Single Nucleus MultiOmics and In Vivo Massively Parallel Reporter Assays Identify Functional Sex-Biased Enhancers Linked to Liver Metabolism and Disease

Ting-Ya Chang and David J Waxman, Departments of Biology, and Biomedical Engineering, and Bioinformatics Program, Boston University, Boston MA 02215

Abstract - Growth hormone (GH) regulates sexual dimorphic programs in liver gene expression impacting lipid metabolism, bile acid synthesis and xenobiotic processing, which contribute to sex differences in metabolic dysfunction-associated steatotic liver disease (MASLD). Despite extensive knowledge of sex-biased transcriptomes, the functional *cis*-regulatory elements orchestrating these sex differences in liver metabolic function are largely uncharacterized. We employed an integrative approach combining single nucleus (sn) multiOmics with in vivo functional enhancer profiling using HDI-STARR-seq to identify and validate GH-responsive, sex-biased enhancers in mouse liver. We profiled ~1,000 differentially accessible regions (DARs) from male, female, and continuous GH-infused (cGH) male livers by snATAC-seq + snRNA-seq, linking chromatin accessibility to sex-biased gene expression. A tiled STARR-seq library comprised of 23,912 unique reporters spanning 1,839 ATAC regions was delivered to mouse liver by hydrodynamic injection (HDI), which enabled functional assessment of enhancer activity under physiologically relevant conditions and with minimal immune response. We identified 4,126 robust active reporters, of which 1,397, representing 840 distinct ATAC regions, showed sex-biased and/or GH-regulated activity, with strong positive correlations between chromatin accessibility and functional enhancer activity. Regulated enhancer regions were significantly enriched for activating histone marks (H3K4me1, H3K27ac) and binding sites for key transcription factors including STAT5, BCL6, HNF4α, and CUX2. Further, *de novo* motif analysis revealed sex-specific transcription factor binding patterns discriminating regulated from non-regulated enhancers. Finally, many of the functionally active sex-biased, cGH-responsive genes were linked to MASLD-enabling or MASLD-protective sex-biased genes. This work establishes HDI-STARR-seq as a powerful platform for dissecting hormone-regulated enhancer function in vivo and gives mechanistic insights into how hormonal control of chromatin accessibility drives sexual dimorphism in hepatic metabolism and differential metabolic liver disease risk between sexes.

Grant support: NIH DK121998 (to DJW)

Introduction

GH-regulated sex differences in hepatic gene expression have been linked to sex difference in lipid and drug metabolism, xenobiotic metabolism, disease susceptibility and disease progression in both humans and mouse models. GH is a key factor affecting mammalian liver in a sex-dependent manner and governs the sex differential expression of hepatic genes through downstream effectors, including sex-specific transcription factors and repressors [1, 2]. Sex-differential temporal patterns of pituitary GH activate intracellular signaling, which leads to distinct expression of CYPs and other metabolic genes, as well as genes involved in cell cycle [3] and lipid metabolism [4, 5]. GH regulates many downstream transcriptional outputs via the transcription factor STAT5b [6-9]. Upon GH binding to its receptor (GHR) in the liver, JAK2 is activated, leading to tyrosine phosphorylation of STAT5 [6, 10]. The phosphorylated STAT5 dimerizes and then translocates to the nucleus, where it binds to DNA motifs with the consensus sequence TTCNNNGAA to increase transcription of linked target genes [9, 11]. The GH-STAT5 signaling axis plays a pivotal role in diverse physiological processes, including body growth [7, 12], cell cycle regulation [3, 12], and hepatic metabolism of lipids, bile acids, steroids and drugs [6, 7, 13]. Dysregulation of STAT5 signaling has been implicated in liver diseases such as fatty liver [4, 13, 14], fibrosis, and hepatocellular carcinoma [5, 15].

In liver, two Stat5 isoforms, Stat5a and Stat5b mediate the actions of GH in regulating the transcription of sex-biased genes, although Stat5b is ~20-fold more abundant than Stat5a. Hepatocyte-specific deletion of Stat5a and Stat5b in adult male liver [16, 17] leads to decreased expression of large numbers of male-biased as well as de-repression of many female-biased liver genes, consistent with earlier findings in global Stat5b-null male mice [18]. Research into downstream transcriptional activators and repressors of STAT5 has shed light on their role in regulating gene expression in the liver. STAT5b can enhance the transcriptional activity of HNF4α [19], particularly toward male-biased liver target genes [19, 20]. This underscores the collaborative nature of

STAT5b and HNF4α in regulating male-biased *Cyp* genes. Furthermore, the transcriptional repressor Bcl6, which shows male-biased expression in liver, has a DNA recognition motif that partially overlaps with the STAT5 consensus site, enabling it to inhibit a subset of liver STAT5 transcriptional responses [21, 22]. These findings illuminate the intricate interplay between transcriptional activators and repressors in modulating gene expression patterns in the liver. DNA accessibility plays a crucial role in modulating sex-dependent gene expression by influencing chromatin accessibility. For instance, in male rat liver, the expression of *Cyp2c12* is affected by chromatin condensation, rendering the GH-responsive element less accessible to STAT5 [23]. Additionally, in female rat livers, where nuclear STAT5 activity is generally low, STAT5 binding to Igf-1 and SOCS2 is restricted to high-affinity sites. In contrast, male-specific gene expression is associated with male-specific STAT5 binding to multiple low-affinity STAT5 sites [24]. These findings underscore the intricate interplay between chromatin structure and STAT5 binding affinity in shaping sex-dependent gene expression patterns.

Liver function is highly dependent on its specified sinusoidal structure and heterogeneity, composed of different cell types with distinct characteristics, functions and topography [25, 26]. Single cell technology is required for dissecting the coordination of different cell types including hepatocytes [26, 27], endothelial cell [28], immune cells [29] and hepatic stellate cells (HSCs) [30] involved in lipid and fat metabolism, xenobiotics metabolism and hepatic carcinogenesis. To better understand how gene regulation specific to different cell types contributes liver function in sex-biased manner, we utilize a cell-type specific analysis to profile chromatin structure and gene expression of specific cell types and investigate gene network under the influence of xenobiotic treatment. With the advent of 10x Genomics MultiOmic ATAC +GEX library, we are able to profile ATAC-seq peaks and RNA expression from a single nucleus, revealing the functional enhancers in response to xenobiotics and growth hormone in a cell type specific manner.

