Physiol Genomics 31: 63–74, 2007.
First published May 29, 2007; doi:10.1152/physiolgenomics.00055.2007.

Role of STAT5a in regulation of sex-specific gene expression in female but not male mouse liver revealed by microarray analysis

Karl H. Clodfelter,[1] Gregory D. Miles,[1] Valerie Wauthier,[1] Minita G. Holloway,[1] Xiaohua Zhang,[2] Paul Hodor,[3] William J. Ray,[4] and David J. Waxman[1]

[1]*Division of Cell and Molecular Biology, Department of Biology, Boston University, Massachusetts; and Merck and Company, Incorporated, Departments of* [2]*Biometrics Research,* [3]*Molecular Profiling, and* [4]*Molecular Endocrinology, West Point, Pennsylvania*

Submitted 7 March 2007; accepted in final form 29 May 2007

Clodfelter KH, Miles GD, Wauthier V, Holloway MG, Zhang X, Hodor P, Ray WJ, Waxman DJ. Role of STAT5a in regulation of sex-specific gene expression in female but not male mouse liver revealed by microarray analysis. *Physiol Genomics* 31: 63–74, 2007. First published May 29, 2007; doi:10.1152/physiolgenomics.00055.2007.—Sexual dimorphism in mammalian liver impacts genes affecting hepatic physiology, including inflammatory responses, diseased states, and the metabolism of steroids and foreign compounds. Liver sex specificity is dictated by sex differences in pituitary growth hormone (GH) secretion, with the transcription factor signal transducer and activator of transcription (STAT)5b required for intracellular signaling initiated by the pulsatile male plasma GH profile. STAT5a, a minor liver STAT5 form >90% identical to STAT5b, also responds to sexually dimorphic plasma GH stimulation but is unable to compensate for the loss of STAT5b and the associated loss of sex-specific liver gene expression. A large-scale gene expression study was conducted using 23,574-feature oligonucleotide microarrays and livers of male and female mice, both wild-type and *Stat5a*-inactivated mice, to elucidate any dependence of liver gene expression on STAT5a. Significant sex differences in expression were found for 2,482 mouse genes, 1,045 showing higher expression in males and 1,437 showing higher expression in females. In contrast to the widespread effects of the loss of STAT5b, STAT5a deficiency had a limited but well-defined impact on liver sex specificity, with 219 of 1,437 female-predominant genes (15%) specifically decreased in expression in STAT5a-deficient female liver. Analysis of liver RNAs from wild-type mice representing three mixed or outbred strains identified 1,028 sexually dimorphic genes across the strains, including 393 female-predominant genes, of which 89 (23%) required STAT5a for normal expression in female liver. These findings highlight the importance of STAT5a for regulation of sex-specific gene expression specifically in female liver, in striking contrast to STAT5b, whose major effects are restricted to male liver.

signal transducer and activator of transcription-5a; knockout; liver sexual dimorphism; growth hormone action; strain-dependent gene expression

GROWTH HORMONE (GH) regulates gene expression in several tissues, most notably liver. GH is secreted by the pituitary gland in a sex-dependent manner under the regulation of gonadal steroids (15, 31). This, in turn, leads to substantial sex differences in GH-regulated liver gene expression (21, 37). In the rat, plasma GH levels are highly pulsatile in males, where hormone peaks every ~3.5 h are followed by a GH-free interval lasting ~2 h, whereas in females, GH is present in plasma in a more continuous manner (15). Mice also show sexually dimorphic GH secretory patterns, with females characterized by more frequent GH pulses and a distinctly shorter GH-free interpulse interval than males (20). These sex-dependent plasma GH profiles control liver gene expression at the level of transcription, as demonstrated for several GH-regulated genes that encode liver cytochrome *P*450 (Cyp) enzymes active in the metabolism of steroids, drugs, and environmental chemicals (34).

GH signaling is initiated by the binding of GH to its cell surface receptor, which activates Janus kinase (JAK)2, a GH receptor-associated tyrosine kinase. JAK2, in turn, phosphorylates GH receptor on multiple cytoplasmic domain tyrosine residues, several of which serve as docking sites for signal transducer and activator of transcription (STAT) transcription factors (6, 12), including STAT5a and STAT5b (7, 9, 25). Each STAT is phosphorylated on a single tyrosine residue, which leads to STAT homo- and heterodimer formation, translocation to the nucleus, binding to STAT5 DNA response elements, and stimulation of gene transcription. Both STAT5a and STAT5b are directly activated in male rat liver in response to each incoming plasma GH pulse (4), whereas in livers of female rats, the persistence of plasma GH stimulation leads to partial desensitization of the STAT5 signaling pathway and substantially lower nuclear STAT5b protein than the peak levels seen in males (3, 4, 35). On the basis of these and other findings, STAT5b has been proposed to serve as a mediator of the sex-dependent effects that GH has on liver gene expression (32). This proposal is supported by the characterization of STAT5b-deficient male mice, which display a reduced body growth rate at puberty and a loss of sex-specific liver expression of *Cyp*s and other genes (11, 28, 30). The effects of STAT5b loss on liver gene expression, but not the effects on body growth rates, are recapitulated in a liver-specific STAT5a-STAT5b double-knockout model. Microarray analysis has further demonstrated that the impact of STAT5b on sex-dependent liver gene expression is global, with 90% of male-predominant genes downregulated and 61% of female-predominant genes upregulated in livers of STAT5b-deficient male mice (5). Studies of STAT5a-deficient mice (19) raise the possibility that STAT5a may be required for expression of certain female-specific mouse Cyp proteins (24); however, this has not been investigated in a systematic way.

Sexual dimorphism of liver gene expression is extensive, affecting at least 1,500 genes (5, 38), including many members of the *Cyp* gene superfamily (34). These sex differences may be impacted by genetic factors that are manifest as strain

Article published online before print. See web site for date of publication (http://physiolgenomics.physiology.org).

Address for reprint requests and other correspondence: D. J. Waxman, Dept. of Biology, Boston Univ., 5 Cummington St., Boston, MA 02215 (e-mail: djw@bu.edu).

differences in sex-dependent gene expression. For example, certain Cyp2b steroid 16α-hydroxylase enzymes show strain-dependent differences that affect sex specificity (23). Moreover, *Cyp2a4* is subject to strain-specific regulation associated with a recessive mutation that derepresses its expression in males (2). Strain-dependent expression also characterizes sex-specific genes regulated by regulators of sex limitation (Rsl), transcriptional repressors that are deficient in certain mouse strains and modulate the expression of *Slp, Mup*, and select other male-specific GH-regulated genes (17, 29). Mutations in *Rsl1* and *Rsl2* serve as examples of strain-specific alleles in transcriptional regulators that impact the sex specificity of GH-responsive genes in the liver (16, 33). It is uncertain whether these or other strain-dependent factors have a widespread effect on the sex specificity of hepatic gene expression.

The present study uses microarray technology to assess the impact of STAT5a deficiency on the sex specificity of liver gene expression. In addition, global assessment of sex specificity was investigated in three different mixed or outbred mouse strains to identify a robust set of strain-independent sex-specific genes, as well as sets of genes whose sex dependence apparently varies with the genetic background. Our findings reveal that *Stat5a* disruption affects the expression of a discrete subset of sex-dependent genes in female mouse liver, in sharp contrast to *Stat5b* disruption, which primarily abrogates sex-specific gene expression in male liver.

MATERIALS AND METHODS

Animals. STAT5a-deficient mice were prepared by targeted disruption of the *Stat5a* gene (19). Livers were collected from 6- to 9-wk-old male and female 129J × Black/Swiss mice, wild-type (WT) mice, and STAT5a knockout (KO) mice. STAT5a KO livers were previously shown to be devoid of detectable STAT5a RNA (and protein), a result that was confirmed by a four- to sixfold downregulation of STAT5a RNA in the present microarray study (see Supplemental Table S2; supplemental data are available at the online version of this article). Moreover, hepatic levels of STAT5b and STAT5b DNA binding activity are unaffected by the loss of STAT5a (24). Livers were also collected from 8- to 10-wk-old male and female WT mice of the ICR strain (Taconic, Germantown, NY). Livers were snap frozen in liquid nitrogen and stored at −80°C until use. All animal studies were carried out using protocols approved by the Boston University Institutional Animal Care and Use Committee.

RNA isolation. Total RNA was isolated from ~0.1 g of frozen mouse liver using Trizol reagent (Invitrogen Life Technologies, Carlsbad, CA) according to the manufacturer's protocol. Liver RNAs prepared from 12 individual 129J × Black Swiss mice (30 μg of RNA per liver, dissolved in diethyl pyrocarbonate-treated water) were used in the present study: three WT males, three STAT5a KO males, three WT females, and three STAT5a KO females. Ten individual ICR mouse liver RNAs (5 WT female livers and 5 WT male livers), kindly provided by Dr. Ekaterina Laz of this laboratory, were used to prepare two pooled liver RNA samples ($n = 2$ livers and $n = 3$ livers, respectively). Two of the individual livers for each sex were also used to prepare single liver RNA samples.

Quantitative PCR analysis. Liver RNA samples were converted to cDNA and used in quantitative PCR (qPCR) assays for individual genes from *groups 13* and *14* (see Table 3) using methods detailed elsewhere (11). Amplification of a single specific product during qPCR cycling was verified by examination of dissociation curves of each amplicon. Relative RNA levels were determined after normalization to the 18S RNA content of each sample. Statistical analysis was carried out by Student's *t*-test using GraphPad Prism software version 4 (San Diego, CA). P values <0.05 were considered significant. qPCR primer design was carried out using Primer Express software (Applied Biosystems), and all primers were verified with respect to their specificity for the target transcript by BLAT [basic local alignment search tool (BLAST)-like alignment tool] analysis of the mouse genome (February 2006 assembly) at http://genome.ucsc.edu/cgi-bin/hgBlat. Primer sequences are shown in Supplemental Table S1.

