

# Unbiased, Genome-Wide *In Vivo* Mapping of Transcriptional Regulatory Elements Reveals Sex Differences in Chromatin Structure Associated with Sex-Specific Liver Gene Expression<sup>∇†</sup>

Guoyu Ling,<sup>1‡</sup> Aarathi Sugathan,<sup>1‡</sup> Tali Mazor,<sup>2</sup> Ernest Fraenkel,<sup>2</sup> and David J. Waxman<sup>1\*</sup>

*Division of Cell and Molecular Biology, Department of Biology, Boston University, Boston, Massachusetts 02215,<sup>1</sup> and Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139<sup>2</sup>*

Received 24 May 2010/Returned for modification 23 July 2010/Accepted 20 September 2010

**We have used a simple and efficient method to identify condition-specific transcriptional regulatory sites *in vivo* to help elucidate the molecular basis of sex-related differences in transcription, which are widespread in mammalian tissues and affect normal physiology, drug response, inflammation, and disease. To systematically uncover transcriptional regulators responsible for these differences, we used DNase hypersensitivity analysis coupled with high-throughput sequencing to produce condition-specific maps of regulatory sites in male and female mouse livers and in livers of male mice feminized by continuous infusion of growth hormone (GH). We identified 71,264 hypersensitive sites, with 1,284 showing robust sex-related differences. Continuous GH infusion suppressed the vast majority of male-specific sites and induced a subset of female-specific sites in male livers. We also identified broad genomic regions (up to ~100 kb) showing sex-dependent hypersensitivity and similar patterns of GH responses. We found a strong association of sex-specific sites with sex-specific transcription; however, a majority of sex-specific sites were >100 kb from sex-specific genes. By analyzing sequence motifs within regulatory regions, we identified two known regulators of liver sexual dimorphism and several new candidates for further investigation. This approach can readily be applied to mapping condition-specific regulatory sites in mammalian tissues under a wide variety of physiological conditions.**

Sexual dimorphism in gene expression is common in mammalian somatic tissues (23) and has broad implications for human health. Sex differences in gene expression may contribute to differences between men and women in the prevalence, extent, and progression of disease, including autoimmune diseases (54), kidney disease (37), cardiovascular disease (45), and liver diseases, such as hepatocellular carcinoma (9, 58). In addition, sex-related differences in pharmacokinetics and pharmacodynamics are common and may affect drug response (52). Sex-related differences in gene expression have been widely studied in liver, where they affect >1,000 transcripts (5, 51, 57) and impact physiological and pathophysiological functions ranging from lipid and fatty acid metabolism to xenobiotic metabolism and disease susceptibility (52). In the liver, sex-related differences in gene expression are primarily determined by growth hormone (GH) signaling (3, 21), which shows important sex-related differences that reflect the sex-related differences in plasma GH profiles seen in many species, including rats, mice, and humans (53).

The underlying mechanisms of sexual dimorphism in mammalian tissues have been only partly elucidated at the molecular level. In the male rat liver, intermittent plasma GH pulses repeatedly activate the latent cytoplasmic transcription factor STAT5b, whose activity is essential for sex-related differences

in the liver (5). The more continuous, female-like pattern of pituitary GH secretion can be mimicked by continuous GH infusion in male mice, which abolishes the normal male, pulsatile plasma GH profile and feminizes liver transcript patterns by suppressing many male-specific genes and inducing many female-specific genes (19). In spite of these findings, the molecular mechanisms whereby STAT5b and other transcription factors regulate sex specificity in the liver have remained elusive (26, 52).

DNase I hypersensitivity (DHS) analysis is a powerful tool to identify functional DNA elements involved in gene regulation. The temporal and spatial association of DHS sites with tissue-specific and developmentally regulated gene expression has long been established (14), and several instances of sex-related differences in DNase hypersensitivity characterizing genes that show sex-dependent transcription have been reported. Early studies identified a DHS site in the mouse liver upstream of *C4a* (the gene encoding sex-limited protein) that is more prominent in males, where an open chromatin structure correlated with a male-predominant pattern of gene expression (16), and examples of sex-regulated DHS sites have been reported for two sex-specific cytochrome P450 (*Cyp*) genes in rat liver (10, 47).

In order to identify sex-related differences in mouse liver chromatin structure on a global scale, we combined DHS analysis with ultrahigh-throughput sequencing (DNase-seq) to probe the open chromatin structure at single-base-pair resolution (1, 6, 40, 46). We show that DNase-seq, whose application until now has been limited to cultured cell lines, can readily be used to map DHS sites in mammalian liver, despite the added complexity of multiple cell types. We obtained high-resolution, genome-wide DHS maps for both male and female mouse

\* Corresponding author. Mailing address: Department of Biology, Boston University, 5 Cummington Street, Boston, MA 02215. Phone: (617) 353-7401. Fax: (617) 353-7404. E-mail: djw@bu.edu.

‡ These authors are joint first authors.

† Supplemental material for this article may be found at <http://mcb.asm.org/>.

∇ Published ahead of print on 27 September 2010.

livers under physiological conditions, and we demonstrate that these maps can be utilized to identify transcriptional regulators of sex-biased liver gene expression. We characterized more than 70,000 DHS sites and show that they encompass a large percentage (approximately 65 to 90%) of binding sites for six liver transcription factors identified earlier by chromatin immunoprecipitation (ChIP)-seq. We identified 1,284 DHS sites for which there are robust sex-related differences and that may contribute to sex-dependent gene expression; ~20% of these sites mapped within 100 kb of a sex-specific liver gene. In addition, a subset of the sex-dependent DHS sites was shown to respond to continuous GH treatment in male mice, likely representing functional DNA elements that mediate hormone-dependent, sex-dependent gene expression. Finally, analysis of the sex-dependent DHS sequences for enriched motifs identified binding sites for two transcription factors, STAT5b and HNF4 $\alpha$ , known to be essential for sex-specific liver gene expression (5, 18, 20), as well as binding sites for other factors suggested to be involved (24) and several novel factors not previously implicated in liver sexual dimorphism. These findings highlight the utility of DNase-seq for elucidating condition-specific transcriptional regulatory sites associated with complex biological processes in mammalian tissue *in vivo* on a genome-wide scale.

#### MATERIALS AND METHODS

**Mouse studies.** Adult male and female ICR mice (CD-1 mice) were purchased from Taconic Farms, Inc. (Germantown, NY), or Charles River Laboratories (Wilmington, MA) and were housed in the Boston University Laboratory Animal Care Facility in accordance with approved protocols. Livers were collected from 8-week-old mice euthanized by CO<sub>2</sub> asphyxiation followed by cervical dislocation. Continuous GH treatment of 7-week-old male mice was achieved using Alzet model 1007D microosmotic pumps (Durect Corporation, Cupertino, CA) implanted subcutaneously (s.c.) under ketamine and xylazine anesthesia and delivering recombinant rat GH (obtained from Arieh Gertler, Protein Laboratories Rehovot, Ltd., Rehovot, Israel) at 20 ng rat GH per h per gram body weight for 7 days (19). RNA was extracted from individual livers by using Trizol reagent (Invitrogen, Carlsbad, CA) followed by reverse transcription using 1  $\mu$ g total RNA and a high-capacity reverse transcription kit (Applied Biosystems, Carlsbad, CA). To verify feminization of liver gene expression by continuous GH infusion, real-time PCR analysis of a panel of established, continuous-GH-responsive genes (19) was performed using Power SYBR green PCR master mix and an ABS 7900HT sequence detection system (both from Applied Biosystems).

For *in vivo* transfection assays, genomic regions corresponding to six individual DHS sites were PCR amplified from ICR mouse genomic DNA and cloned into a modified pGL4.10 vector (Promega, Madison, WI), designated pAlbpmo, that includes a minimal mouse *Alb* promoter driving expression of a firefly luciferase gene. The DHS regions were cloned into pAlbpmo, upstream of the *Alb* promoter. A renilla luciferase reporter vector was constructed in a similar way by inserting an *Alb* enhancer into a modified pGL4.70 reporter vector containing the same mouse minimal *Alb* promoter. Twelve micrograms of firefly reporter vector and 3  $\mu$ g of renilla reporter plasmids were delivered to adult male and female mouse livers by hydrodynamic injection by using the TransIT gene delivery system (Mirus, Madison, WI). Livers were harvested 7 days after injection, homogenized in 1 $\times$  passive lysis buffer (Promega), and firefly luciferase activity normalized to renilla luciferase activity was assayed using a dual-luciferase assay kit (Promega).

**Nucleus isolation.** Nuclei were isolated by using a high-sucrose-based protocol to minimize perturbation of chromatin structure during nucleus isolation (27). Livers were pooled from 2 to 5 mice, minced, and then homogenized in nuclear homogenization buffer (10 mM HEPES, pH 7.9, 25 mM KCl, 0.15 mM spermine, 0.5 mM spermidine, 1 mM EDTA, 2 M sucrose, 10% glycerol, 10 mM NaF, 1 mM orthovanadate, 1 mM phenylmethylsulfonyl fluoride [PMSF], 0.5 mM dithiothreitol [DTT]) supplemented with 1 $\times$  protease inhibitor cocktail (Sigma P8340) by using a motor-driven Potter-Elvehjem homogenizer (5 ml homogenization buffer per g liver tissue). Twenty-five milliliters of homogenate was layered

over a 10-ml cushion comprised of the same buffer and centrifuged at 25,000 rpm for 45 min at 4°C in an SW28 rotor. The pelleted nuclei were suspended in nucleus storage buffer (20 mM Tris-Cl, pH 8.0, 75 mM NaCl, 0.5 mM EDTA, 50% [vol/vol] glycerol, 0.85 mM DTT, 0.125 mM PMSF) by using a Dounce homogenizer, counted by using a hemacytometer after ~100-fold dilution in phosphate-buffered saline, and snap-frozen in aliquots of approximately  $5 \times 10^7$  to  $10 \times 10^7$  nuclei/ml at -80°C.

**DNase I digestion.** DNase I digestion was performed as described previously (40) with some modifications. Frozen nuclei were thawed on ice and washed twice with ice-cold buffer A (15 mM Tris-Cl, pH 8.0, 15 mM NaCl, 60 mM KCl, 1 mM EDTA, 0.5 mM EGTA, 0.5 mM spermidine). DNase digestion was initiated by incubating  $\sim 5 \times 10^6$  nuclei in 1 ml of buffer D (9 volumes of buffer A plus 1 volume of 60 mM CaCl<sub>2</sub>, 750 mM NaCl) for 2 min at 37°C with an optimized amount of RQ1 DNase I (Promega) (see below). Six tubes of nuclei were incubated in parallel. DNase digestion was halted by the addition of 1 ml stop buffer (50 mM Tris-Cl, pH 8.0, 100 mM NaCl, 0.1% SDS, 100 mM EDTA) to each tube, followed by proteinase K digestion (100  $\mu$ g/ml final concentration) overnight at 55°C. The next day, phenol-chloroform extraction was performed; after this, the aqueous phase (approximately 11 to 12 ml combined from the 6 parallel DNase digestion reactions) was removed and adjusted to 0.8 M NaCl by the addition of 5 M NaCl. Control samples were prepared in the same way by incubating  $\sim 5 \mu$ g purified genomic DNA in 0.1 ml digestion buffer with 0.0625, 0.125, 0.25, or 0.5 units of DNase I. The control DNase I digestion sample that yielded a smear of DNA fragments ranging from 100 bp to ~1.5 kb was selected and further purified by sucrose gradient ultracentrifugation.

