Journal of
Emerging Investigators

Activated NF-κB Pathway in an *Irf6*-Deficient Mouse Model for Van der Woude Syndrome

Sanika S. Kulkarni[1], Jungwoo Chang[2], and Inhan Lee[3]

[1]Pioneer High School, Ann Arbor, Michigan

[2]Huron High School, Ann Arbor, Michigan

[3]miRcore, Ann Arbor, Michigan

Summary

Van der Woude syndrome is the fourth most common birth defect in the United States and a major cause of morbidity in countries where corrective surgery is not readily available. This syndrome, generally expressed in the form of cleft lip and palate, is caused by mutations in the *Irf6* gene. The goal of this research is to identify gene networks or pathways that are differentially regulated due to the *Irf6* mutation. To achieve this goal, we analyzed microarray-based gene expression data from *Irf6*-deficient and wild-type mice from the Gene Expression Omnibus database to obtain a systems-level view of the genes involved in Van der Woude syndrome. Because the NF-κB pathway contains gene-gene interactions with multiple other genes from the Irf family, our results suggest that it has gene interactions with *Irf6* as well. Since the up-regulated genes in *Irf6*-deficient mice were related to the NF-κB pathway, they may function as compensatory for *Irf6*.

Received: October 31, 2014; **Accepted:** July 19, 2015; **Published:** September 11, 2015

Introduction

Although Van der Woude syndrome is a relatively unknown disorder, it is the fourth most common birth defect in the United States. Van der Woude is better known as the main cause of cleft lip and palate, a set of facial malformations that forms at a very early stage of fetal development. Cleft lip is the physical split of the two sides of upper lip, whereas cleft palate is a split at the roof of the mouth, involving the soft or hard palate. Van der Woude syndrome is the result of a mutation in the gene *Irf6* (Interferon regulatory factor 6), a DNA-binding transcriptional activator that provides instructions for making a protein that binds to specific regions of DNA and helps control the activity of particular genes. As the IRF6 protein is most active in cells that give rise to tissues in the head and face, a shortage of the IRF6 protein affects the development and maturation of tissues in the skull and face.

To understand global gene expression changes caused by an *Irf6* gene deficiency leading to various Van der Woude phenotypes, we analyzed mouse gene expression data found in conjunction with a research paper by Ingraham and colleagues (1). The data set consists of gene expression values from six mice: three *Irf6*-deficient subjects (using gene trapping) and three wild-type subjects (expressing normal IRF6 protein) (**Figure 1**). Samples of skin were taken from each subject for RNA extraction. Through the comparison of phenotypic differences between knockout and wild-type subjects, as well as gene expression analysis using a microarray platform and various immunohistological analyses, Ingraham and colleagues concluded that *Irf6* plays a significant role in epidermal development, observing that *Irf6*-deficient mice were not fully developed, as compared to wild-type mice at the same stage of growth. *Irf6*-deficient mice had a general appearance of stunted growth, with shorter limbs and digits. The most significant histological feature of the *Irf6*-null mice was the lack of a normal stratified epidermis. The *Irf6*-null epidermis was thicker than that of the wild-type mice. Considering these features in light of gene expression analyses, they concluded that the primary defect resulted from suprabasal keratinocytes failing to stop proliferating and to terminally differentiate, attributing to *Irf6* a new role in regulating proliferation and differentiation of keratinocytes.

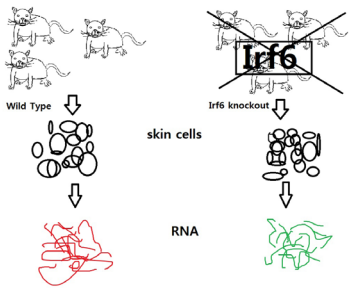

Figure 1: *Irf6* Knockout Mouse. Mouse model that produced data used in this paper. Red – wild-type mouse RNA. There are no visual abnormalities in this embryo. Green - Irf6 gene trap mouse (*Irf6*-deficient mouse) RNA. Prominent characteristics include lack of visible digits and external ears, shorter limbs, lack of distinct hindlimbs, epidermal adhesion in oral cavity, and cleft lip in the secondary palate.