Here, present an integrated functional genomics approach combining single nucleus multiOmic technology applied to adult mouse liver [31] with HDI-STARR-seq to functionally assay sex-biased, GH-responsive enhancers in intact mouse liver. We selected ~1,000 sex-biased and GH-regulated hepatocyte DARs, many linked to sex-biased genes involved in lipid homeostasis, bile acid synthesis, and pathways implicated in fatty liver disease. These regions, as well as several hundred non-regulated control ATAC-seq regions, were synthesized as 23,912 tiled reporters and delivered to mouse liver by hydrodynamic injection (HDI) to assay enhancer activity in vivo across three biological conditions. Our experimental design enabled us to discover physiologically relevant enhancers linked to liver metabolic genes and functions in an endogenous cellular, hormonal and metabolic environment—features that are absent from cell culture-based reporter assays. We systematically identified DARs whose pattern of regulated STARR-seq activity matched their endogenous liver chromatin accessibility patterns, characterized epigenetic signatures enriched at functional enhancers, and discovered transcription factor motifs discriminating regulated from constitutively active regulatory elements. This work advances our understanding of how GH-regulated chromatin accessibility translates into sex-dependent enhancer function driving sexual dimorphism in liver metabolism and sex differential liver disease.

Methods

Animal procedures and tissue procurement. All mouse work was carried out in compliance with procedures approved by the Boston University Institutional Animal Care and Use Committee (protocol # PROTO201800698) and in compliance with ARRIVE 2.0 Essential 10 guidelines [32], including study design, sample size, randomization, experimental animals and procedures, and statistical methods. Male and female CD-1 mice (ICR strain) were purchased from Charles River Laboratories (Wilmington, MA), fed standard rodent chow, and housed in a temperature and humidity-controlled facility on a 12 h light cycle (lights on at 7:30 AM). Livers used to prepare nuclei for the 10X Genomics MultiOme datasets were obtained from Cameron Vergato and Dr. Hong Ma of this laboratory. Briefly, male mice (8 wk of age) were anesthetized with isoflurane and implanted subcutaneously with an Alzet osmotic minipump (cat. # 2002; Durect Corp., Cupertino, CA) to deliver recombinant rat growth hormone (GH) as a continuous infusion (cGH) at 25 ng/g BW/h then euthanized 12 days later. Recombinant rat GH (bacterial lipopolysaccharide depleted (7.7 to 15.5 EU/mg); purchased from Dr. Arieh Gertler, Protein Laboratories Rehovot, Ltd., Rehovot, Israel) was dissolved in 30 mM $NaHCO_3$ and 0.15 M NaCl containing 0.1 mg/ml rat albumin (Sigma-Aldrich, catalog no. A6414), essentially as

107 described [33]. Mice were euthanized under CO_2 anesthesia by cervical dislocation at 8-10 wk of age, and
108 livers were collected at a consistent time of day (between 10:30 AM and 11:30 AM) to minimize the impact of
109 circadian rhythms on gene expression [34]. Livers were excised and flash frozen in liquid N_2 then stored at -
110 80°C. For livers used for HDI-STARR-seq assays, reporter plasmids and STARR-seq plasmid libraries were
111 transfected into livers of male and female mice at 7-9 weeks of age by hydrodynamic injection (HDI) on day 0.
112 For the cGH-treated male mouse group, mice were implanted on day 1 with an Alzet osmotic minipump set to
113 deliver recombinant rat GH at 25 ng/g of BW/h. For the male vehicle control group, the Alzet osmotic
114 minipump was loaded with vehicle (30 mM $NaHCO_3$) and implanted on day 1. Ten days later, on day 11 after
115 HDI injection, livers were harvested and stored at -80°C.
116
117 **In vivo transfection by HDI and validation by luciferase assay.** Naked plasmid DNA was delivered to mouse
118 liver by HDI via the tail vein, as described [35]. Briefly, 30 μg of pooled STARR-seq plasmid was mixed with 3 μg
119 of plasmid pGYL18 encoding Renilla luciferase and delivered by rapid tail vein injection (HDI) into male and
120 female mice on day 0. The effectiveness of HDI and of plasmid delivery to the liver was verified for each mouse
121 by Renilla luciferase assay, as follows. A frozen piece of each liver (50 mg) was homogenized in 500 μl ice cold
122 1x Passive lysis buffer (Promega, cat. #E1910) using a rubber-tipped micropestle that fits the bottom of a 1.7
123 ml centrifuge tube. Liver lysate (4 μl) was assayed in a 96-well plate using Renilla luciferin reagent (25 μl of
124 Stop & Glo Reagent; Promega Dual-Luciferase Reporter Assay System, cat. #E1910) injected into each well via
125 one of the injectors of a CentroXS3 instrument (Berthold Technologies). For each mouse group, liver extracts
126 were prepared at the same time to minimize batch effects and assayed in technical triplicates. The absolute
127 Renilla luciferase activity of each liver was used as a general indicator of the success of HDI in that mouse (Fig.
128 3A).
129
130 **Isolation of single nuclei from frozen mouse liver.** Frozen mouse liver tissue (a total of ~150 mg) was
131 combined from each of n=3-4 biological replicate livers per biological condition (male liver, female liver, and
132 12 day cGH-treated male liver) and processed as a single sample to obtain liver nuclei representing the 3-4
133 livers included in each group. We adapted a protocol developed to extract nuclei from frozen mouse brain
134 tissue [36-38] and optimized it to obtain high-quality nuclei from frozen mouse liver tissue, as follows. Mouse
135 liver (a total of ~ 150 mg frozen liver, pooled from n=3-4 livers/group) was minced in the cold room, then
136 added to 1 ml of ice-cold Nuclei EZ lysis buffer (cat. # NUC101, Sigma-Aldrich) in a 2-ml glass Dounce
137 homogenizer. The sample was placed on ice for 1 min and then homogenized on ice using 10 strokes of
138 Dounce pestle A, followed by another 10 strokes with Dounce pestle B. Samples were kept on ice during
139 homogenization to minimize RNA degradation. To reduce the concentration of nuclei and minimize their
140 aggregation during the subsequent washing steps, another 1 ml of ice-cold Nuclei EZ lysis buffer was added to
141 the lysate in the Dounce homogenizer tube and mixed gently, followed by a 5 min incubation. The lysate was
142 then filtered through a 40-μm cell strainer mesh (cat. #431750, Corning Life Sciences) and collected in a 50-ml
143 conical tube. The lysate from each sample was then split into two Eppendorf tubes, which were kept separate
144 until the final nuclei permeabilization step, below. The lysed nuclei in each tube were pelleted by
145 centrifugation in a swinging bucket rotor at 500 x g for 5 min at 4°C, after which the pellet was gently
146 resuspended by adding 1 ml of fresh EZ lysis buffer to each Eppendorf tube. The 5-min incubation and 5 min
147 centrifugation steps were then repeated a second time. After removing the supernatant, an 0.5 ml aliquot of
148 Nuclei Wash and Resuspension Buffer (2% FBS in phosphate-buffered saline containing 1 U/μL of Protector
149 RNase Inhibitor (Millipore-Sigma, cat. #3335399001) added immediately before use) was added to each tube
150 without disturbing the pellet, followed by a 5 min incubation on ice to allow for buffer exchange. The pellets
151 were then resuspended after adding an additional 0.5 ml of Nuclei Wash and Resuspension Buffer to give a
152 total volume of 1 ml in each tube, and then centrifuged at 500 x g for 5 min at 4°C. The supernatant was
153 removed and the pelleted nuclei were resuspended in 1 ml of Nuclei Wash and Resuspension Buffer. A 10 μl
154 aliquot was removed to visualize and quantify the nuclei on a Cellometer Vision instrument (Nexcelom
155 Bioscience, Lawrence, MA) using the AOPI Staining Kit (cat. #CS20106, Nexcelom Bioscience, Lawrence, MA).
156 The final concentration of nuclei typically ranged from 5,000-6,000 nuclei/μl when starting with 150 mg liver
157 tissue, with viability typically < 10%, as determined by AOPI staining. The high-quality nuclei obtained showed
158 an intact nuclear membrane without peripheral blebbing (Fig. S1).