Small DNA fragments (<1.5 kb) released during DNase I digestion were isolated by sucrose step gradient ultracentrifugation of the DNase-digested nuclei and also from the DNase-digested genomic DNA and the sonicated genomic DNA control samples. The small DNA fragments released from the digested nuclei correspond to DNase-sensitive regions containing multiple cut sites in close proximity, whereas larger DNA fragments present in the samples primarily result from random, nonspecific DNA degradation. The released fragments were size selected on a sucrose step gradient prepared by sequentially layering 3 ml of each sucrose concentration (20 mM Tris-Cl, pH 8.0, 5 mM EDTA, 1 M NaCl containing 40%, 35%, 30%, 25%, 20%, 17.5%, 15%, 12.5%, or 10% sucrose) in an SW28 tube. Half of each sample (~5.7 ml) was loaded on the top of each gradient, and the gradient was centrifuged for 24 h at 25,000 rpm at 25°C in an SW28 rotor. Fractions of 1.9 ml were collected from the top and assayed by agarose gel electrophoresis. Gels were stained with 1 $\times$  SYBR green I nucleic acid gel staining solution (Invitrogen) in 1 $\times$  TAE buffer (40 mM Tris-acetate, 2 mM EDTA, pH 8.5) at room temperature for 30 min and visualized on a Typhoon imager (GE Healthcare, Piscataway, NJ). Fractions that primarily contained DNA fragments of <1.5 kb (typically fraction 7) were purified on Qiagen columns (catalog no. 28704) and assayed by quantitative PCR (qPCR) using primers designed to amplify known hypersensitive genomic regions (see Table S1 in the supplemental material). The typical yield was 1 to 2 ng released DNA fragments per 10<sup>6</sup> nuclei, as determined using the Quanti-IT kit (Invitrogen). Illumina sequencing (see below) was carried out for each of two independent pools of biological replicates (4 to 6 livers for each sex in each replicate). Each sequenced sample was comprised of ~30 ng DNA pooled from at least 3 independent batches of DNase-digested nuclei, to minimize the impact of inter-sample variability in DNase digestion.

**Optimization of DNase digestion conditions.** qPCR primers used to optimize DNase digestion conditions and to assess the quality of DNase-released fragments are listed in Table S1 in the supplemental material: Intergenic-1 primers were used to amplify a genomic region distant (>100 kb) from any known genes, where no DHS regions were expected, and *Rassf6* primers were selected to flank a strong DHS site near the *Rassf6* promoter. For each batch of DNase I enzyme, initial experiments were carried out over a range of DNase I concentrations (5 to 120 units/ml) to identify a DNase concentration (typically ~40 units DNase I/ml) that resulted in <20% gene copy number loss in the Intergenic-1 region but >50% gene copy number loss in the *Rassf6* promoter region. Purified, released DNase fragments were routinely tested for enrichment of *Alb* promoter sequences over intergenic region sequences to verify the quality of the DHS samples prior to Illumina sequencing analysis. This was done by qPCR using *Alb* primers, located within a strong DHS site within the *Alb* promoter, and Intergenic-2 primers, which are in a region near Intergenic-1 primers. Typically, the DNase-released *Alb* promoter fragments showed >16-fold enrichment compared to control genomic DNA. Sex-dependent release of genomic DNA fragments was also routinely verified for a male-specific DHS region near *Ttc39c* and a female-specific DHS region near *Cyp2b9* (see Table S1).

**Illumina sequencing.** Sequences of DNase I-released DNA fragments were sequenced using a Genome Analyzer II instrument (Illumina, Inc., San Diego,

CA). Briefly, about 30 ng of DNase-released DNA fragments was subjected to end repair, adaptor ligation, and PCR enrichment using an Illumina sample preparation kit following the manufacturer's recommendations. A DNA library comprised of approximately 100- to 300-bp fragments was size purified by gel electrophoresis. The concentration of properly ligated samples was estimated by qPCR, and this was followed by cluster generation and sequencing. The sequencing reads were aligned to mouse genome mm9 by using Illumina's Eland extended software, with a maximum of 2 mismatches allowed in the first 25 bp. Totals of 36 million, 32 million, and 28 million reads were sequenced from the male, female, and GH-treated male samples, respectively, with ~82% mapped to unique genomic positions (see Table S2 in the supplemental material).

**Identification of DHS sites.** PeakSeq (39) was modified as outlined below and used to identify DHS sites in male, female, and GH-treated male mouse liver nuclei in comparison to control samples; the latter consisted of sonicated mouse genomic DNA, as well as DNase I-digested mouse genomic DNAs, and were processed in parallel to the DNase I-digested nuclei. Sex-specific DHS sites and sites induced or suppressed in male mouse livers by continuous GH treatment were also identified. The PeakSeq algorithm was modified for identification of DHS peaks as follows. (i) Sequence read numbers from the two samples being compared were square-root normalized to improve the linearity of the data for calculation of the scaling factor by linear regression. (ii) When comparing sequence reads from male and female mouse livers, or from male and continuous GH-treated male mouse livers, only putative peak regions were included in linear regression due to the presence of differing amounts of background in each sample (see Table S3 in the supplemental material). (iii) Minimum thresholds of 7 sequence reads for autosomes and 5 sequence reads for sex chromosomes were applied to all putative DHS peaks to eliminate unusually long ( $\geq 10$ -kb) peaks with few sequence reads. (iv) Finally, putative peaks that were  $< 100$  bp in length were extended to 100 bp and then evaluated for statistical significance. Further details are provided in the supplementary methods and results in the supplemental material.

**Association of DHS sites with genes, H3K4 methylation, and FoxA2 binding sites.** DHS sites were classified as intergenic, coding, or associated with promoter regions based on mapping to known genes, mRNAs, and spliced and unspliced expressed sequence tags (ESTs) in the University of California—Santa Cruz (UCSC) genome browser. To compare DHS sites with liver gene expression, the locations of sex-independent and sex-specific DHS sites were mapped to sex-specific and sex-independent genes expressed in mouse liver (50). ChIP-seq data for histone H3 lysine 4 monomethylation (H3K4-me1), histone H3 lysine 4 trimethylation (H3K4-me3), and FOXA2 binding in female mouse livers (38) were compared to locations of sex-independent and sex-specific DHS sites. The association of DHS sites with the two histone modifications listed above was calculated by determining the numbers of DHS sites that have a histone modification site within 150 bp of either side of the DHS site. To generate distribution plots, distances from the midpoints of DHS sites to sequence tags in peaks previously determined by ChIP-seq for H3K4-me1, H3K4-me3, and FOXA2 were computed.

**Gene expression data.** Liver gene expression data were obtained from an earlier study (50), in which a total of 1,380 transcripts showed significant differences between expression levels in male and female mouse livers. After the removal of microarray probes that do not map to any known gene, probes that map to the same transcript as another probe, and probes mapping to chromosomes for which no DNase hypersensitivity data are present (chrY, chrUn, and chrN\_random), a total of 1,209 genes showing  $> 2$ -fold sex-related differences between expression levels at a  $P$  value of  $< 0.005$  (i.e., sex-specific genes) remained, as did 21,153 sex-independent genes. Of these, 343 sex-specific genes and 7,341 sex-independent genes met the additional criterion of a microarray signal intensity of  $\geq 500$  in the liver.

**Identification of broad DHS regions.** SICER (59), an algorithm that uses a clustering approach to identify extended enriched domains from histone modification ChIP-seq data, was applied to detect broad regions of the genome that are enriched for DNase-seq reads in male or female liver samples compared to controls. Genomic regions that show sex-dependent DNase hypersensitivity were defined as those that were significantly enriched in DNase-digested liver nuclei of males compared to those of females or vice versa. Similarly, genomic regions that responded to continuous GH treatment were identified by comparison of the untreated male samples to the GH-treated male samples. A window size of 200 bp and a gap size of 1,200 bp were used, and significant regions that had a false-discovery rate (FDR) of  $< 10^{-3}$  and  $\geq 2$ -fold difference in expression levels of the pair of liver samples being compared were chosen.

**Enriched transcription factor binding sites in sex-specific DHS sites.** THEME, a hypothesis-based algorithm that tests whether a given motif separates a foreground set of sequences from a background set (29), was used to identify

enriched motifs in 18 sets of sex-specific DHS sites compared to sex-independent DHS sites, as follows: (i) male-specific sites either within 10 kb or up to 50 kb from the transcription start site (TSS) of a sex-specific gene, (ii) male-specific sites within 10 kb or 50 kb of a sex-independent gene, and (iii) male-specific sites distant from any gene. Each set of sex-specific DHS sites was divided into subgroups that respond and subgroups that do not respond to continuous GH treatment in males with a  $> 2$ -fold difference between expression levels at a  $P$  value of  $< 0.01$ , and similar criteria were used for female-specific DHS sites. Transcription factor binding profiles for 97 families of transcription factors were generated by clustering the vertebrate transcription factor position-specific scoring matrices (PSSMs) from the TRANSFAC and JASPAR databases (2, 33). The corresponding 97 motifs were considered enriched if they met the following conditions: a cross-validation error of  $< 0.4$ ,  $P$  value of  $< 0.001$ , normalized log-likelihood ratio score of  $> 0.4$ , and  $\geq 2$  enrichment compared to sex-independent sites. Motifs with a PSSM total information content (relative entropy)  $< 8$  bits were eliminated from further consideration. To identify the transcription factor(s) associated with each motif, the refined family binding profiles were matched back to the TRANSFAC and JASPAR databases by using STAMP (32), and the top factor(s) that matched with an  $E$  value of  $< 10^{-8}$  was identified. One exception was motif 44, for which the best matches, to Fox factors, had an  $E$  value of  $< 10^{-6}$ . Discovered motifs were clustered by hierarchical clustering by Pearson correlation of fold enrichment over sex-independent sites in each of the 18 sets of sex-specific sites by using the hierarchical clustering module of the GenePattern suite of tools (36).

**Microarray data accession numbers.** All high-throughput sequencing data are available in the GEO database (accession no. GSE-21777), in the Sequence Read Archive of NCBI (accession no. SRP002445), and in custom tracks submitted to the UCSC genome browser.