While Ingraham and colleagues focused on the gene expression changes in markers representing epidermal differentiation in skin, this study used a systems-biology approach to understand global gene expression changes due to *Irf6* deficiency. We hypothesized that IRF6 has multiple functions that can affect many cellular molecular pathways and that such affected pathways may contribute to the various phenotypes of the *Irf6* mutation. To identify the global molecular pathways related to *Irf6*, we analyzed the gene expression data from mouse experiments by Ingraham and colleagues, and identified genes that were differentially expressed when *Irf6* is deficient. Comparing the deficient and wild-type mice data, we found that the Gene Ontology (2) term organ morphogenesis was the second most enriched term, after the more general term cellular process, for 13 up-regulated genes in *Irf6*-deficient mice. We found that the Toll-like-receptor-activated NF-κB pathway was strongly related to these 13 up-regulated genes as well. Since other genes in the Irf family have been identified in the Toll-like receptor pathway, it is possible that the NF-κB pathway is activated to compensate for the *Irf6*-deficiency and that *Irf6* may be a downstream link of NF-κB like *Irf3*, *Irf5*, and *Irf7*.

Results

Differently Expressed Genes

After downloading the microarray expression data from GEO for the mouse skins profiled using the Affymetrix Mouse Genome 430 2.0 Array (**Figure 1**), we analyzed the expression data using the GenePattern (3) programs. A Comparative Marker Selection module graph from GenePattern shows the up-regulated genes in *Irf6*-deficient mice, compared to the wild-type mice

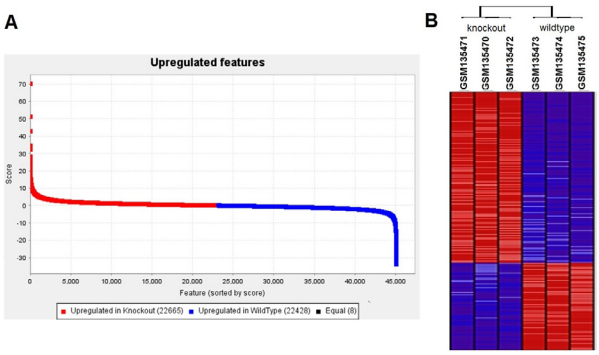

Figure 2: Gene Expression Differences between Irf6-Deficient and Wild-type Mice. A) Red: up-regulated genes in *Irf6*-deficient mice. Blue: up-regulated genes in wild-type mice (down-regulated genes in the deficient mice). A high score (*y*-axis) indicates correlation with the knockout mice (up-regulated in deficient mice) and a low score indicates negative correlation. B) Heat map of gene expression values of three *Irf6*-deficient mice and three wild-type mice. Red: highly expressed, blue: lowly expressed genes.

in red and down-regulated genes in blue labeled "up-regulated in wild-type mice" (**Figure 2A**). The probe-ID of each gene was used to define its feature. These probe-IDs were then sorted by score. The score column shows the value of the metric used to correlate gene expression and phenotype. A high score indicates a correlation with knockout mice and a low score indicates correlation with wild-type mice. As seen in the graph, the up-regulated genes in *Irf6*-deficient mice reached a maximum score of 70, whereas down-regulated genes in *Irf6*-deficient mice had a minimum score of -30. This indicates that there were more up-regulated genes highly correlated with the deficient mice than down-regulated genes correlated with wild-type mice.