Nuclei permeabilization and MultiOme ATAC + Gene Expression library preparation. The resuspended nuclei were kept on ice and transferred within 1 h to the MIT BioMicro Center (Cambridge, MA), where the nuclei were permeabilized immediately prior to 10x Genomics Gel Bead-In EMulsion (GEM) formation and then 10X Genomics MultiOme ATAC + Gene Expression library preparation, to help maintain RNA and DNA integrity. Nuclei were permeabilized as described in 10x Genomics protocol # CG000375, Rev B: Nuclei Isolation from Complex Tissues for Single Cell MultiOme ATAC + Gene Expression Sequencing. Briefly, the nuclei were repelleted at 500 x g for 5 min at 4°C and then resuspended in 100 μl of freshly prepared ice cold 0.1x Lysis Buffer (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 0.01% Tween 20, 0.01% NP40, 0.001% Digitonin, 1% BSA, 1 mM DTT and 1 U/μl Protector RNase Inhibitor (cat. #3335399001, Millipore-Sigma) by pipetting the sample up and down gently 5 times, followed by a 2 min incubation on ice. Freshly prepared ice-cold Wash Buffer (1 ml of 10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$, 1% BSA, 0.1% Tween 20, 1 mM DTT and 1 U/μl Protector RNase Inhibitor) was then added and the sample was mixed by pipetting up and down gently 5 times. After centrifugation at 500 x g for 3 min at 4°C, the supernatant was removed and replaced with the diluted Nuclei Buffer provided in the Chromium Next GEM Single Cell MultiOme ATAC + Gene Expression Reagent Bundle kit (cat. # PN-1000285, 10X Genomics). Immediately prior to the GEM formation procedure, the permeabilized nuclei were filtered through a 40 μm Flowmi® Cell Strainer (cat. # BAH136800040, Millipore-Sigma) to removed aggregated nuclei. The filtered nuclei were then quantified using trypan blue staining and diluted to an estimated concentration of 2,500 nuclei/μl. To prepare GEMs, nuclei were loaded onto a 10x Chromium instrument at the MIT BioMicro Center with a targeted recovery of 6,000-7,500 nuclei per sample, followed by MultiOme ATAC + Gene Expression library preparation following the manufacturer's protocol (Chromium Next GEM Single Cell MultiOme ATAC + Gene Expression Reagent Bundle, cat. # PN-1000285).

Single-cell transcriptional and chromatin accessibility profiling of mouse liver. Single cell-based mouse liver gene expression and chromatin accessibility profiles were determined for each biological group using 10X Genomics MultiOme technology, which implements both snRNA-seq and snATAC-seq analysis in the same individual nuclei. Processing of the 10X Genomics MultiOmic datasets was carried out by Dr. Kritika Karri and Dr. Maxim Pyatkov of this laboratory, using Cell Ranger ARC (v.2.0.0) for initial sample processing, followed by analysis using Signac (v1.1.0) [39] and Seurat (4.1.0) [40]. A total of 127,957 ATAC-seq peaks were identified across all samples and biological groups using the MACS2 peak caller within the Signac package, which provided ATAC-seq sequence read density data for each biological condition and in each cell type, including three hepatocyte subclusters: periportal (PP), midlobular (Mid) and pericentral (PC). This analysis was implemented using the CallPeaks function of Signac, which was executed with the arguments 'group.by' = 'celltype_and_condition,' followed by keepStandardChromosomes( ) with 'coarse' pruning, and subsetByOverlaps( ) utilizing Encode blacklist sites (blacklist_mm10.bed). Following this peak identification process, the CreateChromatinAssay function within Signac was employed to combine peaks and thus aggregate the genomic ranges of all peaks from the individual sets of liver samples to generate a 'peaks' assay.

Identification of differentially accessible regions (DARs) and differentially expressed genes. Differential chromatin accessibility, and differential gene expression, between liver cell types and between biological conditions (male, female and cGH-treated male) was computed by Dr. Kritika Karri of this laboratory using the Loupe Browser (10x Genomics Loupe Browser 6.0 (https://www.10xgenomics.com/), which uses the exact negative binomial test to test for differences in mean ATAC-seq sequence read density (cut sites per cell) in a given ATAC region, or differences in gene expression between groups of cells. Differential ATAC-seq peaks between cell types were initially identified using the Loupe Browser's integrated Global Distinguishing function with a differential expression threshold of Log2 (fold change) >1 at FDR (Benjamini-Hochberg) < 0.05. The ATAC-seq peak regions identified were filtered to remove ATAC-seq peaks found in 3% or fewer of nuclei, in the case of non-hepatocyte clusters, or in fewer than 5% or nuclei in the case of hepatocytes. Cell type-specific DARs were then identified by removing any peaks found to be cell type-specific DARs for 2 or more clusters. We used the Loupe Browser's integrated Locally Distinguishing function to calculate condition-specific DARs across five different pairwise comparisons: Male vs. Female-F1, Male vs. Female-F2, Male vs. cGH-treated

male (M-cGH), Female-F1 vs. M-cGH, and Female-F2 vs. M-cGH, with a significance threshold of adjusted P-value (Bonferroni-corrected FDR) <0.05, except as noted (and where Female-F1 and Female-F2 represent two independent pools of nuclei from control female mouse livers). We used the same approach, based on the Loupe Browser's Locally Distinguishing function, as described previously [41], to identify differentially expressed genes in each liver cell type for the same five comparisons.