## RESULTS

**Generation of liver DHS maps.** DNase-seq was used to generate genome-wide DHS maps for liver tissue obtained from male and female mice and from male mice given a continuous infusion of GH for 7 days, which feminizes the pattern of liver gene expression (19). Mouse liver nuclei prepared from two independent pools of biological replicates were digested with DNase I under optimized conditions (see Materials and Methods), and fragments released from hypersensitive regions were separated from randomly cut DNA fragments, which tend to be much larger (40). DNase-released fragments ranging from approximately 100 to 300 bp were sequenced by using Illumina sequencing technology. The final combined data set is comprised of 29 million sequence reads mapped to unique locations in the mouse genome for male livers and 26 million reads for female livers; 23 million additional sequence reads were obtained for continuous (7-day)-GH-treated male livers (~82% uniquely mappable reads; see Table S2 in the supplemental material). The resultant DHS maps are of high quality, as seen in Fig. 1A for the *Alb* gene region. Eight DHS regions were identified within ~47 kb of the *Alb* gene TSS, with very low background between peaks of hypersensitivity. In addition to identifying the DHS sites at kb  $-0.1$ ,  $-3.5$ ,  $-10.8$ , and  $-13.7$  relative to the *Alb* TSS, previously identified by using classical Southern blotting methods (28), we identified DHS sites at four upstream locations, from kb  $-22$  to  $-47$  (Fig. 1A). Closer examination of the kb  $-13.7$  DHS site revealed a typical structure for a DHS peak, with a roughly symmetric distribution of positive- and negative-strand digestion sites that clearly define the DHS peak boundary (see Fig. S1B in the supplemental material). Using DNase I-digested genomic DNA as a control, we identified 71,264 DHS sites in male and female livers, covering 1.8% of the mouse genome. A total of 48,762 of the DHS sites were high-stringency sites, and the total number of DHS sites increased to 110,785 when the com-



TABLE 1. DHS sites in male and female mouse livers<sup>a</sup>

Sites	No. of sites		
	High stringency <sup>b</sup>	Standard stringency <sup>c</sup>	Low stringency <sup>d</sup>
Total	48,762	71,264	110,785
Male specific	850	2,315	2,800
Female specific	434	1,064	1,382

<sup>a</sup> Male and female DHS sites (compared to controls) determined by PeakSeq analysis were merged to generate a single list of total DHS sites in liver. Male-specific and female-specific DHS sites were determined by comparing male and female samples to each other. Two biological replicates each from male and female mice were combined for each comparison. Stringency differences are as indicated and as described in Methods in the supplemental material, and full lists of sites are presented in Table S5 in the supplemental material.

<sup>b</sup> Sites for which the *P* value was <0.01 and there was a >2-fold difference in expression levels and which were confirmed separately in both sets of biological replicates.

<sup>c</sup> Sites for which the *P* value was <0.01 and there was a >2-fold difference in sequence reader and which were confirmed in at least 2 of the 3 combined samples (male, female, and GH-treated male) for sex-independent DHS sites or sites which were found in at least one individual set of biological replicates for sex-specific DHS sites.

<sup>d</sup> All sites for which the *P* value was <0.01 and there was a >2-fold difference in expression levels in the either the combined male replicate or the combined female replicate sequencing data sets.

bined data sets were used (Table 1; see Table S4 in the supplemental material). There is a high degree of overlap between DHS sites and transcription factor binding sites identified by chromatin immunoprecipitation (ChIP)-chip or ChIP-seq analysis of mouse livers (Table 2), ranging from 67% for that corresponding to CEBP4A (42) to 93% for that corresponding to FXR/NR1H4 (49). Thus, the DHS sites identified here likely include a large fraction of the active regulatory elements in liver tissue.

#### DHS sites showing sex specificity and responsiveness to GH.

We hypothesized that liver chromatin is characterized by sex-associated differences between accessibilities to DNase and that these differences relate to the observed sex-associated differences in liver gene expression. We further anticipated that continuous GH treatment of male mice, which feminizes the overall pattern of liver gene expression (19), would alter the sex-dependent patterns of chromatin accessibility. By comparing the DHS profiles of male and female mouse livers, we identified genomic regions showing DNase I fragment release from male mouse liver nuclei that was significantly greater than that from female mouse liver nuclei, i.e., male-specific DHS sites. Correspondingly, female-specific DHS sites showed significantly greater DNase I cleavage in female mouse liver.

TABLE 2. Overlap between ChIP-chip/ChIP-seq transcription factor binding sites and DHS sites in the liver<sup>a</sup>

Transcription factor	All binding regions		Binding regions containing specific motif	
	No. of sites	% overlap with DHS sites	No. of sites	% overlap with DHS sites
FOXA2	10,958	68	4,111	76
ER $\alpha$ /ESR1	6,272	86	3,292	89
FXR/NR1H4	7,794	93	3,388	92
SREBP-1	426	91	293	94
CEBPA	29,188	67	21,091	67
HNF4 $\alpha$	20,355	90	13,679	91

<sup>a</sup> Sixty-seven to 93% of experimentally determined transcription factor binding sites in mouse liver overlap standard-stringency liver DHS sites. The extents to which these sites overlap with DHS sites were very similar when only the subset of sites that contain the respective binding motif for the factor used for ChIP was considered (last column). Data for FOXA2 (38), ER $\alpha$ /ESR1 (11), FXR/NR1H4 (49), SREBP-1 (43), and CEBPA and HNF4 $\alpha$  (42) are based on ChIP-chip or ChIP-seq in various mouse strains. Binding motifs were obtained from TRANSFAC and also identified *de novo* in each set of binding sites using the motif discovery module of CisGenome.

Totals of 850 male-specific peaks and 434 female-specific peaks were identified as high-stringency sex-specific DHS sites based on their confirmation in each of two independent biological replicates (Table 1); examples are shown in Fig. 1B (*Cyp2d9* gene) and C (*Cux2* intron 2) and in Fig. S4 in the supplemental material. A total of 4,182 sex-specific DHS sites were identified at lower stringencies (Table 1; see Tables S5B to E in the supplemental material). Continuous GH treatment of male mice suppressed 82% of the high-stringency male-specific DHS sites and induced 26% of the high-stringency female-specific DHS sites, whereas <3% of sex-independent DHS sites were GH responsive at the stringencies applied (Table 3). When weaker GH responses were included, 98% of male-specific DHS sites were suppressed and 44% of female-specific DHS sites were induced (see Table S5H).

Examination of the distribution of the 1,284 high-stringency sex-specific sites across chromosomes revealed significant enrichment of female-specific DHS sites on chromosomes 5 and X and enrichment of male-specific DHS sites on chromosomes 3 and 18 compared to the overall list of DHS sites (see Table S6 in the supplemental material). Overall, 65% of sex-specific DHS sites are in the coding region or within 5 kb of the TSS of a known transcript, compared to 78% of sex-independent DHS sites (see Fig. S5). The median lengths of sex-specific and sex-independent DHS were similar, 466 to 575 bp and 437 to

FIG. 1. DHS sites associated with the *Alb* gene (A), *Cyp2d9*, a male-specific gene (B), and *Cux2*, a female-specific gene (C). In panel A, DHS sites appear as sharp, narrow peaks and are marked by distance upstream of the *Alb* gene. The four sites marked in red are the same as those discovered earlier using conventional methods (28). Green arrows at the bottom mark regions of high species conservation (Cons) that coincide with DHS sites. In panel B, male-specific DHS sites in the region of *Cyp2d9* are marked as horizontal bars in the merged DHS sites track, with dark blue used for high-stringency male-specific DHS sites and lighter shades of blue for standard- and low-stringency male-specific DHS sites, numbered as described in Table S5A in the supplemental material. Sex-independent DHS sites are shown in gray. Continuous GH (GH<sub>contin</sub>) treatment of male mice suppressed each of the four high-stringency male DHS sites down to female levels (bottom track). Panel C shows data for intron 2 of *Cux2*, a highly female-specific gene (female/male expression ratio, ~100 [24]), showing female-specific DHS sites (marked as pink horizontal bars), four of which are markedly induced in continuous-GH-treated male livers, as indicated by red asterisks. In all panels, individual sequence reads (35 nt) are represented as a single-nucleotide-wide bar graphed at the chromosomal location of the DNase cut site; this presentation is clearest in panel B, where individual sequence reads can be seen at the resolution presented. Green and red indicate sequence reads from DNase-digested female livers and blue and yellow indicate sequence reads from DNase-digested male livers or continuous GH-treated male livers, as marked. Green and blue represent positive-strand reads, and red and yellow represent negative-strand reads.

TABLE 3. Effect of continuous GH treatment on DHS sites in the male liver<sup>a</sup>

DHS sites ( <i>n</i> )	No. (%) of sites affected by continuous GH treatment in males	
	Induced	Suppressed
Male specific (850)	0	693 (82)
Female specific (434)	114 (26)	1 (0.2)
Sex independent (68,683) <sup>b</sup>	102 (0.1)	1,848 (3)

<sup>a</sup> Numbers of high-stringency sex-specific and standard-stringency sex-independent DHS sites that were induced or suppressed in continuous GH-treated male livers ( $P < 0.01$ ;  $>2$ -fold difference between expression levels of GH-treated males and untreated males). Values in parentheses are percentages of numbers of DHS sites shown in the leftmost column. The GH responses for lower-stringency sex-specific DHS sites and GH responses at lower stringency ( $P < 0.05$ ) are summarized in Table S5H in the supplemental material.

<sup>b</sup> This represents the 71,264 standard-stringency DHS sites after removal of sites that show sex specificity at low stringency.

483 bp, respectively (see Table S7), corresponding to the depletion of  $\sim 2$  nucleosomes.

**Association of liver DNase hypersensitivity with liver gene expression.** In CD4<sup>+</sup> T cells, the probability that a given gene harbors a 5'-end-proximal DHS site increases with the level of gene expression (1). We observed the same trend in mouse liver, where the proportion of sex-independent genes that have a DHS site within 200 bp of the TSS increased with increasing intensity of gene expression, leveling off at  $\sim 90\%$  (Fig. 2A). In contrast, the proportion of genes that show sex-specific expression (50) and have a 5'-end-proximal DHS site increased more gradually with increasing expression ( $P = 0.0006$ ; Fig. 2B). The overall lower percentage of sex-specific genes with a 5'-end-proximal DHS site might indicate that these genes are more commonly regulated by distal elements. Alternatively, the sex-independent genes might simply be close to more nonfunctional DHS sites than are sex-specific genes.