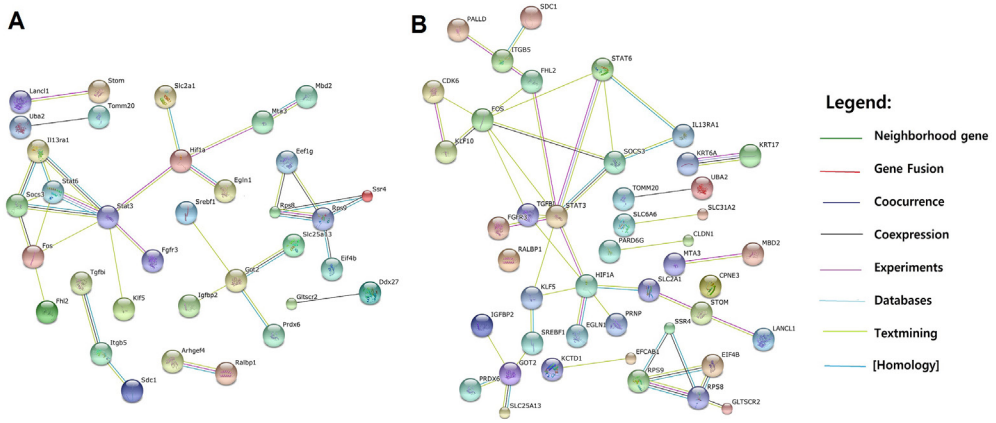

Figure 3: STRING Database Analysis. Both figures used the list of 142 up-regulated genes with a student's t-test *p*-value < 0.001. A) Visual representation of gene interactions among 139 up-regulated mouse genes. B) The interactions of 125 up-regulated human genes. The size of each node reflects structural information associated with the protein. The lines in this graph connect genes of interest to predicted functional partners through direct or indirect interactions. Their thickness is correlated to the strength of the predicted gene-gene interaction. If there is a higher statistical likelihood of a specific gene-gene interaction, the line between their respective circles is thicker. The color of each line represents the source of evidence suggesting the interaction shown.

We constructed an expression heat map of genes with a student's t-test *p*-value < 0.001 after processing through the Hierarchical Clustering module using the Pearson correlation distance (**Figure 2B**). We chose this particular *p*-value because it yielded a sufficient amount of genes for later tests. The two groups were clearly differentiable, as seen in the distinct separation of blue and red in the image. It is also evident that the *Irf6*-deficient mice had many more genes that were up-regulated than down-regulated.

We identified all genes with a *p*-value < 0.001 and a fold change (average gene expression value of knockout divided by average gene expression of wild type) >2 or <0.5 (representing up-regulated and down-regulated genes, respectively) (**Table 1**). This table lists the names, *p*-values, and fold changes associated with the genes shown in the heat map (**Figure 2B**). There were 39 more up-regulated knockout mouse genes than down-regulated genes in the dataset, indicating once again that there were many more up-regulated genes in *Irf6*-deficient mice than up-regulated genes in wild-type mice.

Gene-Gene Interactions

We investigated the connections among the up-regulated genes identified in the previous section. A dataset of 142 up-regulated genes was entered into the STRING database site (4) to form a visual representation of gene-gene interactions (**Figure 3**). All genes inputted into the STRING program had a *p*-value less than 0.001, and all connected genes were visualized using a STRING default parameter with a medium confidence score of 0.4 (**Figure 3A**). To compare mouse and human gene networks, we used the same 127 genes, but selected "human" as the organism type in order to get gene-gene interactions orthologous to mouse. Interestingly, the gene network of human orthologs had more network connections (**Figure 3B**). This may be because more studies have been done with human genes.

Stat3 was a strong central gene in both mice and humans, interacting with numerous genes related to various molecular pathways including hypoxia (*Hif1a*), growth factor receptors (*Fgfr3*), the maintenance of embryonic stem cells (*Klf5*), and the cytokine response (*Stat6*, *Il13ra1*). Humans had homologous genes and connections similar to those in mice, such as the FGFR3 and STAT6 networks, and both had the NF-κB pathway in common, as described in the pathway analysis results.

Enriched GO Biological Processes

In addition to direct gene-gene interactions, we wondered if there were over-represented features among the 142 up-regulated genes. We used Biological Process in the Gene Ontology (GO) database to obtain enriched features. Twenty-four GO terms were found using the data output of BiNGO (Biological Network Gene Ontology Tool, 5) (**Table 2**). **Table 2** lists the gene ontology description, *p*-value, and the genes in the test set. After sorting ontologies by the smallest *p*-value, the five most significant ontologies were determined to be cellular process, organ morphogenesis, tissue morphogenesis, epithelial morphogenesis, and intermediate filament, including three morphogenesis terms. The most significant term, cellular process, was a top ancestor term under Biological Process and is rather general, making it difficult to obtain specific features from the categorized genes. The second most significant term, organ morphogenesis, containing 13 up-regulated genes, had multiple ancestor terms (**Figure 4**). In addition, the third and fourth enriched terms consisted of a subset of these 13 genes. Therefore, we decided to look further into these 13 genes' functions in detail.

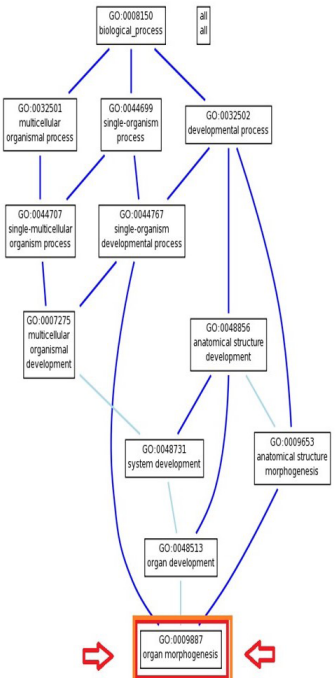

Figure 4: Gene Ontology Structure, showing ancestor terms of organ morphogenesis.