ATAC-seq peak to target gene feature linkages. ATAC-seq peak-to-gene correlations were used to associate sex-biased DARs with their inferred gene targets using feature linkage analysis performed by Dr. Kritika Karri of this laboratory. Briefly, we used the peak-to-gene linkage method embedded in 10x Genomics Cell Ranger ARC (v2.0.1) to link ATAC-seq peaks to specific genes based on Pearson correlations between ATAC-seq peaks and genes, which gives a direct and interpretable measure of the linkage strength (linkage score, values ranging from -1 to 1). We used the set of 127,957 ATAC-seq peaks and 75,798 genes, including > 52,000 mouse lncRNA genes represented by the custom GTF file used in our snRNA-seq analysis [42], to identify 226,841 ATAC-seq peak–gene linkages across the genome, 77% of which involve positive correlations. The subset of these linkages relevant to the ATAC-seq regions included in this study is presented in Table S8, with feature linkage characteristics starting in column AK. Individual linkages were characterized by the ATAC peak to gene TSS distance, the linkage score (Pearson correlation) and the linkage significance (-log(p-value)).

Selection of ATAC regions for STARR-seq reporter library. We selected 1,839 of the 127,957 ATAC-seq peaks regions identified by 10X Genomics MultiOme analysis for HDI-STARR-seq assay of enhancer activity in mouse liver (Table S4A). The ATAC regions selected are in 3 groups: **1)** ATAC regions whose chromatin accessibility is sex-dependent and/or responsive to cGH infusion (Table S2; n= 1,031 ATAC regions, of which 64 also below to the TAD-based group and 1 is in the pericentral hepatocytes-enriched group), as detailed below; **2)** 749 ATAC regions grouped within 10 specific TAD regions (Table S3A); and **3)** 124 ATAC regions that show significant liver cell type specificity (Table S3B), as determined in separate analyses performed by Dr. K Karri.

Selection of sex-biased and/or cGH-regulated DARs. A total of 1,031 sex-biased and/or cGH-regulated DARs (Table S2A) were selected for cloning into the STARR-seq library, of which 612 DARs were sex-biased and cGH-regulated DARs: 360 of the 612 DARs were male-biased and cGH-repressed, and were linked to 235 male-biased, cGH-repressed genes (Table S2B) but were not also linked to any genes that were male-biased but not cGH-regulated, or to genes that were female-biased genes regardless of GH responsiveness; and 252 of the 612 DARs were female-biased, cGH-induced DARs, and were linked to 166 female-biased, cGH-induced genes (Table S2C) but were not also linked to other sex-biased genes (Table S4B, columns W, AD, AK, AR). Subsets of both DAR sets also mapped to non-sex-biased genes. 611 of these 612 ATAC regions are found on autosomes (Table S4A). Other sets of DARs included in the STARR-seq library: **1)** 102 male-biased DARs that were stringently not cGH-responsive DARs, 20 of which are on chrY; **2)** 186 female-biased DARs that were stringently not cGH-responsive, 41 of which are on chrX (Table S4A); **3)** 40 male-biased DARs and 20 female-biased DARs that are regulated by cGH but were not significantly linked by the MultiOme analysis to any protein-coding gene or liver-expressed lncRNA gene, likely due to their large distance (> 80 kb) from the nearest gene (Table S4E, columns J-K) [42], (Pick6 group; Table S2A, Table S4B). These DARS showed STARR-seq activities that were similar to those of the sex-biased and cGH-responsive DARs linked to sex-biased genes (Fig. S11), including regarding the frequency of regulated reporters (Table S2A); and finally, **4)** 71 sex-biased, cGH-regulated DARs with potential to serve as repressors of sex-biased gene expression (Table S2A, Pick4 and Pick5 DARs). Thus, we selected 62 male-biased and cGH-repressed DARs that were linked to either female-specific and cGH-induced genes (n=34 genes) or to female-specific but GH-independent genes (n=22 genes) (Table S2D); this pattern of hormone regulated chromatin accessibility displays a sex bias and cGH response opposite from that expected for enhancers that regulate expression of female-specific genes, suggesting these DARs may repress their linked female-specific genes. Similarly, we selected 9 female-biased and cGH-induced DARs with repressor potential, i.e., they were linked to male-biased and cGH-repressed genes (n=10 genes) or to male-specific and cGH-unresponsive genes n=3 genes), displaying the opposite sex response patterns based on their linkage (Table S2E).

263 **STARR-seq reporter design and synthesis.** The 1,839 ATAC regions selected for inclusion in the STARR-seq
264 library were tiled to generate a total of 23,950 STARR-seq reporter sequences, of which 23,912 reporters were
265 included in the final design (Table S5) owing to issues relating to the liftOver between mouse genome mm10
266 and mouse genome mm9. Each ATAC region was extended by 70 bp on either end, and then tiled using a
267 series of 180 nt reporter sequences, arranged to include 125 nt overlap and a 55 nt shift between neighboring
268 reporters (Fig. 2A). This tiling design minimizes the risk of gaps in reporter coverage of an ATAC region that can
269 arise due to a loss of reporter sequence representation during cloning. Each reporter was synthesized as a 210
270 nt oligonucleotide using SurePrint HiFi oligo synthesis (Agilent, Inc) and is comprised of the 180 nt reporter
271 sequence flanked on each end by a 15 nt adapter sequence. The pool of 23,912 reporters was then cloned into
272 pPromALB_STARR-seq (plasmid pTYC10; Addgene # 220289) [35], as described below. Of note, each 15 nt
273 adaptor sequence serves as a PCR seed and facilitates the generation of the 210 base pair double-stranded
274 DNA library required for STARR-seq library preparation. Illumina primer sequence #1 is 5'-ACGCTCTTCCGATCT-
275 3', and Illumina primer sequence #2 is 5'-AGATCGGAAGAGCAC-3'. Sequences of ATAC regions along with their
276 mouse mm10 genomic coordinates (mouse genome assembly GRCm38.p6) are shown in Table S5.
277
278 **STARR-seq plasmid library construction.** Each reporter sequence was PCR amplified from the SurePrint HiFi
279 oligo synthesis product using specific PCR primers (Table S1) as described below. We set up 6 individual 25 μl
280 PCR reactions (10 ng of oligo pool, 12.5 μl of 2x KAPA HiFi HotStart master mix (Cat No. KK2601, Roche), 2.5 μl
281 of 5 μM forward primer, 2.5 μl of 5 μM reverse primer per 25 ul reaction) using PCR primers ON8352 and
282 ON8353 (Table S1). PCR reactions were set to generate an HA-adaptor-ligated DNA library under low
283 amplification conditions using the following cycling conditions: 98°C for 45 s, followed by 10 cycles of 98°C for
284 15 s, 65°C for 30 s, 72°C for 45 s, and final elongation step for 1 min at 72°C. The HA-adaptor-ligated DNA was
285 purified in two rounds of 1x SPRI bead selection and then cloned into pPromALB_STARR-seq vector (Fig. S3).
286 Next, we used a 2:1 molar ratio of insert to 125 ng SalI/AgeI-digested pPromALB_STARR-seq plasmid in two 10
287 μl In-Fusion® HD Cloning Plus reactions (TaKaRa, cat. #638910) to assemble STARR-seq reporter plasmid
288 libraries via the homology arms. The reaction products were pooled and precipitated using sodium
289 acetate/ethanol, resuspended in 10 μl EB buffer (10 mM Tris HCl, pH 8.5) and stored at -80°C prior to use [43].
290 Two 2.5 μl aliquots of the precipitated material were transformed separately into 20 μl of electrocompetent
291 MegaX DH10B bacterial cells (ThermoFisher Scientific, cat. # EC0113) using Gene Pulser Xcell Total System
292 (Bio-Rad) and Gene Pulser Electroporation Cuvettes (0.1 cm gap, Bio-Rad; cat No. 1652089) under
293 these conditions: 2 kV, 25 μF, 200 ohms. LB bacterial culture medium (1 L) was used to amplify the STARR-seq
294 plasmid library, which was then purified using PureLink™ Endotoxin-Free Maxi Plasmid Purification Kit
295 (ThermoFisher Scientific, cat. #A31231) to obtain ~ 2 mg of the final STARR-seq plasmid library, STARR-TYC8.