Next, we tested the hypothesis that genes that show sex-specific expression in mouse liver are more likely to be associated with sex-specific DHS sites than are sex-independent genes. Supporting this hypothesis, we observed that sex-specific genes are 8.1-fold more likely than sex-independent genes to have a sex-specific DHS site in the coding region and 3.1-fold more likely to be within 100 kb; furthermore, the proportion of genes with a sex-specific DHS site rises with distance more steeply for sex-specific genes than for sex-independent genes (Fig. 3A, left panel). Conversely, sex-specific DHS sites are more likely than sex-independent DHS sites to be located near a sex-specific gene (Fig. 3B, left panel). Finally, the proportion of male-specific genes whose nearest DHS site is also male specific (20%) is  $\sim 10$ -fold greater than the proportion whose nearest DHS site is either female specific or sex independent (2% in both cases), and the case is similar for female-specific genes and female-specific DHS sites (Fig. 3C). Thus, there is a strong association between sex-specific DHS sites and sex-specific gene expression. However, this association is seen for only a subset of sex-specific genes, insofar as only 20% of sex-specific genes have a high-stringency sex-specific DHS site in the coding region and only 43% have a sex-specific DHS site within 100 kb. This compares to 90% of liver-expressed sex-independent genes with a sex-independent DHS site in the coding region and 99% with at least one sex-independent DHS

site within 100 kb (Fig. 3A, right panel). Moreover, only 23% of sex-specific DHS sites are within 100 kb of a sex-specific gene, while 76% of sex-independent DHS sites are within 100 kb of a sex-independent gene (Fig. 3B, right panel). Sex-specific DHS sites may therefore act as distant regulators. Alternatively, this finding may reflect more complex regulatory mechanisms of sex-specific genes, involving interactions between multiple regulatory sites and multiple genes (30), regulatory changes that have no effect on chromatin structure, or posttranscriptional regulation. The subsets of sex-specific genes that do and do not have a sex-specific DHS site within 100 kb include equal proportions of male-specific and female-specific genes; however, the female-specific genes that are within 100 kb of a sex-specific DHS site are 2.7-fold enriched ( $P < 10^{-4}$ ) for the subset of female-specific genes that are suppressed in the female liver upon ablation of pituitary and GH stimulation (class I female-specific genes [50]). The extensive loss of male-specific DHS sites and the induction of a substantial fraction of female-specific DHS sites in livers of continuous-GH-treated male mice (Table 3), in which the gene expression profile is feminized (19), support the conclusion that these sex-specific DHS sites play a functional role in the sex-specific expression of the genes associated with these sites. Indeed, the subset of female-specific DHS sites that respond to

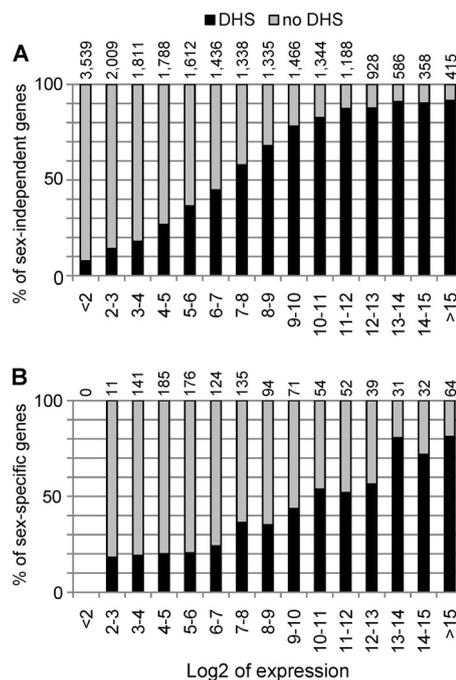


FIG. 2. Relationship between the level of gene expression and the presence of a 5' DHS site. Shown is the fraction of sex-independent genes (A) or the fraction of sex-specific genes (B) that have at least one 5' DHS site within 200 bp of the TSS, graphed as a function of log<sub>2</sub> liver expression level (microarray signal intensity). Sex-specific genes are those that show  $>2$ -fold sex-related expression differences at a  $P$  value of  $<0.005$  as determined by microarray analysis (50), using expression intensities for the sex showing higher expression. The total number of genes at each range of expression intensity is shown above each bar. The proportions of sex-independent and sex-specific genes that have a 5' DHS site at increasing expression levels are significantly different (Wilcoxon signed-rank test,  $P = 0.0006$ ).

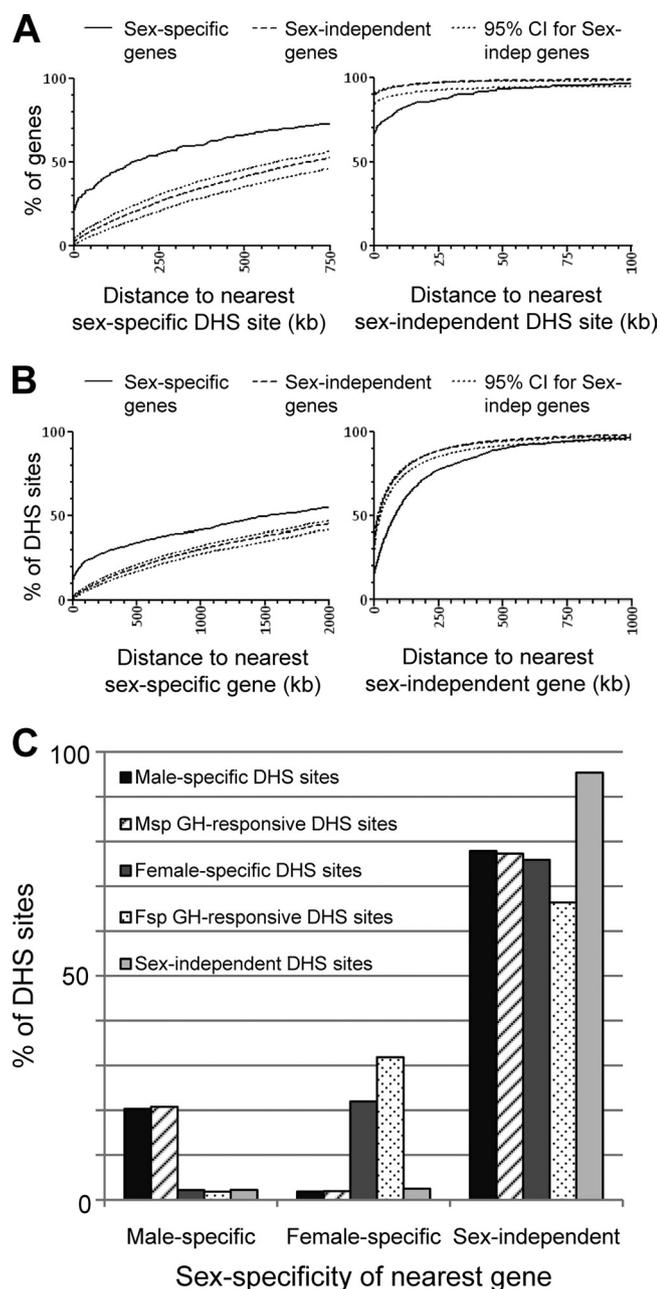


FIG. 3. Proximity of sex-specific and sex-independent (indep) DHS sites to genes expressed in mouse liver. CI, confidence interval. (A) Cumulative distances from the gene body of sex-specific and sex-independent genes to the nearest sex-specific and sex-independent DHS sites. (B) Cumulative distances from sex-specific and sex-independent DHS sites to the nearest sex-specific or sex-independent gene. For both panels A and B,  $x$  axes are on different scales in the left and right panels. For both panels, gene expression in the liver is defined as a microarray signal intensity of  $\geq 500$  in either male or female liver, and sex-specific genes are those that met the combined criteria of  $>2$ -fold sex-related difference in expression at a  $P$  value of  $<0.005$  in a microarray comparison of male and female mouse livers (50). (C) Sex specificity of the genes closest to the indicated categories of sex-specific and sex-independent DHS sites. Shown are the percentages of DHS sites whose closest gene shows male-specific, female-specific, or sex-independent expression, as indicated, using the criteria described above for sex specificity but independent of microarray signal intensity. Msp, male specific; Fsp, female specific.

continuous GH is even more frequently associated with female-specific genes than the full set of female-specific DHS sites (Fig. 3C).

**Enhancer-like activity of sex-specific DHS sites.** Our finding described above that sex-specific DHS sites are more likely to be associated with genes of the same sex specificity suggests that sex-specific DHS sites serve as enhancers of sex-specific gene expression. This possibility is supported by a comparison of our DHS map with maps of histone H3 lysine 4 mono- and trimethylation (H3K4-me1 and H3K4-me3, respectively) reported for female mouse livers (38): 80% of high-stringency female-specific DHS sites are within 150 bp of nucleosomes marked by H3K4-me1 but not H3K4-me3, whereas only 15% are associated with a H3K4-me3 mark (Fig. 4A). This pattern—the presence of H3K4-me1 in the absence of H3K4-me3—is indicative of an enhancer (15). A smaller proportion of sex-independent DHS sites exhibit an enhancer-like H3K4 methylation profile, with 32% of these DHS sites containing the H3K4-me3 mark and only 61% showing the H3K4-me1-only pattern (Fig. 4A). The frequency of the H3K4-me1–H3K4-me3 double mark decreased dramatically with increasing distance from the promoter, as was seen for both female-specific and sex-independent DHS sites (Fig. 4A). Peaks for both histone marks exhibited a trough at the midpoint of female-specific and sex-independent DHS sites, indicating nucleosome depletion (Fig. 4B and C).

To assay for enhancer activity, we selected 6 sex-specific DHS sites, 5 of which were responsive to continuous GH treatment (see Table S8 in the supplemental material), and cloned them into a reporter vector containing a modified *Alb* promoter linked to a luciferase reporter (56). The 6 sex-dependent DHS sites were assayed for their intrinsic ability to enhance the *Alb* promoter following *in vivo* liver transfection by hydrodynamic injection. Five of the six sites exhibited enhancer activity when assayed 7 days after liver transfection (Fig. 5). This time point was selected to allow for decay of the transiently high activity of the *Alb* promoter by using this transfection method (56). The most active DHS fragment, from intron 2 of the highly female-specific *Cux2* gene (see Table S8) (24), was  $>200$ -fold more active than the *Alb* promoter alone but showed similar activities in male and female mouse livers. Two male-specific DHS sites showed 8- to 17-fold higher activity than the *Alb* promoter, with the activity seen for a *Cyp2d9* DHS site in male livers being 3-fold higher than that in female livers (Fig. 5). A female-specific DHS site adjacent to *Acot4* showed female-enriched enhancer activity, albeit at a modest level. One of the six DHS sites (*Cyp2c39*) was inactive.