Pathway Analysis

While the 13 genes all fall under the category of organ morphogenesis, their functions are diverse. As seen in the STRING analysis, *Stat3* interacts with many genes, since it can stimulate transcription of various downstream genes in response to diverse upstream signals. *Hif1a* is also individually responsible for regulating the transcription of many genes. *Fgfr3* is an upstream gene leading to various signal transductions. We decided to focus on one pathway involved in these 13 genes,

individually checking each of them for pathways using the human gene compendium GeneCards (6), since all 13 genes have human orthologs. After identifying characteristic pathways for each gene, we searched for patterns and commonalities among these pathways. Ultimately, we found common pathways/super pathways between genes, and compiled results for SuperPaths (according to PathCards) containing at least 2 genes among the 13 genes (**Table 3**).

We found that MAPK signaling had the most genes, containing HMGN1, KRT6A, ENAH, STAT3, FGFR3, HIF1A, and KRT17. Among these, STAT3 was found to be involved in the "Development PDGF signaling via STATs and NF-κB" pathway, a part of the SuperPath "Immune response IL-23 signaling" pathway. We also found that FGFR3 is a part of the "NF-κB family pathway" (**Table 3**). Recent studies show that FGFR3 activates NF-κB signaling in multiple myeloma and bladder cancer through TGFbeta-activated kinase 1 tyrosine phosphorylation (7) and that increased STAT3 activation in myeloid-derived suppressor cells is correlated with the activation of the noncanonical NF-κB pathway (8). Considering the up-regulated *Stat3* and *Fgfr3* in the *Irf6*-deficient mice, we conclude that the NF-κB pathway is activated in *Irf6*-deficient mice.

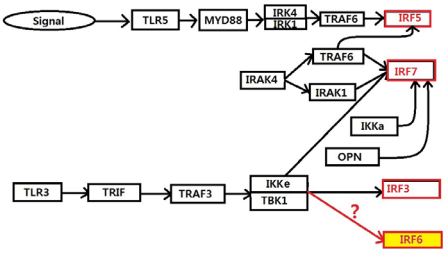

Figure 5: Predicted *Irf6* Pathway. Based on KEGG mouse Toll-like receptor signaling pathway (ID: mmu04620). Diagram showing the network of genes, including our own predicted *Irf6* pathway indicated by the question mark and red arrow. Black arrows are signaling pathways from the KEGG pathway. Any red boxing around a gene indicates that it is a member of the *Irf* family of genes. The bright yellow box indicates the *Irf6* gene.

Discussion

By analyzing gene expression data from *Irf6*-deficient and wild-type mice, we found that diverse molecular pathways beyond epidermal development involve the *Irf6* gene. We had anticipated that *Irf6*-deficient mice would have down-regulated genes, since IRF6 functions as a transcription activator. However, we found that more genes were up-regulated in *Irf6*-deficient mice and that the most enriched GO terms among the up-regulated genes related to morphogenesis. Ingraham and colleagues reported the role of *Irf6* in epidermal differentiation and suggested that *Irf6* deficiency led to the lack of epidermal differentiation. From the genome-wide perspective, however, it seems that overall *Irf6*-deficiency signals compensation. Though *Irf6* mutation causes Van der Woude syndrome, the resultant phenotypes vary among individuals. *Irf6*-deficient mice had many up-regulated genes related to organogenesis, possibly compensating for lost IRF6 function.

Through pathway analysis, we found that JAK, immunity, ERK signaling, and cytoskeletal remodeling pathways, as well as the NF-κB pathway, were related to organ morphogenesis. These pathways might contain viable replacements for a suppressed *Irf6* gene. We found that STAT3 and FGFR3 have NF-κB pathways in common. In addition, the STRING result showed that the same proteins found with BiNGO related to organ morphogenesis were connected: SOCS3, FGFR3, and STAT3. Although up-regulated genes went through different processes, they finally yielded the same result of STAT3 and FGFR3 connections, further confirming that both STAT3 and FGFR3 are crucially related to Irf6 function.