296 **Preparation of plasmid, transfected DNA and transcribed RNA reporter Illumina sequencing libraries.** The
297 following methods, adapted from published STARR-seq protocols [43] and detailed previously [35], were used
298 in a nested PCR amplification protocol to prepare 3 types of Illumina sequencing libraries used to determine
299 the transcriptional enhancer activity of each ATAC region-associated reporter sequence (putative regulatory
300 region) to be assayed by HDI-STARR-seq. 1) A plasmid library was prepared from the same batch of purified
301 plasmid DNA injected into each mouse by HDI. Plasmid DNA was used directly as a PCR template in a nested
302 PCR reaction for library construction (primers ON8381, ON8382; Table S1); 2) a DNA library was prepared from
303 DNA extracted from flash frozen liver tissue collected 11 days after in vivo transfection by HDI (primers
304 ON8381, ON8382); and 3) an RNA library was prepared from the transcribed RNA (reporter sequences)
305 extracted from each HDI-transfected liver (primers ON8379, ON8380). STARR-seq reporter plasmid DNA
306 remaining in the liver 11 days after HDI was extracted using a commercial QIAprep spin miniprep kit (QIAgen,
307 cat. # 27106). The liver tissue (200 mg frozen liver) was initially homogenized in 750 ul of QIAgen Buffer P1
308 using a rubber-tipped micropestle and passed through a 20G needle three times to homogenize large tissue
309 pieces. The extracted plasmid DNA was bound to a QIAprep spin column, eluted in EB buffer and used as a PCR
310 template in a subsequent nested PCR reaction using primers ON8381 and ON8382. To extract transcribed
311 reporter RNA, 120 μg of total liver RNA was isolated from 60-100 mg liver tissue from each mouse by TRIzol
312 guanidinium thiocyanate-phenol-chloroform extraction followed by polyA selection using oligo(dT) beads
313 (New England Biolabs, cat. #E7490L). The transcribed RNA was digested with Turbo DNase I (Ambion, cat. #

314 AM2238) at 37°C for 30 min, reverse transcribed using primer ON8731 for STARR-seq RNA first strand
315 synthesis and Superscript III first-strand synthesis supermix (Invitrogen, cat. # 18080-400) and treated with 10
316 μg RNase A per ml (Fermentas, cat. # EN0531) for 1 h at 37°C. The cDNA obtained was purified using 1.8x SPRI
317 beads and then used as a PCR template in the nested PCR reaction.

318 The nested PCR reaction requires one set of primers specifically targeting STARR-seq reporter DNA or RNA
319 followed by amplification using generalized Illumina dual index primers. Primers ON8381 and ON8382 (Table
320 S1) were designed to amplify the STARR-seq plasmid DNA libraries by specifically targeting the unspliced
321 intron incorporated into the original plasmid sequence and present in the plasmid and DNA libraries but
322 spliced out of the transcribed reporter RNA sequences. Primers ON8379 and ON8380, which straddle the
323 splice junction, were designed to specifically amplify only the transcribed and spliced RNA molecules [35]. The
324 second, nested PCR reaction was performed for all three library types (plasmid, DNA and RNA) using NEBNext®
325 Multiplex Oligos for Illumina (NEB cat. # E7600S) to further amplify the material and attach Illumina i5 and i7
326 barcoded sequences used for Illumina multiplex sequencing, performed at Novogene, Inc. (Sacramento, CA).