**Broad genomic regions of DNase hypersensitivity.** While 87% of the above-identified DHS sites are  $<1$  kb in length, we observed genomic regions with considerably longer sex-dependent hypersensitivity regions, some extending up to  $\sim 100$  kb. To identify such extended DNase hypersensitivity regions, we used SICER (59), a clustering-based algorithm designed to identify diffuse domains of ChIP-enriched regions. We found 3,971 DHS regions  $>10$  kb in length, 58 of which showed significant sex-related differences (Table 4). Continuous GH treatment suppressed 47% of the extended male-specific regions and induced 50% of the female-specific regions, compared to  $<0.2\%$  of the sex-independent regions; the proportion of these  $>10$ -kb female-specific regions that are induced

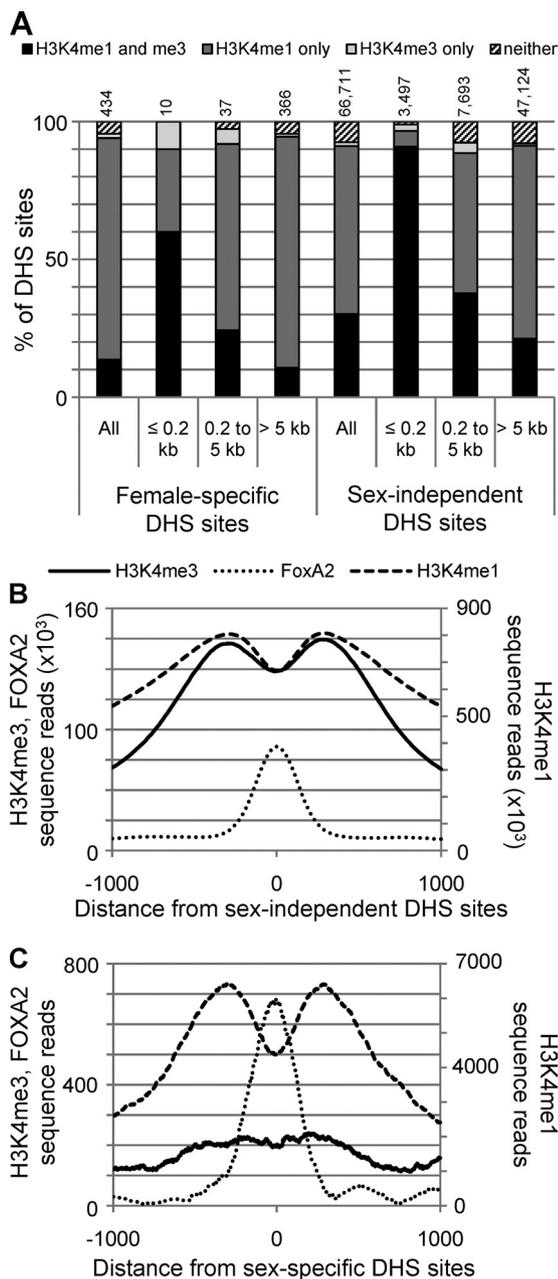


FIG. 4. Histone modifications at sex-independent and sex-specific DHS sites in female mouse livers. (A) Percentages of sex-independent and sex-specific DHS sites within 150 bp of an H3K4-me1 and/or H3K4-me3 mark in female mouse livers. These data are shown for all DHS sites and for the subsets comprised of sites that are proximal (within 200 bp of or between 0.2 and 5 kb from an active sex-specific or sex-independent TSS) to liver-expressed genes (microarray signal intensity of  $\geq 500$ , as in Fig. 3) and those that are distal to genes ( $>5$  kb from the TSS), regardless of liver expression intensity level. Proportions of DHS sites associated with the different patterns of H3K4 methylation marks are statistically significant (between all sex-specific and all sex-independent DHS sites, the  $P$  value was  $<10^{-14}$ ; between sites  $\leq 200$  bp, 0.2 to 5 kb, and  $>5$  kb from the TSS, the  $P$  values were  $<10^{-21}$  for sex-independent sites and  $<10^{-5}$  for sex-specific sites [ $\chi^2$  test]). Distributions of H3K4-me1, H3K4-me3, and FoxA2 ChIP-seq reads in peaks relative to sex-independent (B) and female-specific (C) DHS sites are also shown.

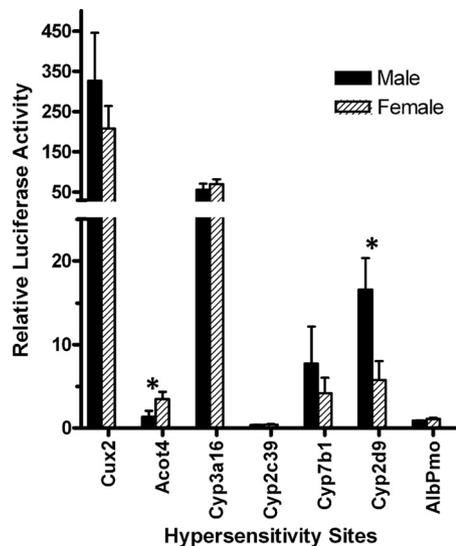


FIG. 5. Enhancer activity of six sex-specific DHS sites assayed in mouse liver. Four high-stringency female DHS sites (Cux2, Acot4, Cyp3a16, and Cyp2c39) and two male DHS sites (Cyp7b1 and Cyp2d9), named after the closest gene to each site, were assayed for reporter gene activity after *in vivo* transfection of mouse livers by hydrodynamic injection. All 4 female-specific DHS sites are associated with H3K4-me1 and not H3K4-me3 marks (see Table S8 in the supplemental material). Data shown represent normalized luciferase activity for at least 3 individual mice (mean  $\pm$  standard error [SE]), and each value is based on the average reading from 3 separate pieces of the liver. Relative luciferase activity of the parental vector pAlbpmo (AlbPmo) in male and female livers (mean value) was set to 1. \*, significant difference in activity between male and female liver transfections at a  $P$  value of  $<0.05$ .

by GH is even higher than that for the short female-specific DHS peaks. Some of the extended DHS regions are comprised of clusters of the short DHS peaks identified above (Fig. 6A and B; see also the supplemental text and Fig. S6 in the supplemental material), while other extended DHS regions contain few sites identified as DHS peaks by PeakSeq, which is

TABLE 4. Broad DHS regions and their response to continuous GH treatment<sup>a</sup>

DHS sites	No. of >10-kb regions	No. (%) of regions affected by continuous GH treatment in males	
		Induced	Suppressed
Male specific	34	0	16 (47)
Female specific	24	12 (50)	0
Sex independent <sup>b</sup>	3,913	0	7 (0.2)

<sup>a</sup> Numbers of sex-specific and sex-independent hypersensitivity regions that are  $>10$  kb in length and numbers of each that were induced or suppressed in continuous-GH-treated male mouse livers, as determined by SICER analysis. All DHS regions are significant, with an FDR of  $<10^{-3}$ , and were  $>2$ -fold enriched for sequence reads in the relevant comparison (male versus control or female versus control for sex-independent regions, male versus control and male versus female for male-specific regions, female versus control and female versus male for female-specific regions, and male versus male plus continuous GH). Values in parentheses are percentages of numbers of regions shown in the leftmost column. Complete lists of the SICER-determined DHS regions are in Table S9 in the supplemental material.

<sup>b</sup> DHS regions significant in comparisons of males and controls or females and controls merged into a single list, excluding those regions that are sex specific.

optimized for identification of short, well-defined discrete peaks (track marked "All DHS sites" in Fig. 6C and D). Additional examples, including GH responses, are shown in Fig. S6. The full list of SICER-identified regions is provided in Table S9.

**Transcription factor binding sites enriched in sex-specific DHS sites.** THEME, a hypothesis-based algorithm that tests for enrichment of predefined motifs (29), was used to examine sex-specific DHS sites for enrichment of transcription factor binding site motifs by using sex-independent DHS sites as a background. Given the expected heterogeneity of sex-specific DHS sites, we carried out these analyses by using subsets comprised of male- and female-specific DHS sites that are (i) within 10 kb or within 50 kb of the TSS of a sex-specific gene, (ii) within 10 kb or within 50 kb of a sex-independent gene, and (iii) distant (>50 kb) from any gene. Each set of DHS sites was further divided into sites that respond and sites that do not respond to continuous GH treatment in males with a >2-fold difference in intensity levels at a  $P$  value of <0.01 (see Table S5A in the supplemental material). Starting with motif families derived from the TRANSFAC and JASPAR databases, we identified 16 enriched motifs (see Table S10). The sets of sex-specific sites were then scanned for each of the 16 motifs, which were then clustered according to fold enrichment in each of the sets of DHS sites (Fig. 7).

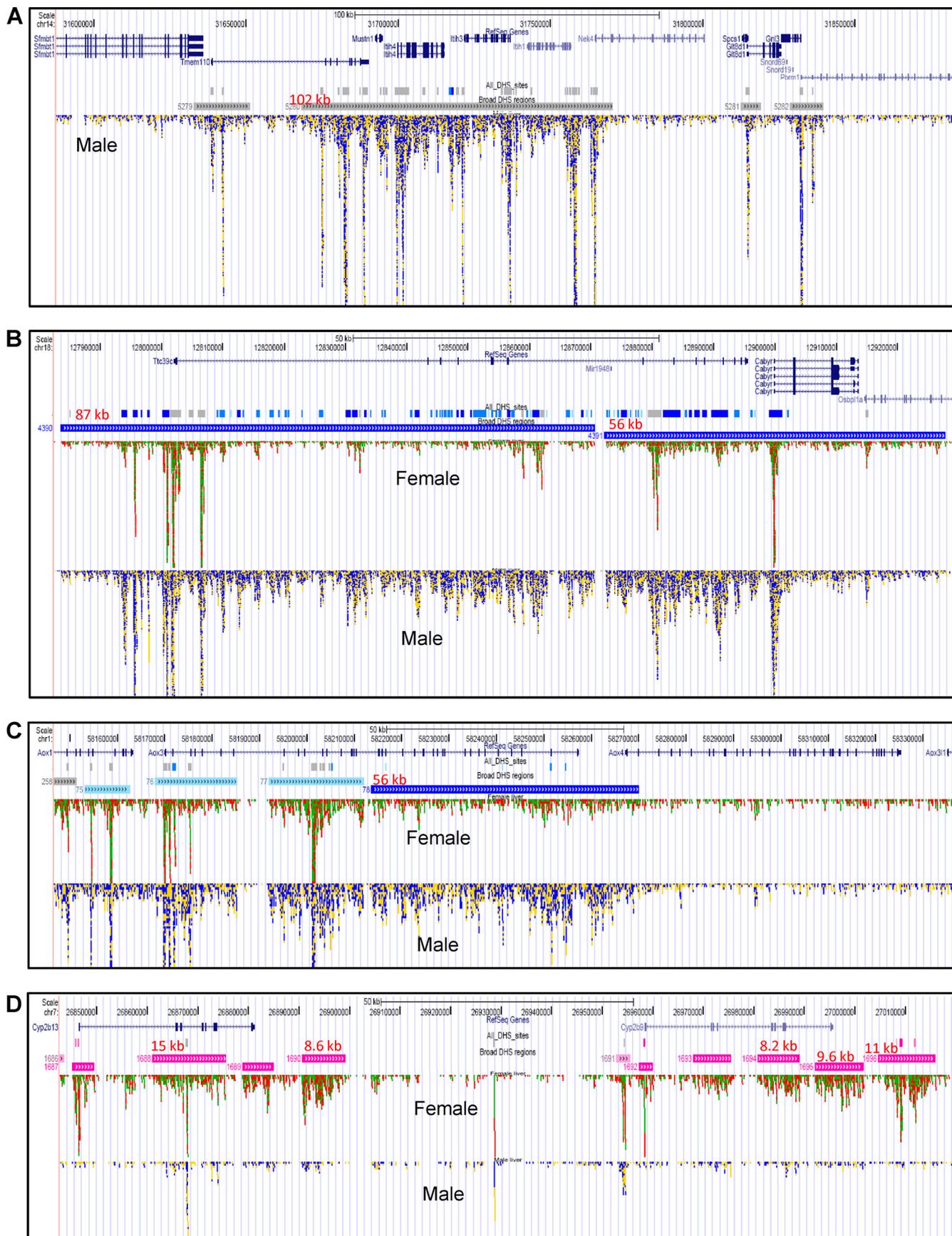
The discovered motifs include binding sites for two factors known to be required for sex-specific liver gene expression. Thus, a motif matching the binding site for STAT5b (motif 28) is enriched in male-specific GH-responsive sites, as is a motif matching HNF4 $\alpha$  (motif 70), consistent with the earlier findings that these two transcription factors are essential for GH-regulated sex-specific gene expression in male mouse liver (5, 18, 20). While the HNF4 $\alpha$ -like motif is most highly enriched in sites within 10 kb of a sex-specific gene, the STAT5 motif shows the highest enrichment in more distal sites, including sites proximal to sex-independent genes, consistent with other studies on STAT5 binding (8, 25, 34). The STAT5 motif clusters together with motifs that match 9 other transcription factors (or transcription factor families), all of which exhibit a common pattern of enrichment in male-specific, GH-responsive DHS sites (motif cluster A) (Fig. 7). Eight of these 10 motifs are underrepresented in female-specific GH-responsive sites (cluster A<sub>1</sub>). These 10 motifs include binding sites for CUX2, a highly female-specific, GH-regulated transcription factor (24), GFI1, a STAT-inducible transcriptional repressor (22, 60), OCT1 (POU2F1), which interacts with STAT5 in binding to the cyclin D1 promoter (31), PBX1, which interacts with OCT1 (35, 48) and may help penetrate repressive chromatin (41), and EVI1 (corresponding to the gene *Mecom*), a positive regulator of PBX1 (44) that interacts with the histone methyltransferase SUV39H1 (13). Two motifs, binding sites for MYC and MAX, were enriched in male-specific DHS sites not responsive to GH. Finally, motifs corresponding to binding sites for VDR, TCFAP2A, and TAL1 were most highly enriched at female-specific DHS sites. VDR activates the gene *CYP3A4* (GH responsive and predominantly expressed in females) in human hepatocytes (4, 7, 55), and a female-specific DHS site containing the VDR motif is associated with a female-specific mouse homolog, *Cyp3a41a* (19).