Microarray analysis. Global expression analysis was determined using the 23,574-feature mouse Rosetta/Merck Mouse TOE 75k Array 1 [Gene Expression Omnibus (GEO) Platform: GPL 3562; Agilent Technology, Palo Alto, CA], with each feature corresponding to a single 60-mer oligonucleotide antisense probe. Each probe is herein referred to as representing a distinct gene, although the actual number of genes analyzed is likely to be smaller than this number because of the presence of nonannotated sequences, some of which may duplicate results for other genes represented on the chip. Liver RNA samples ($n = 3$ independent biological replicates for each of 4 sex-genotype combinations) were used in four separate competitive hybridization experiments in a loop design: male (M) WT vs. female (F) WT (M-WT:F-WT), male WT vs. male STAT5a KO (M-WT:M-KO), female WT vs. female STAT5a KO (F-WT:F-KO), and male STAT5a KO vs. female STAT5a KO (M-KO:F-KO). Sample labeling, hybridization, and array scanning were performed as described (13). Briefly, total RNA was reverse transcribed using an oligo(dT) primer, followed by second-strand cDNA synthesis. Labeled cRNA was generated in a two-step process by derivatization of the transcribed products with either Cy3 or Cy5 dyes. For each microarray hybridization, a mixture of equal amounts of a Cy3-labeled and Cy5-labeled sample was used. Each sample pair was hybridized on two separate microarrays, with fluor reversal. All four comparison groups consisted of three biological replicates, giving a total of 24 microarrays for the STAT5a data set. Four fluor-reversed pairs, corresponding to the two pooled RNA and two single liver RNA samples, were hybridized for the WT male and WT female comparison of the ICR mice.

Microarrays were scanned, and individual feature intensities were preprocessed in a series of steps, consisting of background subtraction, normalization to mean intensities of the Cy3 and Cy5 channels, and detrending to fit a linear relationship between channels (8). Normalized intensities from fluor-reversed pairs of arrays were used to derive expression ratios using the Rosetta error model (8, 36). Probes with intensities resulting from oversaturation were excluded from downstream statistical analysis. The microarray platform and analytical methods used were highly sensitive, and were able to reliably detect gene expression fluorescent signal intensities over a >600-fold dynamic range (0.085–53.2 intensity units), with 18,204 of the 23,574 probes deemed to be reliably expressed in the mouse liver samples analyzed (signal intensities >5 times the whole chip background value for either channel). A similar result was obtained, based on the minimum signal intensities, that reliably allowed for discrimination of sex-specific genes. Thus, for each 129J × Black/Swiss M-WT:F-WT fluorescent reverse pair, the two average intensity values for each gene were compared by *t*-test to determine whether the two values were significantly different at $P < 0.0001$. For each channel, the lowest signal intensity at which gene expression was reliably demonstrable (the cut-off intensity value) is given by the minimum value of the greater of the two intensity values for those genes that met the threshold for significant differential expression. A total of 16,876 genes with intensities above the cut-off intensity value for either channel were thus found to be expressed in WT mouse livers of either sex.

Expression ratios obtained in this study are included in Supplemental Table S2. The data are also available for query or download from the GEO website at the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/geo) as GEO series GSE7169 and GSE7170. GenBank accession numbers and associated gene names, gene descriptions, and Unigene numbers based on Unigene build 161 were obtained for unassigned features using the

BLAT analysis tool located on the University of California Santa Cruz Genome Browser website. Assignments were based on the best matches (identity of ≥56 of 60 nucleotides) to the 60-nucleotide probe sequences of each probe. Probes that returned multiple accession numbers were verified to make sure that all accession numbers represented the same gene. Probe sequences are available on request.

Statistical analysis. Moderated *t*-statistics, using standard errors moderated by a simple Bayesian model, were generated using the linear models for microarray data (LIMMA) package as part of the Bioconductor project within the R statistical program (26). This is the same *P* value used in our prior microarray study on the effect of STAT5b in mouse liver gene expression (5). A filter ($P < 0.05$) was applied to the *P* values to determine the statistical significance of each gene's differential expression for each of the four microarray comparisons (M-WT:F-WT, M-WT:M-KO, F-KO:F-WT, F-KO:M-KO). A fold-change filter of 1.5-fold was combined with the above *P* value filter to reduce the false discovery rate (FDR) to <6%, as follows. Of the 23,574 features represented on the array, 4,597 met the 1.5-fold expression filter for at least one of the four microarray comparisons. The number of genes expected to meet the combined threshold ($P < 0.05$ and >1.5-fold change in expression) by chance (type I errors) is $0.05 \times 4{,}597$, or 230 genes. The actual number of genes meeting the combined threshold was 3,905 genes (see RESULTS), corresponding to an FDR of 230/3,905, or 5.9%. Multiple testing correction methods, such as Bonferroni or Holm step-down, were not applied because these options depend heavily on the independence of each gene's expression and thus filter out many bona fide regulated genes (e.g., genes validated by qPCR) to avoid all type I errors; these methods are thus too restrictive in their effort to avoid false positives, as noted elsewhere (1). Fisher's exact test (http://www.exactoid.com/fisher/) was used to compare the effects of STAT5a deficiency on gene expression in males vs. females.

To compare sex specificity across strains, microarrays comparing WT male vs. WT female gene expression in three distinct genetic backgrounds were averaged ($n = 10$ arrays) to determine the robustness of sex specificity across all three strains and experiments. These three data sets were based on analysis using the same microarray platform for WT males and WT females from the present STAT5a KO study, from the prior study of STAT5b KO mice (5), and from a set of ICR mouse livers, as described above. A differential expression filter (mean ratio >1.5) was applied to the gene expression values deemed statistically significant by the $P < 0.05$ filter. Analysis of variance (ANOVA) with a Benjamini and Hochberg FDR of 0.001 was used, as implemented in GeneSpring GX 7.3.1 software (Agilent Technologies, Santa Clara, CA), to determine strain specificity for the genes that met the threshold criteria for only one of the three independent data sets for M-WT:F-WT comparison.

A system of binary and decimal flags was applied for clustering genes based on expression ratios obtained in all four microarrays, as previously described in our STAT5b KO microarray study (5). For the purpose of this clustering, threshold ratios for differential expression were reduced to values of >1.25 and <0.8 for the three arrays that involved STAT5a KO liver RNAs, with retention of the $P < 0.05$ threshold for statistical significance. Average ratios meeting these threshold and significance criteria contributed to the binary- and decimal-based flag. Thus genes with a M-WT:F-WT microarray ratio meeting the criteria were assigned a binary flag value of 1, while genes meeting the criteria for the M-WT:M-KO, F-KO:F-WT, and F-KO:M-KO microarrays were, respectively, assigned binary flag values of 2, 4, and 8. Genes not meeting these criteria were assigned flag values of 0. The sum of these binary-based flag values defines the whole number portion of the flag and was used as a simple method to identify which of the four microarrays met our criteria for inclusion for any given gene of interest, regardless of the direction (up or down) of the regulation. The flag value was then extended using decimal values of 0.1, 0.01, 0.001, and 0.0001 or 0.2, 0.02, 0.002, and 0.0002 for each of the four microarrays to indicate the direction of regulation between the two conditions on the microarray. Thus average ratios for the M-WT:F-WT microarray >1.5 were assigned a decimal value of 0.2 to indicate upregulation, while average ratios <0.66 were assigned a value of 0.1 to indicate downregulation. The three other microarray ratios were similarly flagged based on threshold values of >1.25 and <0.8, as indicated above, by advancing to a new decimal position for each microarray (i.e., the M-WT:M-KO flag is in the hundredths position, and so on). For each gene, the resulting binary sum describes which microarray ratios met the selection criteria, and the four-digit decimal value describes the direction of regulation (total flagging sum).

Gene Ontology term enrichment analysis. Gene Ontology (GO) molecular function annotations were analyzed for term enrichment using the GO Browser within GeneSpring GX software. Because the GO term assigned to each gene represents a subclass of its parent class within the ontology, each gene was also assigned all parent GO class terms. For each of the GO categories present among each list of genes with GO annotations, the number of genes assigned that term was calculated along with a count of all genes on the microarray

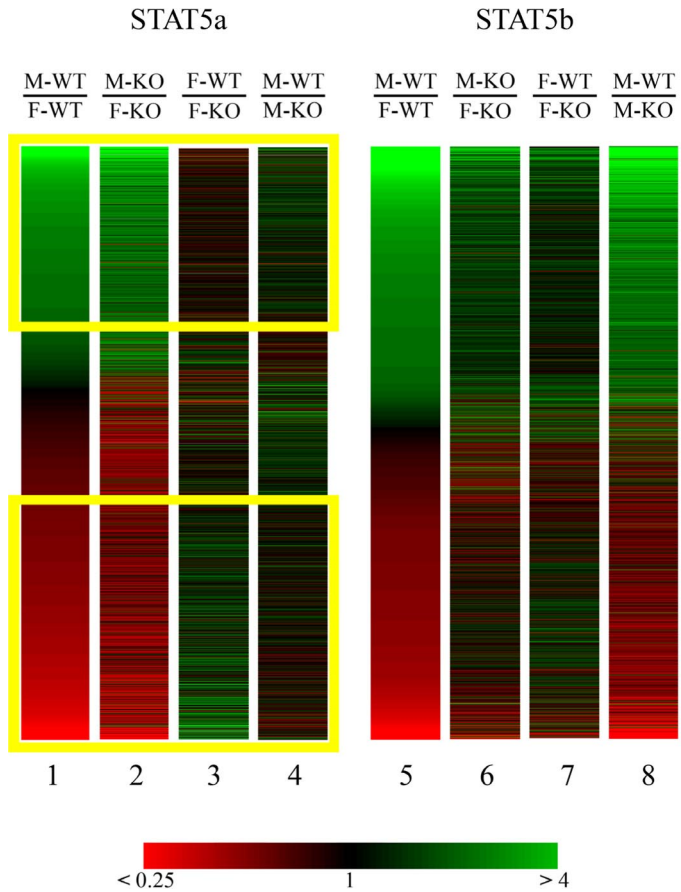

Fig. 1. False-color heat maps for expression profiles of 3,905 differentially expressed genes from the signal transducer and activator of transcription (STAT)5a knockout (KO) data set and 2,231 differentially expressed genes from the STAT5b KO data set. Genes are depicted based on their mean expression ratios across 4 experimental pairings, as indicated at *top* (WT, wild type; M, male; F, female). Genes are colored according to the color bar at *bottom*, ranging from bright green for an expression ratio >4 to bright red for an expression ratio <0.25, with black corresponding to a ratio of 1. In *lanes 1–4*, 3,905 genes from the present STAT5a data set are sorted according to the mean M-WT:F-WT ratio. The *top* and *bottom* yellow boxes enclose genes determined to be male specific and female specific, respectively (M-WT:F-WT >1.5 or <0.66 at $P < 0.05$). In *lanes 5–8*, 2,231 differentially expressed genes from the STAT5b KO data set (5) are sorted according to the M-WT:F-WT ratio.