327 **Enhancer identification in TAD regions with sex-biased genes.** Topologically associating domains (TADs) are
328 fundamental units that compartmentalize genomic features and help constrain interactions of cis-regulatory
329 sequences to their within-TAD genic targets. We examined 6 TAD regions showing a strong tendency toward
330 sex-bias that contain genes encoding important liver metabolic functions, such as cytochromes P450 (CYPs),
331 solute carrier (SLC) family members, cytosolic sulfotransferases (SULTs), UDP glucuronosyltransferases (UGTs),
332 and 3β-hydroxysteroid dehydrogenases (Hsd3b). As controls, we examined ATAC regions from 4 TADs that do
333 not contain any sex-biased genes or sex-biased gene linkages (TAD_203, TAD_1519, TAD_4080, TAD_4241).
334 ATAC regions within the selected TADs were incorporated in the STARR-seq library (Table S3A).
335
336 **Liver cell type-specific ATAC regions.** We used the liver MultiOme dataset to profile open chromatin regions
337 for each liver cell type or subtype: hepatocytes (including periportal, midlobular and pericentral hepatocytes),
338 liver endothelial cells, hepatic stellate cells (HSC), immune cells, Kupffer cells and cholangiocytes. For each
339 liver cell type (cell cluster), we selected 20 ATAC regions showing the highest specificity for that cell type and
340 defined them as cell-type specific ATAC regions based on the following criteria for the ATAC-seq signal: log2
341 (fold change) > 1 using the Loupe Browser's integrated Global Distinguishing function, FDR (Benjamini-
342 Hochberg) < 0.05, and the ATAC-seq signal detected in > 3% of nuclei (> 5% for hepatocytes). ATAC regions
343 found to be cell type specific for 2 or more liver cell clusters were removed. Very few liver cell type-specific
344 ATAC regions were found for midlobular hepatocytes and cholangiocytes (Table S4A).
345
346 **Sequencing analysis of STARR-seq libraries.** Paired-end sequence reads, each 150 base pairs long, were
347 generated for the STARR-seq plasmid, DNA, and RNA libraries. A custom pipeline was developed to handle the
348 raw sequencing data [35]. This pipeline takes raw FASTQ files as input and produces various quality control
349 metrics, such as FASTQC reports (version 0.0.13.2 of FASTXToolkit), identification and removal of
350 contaminating adapter sequences (using Trim_galore version 0.4.2), and determination of insert-size length
351 distributions for paired-end read sequencing using Picard (version 1.123). The reads were aligned to the
352 mouse genome (version mm9) using Bowtie2 (version 2.2.6) [44]. Subsequently, all identified peak sets were
353 filtered to eliminate peaks overlapping ENCODE blacklisted regions [45]. An average of 9.5 million fragments
354 were sequenced for the RNA libraries and 27.1 million fragments for the DNA and plasmid libraries (Table S6).
355 Almost all (96-99%) of the plasmid, DNA and RNA library reads mapped to the 23,912 predesigned reporters
356 for all samples except for library G202M16, a negative control for HDI in the absence of a STARR-seq reporter
357 library. Five RNA sequencing samples were excluded from the downstream analysis (G202M1, G202M2,
358 G202M6, G202M10, G202M12) due to their low Pearson's correlation with other RNA libraries from the same
359 biological condition (Table S6). Similarly, DNA sequencing sample G202M18 was removed due to its low
360 Pearson's correlation with the other 3 DNA libraries and the 2 plasmid libraries (Table S6). Raw fastq
361 sequencing files are available at GEO under accession numbers to be assigned by GEO.
362

363 **Characterization of STARR-seq library.** STARR-seq plasmid sequencing reads, and liver-extracted STARR-seq
364 DNA library sequencing reads, were mapped to the set of 23,912 reporters. Since adjacent reporter sequences
365 overlap by 125 bp of sequence, we used Bedtools (v2.31.0) to set > 95% overlap (>171/180 nt) as the
366 threshold for assigning a sequence read to a given reporter to avoid counting partial overlaps from
367 neighboring reporters. Sequence reads obtained for each plasmid library mapped with high frequency (98.8%)
368 to the 23,912 designated reporters (Table S6), indicating very low non-specific PCR amplification and minimal
369 sequencing errors. The two technical replicate plasmid libraries were highly correlated based on normalized
370 counts across the 23,912 reporter regions (R^2=0.984, Fig. S4D). Almost all (>98%) of the inserts were 180 nt
371 long, showing that intact oligos were cloned into the library. The median number of normalized plasmid
372 sequence reads in each reporter region was 322, and 81.9% of the 23,912 reporters were represented within
373 6.4x of the median frequency (Fig. 3C). HDI liver-extracted STARR-seq DNA libraries were highly similar to the
374 STARR-seq plasmid libraries (R^2=0.982), indicating no significant reporter sequence-dependent bias in plasmid
375 uptake or retention by hepatocytes following HDI transfection. The GC content of the 23,912 reporters
376 included in our library design ranged from 19-87%, with the highest frequency observed for reporters at ~45-
377 50% GC content (Fig. S4A). The PCR-amplified plasmid and liver-extracted DNA libraries average sequence read
378 counts peaked at ~55-70% GC content (Fig. S4B-S4C), consistent with the known bias against both GC-rich and
379 AT-rich in sequencing libraries prepared using a PCR-amplification step [46].
380
381 **Qualification of STARR-seq library reporters for downstream analysis.** A total of 19,136 STARR-seq reporters
382 (Table S6) were qualified for inclusion in downstream analysis based on the following stringent threshold for
383 robust representation in the cloned and bacterially amplified STARR-seq plasmid library, STARR-TYC8: > 40
384 mapped Illumina sequence read pairs in each of the 3 liver-extracted STARR-seq DNA libraries and in both
385 individual plasmid libraries (Fig. S5B, first 2 columns). 4,776 other reporters did not meet this threshold for
386 consistent STARR-seq library representation and were excluded from downstream analysis. A very similar list
387 of qualified reporters was obtained when the threshold of > 40 mapped Illumina sequence read pairs was
388 normalized to the total sequence depth of each plasmid or DNA library.
389
390 **Determination of HDI-STARR-seq enhancer activity.** Illumina sequencing libraries prepared from RNA
391 extracted from each HDI-STARR-seq-transfected liver ('RNA library') were analyzed using Bedtools to count the
392 number of RNA sequence reads that overlapped each of the 19,136 qualified reporters. Qualified reporters
393 with > 30 mapped Illumina RNA sequence read pairs in all 3 biological replicates livers from untreated male
394 mice, in both biological replicate livers from cGH-treated male mice, or in both biological replicates from
395 untreated female mice were deemed active enhancers (a total of 11,966 of the 19,136 qualified STARR-seq
396 reporters; ranging from n=7,099-8,922 for each biological condition; Fig. S5E). Further, a total of 4,126 of the
397 19,136 qualified STARR-seq reporters met these criteria for active enhancers under all 3 biological conditions
398 and were designated robust active enhancers (Table S7, column AT). The robust active enhancers were
399 somewhat more highly represented in the STARR-seq plasmid library than the qualified reporters that showed
400 activity under only 1 or 2 of the 3 biological conditions (Fig. S6A), suggesting a higher false positive rate for the
401 latter sets of, non-robust STARR-seq enhancers owing to noise associated with genomic regions with lower
402 representation in the STARR-seq plasmid library. We therefore limited our analysis of differences in enhancer
403 activity between biological conditions (see below) to the set of 4,126 robust active reporters. 982 of the
404 19,136 qualified reporters were designated stringently inactive reporters based on these criteria: ≤ 30 mapped
405 Illumina sequence read pairs in each RNA library across all 3 biological conditions (i.e., in all 7 RNA libraries
406 included in our analysis; Table S7, column AT). Since many of these inactive reporters were poorly represented
407 in the STARR-seq library, we subdivided the 982 stringently inactive reporters into two sets based on mean
408 plasmid library read counts: stringently inactive-High (top 400 reporters), which were used as a background
409 set for downstream enrichment analyses; and stringently inactive-Low (bottom 582 reporters), which were
410 excluded from the analysis (Fig. S6A). A further 6,188 qualified reporters, designated 'Other', showed
411 inconsistent STARR-seq enhancer activity across biological replicates, as well as low representation in the
412 plasmid library (Fig. S6A) and were excluded from further analysis. Of note, the lists of robust active enhancers
413 and stringently inactive reporters were highly similar or nearly identical (e.g., 4,125 of the 4,126 robust active
414 reporters were retained) when the thresholds for minimum numbers of raw mapped read counts stated above
415 were normalized for the total sequence depth of each RNA library.