## DISCUSSION

We present a set of detailed, high-quality, genome-wide hypersensitivity maps comprised of more than 70,000 DHS sites, which encompass the transcriptional regulatory elements in the mouse liver *in vivo*. DHS maps were generated for both male and female mouse livers, from which we were able to identify 1,284 high-stringency sex-specific DHS sites, a subset of which was responsive to changes in plasma GH status, the major determinant of sex differences in liver gene expression. We demonstrate the utility of these maps for identifying binding sites for transcription factors previously shown to be essential for GH-regulated sex-specific gene expression (STAT5b and HNF4 $\alpha$  [5, 18, 20]), as well as binding sites for several novel factors not previously implicated in this process. These DHS sites encompass 1.8% of the mappable mouse genome, which substantially narrows down the sequence space in searches for gene regulatory sequences, including binding site motifs important for liver gene expression. The fine structure of DHS sites with a high density of sequence reads (see Fig. S1B in the supplemental material) suggests that it might be possible to visualize transcription factor binding directly in the form of digital footprints within DHS sites (17). Further analysis of hypersensitivity data collected at greater sequencing depth will be required to establish the feasibility of this approach in mammalian tissues. DHS sites are expected to encompass key regulatory elements, including promoters, enhancers, silencers, and insulators associated with the expression of thousands of genes in their native chromatin structure under physiological conditions. The DHS maps presented here for mouse liver tissue, in combination with corresponding sets of genome-wide histone modification and transcription factor binding maps (11, 38, 42, 43, 49), can be expected to serve as a valuable resource for elucidation of transcriptional networks controlling a wide range of physiological and pathophysiological processes.

Most sex-dependent DHS sites were short and highly localized (median length, ~500 bp), but in several cases, sex-specific hypersensitivity extended over broad regions, up to ~100 kb in length (Fig. 6; see also Fig. S6 in the supplemental material). The accessibility of many of these sex-dependent DHS sites and regions was altered by continuous GH infusion in male mice, which both feminizes the overall pattern of liver gene expression (19) and rendered the vast majority of the male-specific DHS sites less accessible to DNase while increasing the hypersensitivity of a substantial subset of the female-specific DHS sites and extended regions. These findings support the proposal that these GH-responsive DHS sites play a functional role in liver sexual dimorphism and suggest that a common upstream pathway responsive to GH, such as the activation of STAT5b (26, 53), regulates their differential chromatin accessibility in male and female livers.

We also observed a strong association between sex-specific DHS sites and sex-specific gene expression, with sex-specific genes more likely than sex-independent genes to have a nearby sex-specific DHS site. Moreover, sex-specific DHS sites were more likely than sex-independent DHS sites to have a nearby sex-specific gene. In some cases, multiple DHS sites, or extended hypersensitivity regions (discussed above), were associated with sex-specific genes. These may act in concert to



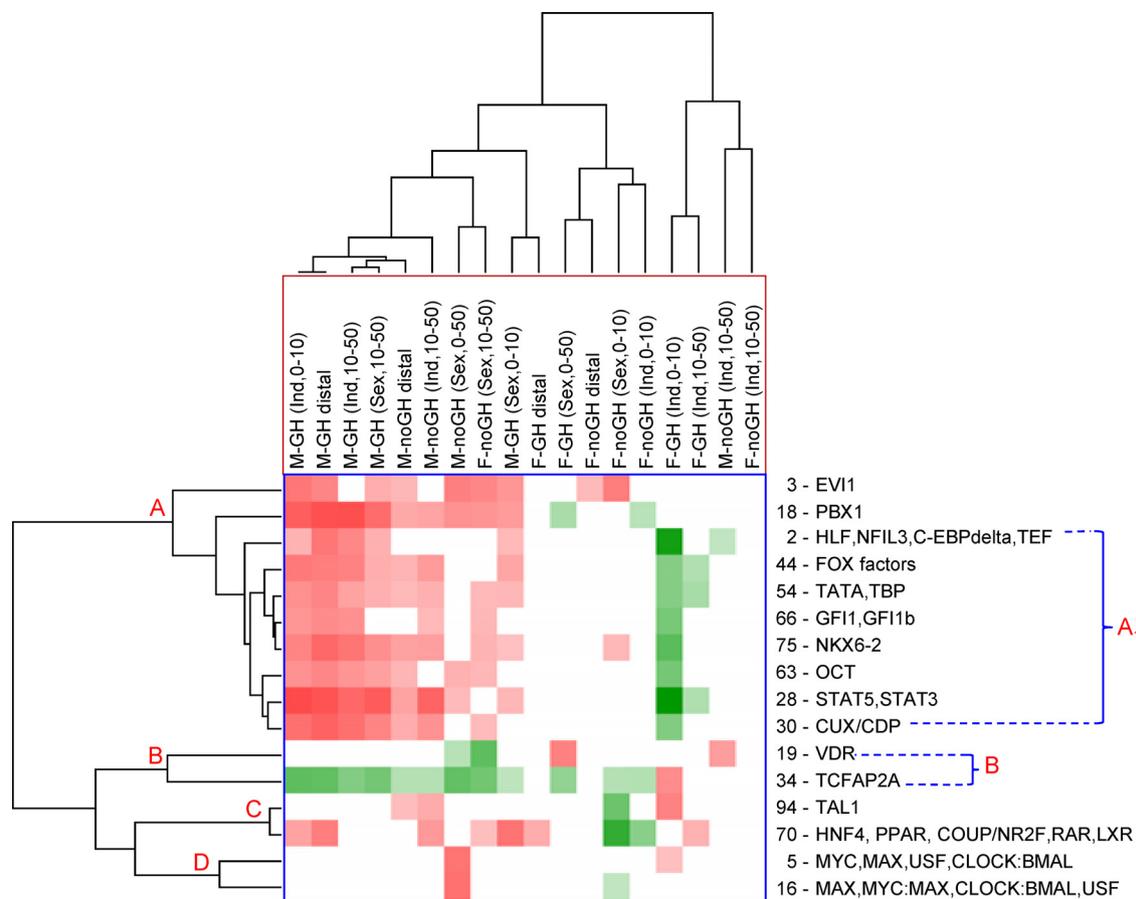


FIG. 7. Heat map of transcription factor binding motifs enriched in sex-specific DHS sites compared to sex-independent DHS sites. Motifs are clustered by fold enrichment in sets of male-specific (M) and female-specific (F) DHS sites that respond (GH) or do not respond (noGH) to GH with a >2-fold difference in hypersensitivity at a *P* value of <0.01 and are either 0 to 10 kb or 10 to 50 kb from the TSS of either sex-specific (Sex) or sex-independent (Ind) genes or >50 kb from any gene (distal), as marked. Combined data for 0 to 50 kb are presented for the M-noGH and F-GH sets, which both contained <15 DHS sites in the 10- to 50-kb subset. Red indicates overrepresentation and green indicates underrepresentation >1.45-fold compared to sex-independent DHS sites, with darker color indicating greater enrichment or depletion. The full data set is shown in Table S10A in the supplemental material. Where multiple factors are listed at the right, they are in order of highest to lowest match to the motif matrix. Major clusters of motifs (A to D) are identified on the left. Cluster A (10 motifs) is enriched in male-specific DHS sites; 8 of these motifs (subcluster A<sub>1</sub>) are enriched in several male-specific GH-responsive (M-GH) DHS sets and depleted in the F-GH, sex-independent (0- to 10-kb) set. Motif 30 (CUX/CDP) is a binding site for CUX2 (12). Motifs in cluster B are most highly enriched in F-GH DHS sets and are most highly depleted in various male DHS sets and in an F-noGH set. Motif 94 (TAL1) is most highly enriched in the same F-GH set as TFAP2A. Motif 70 is enriched in several M-GH sets, but in contrast to the motifs in cluster A, it is most highly enriched in the DHS set that is within 10 kb of sex-specific genes. Two very similar motifs for MYC/MAX are enriched in male-specific non-GH-responsive sites.

increase the magnitude of differences in gene expression between male and female livers. However, in other cases, we observed groups of sex-specific DHS sites not located near sex-specific genes: one striking example is a cluster of female-specific DHS sites on the X chromosome (see Fig. S6B in the supplemental material), and another is a cluster of male-spe-

cific sites on chr13, whose nearest sex-specific genes (the *Cd180* and *Sgtb* genes) are weakly female specific and located 800 kb and 500 kb from the cluster, respectively (see Fig. S6A). Indeed, for a majority of sex-specific DHS sites, the closest gene was not a sex-specific gene, and only a subset of sex-specific genes have a sex-specific DHS site within 100 kb of the

FIG. 6. Broad regions of DNase hypersensitivity. Tracks in each panel show (in order, from the top) RefSeq genes, DHS peaks identified using PeakSeq (all DHS sites), and DHS regions identified using SICER (broad DHS regions), with the length of the SICER-detected region marked in red. Gray, blue, and pink indicate sex-independent, male-specific, and female-specific SICER regions, respectively, with lighter blue or pink indicating a <2-fold difference in sex specificity. Panel A depicts a 102-kb sex-independent region overlapping with six sex-independent genes. Panels B and C show 56- to 87-kb male-specific DHS regions in the vicinity of male-specific genes *Ttc39c* and *Aox3*, respectively, and panel D shows several female-specific DHS regions at the female-specific genes *Cyp2b9* and *Cyp2b13*. While the broad DHS regions in panels A and B contain multiple individual DHS peaks, as identified by PeakSeq, the broad DHS regions shown in panels C and D were not identified by PeakSeq, which is optimized for identification of short, well-defined discrete peaks.

gene. These findings suggest the importance of long-range DNA interactions for sex-specific gene expression, as well as more complex interactions between multiple regulatory sites and multiple genes (30). While our results are consistent with regulation by distal sex-specific DHS sites, it is also possible that some sex-specific genes are regulated via sex-independent DHS sites whose cognate transcription factors are expressed or regulated in a sex-dependent manner. Other sex-specific genes may be regulated posttranscriptionally, i.e., by a mechanism that does not involve sex-related differences in chromatin accessibility.