Table 1. *Changes in gene expression in STAT5a KO mice compared with WT mice of the same sex*

	Gene Expression Changes in STAT5a KO Males Only		Gene Expression Changes in STAT5a KO Females Only		Gene Expression Changes in STAT5a KO Mice of Both sexes			
Comparison	Gene Count	Percentage, %	Gene Count	Percentage, %	Gene Count	Percentage, %	Total	Percentage, %
WT < KO	116	12	231	25	22	2	369	40
WT > KO	173	19	348	37	25	3	546	59
Total	289	31	579	62	47	5	930*	

Changes in gene expression in signal transducer and activator of transcription (STAT)5a knockout (KO) mice compared with wild-type (WT) mice of the same sex; 930 liver-expressed genes meeting the criteria described in MATERIALS AND METHODS in either males or females of the STAT5a KO strain are listed according to the direction of expression change in KO compared with WT mouse liver and whether the change in expression occurs in males, females, or both sexes. Sixty-two percent of all genes showing a change of expression in any of the STAT5a KO mice compared with WT showed a change in expression in females only. *Fifteen genes showed differential expression in both sexes, but regulation of these genes by STAT5a was in opposite directions depending on the sex. They are included in the 930 gene count, but they are not represented in any of the columns.

assigned that term. The *P* value for each category was calculated based on the total number of genes on the microarray within the category and the number of genes in the group being tested within the category. The *P* value represents the likelihood that at least as many genes would occur if a list of equal size were selected by chance from the total gene count.

RESULTS

Experimental design. A large-scale expression study was conducted to investigate the role of STAT5a in the sex specificity of mouse liver gene expression. RNA was isolated from livers of male and female mice that were either WT or contained a targeted disruption of the *Stat5a* gene (KO). RNA samples representing each sex-genotype combination were analyzed in four sets of competitive hybridization to 23,574-feature oligonucleotide microarrays: *1*) M-WT vs. F-WT, *2*) M-WT vs. M-KO, *3*) F-WT vs. F-KO, and *4*) M-KO vs. F-KO. Normalized hybridization intensities were used to calculate mean expression ratios based on $n = 3$ biological replicates for each sex-genotype combination, and *P* values were calculated using the LIMMA tool set within the R statistical program (26). Probes representing 3,905 genes met the threshold criteria for differential expression (average expression ratio >1.5-fold and a significance level of $P < 0.05$) for at least one of the four sex-genotype combinations. Thus these 3,905 genes (21% of the 18,204 probes giving a liver expression measurement >5-fold background in at least 1 sex) were expressed in a sex-specific manner in either WT or STAT5a KO mice or responded to the loss of STAT5a in either males or females. Average expression ratios for the 3,905 genes of interest are listed in Supplemental Table S2.

Overview of sex specificity and impact of STAT5a deficiency. Gene expression differences between WT male and WT female liver were found for 2,482 of the 3,905 genes of interest (64%), indicating that these genes are sexually dimorphic in WT liver. These genes are colored green (M-WT > F-WT) or red (M-WT < F-WT) in Fig. 1, *lane 1*, where they are displayed at the far ends of a false-color heat map containing all 3,905 genes sorted by average M-WT:F-WT ratio. Expression of 1,045 of the 2,482 genes was male predominant in WT mouse liver (M-WT:F-WT >1.5; Fig. 1, *lane 1*, *top* yellow box), while 1,437 genes were female predominant (M-WT:F-WT <0.667; Fig. 1, *lane 1*, *bottom* yellow box).

Hierarchical clustering of the four microarray data sets grouped the male-female comparisons for WT and KO mice together (Pearson correlation coefficient = 0.726). Thus genes found to be sex specific in WT mice generally retained sex specificity in the STAT5a KO mice. Additionally, no discernable overall correlation was found between sex specificity and the loss of STAT5a in either sex. For 930 of the 3,905 genes of interest (24%), the threshold criteria for significant change in expression was met when comparing WT and STAT5a KO mice of the same sex (Table 1). For 579 of these 930 genes (62%), the loss of STAT5a affected gene expression in female liver only, whereas for 289 of the 930 genes (31%), the loss of STAT5a affected gene expression in males only (Table 1). In male mouse liver, STAT5a deficiency had no effect on 93–95% of the sex-predominant genes. STAT5a deficiency also had no effect on 93% of male-specific genes in female liver (Table 2). By contrast, in female liver, 219 of 1,437 female genes (15%) were decreased in expression in the absence of STAT5a (Table 2). The impact of STAT5a deficiency on female-specific gene expression in female liver was significantly greater than the effects of STAT5a deficiency on either *1*) male-predominant or female-predominant gene expression in male liver ($P < 10^{-12}$; Fisher's exact test) or *2*) male-predominant gene expression in female liver ($P < 10^{-10}$; Fisher's exact test). The impact of STAT5a deficiency on these female-predominant genes is discussed further, below.

Table 2. *Impact of STAT5a deficiency on expression of male-specific and female-specific gene expression*

	Response to KO in Males		Response to KO in Females	
	Gene Count	Percentage, %	Gene Count	Percentage, %
Male-specific genes				
Increase	14	1	56	5
Decrease	61	6	16	2
No change	970	93	973	93
Total	1,045		1,045	
Female-specific genes				
Increase	50	3	22	2
Decrease	23	2	219	15
No change	1,364	95	1,196	83
Total	1,437		1,437	

Impact of STAT5a deficiency on male-specific and female-specific gene expression; 2,482 liver-expressed genes meeting the criteria for sex specificity in WT mice are listed according to the sex specificity of their expression in STAT5a KO mice; 93–95% of sex-specific genes do not respond to the loss of STAT5a in either sex, but male genes that do respond tend to be induced in the absence of STAT5a in female liver and repressed in the absence of STAT5a in male liver. In contrast, female genes that respond to the loss of STAT5a tend to be repressed in the absence of STAT5a in female liver. Fifteen percent of the female-specific genes show the latter response.

Clustering by significance and differential expression. The 3,905 genes that met the threshold criteria for at least one of the four sex-genotype expression data sets were clustered into subgroups using a "flagging" system (5), whereby each gene is assigned to a specific category based on its expression ratio in each of the four sex-genotype combinations investigated (see MATERIALS AND METHODS); 3,820 of the 3,905 genes were thus classified into 28 groups of coexpressed genes, as shown in Table 3. Table 3 also presents the distribution of gene counts in each of the corresponding categories determined earlier based on sex specificity and the loss of STAT5b (5). Groups comprised of <15 genes in both the present study and the previous study of STAT5b KO mice (5) are collected into a single category named "other." Striking differences in these two profiles are apparent from the comparisons presented in Supplemental Fig. S1. Six of the seven largest gene groups identified in the present study, comprising 2,975 of the 3,905 genes (76%) (*groups 4A*, *4B*, *5A*, *5B*, *7A*, and *7B*), did not show a significant change in expression in either male or female STAT5a KO mice (i.e., no regulation by STAT5a in either males or females; Table 3). The remaining 930 genes (24%) met the threshold criteria for a response to STAT5a deficiency in either males or females, as noted above. These latter genes include *group 13B*, the sixth largest gene group, which is comprised of 164 female-predominant genes that were downregulated with STAT5a deficiency in female but not male liver, leading to the loss of sex specificity in the STAT5a KO strain (Table 4). A corresponding group of 26 male-predominant genes (*group 13A*) was specifically upregulated in STAT5a-deficient female liver, leading to a loss of sex specificity.

Female-specific genes that respond to STAT5a deficiency in females. The loss of STAT5a in female liver decreased the sex specificity of 275 genes, with 219 female-predominant genes decreased and 56 male-predominant genes increased (Table 2). Of these 275 genes, 190 (69%) belong to *groups 13A* and *13B* (Table 3), indicating a loss of sex specificity in the KO mice; 79 of the 275 genes (29%) comprise *groups 14A* and *14B*, whose genes display the same pattern of response to STAT5a deficiency in female liver as *groups 13A* and *13B*, albeit with partial retention of sex specificity in the STAT5a KO strain (Table 4 vs. Table 5). The quantitative relationships between M-WT:F-WT ratios and F-KO:F-WT ratios for the 190 genes in *groups 13A* and *13B* and for the 79 genes in *groups 14A* and *14B* are shown in Fig. 2, where mean $\log_2$ ratios for each comparison are plotted on a log-log scale. A linear relationship between the magnitude of sex specificity and the response to

Table 3. *Distribution of the 3,905 from the STAT5a KO study and 2,231 differentially expressed genes from the STAT5b KO study within 28 flag-based coexpression groups*