In connection with NF-κB pathways, we located the Toll-like receptor signaling pathway (ID: mmu04620) from a KEGG pathway (9, 10) that contains members of the Irf gene family (*Irf3*, *Irf5*, and *Irf7*). Part of the mmu04620 pathway utilizing IRF genes includes NF-κB genes (IKKs) (**Figure 5**). Considering the activated NF-κB pathway in *Irf6*-deficient mice and that IRF genes are inducible by IKK genes, we hypothesized that the NF-κB pathway compensates for the *Irf6* deficiency and that an element in the mmu04620 pathway activates *Irf6*. We added *Irf6* as a gene potentially downstream of IKKe (shown in **Figure 5**).

In the future, we will try to pick different gene ontologies with less significant *p*-values and test the statistical significance of their connections to *Irf6*, thus testing for "linker genes," or genes that appear in multiple pathways of the genes of interest, and expand this test to different genes underlying genetic diseases. In general, this method could also be used to shed light on shared functions of certain genes and to test which genes could potentially fill the roles of others. This systems biology approach to genomics focuses on the analysis of entire gene pathways, which may prove more effective in identifying genes with changes in gene expression. We believe that this global outlook on gene networks holds promise.

By analyzing the genes in the NF-κB pathway, as well as other pathways that we found, we may further isolate specific genes able to function in *Irf6*'s role. If we end up finding a definite gene in a pathway, how would we use that information to prevent Van Der Woude Syndrome? We would like to look into this question for our next project.

Methods

Data

Gene expression profiles of three wild-type mice (normal) and three *Irf6*-deficient mice (through gene trapping) were obtained from Gene Expression Omnibus (GSE5800). The Affymetrix Mouse Genome 430 2.0 Array was used to obtain expression profiling by array.

RNA Biomarkers

Differently expressed genes were found through the GenePattern software package and up-regulated genes with a student's t-test *p*-value less than 0.001 were further analyzed as RNA biomarkers. Specifically, a table grouping the individual data for genes was created from the gene expression profiles of mice and downloaded from the GEO database, separating *Irf6*-deficient and wild-type mice. That data was run through the Comparative Marker Selection module of the GenePattern suite, yielding a compilation of all the up-regulated genes (over-expressed genes) in the deficient and wild-type mice. We extracted genes with *p*-values less than 0.001 from the Comparative Marker Selection results. In order to visualize a clustering output, we used Hierarchical Clustering with Pearson correlation distance to confirm differentiation of control and deficient groups.

Gene Ontology and Pathways

The Affymetrix IDs were matched up to actual gene symbols using the table provided by the dataset in the GEO database. To find the types of gene ontologies most enriched among those up-regulated genes, the genes were then run through BiNGO (Biological Network Gene Ontology Tool) with a *p*-value cutoff of < 0.05. The gene ontology term of the second most significant one was selected for further investigation, since the most significant term was too general. Each of the genes within the selected gene ontology term was individually checked for pathways/super pathways in the GeneCards database. These same genes were then cross-referenced to find common pathways. The shared pathways identified among these genes and the pathways with the largest number of these genes were subsequently isolated. Finally, these pathways were analyzed further and a single pathway containing *Irf* genes with the selected pathways was identified.

Acknowledgment

We appreciate the generous support of University of Michigan WISE (Women In Science and Engineering).

References

1. Ingraham, Christopher R, Akira Kinoshita, Shinji Kondo, Baoli Yang, Samin Sajan, Kurt Trout, Margaret Malik, Martine Dunnwald, et al. "Abnormal skin, limb and craniofacial morphogenesis in mice deficient for interferon regulatory factor 6 (Irf6)." *Nat Genet*. 38.11 (2006): 1335-1340.
2. The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat Genet*. May 2000;25(1):25-9.
3. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP GenePattern 2.0 *Nat Genet* 38 no. 5 (2006): pp500-501
4. Jensen et al. *Nucleic Acids Res.* 2009, 37(Database issue):D412-6
5. Maere S, Heymans K, Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics* 21, 3448-3449.
6. "Search the GeneCards Human Gene Database." GeneCards. Weizman Institute of Science, n.d. Web. 20 June 2013.
7. Salazar, Lisa et al. "Fibroblast Growth Factor Receptor 3 Interacts with and Activates TGFβ-Activated Kinase 1 Tyrosine Phosphorylation and NFκB Signaling in Multiple Myeloma and Bladder Cancer." Ed. Hari K. Koul. *PLoS ONE* 9.1 (2014): e86470. PMC.
8. Yu, J et al. "Noncanonical NF-κB Activation Mediates STAT3-stimulated IDO Upregulation in Myeloid-derived Suppressor Cells in Breast Cancer." *J Immunol* 193.5 (2014): 2574-86. PMC.
9. Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M.; Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* 42, D199–D205 (2014).
10. Kanehisa, M. and Goto, S.; KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28, 27-30 (2000).