Histone methylation marks, such as H3K4-me1 and H3K4-me3, are associated with active regions of chromatin, including enhancers and promoters, which are anticipated to coincide with DHS regions. Indeed, based on H3K4 methylation ChIP-seq maps for female mouse livers (38), we found that the fraction of H3K4-me1 marks associated with DHS sites was much higher than that of H3K4-me3 marks, particularly for female-specific DHS sites. As H3K4-me1 in the absence of H3K4-me3 is a characteristic of enhancers (15), we surmise that many liver DHS sites function as enhancers, some of which may exhibit sex-specific activities. This is supported by our *in vivo* reporter gene assays, in which 5 out of 6 DHS sites investigated demonstrated intrinsic enhancer activity when delivered to mouse livers by hydrodynamic injection. Moreover, several of the enhancer sequences tested showed sex-related differences in activity that match the sex specificity of the DHS site and the associated genes. The sex-related differences in *in vivo* enhancer activities seen here, were, however, considerably smaller than the sex-related differences in levels of expression of the genes themselves, suggesting that multiple DHS sites may be required to confer a high degree of sex specificity to gene expression. Indeed, multiple sex-dependent DHS sites are associated with the three genes whose enhancers showed sex-related differences in levels of activity (*Acot4*, *Cyp7b1*, and *Cyp2d9*). In the case of the *Cux2* intron 2 DHS site tested, no sex-related difference between enhancer activities was observed, indicating that the cloned fragment does not recapitulate the strong sex-related difference in DNA accessibility seen in intact liver chromatin (female/male DHS site sequence read ratio, 7.3). Nevertheless, given the very strong enhancer activity of this genomic fragment (Fig. 5), coupled with its 7-fold-lower accessibility in the male liver, this DHS site could make a substantial contribution to the strong (~100-fold) female specificity that characterizes *Cux2* gene expression (24). Together, these findings suggest that some sex-dependent DHS sites exhibit intrinsic sex-related differences in enhancer activity, e.g., due to the binding of transcription factors that are expressed or activated in a sex- and plasma GH pattern-dependent manner (e.g., STAT5b), while other sex-dependent DHS sites (e.g., the *Cux2* intron 2 enhancer) impart strong sex-related differences to gene transcription by virtue of the large sex-related differences in their accessibility in intact liver chromatin *per se*, even though they might not directly bind sex-specific transcription factors. Further studies will be required to identify the factors and establish the underlying mechanisms that initiate and maintain these sex-related differences in chromatin structure.

Motif analysis identified 13 transcription factor binding motifs that are enriched in one or more subsets of male-specific

DHS sites compared to sex-independent DHS sites (Fig. 7; see Table S10 in the supplemental material). Three other motifs were enriched in female-specific DHS sites, and one of these, the motif for TCFAP2A, was depleted in subsets of male-specific DHS sites. The male DHS site-enriched motifs include motifs that bind liver-expressed transcription factor families from the FOX and nuclear receptor families (e.g., HNF4 $\alpha$ ), as well as the binding site for STAT5b, which exhibits important sex-related differences in responsiveness to GH stimulation *in vivo* (52). Another male DHS-enriched motif, CDP, matches the binding site for CUX2 (12), a transcriptional repressor expressed in female livers at a level 100-fold higher than that in male livers (24), suggesting that CUX2 may enforce male specificity by binding to male-specific DHS sites in female livers, thereby suppressing the residual activity of enhancers that are partially accessible in females. A subcluster of 8 male DHS site-enriched motifs (subcluster A<sub>1</sub>) (Fig. 7), which includes motifs corresponding to binding sites for STAT5b and CUX2, was depleted in a subset of female-specific DHS sites responsive to GH. Given the high frequency of these 8 male-specific DHS site-enriched motifs, it is not surprising that many male-specific DHS sites contain matches for 6 or more of the 8 motifs (see Fig. S7). Further work will be needed to determine whether or not particular combinations of these 8 motifs have distinct functions and to identify the specific factors that actually bind to their cognate sequences at DHS sites in male and female livers.

Our finding that male-specific, GH-responsive DHS sites are enriched for both STAT5b-like and HNF4 $\alpha$ -like (nuclear receptor) motifs is consistent with our earlier observation that these factors are both essential for sex-specific gene expression in mouse livers. STAT5b is one of the major direct effectors of GH signaling in liver, and its deletion downregulates ~90% of male-specific genes in the male mouse liver (5, 18). Similarly, knockout of the nuclear receptor HNF4 $\alpha$ , enriched in the liver, abolishes sex-related differences in the liver (20). By carrying out the motif analysis separately for DHS sites that are near sex-specific genes, near sex-independent genes, and distant from genes, we showed that the HNF4 $\alpha$ -like motif is most highly enriched at DHS sites within 10 kb of sex-specific genes, while the STAT5-like motif is highly enriched at distal sites, as well as at sites proximal to sex-independent genes. The latter finding is consistent with the occurrence of functional STAT5 binding sites at large distances from target genes (8).

In conclusion, the present investigation of sex-related differences in chromatin accessibility has identified condition-specific transcriptional regulatory sites in the mouse liver on a genome-wide scale. These differences are manifested as sex-specific DHS sites, which in some cases encompass extended chromatin regions. A subset of these sex-specific genomic sites and regions is associated with genes expressed in a sex-dependent manner, strongly suggesting that they play a functional role in liver sexual dimorphism; however, a majority of sex-specific DHS sites are distal to sex-specific genes, making it more difficult to establish their significance. Transcription factor binding motifs identified as enriched in these sites serve as candidates for further study of the molecular mechanisms that govern sex-specific liver gene transcription. Further study will be required to determine how sex-related differences in chromatin accessibility are established and maintained in response

to sex-related differences in plasma GH patterns, which are programmed by early androgen exposure and first emerge at puberty.

#### ACKNOWLEDGMENTS

This work was supported in part by NIH grant DK33765 (to D.J.W.). A.S. received Training Core support from the Superfund Research Program at Boston University (NIH grant 5 P42 ES07381).

We thank Minita Holloway of this laboratory for contributions to initial optimization of the DHS protocol and X. Shirley Liu (Dana-Farber Cancer Institute) for initial discussions about experimental design.

We have no financial interests to declare.