						Sex Specificity		Response to the Loss of STAT5a (or of STAT5b)	
Group	TFS	STAT5a	STAT5a, %	STAT5b	STAT5b, %	WT	KO	Males	Females
1A	3.2200	27	0.7	560	25.1	M	—	Down	—
1B	3.1100	26	0.7	357	16.0	F	—	Up	—
2A	15.2211	6	0.2	44	2.0	M	M	Down	Down
2B	15.1122	3	0.1	39	1.7	F	F	Up	Up
3A	11.2201	25	0.6	139	6.2	M	M	Down	—
3B	11.1102	18	0.5	36	1.6	F	F	Up	—
4A	1.2000	123	3.1	58	2.6	M	—	—	—
4B	1.1000	280	7.2	229	10.3	F	—	—	—
5A	9.2001	783	20.1	20	0.9	M	M	—	—
5B	9.1002	851	21.8	43	1.9	F	F	—	—
6A	7.2210	0	0.0	15	0.7	M	—	Down	Down
6B	7.1120	0	0.0	15	0.7	F	—	Up	Up
7A	8.0001	353	9.0	39	1.7	—	M	—	—
7B	8.0002	585	15.0	127	5.7	—	F	—	—
8A	2.0200	61	1.6	86	3.9	—	—	Down	—
8B	2.0100	24	0.6	133	6.0	—	—	Up	—
9A	4.0010	83	2.1	30	1.3	—	—	—	Down
9B	4.0020	88	2.3	17	0.8	—	—	—	Up
10A	6.0210	16	0.4	36	1.6	—	—	Down	Down
10B	6.0120	14	0.4	21	0.9	—	—	Up	Up
11A	10.0101	36	0.9	51	2.3	—	M	Up	—
11B	10.0202	37	0.9	13	0.6	—	F	Down	—
12A	12.0022	71	1.8	40	1.8	—	F	—	Up
12B	12.0011	41	1.0	11	0.5	—	M	—	Down
13A	5.2020	26	0.7	2	0.1	M	—	—	Up
13B	5.1010	164	4.2	14	0.6	F	—	—	Down
14A	13.2021	28	0.7	0	0.0	M	M	—	Up
14B	13.1012	51	1.3	1	0.0	F	F	—	Down
Other	Other	85	2.2	55	2.5	M/F	M/F	Up/Down	Up/Down

Distribution of 3,905 differentially expressed genes from the present STAT5a KO study and 2,231 differentially expressed genes from the STAT5b KO study (5) within coexpressed gene groups. Genes are grouped based on the total flagging sum (TFS) system (see MATERIALS AND METHODS). This is indicated by the direction of response for each ratio as detailed in the 4 *right* most columns, as indicated by male (M), female (F), upregulation (up), or downregulation (down) in WT and in STAT5a KO (or STAT5b KO) males or females, or no response (—). Twenty-eight gene groups, each containing at least 15 genes in 1 of the 2 data sets, were assigned group nos. Fifty-five genes from the STAT5b KO data set, representing <2.5% of the genes, were distributed among 20 additional gene groups and are classified as "other." Eighty-five genes from the STAT5a KO data set, representing ~2% of the genes, were among 18 additional gene groups also classified as other. Percentages are based on 3,905 and 2,231 genes.

Table 4. *Sex-specific genes that are up- or downregulated in STAT5a-deficient females but not males and are not sex-specific in the STAT5a KO strain*

RefSeq/GenBank ID	Common Name(s)	Description	Expression Ratio (Average) M-WT: F-WT	M-WT: M-KO	F-KO: F-WT	F-KO: M-KO
		Group 13A. Male predominant and increased in STAT5a-deficient females (total of 26 genes)				
NM_146245	Lrrc21	Leucine-rich repeat containing 21	**3.34**	0.60	**4.76**	0.79
NM_027907	Agxt2l1	Alanine glyoxylate aminotransferase 2-like 1	**3.55**	0.81	**3.12**	0.70
NM_198967	Tmtc1	Transmembrane and tetratricopeptide repeat containing 1	**2.73**	0.93	**2.72**	0.91
NM_025308	1810007E14Rik	hypothetical protein LOC53906	**2.85**	1.23	**2.66**	0.78
MMT00033789	MMT00033789	Similar to serine (or cysteine) proteinase inhibitor, clade A, member 3N	**2.34**	1.09	**2.23**	1.02
NM_010406	Hc	Hemolytic complement	**3.44**	2.02	**2.11**	1.45
NM_020581	Angptl4	Angiopoietin-like 4	**2.18**	0.82	**2.06**	0.74
MMT00033197	MMT00033197	Similar to serine (or cysteine) proteinase inhibitor, clade A, member 3N	**2.15**	1.09	**1.99**	1.00
NM_008458	Serpina3c	Serine (or cysteine) proteinase inhibitor, clade A, member 3C	**2.00**	0.86	**1.96**	0.79
NM_019748	Sae1	SUMO1 activating enzyme subunit 1	**2.12**	1.06	**1.76**	0.88
NM_145743	Lace1	Lactation elevated 1	**1.96**	0.97	**1.74**	0.87
NM_010012	Cyp8b1	Cytochrome *P*450, family 8, subfamily b, polypeptide 1	**2.47**	1.08	**1.74**	0.70
NM_025576	Ptpmt1	Protein tyrosine phosphatase, mitochondrial 1	**1.83**	1.30	**1.60**	1.14
NM_145940	Wipi1	WD repeat domain, phosphoinositide interacting 1	**1.70**	0.83	**1.59**	0.78
NM_172142	IkappaBNS	NF-kappaB inhibitor	**1.66**	1.05	**1.51**	0.96
NM_134112	Kctd1	Potassium channel tetramerisation domain containing 1	**1.75**	1.05	**1.44**	0.80
NM_019823	Cyp2d22	Cytochrome *P*450, family 2, subfamily d, polypeptide 22	**1.66**	0.92	**1.43**	0.76
NM_133902	Sdsl	Serine dehydratase-like	**1.53**	0.94	**1.40**	0.83
NM_019911	Tdo2	Tryptophan 2,3-dioxygenase	**2.08**	1.05	**1.39**	0.82
NM_008391	Irf2	Interferon regulatory factor 2	**1.56**	0.99	**1.36**	0.94
NM_001040700	BC051019	RefSeq Gene BC051019	**1.74**	1.07	**1.34**	0.82
NM_172479	Slc38a5	Solute carrier family 38, member 5	**1.53**	1.01	**1.34**	0.92
NM_133906	Zkscan1	Zinc finger with KRAB and SCAN domains 1	**1.60**	1.23	**1.30**	1.01
NM_145365	Creb3l3	cAMP-responsive element binding protein 3-like 3	**1.53**	0.99	**1.30**	0.83
NM_008933	Prm2	Protamine 2	**1.69**	1.06	**1.27**	0.85
NM_009255	Serpine2	Serine (or cysteine) proteinase inhibitor, clade E, member 2	**1.51**	1.08	**1.27**	0.85
		Group 13B. Female predominant and decreased in STAT5a-deficient females (total of 164 genes)				
NM_175270	Ankrd56	Ankrd56 ankyrin repeat domain 56	**0.27**	0.76	**0.13**	0.40
NM_010766	Marco	Macrophage receptor with collagenous structure	**0.27**	0.36	**0.19**	0.32
NM_007836	Gadd45a	Growth arrest and DNA damage-inducible 45-alpha	**0.25**	1.07	**0.19**	0.89
NM_010654	Klrd1	Killer cell lectin-like receptor, subfamily D, member 1	**0.10**	1.17	**0.28**	1.34
NM_008321	Idb3, Id3	Inhibitor of DNA binding 3	**0.39**	0.71	**0.29**	0.56
NM_019467	Aif1	Allograft inflammatory factor 1	**0.38**	0.78	**0.32**	0.68
NM_011087	Pira1	Paired-Ig-like receptor A1	**0.31**	0.85	**0.32**	0.98
NM_020008	Clec7a	C-type lectin, family 7, member a	**0.40**	0.93	**0.34**	0.85
NM_021272	Fabp7	Fatty acid binding protein 7, brain	**0.39**	0.98	**0.36**	0.90
NM_010496	Idb2, Id2	Inhibitor of DNA binding 2	**0.44**	0.90	**0.36**	0.75
NM_010724	Psmb8	Proteosome subunit, beta type 8 (large multifunctional protease)	**0.31**	0.62	**0.37**	0.77
NM_010531	Il18bp	Interleukin 18 binding protein	**0.33**	0.63	**0.37**	0.82
NM_008527	Klrb1c	Killer cell lectin-like receptor subfamily B member 1C	**0.49**	0.94	**0.37**	0.99
BC013712	BC013712	basement membrane-induced protein	**0.38**	0.78	**0.37**	0.77
NM_026268	Dusp6	Dual-specificity phosphatase 6	**0.29**	1.58	**0.38**	2.05
NM_019494	Cxcl11	Chemokine (C-X-C motif) ligand 11	**0.45**	0.84	**0.39**	0.90
NM_029620	Pcolce2	Procollagen C-endopeptidase enhancer 2	**0.35**	0.73	**0.40**	0.82
NM_010407	Hck	Hemopoietic cell kinase	**0.47**	0.82	**0.40**	0.72
NM_010370	Gzma	Granzyme A	**0.49**	1.20	**0.41**	1.15
NM_178933	Ostb	Organic solute transporter beta	**0.57**	1.15	**0.41**	0.80
NM_030060	9130211l03Rik	Jun dimerization protein p21SNFT	**0.35**	0.83	**0.41**	1.10
AF065902	Itgal	Integrin alpha L	**0.37**	0.99	**0.41**	1.21
NM_178759	Timd4	T-cell immunoglobulin and mucin domain containing 4	**0.22**	0.65	**0.41**	1.23
NM_144559	Fcrl3	Fc receptor-like 3	**0.41**	0.91	**0.42**	1.04
NM_172659	Slc2a6	Solute carrier family 2 (facilitated glucose transporter), member 6	**0.51**	0.84	**0.42**	0.73
BI407225	Hk3	Hexokinase 3	**0.44**	0.83	**0.42**	0.81

Sex-specific liver genes belonging to *groups 13A* and *13B* that are up- or downregulated in STAT5a-deficient females but not males and whose sex specificity is lost in the STAT5a KO strain. *Top*: the 26 male-specific genes representing *group 13A*. *Bottom*: the top 26 female-specific genes representing *group 13B*. MMT accession nos. are identification (ID) nos. assigned to microarray probes that could not be linked to a specific GenBank accession no.. Expression ratios shown are average values based on 3 independent microarrays, each carried out in duplicate (dye swap). Individual gene data for all other members of *group 13B*, and for all other groups listed in Table 3, are presented in Supplemental Table S2. Values shown in bold are significantly different from a value of 1.0 by *P* value.