Up-regulated in knockout mice				Down-regulated in knockout mice			
Probe ID	Gene Symbol	p[1]	fc[2]	Probe ID	Gene Symbol	p[1]	fc[2]
1427700_x_at	*Krt6a*	0.00096	121.04	1417621_at	*Nfatc1*	0.00045	0.51
1422784_at	*Krt6a*	0.00029	66.19	1436054_at	*Khnyn*	0.00053	0.50
1439661_at	*Slc16a14*	0.00099	14.68	1455056_at	*Lmo7*	0.00029	0.50
1450618_a_at	*Sprr2a2*	0.00020	12.95	1451972_at	*Glcci1*	0.00085	0.48
1451594_s_at	*Serpinb6c*	0.00094	9.91	1435486_at	*Pak3*	0.00004	0.48
1435639_at	*2610528A11Rik*	0.00010	9.58	1443830_x_at	*Rnf103*	0.00089	0.48
1436100_at	*Sh2d5*	0.00067	7.72	1429390_at	*Acpl2*	0.00040	0.46
1453326_at	*C2cd4b*	0.00004	5.84	1417404_at	*Elovl6*	0.00060	0.45
1423271_at	*Gjb2*	0.00025	4.92	1460646_at	*Csnk2a2*	0.00030	0.45
1419591_at	*Gsdmc*	0.00012	4.80	1454889_x_at	*Tmcc3*	0.00073	0.42
1416236_a_at	*Mpzl2*	0.00008	4.46	1455812_x_at	*Vasn*	0.00030	0.40
1427164_at	*Il13ra1*	0.00046	4.20	1434252_at	*Tmcc3*	0.00002	0.40
1426600_at	*Slc2a1*	0.00014	4.04	1439036_a_at	*Atp1b1*	0.00025	0.39
1421346_a_at	*Slc6a6*	0.00021	4.00	1430596_s_at	*Vgll3*	0.00050	0.37
1426599_a_at	*Slc2a1*	0.00016	3.96	1439018_at	*Fhdc1*	0.00040	0.36
1432135_s_at	*Spdya*	0.00091	3.86	1423689_a_at	*Gpsm1*	0.00069	0.31
1435033_at	*Arhgef4*	0.00003	3.81	1427268_at	*Flg*	0.00068	0.30
1448754_at	*Rbp1*	0.00014	3.68	1428864_at	*Dusp8*	0.00090	0.29
1453792_at	*Zfp329*	0.00038	3.62	1421655_a_at	*Ccr4*	0.00011	0.25
1437361_at	*Gm11545*	0.00083	3.48	1447503_at	*BC016495*	0.00058	0.22
1456250_x_at	*Tgfbi*	0.00035	3.44	1441793_at	*Rnf39*	0.00064	0.21
1448529_at	*Thbd*	0.00002	3.43	1455015_at	*Tbc1d9*	0.00045	0.20
1451775_s_at	*Il13ra1*	0.00017	3.27	1452001_at	*Nfe2*	0.00038	0.19
1448123_s_at	*Tgfbi*	0.00032	3.27	1434473_at	*Slc16a5*	0.00049	0.15
1435842_at	*Nat8l*	0.00008	3.26	1452520_a_at	*Chrng*	0.00026	0.13
1455899_x_at	*Socs3*	0.00071	3.24	1436078_at	*Fcho1*	0.00006	0.11
1437463_x_at	*Tgfbi*	0.00003	3.20	1455825_s_at	*Lnx1*	0.00042	0.11
1434773_a_at	*Slc2a1*	0.00016	3.15	1434553_at	*Tmem56*	0.00032	0.05
1419097_a_at	*Stom*	0.00004	3.14	1423476_at	*Slc46a2*	0.00072	0.04
1427011_a_at	*Lancl1*	0.00049	3.07				
1437149_at	*Slc6a6*	0.00081	2.99				
1455256_at	*Tnik*	0.00001	2.96				
1434425_at	*Tchh*	0.00076	2.90				
1422293_a_at	*Kctd1*	0.00078	2.76				
1423227_at	*Krt17*	0.00046	2.72				
1437279_x_at	*Sdc1*	0.00072	2.71				
1423100_at	*Fos*	0.00058	2.71				
1419184_a_at	*Fhl2*	0.00001	2.69				
1449424_at	*Plek2*	0.00009	2.65				
1450782_at	*Wnt4*	0.00006	2.62				
1434851_s_at	*Crb3*	0.00090	2.59				
1425658_at	*Cd109*	0.00068	2.54				
1439887_at	*Rnf152*	0.00072	2.54				
1425073_at	*Plekhg6*	0.00080	2.53				
1434150_a_at	*Mettl7a1*	0.00084	2.52				
1449481_at	*Slc25a13*	0.00063	2.49				
1415944_at	*Sdc1*	0.00029	2.48				
1441687_at	*Wnt4*	0.00038	2.46				
1450626_at	*Manba*	0.00041	2.44				
1426368_at	*Rin2*	0.00024	2.44				