416
417 **Relationship between sex bias and cGH responsiveness of STARR-seq enhancer activity.** A linear regression
418 model [47] was implemented in R Studio to assess the relationship between response to cGH treatment (log2
419 ratio of STARR-seq enhancer activity in cGH-treated male liver to STARR-seq enhancer activity in male liver)
420 and sex bias (log2 ratio of STARR-seq enhancer activity in female liver to STARR-seq enhancer activity in male
421 liver).
422
423 **Regulated active STARR-seq enhancers.** Robust active enhancers that showed > 2-fold differences between
424 biological conditions in their STARR-seq reporter activity (defined as mean RNA library sequence reads for a
425 given STARR-seq reporter across biological replicates, normalized for total mapped sequence reads) were
426 designated regulated active enhancers (Table S7A). Regulated active enhancers whose pattern of STARR-seq
427 enhancer activity regulation matched their pattern of ATAC-seq activity regulation were designated positively
428 acting regulated enhancers. One example is the set of DARs whose ATAC-seq activity was both male-biased
429 and cGH-suppressed (Pick #1 DAR set, Table S2) and whose RNA reporter(s) showed >2-fold increase in STARR-
430 seq enhancer activity in male compared to female liver as well as a >2-fold decrease in STARR-seq enhancer
431 activity in cGH-treated male liver compared to control male liver. Another example is the set of DARs whose
432 ATAC-seq activity was male-biased but not cGH-regulated (Pick3_M DAR set, Table S2) and whose reporter(s)
433 showed >2-fold increase in STARR-seq enhancer activity in male compared to female liver but did not show >2-
434 fold decrease following cGH infusion. In contrast, male-biased and cGH-suppressed DARs whose STARR-seq
435 reporter activity showed < 2-fold male-bias and < 2-fold response to cGH infusion were non-regulated
436 enhancer DARs. We also identified a set of male-biased, cGH-repressed DARs linked to female-biased genes
437 (putative repressors; Pick #4 DAR set, Table S2), whose reporters showed >2-fold increase in STARR-seq
438 enhancer activity in male compared to female liver and also >2-fold decrease in STARR-seq enhancer activity in
439 cGH-treated male compared to control male liver, and correspondingly for as set of female-biased, cGH-
440 induced DARs linked to male-biased genes (Pick #4 and Pick #5 DAR sets; Table S2. Those regulated regions in
441 pick 4 are more active in male, with higher chromatin accessibility, with a linkage to a less expressing gene in
442 male (Female-biased gene). This suggests a mechanism to maintain a minimal gene expression of essential
443 female-biased genes in male even though the expression in male is much lower. On the contrary, the same
444 regulated regions in female livers are less active possibly due to closed chromatin regions. Finally, 452 robust
445 active enhancers whose reporters showed no regulated STARR-seq enhancer activity in any comparison (Table
446 S9A, column AI) were identified as non-regulated DARs.
447
448 **Chromatin state analysis of STARR-seq regions in liver.** Chromatin state maps, previously defined for adult
449 male mouse liver, and separately, for adult female mouse liver [48] using ChromHMM [49], were used to
450 determine the enrichment of individual chromatin states in various sets ATAC regions. These chromatin states
451 are based on a 14-state model (states E1-E14) and were previously used to characterize sex differences in
452 chromatin state and their relationships to sex-biased gene expression [48]. Bedtools overlap was used to
453 assign the chromatin state determined for each sex to each ATAC region. An ATAC region that overlapped two
454 or more different chromatin states was assigned to the state with the greatest overlap in bp. The distribution
455 of chromatin state enrichments (overall chromatin landscape) was determined separately for male and for
456 female liver for each set of functionally active enhancer region DARs in comparison to a background set
457 comprised of 317 DARs whose reporters showed no regulated STARR-seq enhancer activity in any comparison
458 (i.e., 452 DARs mentioned above, after excluding 82 unregulated DARs from the Pick1 set and 53 unregulated
459 DARs from the Pick2 set) (Table S9A column AJ).

460 **Motif analysis.** De novo motif discovery was carried out using MEME-CHIP [50] (https://meme-suite.org) run
461 in the classic mode to discover motifs enriched relative to a random model based on frequencies of each nt in
462 the provided sequences (Table S13A). MEME-CHIP was also run in the discriminative (non-classic) mode
463 comparing an input set of sequences to a background set of sequences, as specified in the text, using default
464 settings for both modes (Table S13B). For each set of regions analyzed (Table S9A, columns U-X) sequences
465 were input using the coordinates for each region shown in Table S9A. For example, STARR-seq analysis of a set
466 of 360 male-biased and cGH-repressed DARs identified 41 functionally active regulated ATAC regions (i.e.,
467 regulated Pick1, Table S9A). Sequences of those 41 ATAC regions were obtained from the mouse mm10

genome using the bedtools getfasta command and used as input. Specified background sets include the set of 65 stringently inactive ATAC regions, defined as ATAC regions that contain 1 or more stringently inactive STARR-seq_High reporters (Table S7A, column AT) but do not include any of the 4,126 robust active reporters (Table S9A, column AK), the set of 82 active but not regulated, male-biased, cGH-repressed DARs and the set of 53 active but not regulated, female-biased, cGH-induced DARs (Table S9A). The de novo discovered motifs were compared with published motifs mouse TF motifs determined by protein-binding microarrays and obtained from the UniPROBE motif database [51]. Separately, Find Individual Motif Occurrences (FIMO) analysis was conducted to identify Stat5a/Stat5b, Bcl6, Cux2 and Hnf6 motif occurrences within sets of regions by scanning the motif for STAT5A/STAT5B (JASPAR motif ID: MA0519.1), Bcl6 (Hocomoco mouse motif ID: BCL6_MOUSE.H11MO.0.A), Cux2 (Hocomoco mouse motif ID: CUX2_MOUSE.H11MO.0.C) and Hnf6 (Hocomoco mouse motif ID: HNF6_MOUSE.H11MO.0.A) [52, 53] (Table S13C Columns AP-BM).