Table 5. *Sex-specific genes that are up- or downregulated in STAT5a-deficient females but not males and are sex specific in the STAT5a KO strain*

RefSeq/GenBank ID	Common Name(s)	Description	Expression Ratio (Average) M-WT: F-WT	M-WT: M-KO	F-KO: F-WT	F-KO: M-KO
		Group 14A. Male predominant in WT and KO and increased in STAT5a-deficient females (total of 28 genes)				
NM_028903	Scara5	Scavenger receptor class A, member 5 (putative)	**50.66**	0.70	**3.67**	**0.03**
AK050313	C8b	Complement component 8, beta subunit	**6.97**	0.96	**1.45**	**0.20**
NM_183278	2200001I15Rik	RIKEN cDNA 2200001I15 gene	**5.96**	0.66	**2.92**	**0.31**
NM_007912	Egfr	Epidermal growth factor receptor	**5.15**	1.14	**2.39**	**0.53**
NM_177101	4833442J19Rik	RIKEN cDNA 4833442J19 gene	**4.73**	0.88	**2.00**	**0.36**
NM_146148	C8a	Complement component 8, alpha polypeptide	**4.44**	ND	**1.44**	**0.24**
BC021325	Serpina1c	Serine (or cysteine) peptidase inhibitor, clade A, member 1c	**4.01**	0.85	**1.59**	**0.32**
NM_021793	Tmem8	Transmembrane protein 8 (five membrane-spanning domain	**3.01**	1.19	**1.42**	**0.61**
NM_025989	Gp2	Glycoprotein 2 (zymogen granule membrane)	**2.79**	1.18	**1.27**	**0.54**
NM_015744	Enpp2	Ectonucleotide pyrophosphatase/phosphodiesterase 2	**2.79**	0.79	**1.68**	**0.49**
BC048380	Es31	Esterase 31	**2.68**	1.07	**1.33**	**0.51**
NM_008645	Mug1	Murinoglobulin 1	**2.66**	0.81	**1.57**	**0.49**
M75720	Serpina1a	Serine (or cysteine) proteinase inhibitor, clade A, member 1	**2.44**	0.89	**1.46**	**0.45**
BF147646	Aak1	AP2-associated kinase 1	**2.44**	0.86	**1.49**	**0.50**
NM_146188	Kctd15	Potassium channel tetramerization domain containing 15	**2.40**	0.86	**1.52**	**0.54**
NM_013601	Msx2	Homeobox, msh-like 2	**2.30**	0.88	**1.45**	**0.53**
NM_009246	Serpina1d	Serine (or cysteine) proteinase inhibitor, clade A, member 1	**2.28**	0.88	**1.50**	**0.54**
NM_009253	Serpina3m	Serine (or cysteine) proteinase inhibitor, clade A, member 3	**2.25**	0.86	**1.53**	**0.60**
NM_009695	Apoc2	Apolipoprotein C-II	**1.90**	0.79	**1.65**	**0.64**
NM_001039376	Pde4dip	Phosphodiesterase 4D interacting protein (myomegalin)	**1.85**	0.97	**1.29**	**0.70**
NM_025680	Ctnnbl1	Catenin, beta like 1	**1.81**	0.98	**1.37**	**0.74**
NM_147217	Gprc5c	G protein-coupled receptor, family C, group 5, member C	**1.71**	0.88	**1.39**	**0.68**
NM_010638	Klf9	Kruppel-like factor 9	**1.70**	0.88	**1.33**	**0.64**
NM_146751	Olfr648	Olfactory receptor 648	**1.65**	0.87	**1.27**	**0.68**
NM_172404	Ccbl1	Cysteine conjugate-beta lyase 1	**1.63**	1.01	**1.27**	**0.80**
NM_172873	Cdcp2	CUB domain containing protein 2	**1.62**	0.93	**1.26**	**0.72**
NM_011453	Serpinb9c	Serine (or cysteine) peptidase inhibitor, clade B, member 9c	**1.60**	0.93	**1.29**	**0.70**
NM_007540	Bdnf	Brain-derived neurotrophic factor	**1.60**	0.89	**1.35**	**0.63**
		Group 14B. Female predominant in WT and KO and decreased in STAT5a-deficient females (total of 51 genes)				
NM_146232	BC014805	hypothetical protein LOC236149	**0.01**	0.85	**0.63**	**33.83**
NM_001081325	EG629219	Predicted gene, EG629219, sulfotransferase related	**0.02**	1.02	**0.29**	**31.19**
NM_009286	C730007P19	Sulfotransferase family 2A, DHEA-preferring, member 2	**0.05**	0.91	**0.53**	**13.85**
NM_011994	Abcd2	ATP-binding cassette, subfamily D (ALD), member 2	**0.07**	0.85	**0.51**	**6.60**
AB056442	OAT6RP	organic ion transporter 6-related protein	**0.08**	1.07	**0.42**	**7.65**
NM_007820	Cyp3a16	Cytochrome *P*450, family 3, subfamily a, polypeptide 16	**0.08**	1.32	**0.51**	**7.53**
MMT00046397	LOC434121	sulfotransferase (mSTa1)-related	**0.09**	1.01	**0.51**	**5.55**
AK149509	EG233005	Cytochrome *P*450, family 2, subfamily a-related, polypeptide 12	**0.10**	1.05	**0.45**	**5.73**
NM_134246	Acot3	Peroxisomal acyl-CoA thioesterase	**0.12**	1.10	**0.56**	**5.92**
NM_054088	Pnp1a3	Patatin-like phospholipase domain containing 3	**0.13**	1.42	**0.27**	**3.33**
NM_007643	Cd36	CD36 antigen	**0.13**	0.72	**0.53**	**2.93**
NM_018784	St3 gal6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	**0.14**	0.87	**0.57**	**4.22**
NM_008706	Nqo1	NAD(P)H dehydrogenase, quinone 1	**0.17**	0.81	**0.73**	**3.94**
NM_178765	5730410E15Rik	Golgi-localized syntaphilin-related protein isoform C	**0.19**	1.11	**0.36**	**3.27**
NM_008935	Prom1	Prominin 1	**0.21**	1.06	**0.46**	**2.67**
NM_008357	Il15	Interleukin 15	**0.22**	0.83	**0.56**	**2.28**
BB153684	Slc16a5	Solute carrier family 16 (monocarboxylic acid transporters),	**0.24**	1.07	**0.65**	**3.54**
NM_019919	Ltbp1	Latent transforming growth factor beta binding protein 1	**0.24**	0.79	**0.65**	**2.40**
NM_013825	Ly75	Lymphocyte antigen 75	**0.24**	0.76	**0.57**	**1.91**
NM_009982	Ctsc	Cathepsin C	**0.29**	0.80	**0.64**	**1.82**
NM_029653	Dapk1	Death-associated protein kinase 1	**0.35**	1.46	**0.65**	**2.89**
NM_025408	Phca	Phytoceramidase, alkaline	**0.38**	1.24	**0.69**	**2.38**
NM_144846	0910001A06Rik	hypothetical protein LOC223601	**0.38**	1.12	**0.66**	**1.94**
NM_023737	Ehhadh	Hydratase/3-hydroxyacyl Coenzyme A dehydrogenase	**0.40**	1.18	**0.75**	**2.25**
NM_007987	Fas	Fas (TNF receptor superfamily member)	**0.41**	0.94	**0.72**	**1.61**
NM_008371	Il7	Interleukin 7	**0.42**	1.04	**0.50**	**1.57**
NM_026082	Dock7	Dedicator of cytokinesis 7	**0.44**	0.97	**0.73**	**1.64**
NM_126166	Tlr3	Toll-like receptor 3	**0.44**	1.03	**0.59**	**1.39**

Sex-specific liver genes belonging to *groups 14A* and *14B* that are up- or downregulated in STAT5a-deficient females but not males and whose sex specificity is retained in the STAT5a KO strain. *Top*: the 28 male-specific genes representing *group 14A*. *Bottom*: the top 28 female-specific genes representing *group 14B*. Other details are as in Table 4. Individual gene data for all other members of *group 14B* are presented in Supplemental Table S2. Values shown in bold are significant by *P* value.

