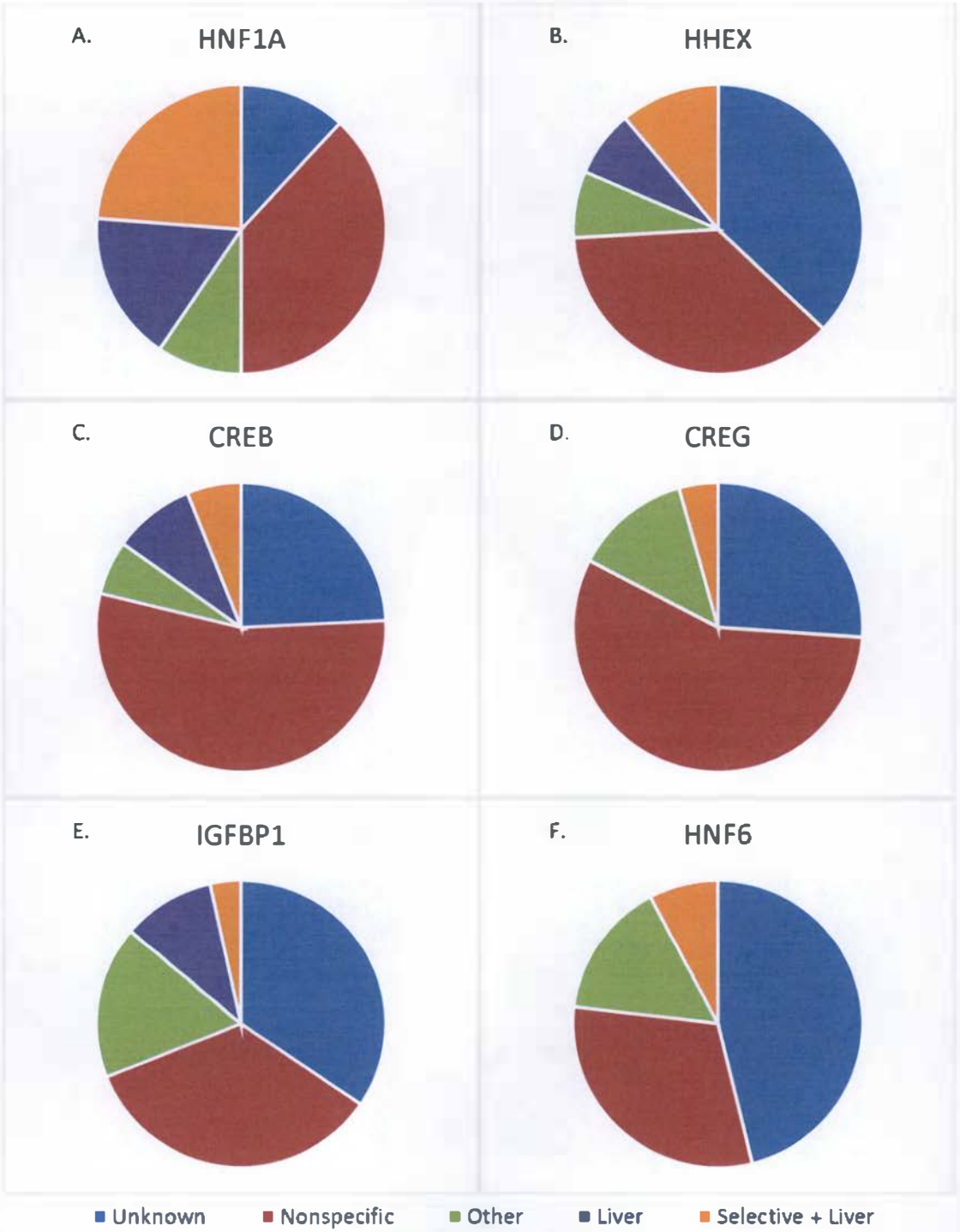

Figure 8. Tissue specificity of genes upregulated ≥2.5 fold in HNF1α (A), HHEX (B), CREB (C), CREG (D), IGFBP1 (E), and HNF6 (D) when transfected into H11 cells compared to non-transfected H11 cells. Only genes upregulated ≥2.5 fold in Fgl4 cells compared to H11 cells were considered.