Table 1: Differently expressed genes in *Irf6*-deficient mice. [1]Student t-test p-value. [2]Fold change.

Description	p-value	Genes in test set
cellular process	8.38E-09	*HMGN1, BACH1, SLC16A14, 1810011O10RIK, ENAH, MPZL2, UQCRC1, NUAK2, PRR13, GOT2, FOS, WNT4, ANG, TGFBI, SLC2A1, DNAJC5, AGPAT3, NMNAT3, TCHH, ZFP329, TNIK, LIMK2, SOCS3, RALBP1, DTX3L, POLB, CDK6, MBD2, PALLD, SPDYA, ATP6V1A, SDC1, HIF1A, MAPK6, KRT17, PRDX6, CLDN1, RIN2, EEF1G, PRNP, RAB10, HS3ST3B1, CROT, DEGS1, KRT6A, FGFR3, GALNT7, RBP1, MAPKAPK3, FHL2, ITGB5, STAT6, COL8A1, PHLDA3, SREBF1, KLF5, KLF10, RPS9, CRB3, SFT2D3, MANBA, STAT3, ITPR2, GJB2, SLC25A13, UBA2, BNC1, SLC6A6, PARD6G*
organ morphogenesis	4.15E-06	*HMGN1, KRT6A, WNT4, ENAH, FGFR3, HIF1A, KRT17, SOCS3, FHL2, CRB3, EGLN1, COL8A1, STAT3*
tissue morphogenesis	7.17E-06	*KRT6A, WNT4, ENAH, FGFR3, HIF1A, KRT17, SOCS3, CRB3, EGLN1*
morphogenesis of an epithelium	9.27E-06	*KRT6A, WNT4, ENAH, FGFR3, HIF1A, KRT17, SOCS3, CRB3*
intermediate filament organization	1.23E-05	*TCHH, KRT6A, KRT17*
tissue development	5.54E-05	*KRT6A, WNT4, ENAH, FGFR3, HIF1A, KRT17, SOCS3, KLF10, FHL2, CRB3, EGLN1, FAM83H*
intermediate filament cytoskeleton organization	7.07E-05	*TCHH, KRT6A, KRT17*
epithelium development	9.96E-05	*KRT6A, WNT4, ENAH, FGFR3, HIF1A, KRT17, SOCS3, CRB3*
intermediate filament-based process	1.11E-04	*TCHH, KRT6A, KRT17*
system development	1.13E-04	*HMGN1, KLF5, KRT6A, ENAH, FGFR3, SOCS3, KLF10, FHL2, CRB3, EGLN1, CDK6, PALLD, SIGMAR1, STAT3, FOS, WNT4, HIF1A, KRT17, ANG, CLDN1, COL8A1, FAM83H*
metabolic process	1.37E-04	*HMGN1, BACH1, UQCRC1, FGFR3, GALNT7, NUAK2, RBP1, PRR13, MAPKAPK3, FHL2, EGLN1, STAT6, GOT2, FOS, WNT4, ANG, DNAJC5, UCK2, AGPAT3, NMNAT3, SREBF1, KLF5, TNIK, ZFP329, LIMK2, KLF10, RPS9, CDK6, POLB, MBD2, SIGMAR1, MANBA, STAT3, ATP6V1A, HIF1A, SLC25A13, MAPK6, PRDX6, UBA2, BNC1, EEF1G, PRNP, CROT, HS3ST3B1, NAT8L, DEGS1*
anatomical structure development	1.61E-04	*HMGN1, KLF5, KRT6A, ENAH, FGFR3, SOCS3, KLF10, FHL2, CRB3, EGLN1, CDK6, PALLD, SIGMAR1, STAT3, FOS, SDC1, WNT4, HIF1A, KRT17, ANG, CLDN1, COL8A1, FAM83H*
anatomical structure morphogenesis	1.72E-04	*KLF5, HMGN1, KRT6A, WNT4, ENAH, FGFR3, HIF1A, KRT17, ANG, SOCS3, FHL2, CRB3, EGLN1, COL8A1, STAT3*
cellular metabolic process	1.76E-04	*HMGN1, BACH1, FGFR3, UQCRC1, GALNT7, RBP1, NUAK2, PRR13, MAPKAPK3, FHL2, STAT6, GOT2, FOS, DNAJC5, AGPAT3, NMNAT3, SREBF1, KLF5, TNIK, ZFP329, LIMK2, KLF10, RPS9, CDK6, POLB, MBD2, MANBA, STAT3, ATP6V1A, HIF1A, SLC25A13, MAPK6, PRDX6, UBA2, BNC1, EEF1G, PRNP, HS3ST3B1, CROT, DEGS1*
cellular response to stimulus	2.05E-04	*SPDYA, SREBF1, HMGN1, FOS, HIF1A, NUAK2, PRDX6, DTX3L, BNC1, POLB, STAT3*
response to chemical stimulus	2.36E-04	*SREBF1, RBP1, LY96, KLF10, EGLN1, STAT3, STAT6, FOS, HIF1A, SPRR2A1, PRDX6, BNC1, PRNP, CROT*
pyrimidine dimer repair	2.70E-04	*HMGN1, POLB*
organ development	2.93E-04	*HMGN1, KLF5, KRT6A, ENAH, FGFR3, SOCS3, KLF10, FHL2, CRB3, EGLN1, CDK6, STAT3, WNT4, HIF1A, KRT17, ANG, COL8A1, FAM83H*
anatomical structure formation involved in morphogenesis	3.33E-04	*KLF5, WNT4, ENAH, HIF1A, ANG, FHL2, EGLN1, COL8A1*
multicellular organismal development	3.54E-04	*HMGN1, KLF5, KRT6A, ENAH, FGFR3, SOCS3, KLF10, FHL2, CRB3, EGLN1, CDK6, PALLD, SIGMAR1, STAT3, SPDYA, FOS, WNT4, THBD, HIF1A, KRT17, ANG, CLDN1, COL8A1, FAM83H*