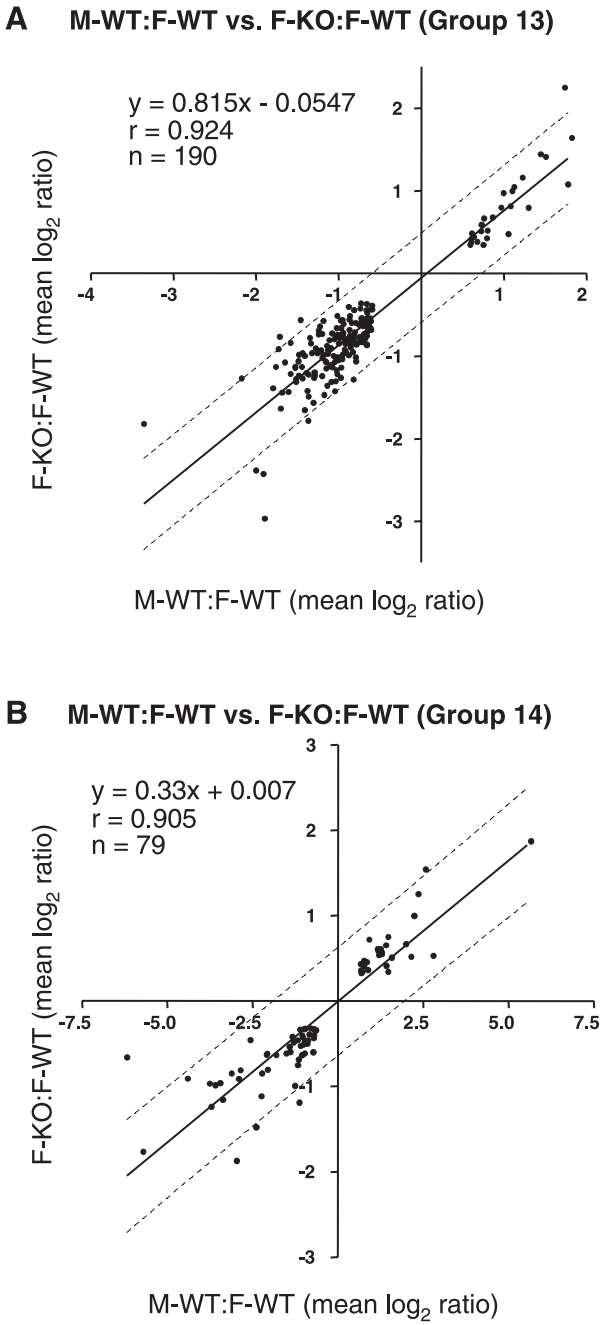

Fig. 2. Scatterplots for quantitative correlation of M-WT:F-WT and F-KO:F-WT ratios for genes in *groups 13* and *14*. *A*: $\log_2$ ratios for M-WT:F-WT (*x*-axis) and the corresponding F-KO:F-WT data (*y*-axis) plotted on a log-log scale for all genes in *groups 13A* and *13B* (Table 3; $n = 190$ genes). The 2 sets of ratios for the 2 experimental comparisons are highly correlated ($r = 0.924$), and the best-fit line ($y = 0.815x - 0.0547$), shown in solid, has a slope close to 1 and an intercept near 0. *B*: $\log_2$ ratios for M-WT:F-WT (*x*-axis) and the corresponding F-KO:F-WT data (*y*-axis) plotted on a log-log scale for all genes in *groups 14A* and *14B* (Table 3; $n = 79$ genes). The 2 sets of ratios for the 2 experimental comparisons are highly correlated ($r = 0.905$), and the best-fit line ($y = 0.33x + 0.007$), shown in solid, has an intercept near 0 but a slope near 1/3. In both *A* and *B*, genes found in the *top right* quadrant (I) are male predominant and were increased in expression in female liver in the absence of STAT5a, while genes found in the *bottom left* quadrant (III) are female predominant and were decreased in expression in STAT5a-deficient female liver; 95% prediction boundaries are shown as dashed lines.

STAT5a deficiency in female liver is evident for the genes in *groups 13A* and *13B* (Fig. 2*A*; slope = 0.815, *y*-intercept = 0.0547, $r = 0.924$), indicating that STAT5a plays a significant role in the sex specificity of these genes. A linear relationship between sex specificity and response to the loss of STAT5a in female liver was also seen for the genes in *groups 14A* and *14B* (Fig. 2*B*; slope = 0.33, *y*-intercept = −0.007, $r = 0.905$), where the slope of 0.33 highlights the partial loss of sex specificity in the absence of STAT5a. These expression profiles were confirmed by qPCR analysis for select genes from *groups 13A*, *13B*, *14A*, and *14B*, as shown in Fig. 3 and in Supplemental Table S3. A majority of the sex-specific genes responsive to the loss of STAT5a in female liver that are presently annotated with GO terms (194 of the 275 genes in this category) are associated with six primary molecular function GO categories: binding, catalytic activity, signal transduction, transporter, enzyme regulation, and transcription regulation (Table 6). Four subcategories of molecular function, signal transducer activity, receptor activity, sugar binding, and carbohydrate binding, were significantly enriched ($P \leq 0.0004$) among the 194 genes and encompass 68 unique genes. Of note, *groups 13A* and *14A* both include multiple members of the Serpina proteinase inhibitor family in the listing of top genes (Tables 4 and 5).

Comparison of STAT5a deficiency with STAT5b deficiency. No overall correlation between sex-specific gene expression and the effect of STAT5a deficiency was seen, either in males or in females, in contrast to our earlier finding with STAT5b-deficient mice (5). Qualitatively, when the regulated genes in the present STAT5a KO study were sorted by sex specificity, only the M-KO:F-KO comparison showed similarity to M-WT:F-WT in the false-color heat map (Fig. 1, *lane 2* vs. *lane 1*), whereas an extensive similarity in the heat map was apparent for the M-WT:F-WT and M-WT:M-KO comparisons in the STAT5b KO study (Fig. 1, *lane 8* vs. *lane 5*). The distribution of genes among the coexpressed gene groups also exhibited marked differences in the requirement of STAT5a and STAT5b for sex-specific gene expression (Table 3 and Supplemental Fig. S1). STAT5b was required for the expression in male liver of 1,205 sex-specific genes, or 75% of the 1,603 sex-specific genes identified in the STAT5b study (5) (genes comprising STAT5b KO study *groups 1A*, *1B*, *2A*, *2B*, *3A*, *3B*, *6A*, and *6B*). By contrast, the corresponding eight groups comprised only ~4% of the 2,482 sex-specific genes in the present STAT5a study (Table 3). Six of the seven largest groups of coexpressed genes identified in the present study are made up of sex-specific genes that did not respond to the loss of STAT5a in either males or females (*groups 4A*, *4B*, *5A*, *5B*, *7A*, and *7B*). Together, these six groups comprise 76% of the genes of interest in the STAT5a study vs. only 23% of the corresponding set of genes in the STAT5b study. Major differences in the impact of STAT5a vs. STAT5b deficiency are thus apparent.

Comparison of sex specificity among three mouse strains. M-WT:F-WT expression ratios, obtained in the present analysis of 129J × Black Swiss mouse livers, were compared with the corresponding ratios determined using the same microarray platform for two sets of male and female mouse livers from different genetic backgrounds: 129J × BALB/c mice [genetic background used in the STAT5b KO mouse study (5)] and ICR mice, an outbred strain. Normalized M-WT:F-WT expression

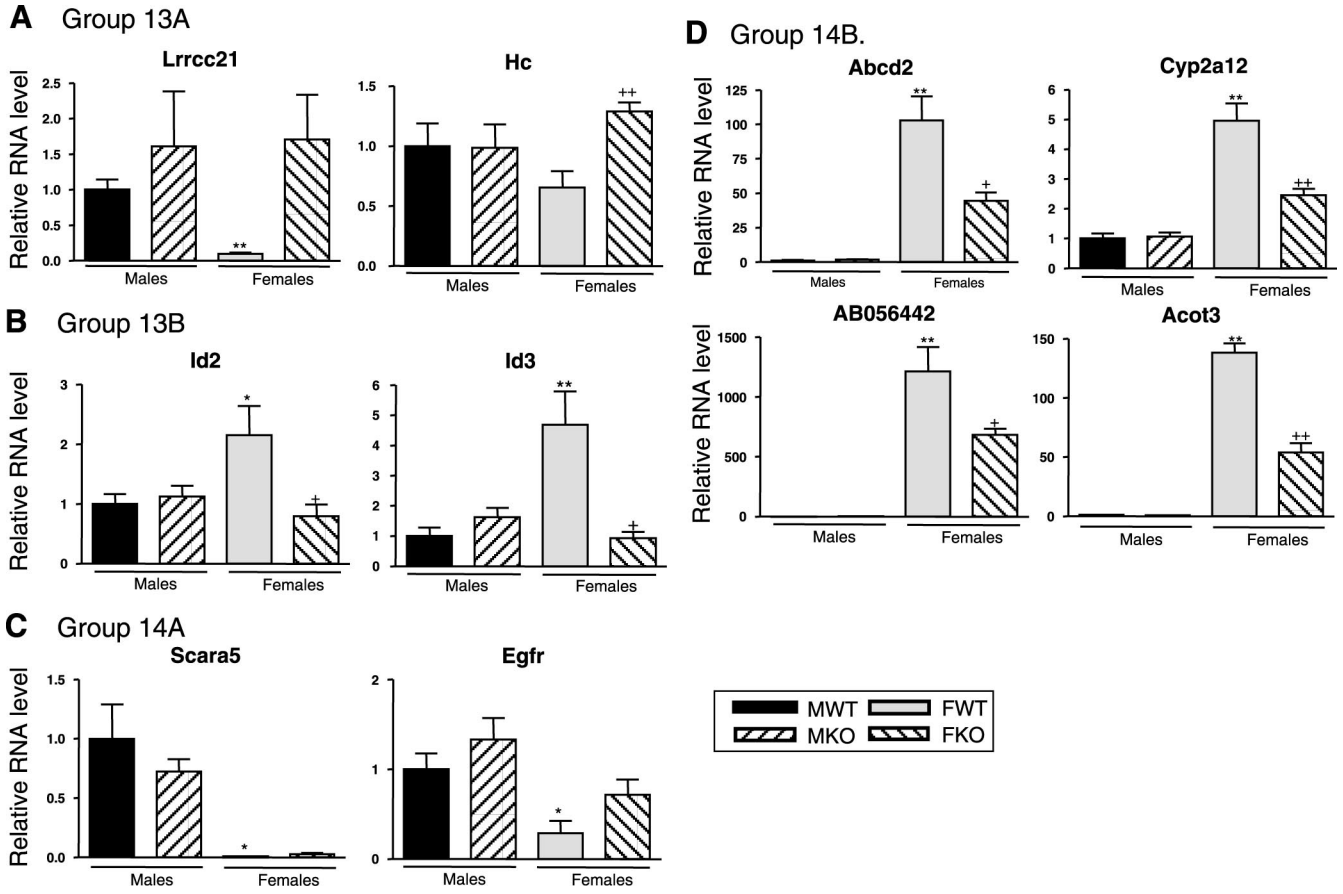

Fig. 3. Quantitative PCR analysis of select genes in *groups 13A*, *13B*, *14A*, and *14B*. RNA samples prepared from individual male and female livers, wild type (MWT and FWT, respectively) and STAT5a knockout (MKO and FKO, respectively), as indicated, were assayed for the indicated RNAs from each group as described in MATERIALS AND METHODS. Significant differences are indicated as follows. For F-WT vs. M-WT, $*P < 0.05$ and $**P < 0.01$; for M-KO vs. M-WT and F-KO vs. F-WT, $+P < 0.05$ and $++P < 0.01$. Results confirm the general trends in expression that are characteristic of each group (c.f., Tables 4 and 5).

ratios for a total of 10 microarrays (3 from 129J × Black Swiss mice, 3 from 129J × BALB/c mice, and 4 from ICR mice) were combined to obtain a single data set, of which 1,028 genes (635 male predominant, 393 female predominant) met the threshold of >1.5-fold differential expression at $P < 0.001$ (Supplemental Table S4). These 1,028 genes include 109 of the 275 of the genes (40%) whose sex specificity decreased in female liver in the absence of STAT5a (c.f., Table 2). Furthermore, 523 of the 1,028 genes (370 male predominant, 153 female predominant) met the threshold of >1.5-fold differential expression and $P < 0.05$ in all three independent data sets, indicating robust sex-specific expression across strains and experiments (Supplemental Fig. S2 and Supplemental Table S5).