#### REFERENCES

- Boyle, A. P., S. Davis, H. P. Shulha, P. Meltzer, E. H. Margulies, Z. Weng, T. S. Furey, and G. E. Crawford. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* **132**:311–322.
- Bryne, J. C., E. Valen, M. H. Tang, T. Marstrand, O. Winther, I. da Piedade, A. Krogh, B. Lenhard, and A. Sandelin. 2008. JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic Acids Res.* **36**:D102–D106.
- Ceseña, T. I., T. X. Cui, G. Piwien-Pilipuk, J. Kaplani, A. A. Calinescu, J. S. Huo, J. A. Iniguez-Lluhi, R. Kwok, and J. Schwartz. 2007. Multiple mechanisms of growth hormone-regulated gene transcription. *Mol. Genet. Metab.* **90**:126–133.
- Cheung, C., A. M. Yu, C. S. Chen, K. W. Krausz, L. G. Byrd, L. Feigenbaum, R. J. Edwards, D. J. Waxman, and F. J. Gonzalez. 2006. Growth hormone determines sexual dimorphism of hepatic cytochrome P450 3A4 expression in transgenic mice. *J. Pharmacol. Exp. Ther.* **316**:1328–1334.
- Clodfelter, K. H., M. G. Holloway, P. Hodor, S. H. Park, W. J. Ray, and D. J. Waxman. 2006. 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.
- Crawford, G. E., S. Davis, P. C. Scacheri, G. Renaud, M. J. Halawi, M. R. Erdos, R. Green, P. S. Meltzer, T. G. Wolfsberg, and F. S. Collins. 2006. DNase-chip: a high-resolution method to identify DNase I hypersensitive sites using tiled microarrays. *Nat. Methods* **3**:503–509.
- Drocourt, L., J. C. Ourlin, J. M. Pascussi, P. Maurel, and M. J. Vilarem. 2002. Expression of CYP3A4, CYP2B6, and CYP2C9 is regulated by the vitamin D receptor pathway in primary human hepatocytes. *J. Biol. Chem.* **277**:25125–25132.
- Eleswarapu, S., Z. Gu, and H. Jiang. 2008. Growth hormone regulation of insulin-like growth factor-I gene expression may be mediated by multiple distal signal transducer and activator of transcription 5 binding sites. *Endocrinology* **149**:2230–2240.
- El-Serag, H. B. 2004. Hepatocellular carcinoma: recent trends in the United States. *Gastroenterology* **127**:S27–S34.
- Endo, M., Y. Takahashi, Y. Sasaki, T. Saito, and T. Kamataki. 2005. Novel gender-related regulation of CYP2C12 gene expression in rats. *Mol. Endocrinol.* **19**:1181–1190.
- Gao, H., S. Falt, A. Sandelin, J. A. Gustafsson, and K. Dahlman-Wright. 2008. Genome-wide identification of estrogen receptor alpha-binding sites in mouse liver. *Mol. Endocrinol.* **22**:10–22.
- Gingras, H., O. Cases, M. Krasilnikova, G. Berube, and A. Nepveu. 2005. Biochemical characterization of the mammalian Cux2 protein. *Gene* **344**:273–285.
- Goyama, S., E. Nitta, T. Yoshino, S. Kako, N. Watanabe-Okochi, M. Shimabe, Y. Imai, K. Takahashi, and M. Kurokawa. 2010. EVI-1 interacts with histone methyltransferases SUV39H1 and G9a for transcriptional repression and bone marrow immortalization. *Leukemia* **24**:81–88.
- Gross, D. S., and W. T. Garrard. 1988. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* **57**:159–197.
- Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching, R. D. Hawkins, L. O. Barrera, S. Van Calcar, C. Qu, K. A. Ching, W. Wang, Z. Weng, R. D. Green, G. E. Crawford, and B. Ren. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* **39**:311–318.
- Hemenway, C., and D. M. Robins. 1987. DNase I-hypersensitive sites associated with expression and hormonal regulation of mouse C4 and Slp genes. *Proc. Natl. Acad. Sci. U. S. A.* **84**:4816–4820.
- Hesselberth, J. R., X. Chen, Z. Zhang, P. J. Sabo, R. Sandstrom, A. P. Reynolds, R. E. Thurman, S. Neph, M. S. Kuehn, W. S. Noble, S. Fields, and J. A. Stamatoyannopoulos. 2009. Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. *Nat. Methods* **6**:283–289.
- Holloway, M. G., Y. Cui, E. V. Laz, A. Hosui, L. Hennighausen, and D. J. Waxman. 2007. Loss of sexually dimorphic liver gene expression upon hepatocyte-specific deletion of Stat5a-Stat5b locus. *Endocrinology* **148**:1977–1986.
- Holloway, M. G., E. V. Laz, and D. J. Waxman. 2006. 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.
- Holloway, M. G., G. D. Miles, A. A. Dombkowski, and D. J. Waxman. 2008. Liver-specific hepatocyte nuclear factor-4alpha deficiency: greater impact on gene expression in male than in female mouse liver. *Mol. Endocrinol.* **22**:1274–1286.
- Hosui, A., and L. Hennighausen. 2008. Genomic dissection of the cytokine-controlled STAT5 signaling network in liver. *Physiol. Genomics* **34**:135–143.
- Ichiyama, K., M. Hashimoto, T. Sekiya, R. Nakagawa, Y. Wakabayashi, Y. Sugiyama, K. Komai, I. Saba, T. Moroy, and A. Yoshimura. 2009. Gfi1 negatively regulates T(h)17 differentiation by inhibiting RORgamma activity. *Int. Immunol.* **21**:881–889.
- Isensee, J., and P. Ruiz Noppinger. 2007. Sexually dimorphic gene expression in mammalian somatic tissue. *Gen. Med.* **4**(Suppl. B):S75–S95.
- Laz, E. V., M. G. Holloway, C. S. Chen, and D. J. Waxman. 2007. Characterization of three growth hormone-responsive transcription factors preferentially expressed in adult female liver. *Endocrinology* **148**:3327–3337.
- Laz, E. V., A. Sugathan, and D. J. Waxman. 2009. Dynamic in vivo binding of STAT5 to growth hormone-regulated genes in intact rat liver. Sex-specific binding at low- but not high-affinity STAT5 sites. *Mol. Endocrinol.* **23**:1242–1254.
- Lichanska, A. M., and M. J. Waters. 2008. How growth hormone controls growth, obesity and sexual dimorphism. *Trends Genet.* **24**:41–47.
- Lichtsteiner, S., J. Wuarin, and U. Schibler. 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. *Cell* **51**:963–973.
- Liu, J. K., Y. Bergman, and K. S. Zaret. 1988. The mouse albumin promoter and a distal upstream site are simultaneously DNase I hypersensitive in liver chromatin and bind similar liver-abundant factors in vitro. *Genes Dev.* **2**:528–541.
- Maclsaac, K. D., D. B. Gordon, L. Nekudova, D. T. Odom, J. Schreiber, D. K. Gifford, R. A. Young, and E. Fraenkel. 2006. A hypothesis-based approach for identifying the binding specificity of regulatory proteins from chromatin immunoprecipitation data. *Bioinformatics* **22**:423–429.
- Maclsaac, K. D., K. A. Lo, W. Gordon, S. Motola, S. Mazor, and E. Fraenkel. 2010. A quantitative model of transcription regulation reveals the influence of binding location on expression. *PLoS Comput. Biol.* **6**:e1000773.
- Magné, S., S. Caron, M. Charon, M. C. Rouyez, and I. Dusanter-Fourt. 2003. STAT5 and Oct-1 form a stable complex that modulates cyclin D1 expression. *Mol. Cell Biol.* **23**:8934–8945.
- Mahony, S., and P. V. Benos. 2007. STAMP: a web tool for exploring DNA-binding motif similarities. *Nucleic Acids Res.* **35**:W253–W258.
- Matys, V., O. V. Kel-Margoulis, E. Fricke, I. Liebich, S. Land, A. Barre-Dirrie, I. Reuter, D. Chekmenev, M. Krull, K. Hornischer, N. Voss, P. Stegmaier, B. Lewicki-Potapov, H. Saxel, A. E. Kel, and E. Wingender. 2006. TRANSFAC and its module TRANSCOMP: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res.* **34**:D108–D110.
- Nelson, E. A., S. R. Walker, J. V. Alvarez, and D. A. Frank. 2004. Isolation of unique STAT5 targets by chromatin immunoprecipitation-based gene identification. *J. Biol. Chem.* **279**:54724–54730.
- Rave-Harel, N., M. L. Givens, S. B. Nelson, H. A. Duong, D. Coss, M. E. Clark, S. B. Hall, M. P. Kamps, and P. L. Mellon. 2004. TALE homeodomain proteins regulate gonadotropin-releasing hormone gene expression independently and via interactions with Oct-1. *J. Biol. Chem.* **279**:30287–30297.
- Reich, M., T. Liefeld, J. Gould, J. Lerner, P. Tamayo, and J. P. Mesirov. 2006. GenePattern 2.0. *Nat. Genet.* **38**:500–501.
- Rinn, J. L., and M. Snyder. 2005. Sexual dimorphism in mammalian gene expression. *Trends Genet.* **21**:298–305.
- Robertson, A. G., M. Bilenky, A. Tam, Y. Zhao, T. Zeng, N. Thiessen, T. Cezard, A. P. Fejes, E. D. Wederell, R. Cullum, G. Euskirchen, M. Krzywiniski, I. Biro, M. Snyder, P. A. Hoodless, M. Hirst, M. A. Marra, and S. J. Jones. 2008. Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. *Genome Res.* **18**:1906–1917.
- Rozowsky, J., G. Euskirchen, R. K. Auerbach, Z. D. Zhang, T. Gibson, R. Bjornson, N. Carriero, M. Snyder, and M. B. Gerstein. 2009. PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nat. Biotechnol.* **27**:66–75.
- Sabo, P. J., M. S. Kuehn, R. Thurman, B. E. Johnson, E. M. Johnson, H. Cao, M. Yu, E. Rosenzweig, J. Goldy, A. Haydock, M. Weaver, A. Shafer, K. Lee, F. Neri, R. Humbert, M. A. Singer, T. A. Richmond, M. O. Dorschner, M. McArthur, M. Hawrylycz, R. D. Green, P. A. Navas, W. S. Noble, and J. A. Stamatoyannopoulos. 2006. Genome-scale mapping of DNase I sensitivity in vivo using tiling DNA microarrays. *Nat. Methods* **3**:511–518.
- Sagerström, C. G. 2004. PbX marks the spot. *Dev. Cell* **6**:737–738.
- Schmidt, D., M. D. Wilson, B. Ballester, P. C. Schwalie, G. D. Brown, A.

- Marshall, C. Kutter, S. Watt, C. P. Martinez-Jimenez, S. Mackay, I. Talianidis, P. Flicek, and D. T. Odom. 2010. Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* **328**:1036–1040.
43. Seo, Y. K., H. K. Chong, A. M. Infante, S. S. Im, X. Xie, and T. F. Osborne. 2009. Genome-wide analysis of SREBP-1 binding in mouse liver chromatin reveals a preference for promoter proximal binding to a new motif. *Proc. Natl. Acad. Sci. U. S. A.* **106**:13765–13769.
  44. Shimabe, M., S. Goyama, N. Watanabe-Okochi, A. Yoshimi, M. Ichikawa, Y. Imai, and M. Kurokawa. 2009. Pbx1 is a downstream target of Evi-1 in hematopoietic stem/progenitors and leukemic cells. *Oncogene* **28**:4364–4374.
  45. Silander, K., M. Alanne, K. Kristiansson, O. Saarela, S. Ripatti, K. Auro, J. Karvanen, S. Kulathinal, M. Niemela, P. Ellonen, E. Vartiainen, P. Jousilahti, J. Saarela, K. Kuulasmaa, A. Evans, M. Perola, V. Salomaa, and L. Peltonen. 2008. Gender differences in genetic risk profiles for cardiovascular disease. *PLoS One* **3**:e3615.
  46. Song, L., and G. E. Crawford. 2010. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb. Protoc.* **2010**:pdb.prot5384.
  47. Ström, A., H. Eguchi, A. Mode, C. Legraverend, P. Tollet, P. E. Stromstedt, and J. A. Gustafsson. 1994. Characterization of the proximal promoter and two silencer elements in the CYP2C11 gene expressed in rat liver. *DNA Cell Biol.* **13**:805–819.
  48. Subramaniam, N., W. Cairns, and S. Okret. 1998. Glucocorticoids repress transcription from a negative glucocorticoid response element recognized by two homeodomain-containing proteins, Pbx and Oct-1. *J. Biol. Chem.* **273**:23567–23574.
  49. Thomas, A. M., S. N. Hart, B. Kong, J. Fang, X. B. Zhong, and G. L. Guo. 2010. Genome-wide tissue-specific farnesoid X receptor binding in mouse liver and intestine. *Hepatology* **51**:1410–1419.
  50. Wauthier, V., A. Sugathan, R. D. Meyer, A. A. Dombkowski, and D. J. Waxman. 2010. Intrinsic sex differences in the early growth hormone responsiveness of sex-specific genes in mouse liver. *Mol. Endocrinol.* **24**:667–678.
  51. Wauthier, V., and D. J. Waxman. 2008. Sex-specific early growth hormone response genes in rat liver. *Mol. Endocrinol.* **22**:1962–1974.
  52. Waxman, D. J., and M. G. Holloway. 2009. Centennial perspective: sex differences in the expression of hepatic drug metabolizing enzymes. *Mol. Pharmacol.* **76**:215–228.
  53. Waxman, D. J., and C. O'Connor. 2006. Growth hormone regulation of sex-dependent liver gene expression. *Mol. Endocrinol.* **20**:2613–2629.
  54. Whitacre, C. C. 2001. Sex differences in autoimmune disease. *Nat. Immunol.* **2**:777–780.
  55. Wolbold, R., K. Klein, O. Burk, A. K. Nussler, P. Neuhaus, M. Eichelbaum, M. Schwab, and U. M. Zanger. 2003. Sex is a major determinant of CYP3A4 expression in human liver. *Hepatology* **38**:978–988.
  56. Wooddell, C. I., T. Reppen, J. A. Wolff, and H. Herweijer. 2008. Sustained liver-specific transgene expression from the albumin promoter in mice following hydrodynamic plasmid DNA delivery. *J. Gene Med.* **10**:551–563.
  57. Yang, X., E. E. Schadt, S. Wang, H. Wang, A. P. Arnold, L. Ingram-Drake, T. A. Drake, and A. J. Lusis. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res.* **16**:995–1004.
  58. Yokoyama, Y., Y. Nimura, M. Nagino, K. I. Bland, and I. H. Chaudry. 2005. Current understanding of gender dimorphism in hepatic pathophysiology. *J. Surg. Res.* **128**:147–156.
  59. Zang, C., D. E. Schones, C. Zeng, K. Cui, K. Zhao, and W. Peng. 2009. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**:1952–1958.
  60. Zweidler-Mckay, P. A., H. L. Grimes, M. M. Flubacher, and P. N. Tschlis. 1996. Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. *Mol. Cell. Biol.* **16**:4024–4034.