Table 2: Enriched gene ontology terms in up-regulated gene set in *Irf6*-deficient mice.

Super Pathways	Gene Names												
	HMGN1	*KRT6A*	*WNT4*	*ENAH*	*FHL2*	*EGLN1*	*COL8A1*	*STAT3*	*FGFR 3*	*HIF1A*	*KRT17*	*CRB3*	*SOCS3*
Angiogenesis						X			X	X			
Common Cytokine Receptor Gamma-Chain Family Signaling Pathways								X					X
Cytoskeletal Signaling		X									X		
Cytoskeleton remodeling Keratin filaments		X									X		
Development HGF signaling pathway						X				X			
Development Leptin signaling via JAK/STAT and MAPK cascades										X			X
Immune response IL-23 signaling pathway								X					X
Development Prolactin receptor signaling								X					X
Disease									X	X			
EGFR1 Signaling Pathway											X		X
HIF-1 signaling pathway						X				X			
HIF-1-alpha transcription factor network						X				X			
HIF1Alpha Pathway						X				X			
IL-2 Signaling pathway								X					X
Immune System				X					X				X
Influenza A	X												X
MAPK signaling pathway	X	X		X				X	X	X	X		
Nanog in Mammalian ESC Pluripotency			X						X				
NF-kB Family Pathway									X				
Osteoclast differentiation					X								X
p70S6K Signaling									X	X			
Pathways in cancer			X			X			X	X			
PEDF Induced Signaling									X				X
Proteoglycans in cancer			X							X			
PTEN Pathway							X		X				
Regulation of Actin Cytoskeleton				X					X				
Rho Family GTPases	X						X		X	X			
Sertoli-Sertoli Cell Junction Dynamics												X	
SIDS Susceptibility Pathways										X			
Signaling by GPCR			X						X	X			
TGF-Beta Pathway								X	X				X
Translation Insulin regulation of translation										X			X
Translational Control						X				X			
Tyrosine Kinases / Adaptors									X				X
WNT Signaling			X		X								

Table 3: Pathways involving more than 2 genes identified in the organ morphogenesis GO terms from Table 2.