Further examination of the three microarray data sets revealed genes that met the criteria for sex specificity in only one of the data sets (Supplemental Fig. S2). ANOVA analysis identified 54 genes showing sex specificity in 129J × BALB/c mice but not in the other two strains at $P < 0.001$ (Supplemental Table S6). Similarly, 49 genes were identified as sex specific in ICR mice only, while 551 genes were sex specific in 129J × Black Swiss mice only. Top candidates for these strain-dependent, sex-specific genes are presented in Supplemental Table S6.

DISCUSSION

The present study investigates the role of STAT5a in the sexual dimorphism of mouse liver gene expression, which was previously shown to be extensive and highly dependent on STAT5b, the most abundant liver STAT5 form. Whereas a majority of the sex-specific hepatic genes identified here were unaffected by the loss of STAT5a, expression of a unique subset of sex-specific genes was specifically altered in STAT5a-deficient female mouse liver. These include four *Cyp* genes, *Cyp2a12*, *Cyp2d22*, *Cyp3a16*, and *Cyp8b1*, each of which responds to the loss of STAT5a in female but not male liver (gene *groups 13A* and *14B*; Table 3). Overall, 219 genes, corresponding to 15% of female-predominant genes, were decreased in expression, whereas 5% of male-predominant genes were increased in expression in livers of female but not male mice deficient in STAT5a. The portion of female-predominant genes that required STAT5a for expression increased to 23% (89 genes) when a strain-independent set of 393 female-specific genes was considered. Thus STAT5a is likely to exhibit transcriptional effects in female mouse liver that are not compensated for by STAT5b. This finding contrasts with the extensive dependence of sex-specific gene expression on STAT5b seen in male liver, where nearly 90% of all sex specificity is lost or significantly reduced in the absence of

Table 6. *Distribution of molecular function-related GO categories and GO enrichment for 194 sex-specific genes that show a dependence on STAT5a in female mouse liver*

	GO Category No.	Gene Count	Percentage
Binding	5,488*	101	35.4
Catalytic activity	3,824	65	22.8
Signal transducer activity	4,871*	63	22.1
Transporter activity	5,215	16	5.6
Enzyme regulator activity	30,234	16	5.6
Transcription regulator activity	30,528	11	3.9
Structural molecule activity	5,198	2	0.7
Antioxidant activity	16,209	1	0.4
Molecular function unknown	5,554	10	3.5

GO Category	Total Genes in Category	%Total Genes in GO Category	Select Genes in Category	%Select Genes in Category	*P* Value
GO:4871: signal transducer activity	2,550	19.48	63	32.47	0.00001
GO:4872: receptor activity	1,960	14.97	51	26.29	0.00003
GO:5488: binding					
GO:30246: carbohydrate binding	232	1.77	13	6.70	0.00004
GO:5529: sugar binding	167	1.28	11	5.67	0.00004

Distribution of molecular function-related Gene Ontology (GO) categories and GO enrichment for 194 sex-specific genes that show a dependence on STAT5a in female mouse liver. *Top*: GO category distribution. Genes showing sex-specific expression (M-WT:F-WT >1.5 or <0.66 and $P < 0.05$) and a response to the loss of STAT5a in female liver (F-KO:F-WT > 1.25 or <0.8 and $P < 0.05$) were grouped within GeneSpring GX on the basis of the molecular function-related GO categories assigned to the genes; 194 genes met the criteria for inclusion and were annotated with the GO categories shown. Genes were also assigned all parent categories within the ontology. *Enriched category nos., presented in more detail at *bottom*. *Bottom*: enriched GO terms. Enriched GO terms were selected for overrepresentation of genes with $P < 0.0001$. The total no. of genes and the percentage of total genes within each GO category relate to all genes represented on the microarray with the indicated GO annotations. The select gene count and the percentage of selected genes in each category are those genes within the 194 genes described above that fall within the indicated GO category. *P* values were assigned by GeneSpring GX for each GO category based on whether the no. of selected genes within each category was greater than expected by random sampling of 194 genes from the total set of genes represented on the microarray, given the subset of the total no. of genes with that GO annotation. Child categories are noted below their parent categories within the GO hierarchy.

STAT5b (5). This finding is highlighted by the markedly different distribution of sex-specific genes among the major STAT5a- and STAT5b-coregulated gene groups (Supplemental Fig. S1). The precise physiological significance of this specific role of STAT5a in female liver is difficult to ascertain. Of note, however, is that four molecular functions as annotated by GO were overrepresented in the set of STAT5a-dependent, sex-specific genes in female liver: signal transduction, receptors, carbohydrate binding, and sugar binding ($P < 0.0001$).

Earlier studies suggested a role for STAT5a in the expression of a female-specific *Cyp2b* family member, based on the loss of a female-specific Cyp2b protein in STAT5a-deficient female mouse liver (24). In that case, however, the Cyp2b protein was also decreased in STAT5b-deficient females, a pattern of regulation distinct from that of the major group of STAT5a-dependent female genes described here. The role of STAT5a as a feminizing factor in female mouse liver (positive regulation of female genes and negative regulation of male genes) is analogous to that of STAT5b in male liver, where STAT5b serves as a masculinizing factor (positive regulation of male genes and negative regulation of female genes). STAT5a but not STAT5b also plays the major role in another female-specific function, mammary gland differentiation and lactogenesis in response to prolactin stimulation (19).

The present study includes an analysis of sex-specific mouse liver genes across three mixed or outbred mouse strains. Combined analysis of the data from all three strains (10 microarrays) taken as a single data set identified 1,028 genes showing sex-dependent expression at $P < 0.001$ across all three studies (Supplemental Table S4). These include liver-expressed Y-linked genes such as *Ddx3y*, *Eif2s3y*, and *Jarid1d* as well as 26 *Cyp* genes and 10 *Sult* (sulfotransferase) genes. A subset comprised of 523 genes met the threshold for sex specificity in all three mouse strain backgrounds and thus corresponds to a robust core of sex-specific hepatic genes. Forty-seven genes annotated as transcriptional regulators by GO were included among the robust sex-specific genes, including cut-like 2 (*Cutl2*), SRY-box containing gene 15 (*Sox15*), and myogenic factors 5 and 6 (*Myf5, Myf6*), a subset of which could potentially mediate the sex-specific actions of GH-activated STAT5a or STAT5b, as discussed elsewhere (34). *Cutl2*, in particular, has recently been characterized as a highly female-specific, continuous GH-regulated nuclear factor in both mouse and rat liver ($F > M \sim 100$) that could contribute to the establishment or enforcement of liver sex specificity (18).

Also identified here are genes whose liver sex specificity was apparent in only one of the three mouse genetic backgrounds investigated. Some of these genes may be subject to strain-dependent genetic regulation. Genes exhibiting significant discrepancies in expression across the three strain backgrounds are listed in Supplemental Table S6. Only a subset of these genes is likely to exhibit true strain-dependent sex specificity. Technical factors inherent in microarray analysis, as well as the fact that individual livers, rather than pools of livers, comprised the biological triplicates for one of the microarray studies, could contribute to the apparent strain-dependent loss of sex specificity for some of these genes. Nevertheless, these findings raise the possibility that the strain dependence of mouse liver sex specificity may be more exten-

sive than was previously recognized. Earlier studies identified *Rsl* alleles that conferred strain-specific regulation of certain sex-specific genes, notably *Slp*, *Cyp2d9*, and several *Mup* family members (17). In other studies, strain-specific regulation of the female-predominant *Cyp2a4* was associated with a recessive mutation at the GH-dependent repression (*Gdr*) locus that abolishes repression of *Cyp2a4* expression in males (2). Further investigation using larger numbers of arrays representing several pure outbred and inbred mouse strains, including strains containing the *Rsl* and *Gdr* mutations, will be necessary to isolate effects due to strain differences from those due to individual variability that is strain independent, as well as from technical limitations of microarray technology.

It is apparent that STAT5a plays a unique role in sexually dimorphic gene expression in female mouse liver. Presumably, these actions of STAT5a are GH regulated, insofar as the vast majority of liver sexual dimorphism, including at least some of the STAT5a-dependent genes presently identified, is dictated by plasma GH profiles (34). It is not known, however, whether the unique actions of STAT5a in female liver reported here reflect a direct action of STAT5a on the genes in *groups 13* and *14* (Table 3), or whether the requirement for STAT5a is indirect. It is also unclear why STAT5b is unable to compensate for the loss of STAT5a, given its 10- to 20-fold higher expression in liver (24) and the very close similarities in biochemical and gene regulatory activities of these two closely related (>90% identical) STAT5 proteins (7, 14). Conceivably, the unique actions of STAT5a could reflect a requirement for tetrameric STAT5a complexes for *trans*-activation via paired STAT5 response elements, insofar as STAT5a, but not STAT5b, readily forms tetrameric STAT-DNA complexes (27). Formation of DNA-bound STAT5a tetramers could be less efficient in male than in female liver owing to competition for DNA binding by the much more abundant dimeric STAT5b complexes that form in response to each incoming plasma GH pulse (3, 4). Further study will be required to investigate these and other possible mechanisms.