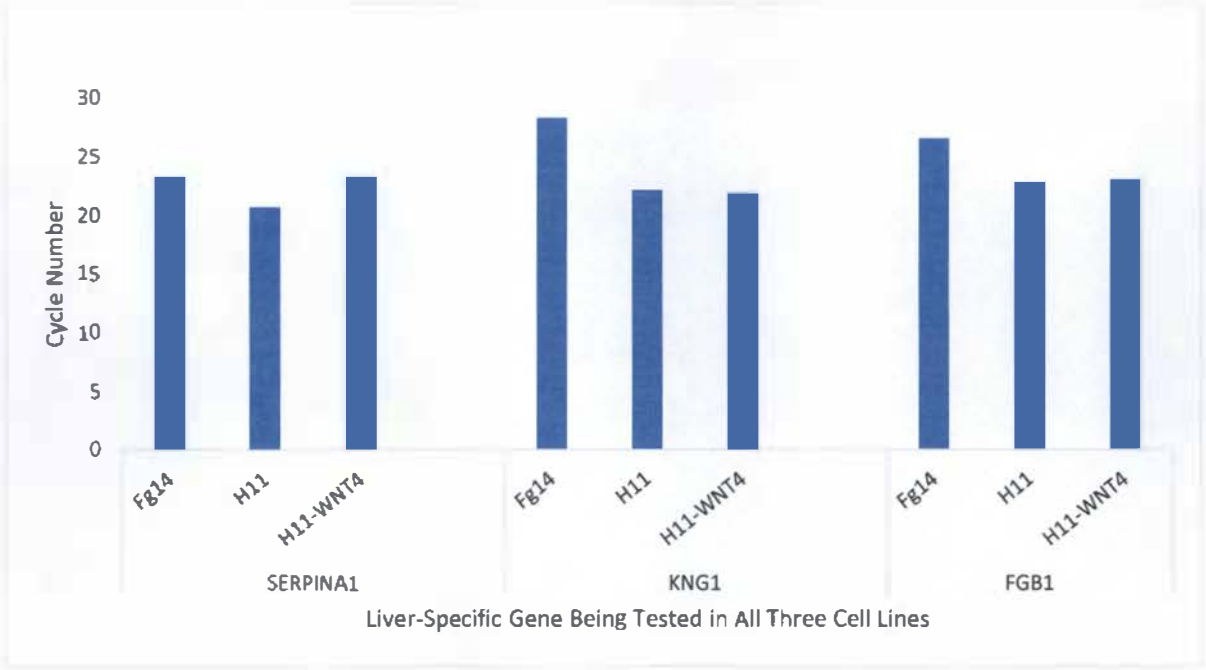

Figure 9. Fg14, H11, and H11-WNT4 Pool cycle numbers when tested for SERPINA1, KNG1, and FGB1. A lower cycle number is a higher expression level.

5. LITERATURE CITED

1. Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M., and Darnell, J. E. (1981). Transcriptional control in the production of liver-specific mRNAs. Cell, 23(3), 731-739.
2. Costa, R. H., Kalinichenko, V. V., Holterman, A. L., and Wang, X. (2003). Transcription factors in liver development, differentiation, and regeneration. Hepatology, 38(6), 1331-1347.
3. Odom, D. T, Zizlsperger, N., Gordon, D. B., Bell, G. W., Rinaldi, N. J., et al. (2004). Control of Pancreas and Liver Gene Expression by HNF Transcription Factors. Science, 303(5662), 1378-1381.
4. Bulla, G. A., Givens, E., Brown, S., Oladiran, B., and Kraus, D. (2000). A common regulatory locus affects both HNF4α /HNF1α pathway activation and sensitivity to LPS-mediated apoptosis in rat hepatoma cells. Journal of Cell Science, 114, 1205-1212.
5. Bulla, G. A. (1997). Hepatocyte nuclear factor-4 prevents silencing of hepatocyte nuclear factor-1 expression in hepatoma x fibroblast cell hybrids. Nucleic Acids Research, 25(12), 2501-2508.
6. Lee, A., Scapa, E. F., Cohen, D. E., and Climcher, L. H. (2008). Regulation of Hepatic Lipogenesis by the Transcription Factor XBP1. Science, 320(5882), 1492-1496.
7. Duncan, S. A., Navas, M. A., Dufort, D., Rossant, J., and Stoffel, M. (1998). Regulation of a Transcription Factor Network Required for Differentiation and Metabolism. Science, 281(5377), 692-695.
8. Odom, D. T., Dowell, R. D., Jacobsen, E. S., Nekludova, L., Rolfe, P. A., et al. (2006). Core transcriptional regulatory circuitry in human hepatocytes. Molecular Systems Biology, 2, 1-5.

9. Deschatrette, J., Moore, E. M., Dubois, M., and Weiss, M. C. (1980). Dedifferentiated variants of a rat hepatoma: Reversion analysis. Cell, 19(4), 1043-1051.
10. Bulla, G. A. (1997). Selective Loss of the Hepatic Phenoytpe Due to the Absence of a Transcriptional Activation Pathway. Somatic Cell and Molecular Genetics, 23(3), 185-201.
11. Bulla, G. A., Luong, Q., Shretha, S., Reeb, S., and Hickman, S. (2010). Genome-wide analysis of hepatic gene silencing in mammalian cell hybrids. Genomics, 96(6), 323-332.
12. Bulla, G. A., Aylmer, C. M., Dust, A. L., Kurewich, J. L., Mire, L. K., and Estanda, A. B. (2012). Genome-wide analysis of hepatic gene silencing in hepatoma cell variants. Genomics, 100(3), 176-183.
13. Mandel, H., Shemer, R., Borochowitz, Z. U., Okopnik, M., Knopf, C., et al. (2008). SERKAL Syndrome: An Autosomal-Recessive Disorder Caused by a Loss-of-Function Mutation in *WNT4*. American Journal of Human Genetics, 82(1), 39-47.
14. Caprioli, A., Villasenor, A., Sylie, L. A., Braitsch, C., Marty-Santos, L., et al. (2015). Wnt4 is essential to normal mammalian lung development. Developmental Biology, 406(2), 222-234.
15. Itäranta, P., Chi, L, Seppänen, T., Niku, M., Tuukkanen, J., et al. (2006). Wnt-4 signaling is involved in the control of smooth muscle fate via Bmp-4 in the medullary stroma of the developing kidney. Developmental Biology, 293(2), 473-483.
16. Strochlic, L., Falk, J., Goillot, E, Sigoillot, S., Bourgeois, F., et al. (2012). Wnt4 Participates in the Formation of Vertebrae Neuromuscular Junction. PLoS One, 7(1)