STAT5b has been proposed to serve as an upstream transcriptional regulator of a much larger downstream transcriptional network that leads to activation of many of the male-predominant genes and repression of the many female-predominant genes whose hepatic expression is dependent on STAT5b (34). Conceivably, STAT5a could play such a role, albeit on a much smaller scale, in female liver. Of note, 11 of the sex-specific genes dependent on STAT5a in female liver were annotated as transcriptional regulators by GO; these include inhibitor of DNA binding 2 (*Id2*) and interferon regulatory factor 2 (*Irf2*). The apparent STAT5a-dependent repression of male-predominant genes in female liver might not be a direct function of STAT5a. For example, an inhibitory nuclear factor, such as Id2, could serve as a direct target of STAT5a and contribute to STAT5a inhibition of select male-specific genes in female liver. Interestingly, Id2 and STAT5a show cooperative activity in development and maturation of mouse mammary tissue, where the loss of Id2 decreased active tyrosine-phosphorylated STAT5a levels by 40% (22). A similar pathway, involving STAT5a and other transcriptional regulators, could be responsible for the limited number of genes that exhibit STAT5a dependence in female liver. Further studies will be required to elucidate the role of STAT5a in female gene expression in liver and other tissues, including mammary gland, where STAT5a is the major STAT5 form.

GRANTS

This work was supported in part by National Institutes of Health (NIH) Grant DK-33765 (to D. J. Waxman). K. H. Clodfelter and G. D. Miles received training core support from the Superfund Basic Research Center at Boston University, NIH Grant 5-P42-ES07381, and V. Wauthier received fellowship support from the Belgian American Educational Foundation.

REFERENCES

1. **Ahluwalia A, Clodfelter KH, Waxman DJ.** Sexual dimorphism of rat liver gene expression: regulatory role of growth hormone revealed by deoxyribonucleic acid microarray analysis. *Mol Endocrinol* 18: 747–760, 2004.
2. **Aida K, Negishi M.** A *trans*-acting locus regulates transcriptional repression of the female-specific steroid 15 alpha-hydroxylase gene in male mice. *J Mol Endocrinol* 11: 213–222, 1993.
3. **Choi HK, Waxman DJ.** Growth hormone, but not prolactin, maintains low-level activation of STAT5a and STAT5b in female rat liver. *Endocrinology* 140: 5126–5135, 1999.
4. **Choi HK, Waxman DJ.** Plasma growth hormone pulse activation of hepatic JAK-STAT5 signaling: developmental regulation and role in male-specific liver gene expression. *Endocrinology* 141: 3245–3255, 2000.
5. **Clodfelter KH, Holloway MG, Hodor P, Park SH, Ray WJ, Waxman DJ.** Sex-dependent liver gene expression is extensive and largely dependent upon signal transducer and activator of transcription 5b (STAT5b): STAT5b-dependent activation of male genes and repression of female genes revealed by microarray analysis. *Mol Endocrinol* 20: 1333–1351, 2006.
6. **Darnell JE Jr.** STATs and gene regulation. *Science* 277: 1630–1635, 1997.
7. **Grimley PM, Dong F, Rui H.** Stat5a and Stat5b: fraternal twins of signal transduction and transcriptional activation. *Cytokine Growth Factor Rev* 10: 131–157, 1999.
8. **He YD, Dai H, Schadt EE, Cavet G, Edwards SW, Stepaniants SB, Duenwald S, Kleinhanz R, Jones AR, Shoemaker DD, Stoughton RB.** Microarray standard data set and figures of merit for comparing data processing methods and experiment designs. *Bioinformatics* 19: 956–965, 2003.
9. **Herrington J, Carter-Su C.** Signaling pathways activated by the growth hormone receptor. *Trends Endocrinol Metab* 12: 252–257, 2001.
10. **Holloway MG, Cui Y, Laz EV, Hosui A, Hennighausen L, Waxman DJ.** Loss of sexually dimorphic liver gene expression upon hepatocyte-specific deletion of Stat5a-Stat5b locus. *Endocrinology* 148: 1977–1986, 2007.
11. **Holloway MG, Laz EV, Waxman DJ.** Codependence of growth hormone-responsive, sexually dimorphic hepatic gene expression on signal transducer and activator of transcription 5b and hepatic nuclear factor 4alpha. *Mol Endocrinol* 20: 647–660, 2006.
12. **Horvath CM.** STAT proteins and transcriptional responses to extracellular signals. *Trends Biochem Sci* 25: 496–502, 2000.
13. **Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR, Kobayashi S, Davis C, Dai H, He YD, Stephaniants SB, Cavet G, Walker WL, West A, Coffey E, Shoemaker DD, Stoughton R, Blanchard AP, Friend SH, Linsley PS.** Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol* 19: 342–347, 2001.
14. **Ihle JN.** The Stat family in cytokine signaling. *Curr Opin Cell Biol* 13: 211–217, 2001.
15. **Jansson JO, Eden S, Isaksson O.** Sexual dimorphism in the control of growth hormone secretion. *Endocr Rev* 6: 128–150, 1985.
16. **Krebs CJ, Larkins LK, Khan SM, Robins DM.** Expansion and diversification of KRAB zinc-finger genes within a cluster including regulator of sex-limitation 1 and 2. *Genomics* 85: 752–761, 2005.
17. **Krebs CJ, Larkins LK, Price R, Tullis KM, Miller RD, Robins DM.** Regulator of sex-limitation (Rsl) encodes a pair of KRAB zinc-finger genes that control sexually dimorphic liver gene expression. *Genes Dev* 17: 2664–2674, 2003.
18. **Laz EV, Holloway MG, Chen CS, Waxman DJ.** Characterization of three growth hormone-responsive transcription factors preferentially expressed in adult female liver. *Endocrinology*. 148:3327–3337, 2007.

19. **Liu X, Robinson GW, Wagner KU, Garrett L, Wynshaw-Boris A, Hennighausen L.** Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes Dev* 11: 179–186, 1997.
20. **MacLeod JN, Pampori NA, Shapiro BH.** Sex differences in the ultradian pattern of plasma growth hormone concentrations in mice. *J Endocrinol* 131: 395–399, 1991.
21. **Mode A, Gustafsson JA.** Sex and the liver–a journey through five decades. *Drug Metab Rev* 38: 197–207, 2006.
22. **Mori S, Nishikawa SI, Yokota Y.** Lactation defect in mice lacking the helix-loop-helix inhibitor Id2. *EMBO J* 19: 5772–5781, 2000.
23. **Noshiro M, Serabjit-Singh CJ, Bend JR, Negishi M.** Female-predominant expression of testosterone 16 alpha-hydroxylase ("I"-P-450(16) alpha) and its repression in strain 129/J. *Arch Biochem Biophys* 244: 857–864, 1986.
24. **Park SH, Liu X, Hennighausen L, Davey HW, Waxman DJ.** Distinctive roles of STAT5a and STAT5b in sexual dimorphism of hepatic P450 gene expression. Impact of STAT5a gene disruption. *J Biol Chem* 274: 7421–7430, 1999.
25. **Schwartz J, Huo JS, Piwien-Pilipuk G.** Growth hormone regulated gene expression. *Minerva Endocrinol* 27: 231–241, 2002.
26. **Smyth GK, Speed T.** Normalization of cDNA microarray data. *Methods* 31: 265–273, 2003.
27. **Soldaini E, John S, Moro S, Bollenbacher J, Schindler U, Leonard WJ.** DNA binding site selection of dimeric and tetrameric Stat5 proteins reveals a large repertoire of divergent tetrameric Stat5a binding sites. *Mol Cell Biol* 20: 389–401, 2000.
28. **Teglund S, McKay C, Schuetz E, van Deursen JM, Stravopodis D, Wang D, Brown M, Bodner S, Grosveld G, Ihle JN.** Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. *Cell* 93: 841–850, 1998.
29. **Tullis KM, Krebs CJ, Leung JY, Robins DM.** The regulator of sex-limitation gene, rsl, enforces male-specific liver gene expression by negative regulation. *Endocrinology* 144: 1854–1860, 2003.
30. **Udy GB, Towers RP, Snell RG, Wilkins RJ, Park SH, Ram PA, Waxman DJ, Davey HW.** Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. *Proc Natl Acad Sci USA* 94: 7239–7244, 1997.
31. **Veldhuis JD, Anderson SM, Shah N, Bray M, Vick T, Gentili A, Mulligan T, Johnson ML, Weltman A, Evans WS, Iranmanesh A.** Neurophysiological regulation and target-tissue impact of the pulsatile mode of growth hormone secretion in the human. *Growth Horm IGF Res*11, *Suppl* A: S25–S37, 2001.
32. **Waxman DJ.** Growth hormone pulse-activated STAT5 signalling: a unique regulatory mechanism governing sexual dimorphism of liver gene expression. *Novartis Found Symp* 227: 61–74; discussion 75–81, 2000.
33. **Waxman DJ, Celenza JL.** Sexual dimorphism of hepatic gene expression: novel biological role of KRAB zinc finger repressors revealed. *Genes Dev* 17: 2607–2613, 2003.
34. **Waxman DJ, O'Connor C.** Growth hormone regulation of sex-dependent liver gene expression. *Mol Endocrinol* 20: 2613–2629, 2006.
35. **Waxman DJ, Ram PA, Park SH, Choi HK.** Intermittent plasma growth hormone triggers tyrosine phosphorylation and nuclear translocation of a liver-expressed, Stat 5-related DNA binding protein. Proposed role as an intracellular regulator of male-specific liver gene transcription. *J Biol Chem* 270: 13262–13270, 1995.
36. **Weng L, Dai H, Zhan Y, He Y, Stepaniants SB, Bassett DE.** Rosetta error model for gene expression analysis. *Bioinformatics* 22: 1111–1121, 2006.
37. **Wiwi CA, Waxman DJ.** Role of hepatocyte nuclear factors in growth hormone-regulated, sexually dimorphic expression of liver cytochromes P450. *Growth Factors* 22: 79–88, 2004.
38. **Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ.** Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res* 16: 995–1004, 2006.