Results

Sex-biased and GH-responsive enhancers predicted by snATAC-seq + snRNA-seq MultiOmic analysis. We implemented snATAC-seq in combination with snRNA-seq (10X Genomics MultiOme technology) to identify open chromatin regions in mouse hepatocytes and non-parenchymal cells (NPCs) that show sex-dependent and/or GH-regulated accessibility and link them to specific genes [31]. Livers from male, female, and cGH-infused male mice (n=3-4 livers per group) were harvested and pooled for nuclei isolation (Fig. S1), followed by snATAC-seq + snRNA-seq MultiOmic analysis. 46,188 nuclei across the 3 biological conditions passed quality control filters for both the snATAC-seq and snRNA-seq modalities [31]. The high quality of the ATAC-seq data was indicated by TSS enrichment scores ranging from 5.5 to 6.4 (Fig. S1B). Clusters of nuclei representing 5 major liver cell types were resolved in a UMAP and identified based on established marker genes (Fig. 1A). A total of 127,957 ATAC regions were identified across all liver cell and characterized with respect to sex-differences and the impact of cGH infusion on chromatin accessibility types [31]. Cell Ranger ARC was used to link individual ATAC regions to their predicted gene targets based on correlations between chromatin accessibility and the expression of nearby genes. The linked ATAC regions and their predicted gene targets are said to be 'co-expressed' across the population of 46,188 nuclei, i.e., enriched for a shared regulatory mechanism, inferring an enhancer–gene targeting relationship. For example, a sex-biased DAR (ATAC_39124) located 46.1 kb downstream of the TSS of *Cux2*, which codes for a female-biased transcriptional repressor gene, makes highly significant linkages to 4 nearby genes (Fig. 1B). Two of 3 positively correlated linkages involving this ATAC region were made to female-biased genes (*Cux2*, *lnc4593*), which implicates the ATAC site in positive gene regulation, while one negatively correlated linkage was made to a male-biased gene (*P2rx4*), implicating the ATAC site in negative gene regulation (Fig. 1C). The linkage of this ATAC region to *Cux2* showed the highest significance (-log(p-value) = 22.2) and the highest correlation (correlation score = 0.30; Table S8).

Selection of DARs for HDI-STARR-seq library. We selected liver ATAC regions for analysis of functional enhancer activity in mouse liver by HDI-STARR-seq, focusing on ATAC regions showing cGH-regulated sex differences in chromatin accessibility in hepatocytes, the major liver cell type exhibiting sex differences in gene expression [41] and chromatin accessibility [31]. We selected 360 differentially accessible regions (DARs) whose accessibility was male-biased and cGH-repressed and that were linked to cGH-repressed male-biased genes, as well as 252 cGH-induced female-biased DARs linked to cGH-induced female-biased genes (Fig. 2B). We also selected 102 male-biased and 186 female-biased DARs whose chromatin accessibility was unresponsive to cGH infusion, subsets of which were on chrY or chrX (Table S2C). We also included in our STARR-seq library 749 ATAC regions, selected from a total of 10 TAD regions [54], which encompass well studied sex-biased genes important for liver metabolic function, including cytochromes P450 (CYPs), solute carrier (SLC) family members, sulfotransferases (SULTs), UDP glucuronosyltransferases (UGTs), and 3β-hydroxysteroid dehydrogenases (Hsd3b family) (Table S3A), including genes that impact lipid metabolism and fatty liver disease. Finally, we included in the STARR-seq library 61 ATAC-regions showing hepatocyte-specific chromatin accessibility and 63 regions showing specificity for liver NPCs, to determine whether HDI-STARR-seq can be used to assay ATAC regions associated with liver NPCs. Overall, we assayed 1,839 ATAC regions (634 nt, mean length), which were tiled using a series of overlapping STARR-seq reporters spaced 55 bp apart, which allows for finer mapping of the transcriptionally active sequences within each ATAC region (Fig. 2A).

518 **Identification of active enhancers by HDI-STARR-seq.** A STARR-seq plasmid library representing the 1,839
519 ATAC regions selected for analysis was prepared by cloning *en masse* into the STARR-seq reporter library
520 vector pPromALB_STARRseq [35] a pool of 23,912 unique single-stranded oligonucleotides (Table S7A) tiled
521 across the ATAC regions (Fig. 2A). Sequence analysis established that the STARR-seq plasmid library was of
522 high quality, with 19,136 of the 23,912 unique STARR-seq reporters (80%) qualified for inclusion in
523 downstream analysis based on stringent thresholds applied for robust representation in the library (see
524 Methods). Next, we employed HDI-STARR-seq [35] to introduce the STARR-seq reporter plasmid library into
525 mouse liver to assay the functional transcriptional enhancer activity of each ATAC region. HDI-STARR-seq
526 activity was assayed under 3 distinct biological conditions: male liver, female liver and 10-day cGH-infused
527 male liver. Livers were harvested 11 days after HDI, and the transfection efficiency of each liver was
528 determined by assaying a co-transfected Renilla luciferase reporter plasmid (Fig. 3A). Five of the 12 HDI-
529 transfected livers gave comparatively low Renilla luciferase activity, which indicates poor transfection [35],
530 and were excluded from downstream analysis. RNA sequence reads extracted from the 7 robustly HDI-
531 transfected livers (n=3 males, n=2 females, n=2 cGH-infused males) were mapped to the 23,912 unique
532 STARR-seq library reporters, quantified, and verified regarding their overall consistency (Pearson's correlation;
533 Fig. 3B). 11,966 of the 19,136 STARR-seq reporters (62.5%) showed STARR-seq enhancer activity under at least
534 one biological condition and were deemed active (Fig. S5E); 4,126 reporters (21.6%) were active under all 3
535 conditions and were designated *robust active enhancers* (Table S7, column AT).
536
537 **Impact of cGH infusion on functional activity of active sex-biased enhancers identified by HDI-STARR-seq**.
538 Subcutaneous infusion of GH in male mice using osmotic minipumps (cGH infusion) overrides the endogenous
539 pulsatile male pattern of plasma GH and substantially feminizes the expression of a large fraction of sex-biased
540 genes in mouse liver within a 1 week period [33, 55]. Here, we employed cGH infusion to identify ATAC regions
541 whose functional enhancer activity in liver is regulated by the sex-dependent plasma GH pattern. cGH infusion
542 for 10 days was effective in feminizing the livers, as confirmed by the increased expression of female-biased
543 genes to control female liver levels (*A1bg*, *Sult3a1*, *Cux2)* and by the repression of male-biased genes down to
544 the low-level characteristic of female liver (complete repression, *Elovl3*; partial repression, *Cyp7b1*) (Fig. S2B).
545 We also examined the expression of several stress genes, which were previously found to be elevated in
546 expression in liver for up to 4 days following the surgical procedure used for osmotic minipump implantation
547 [33]. No significant induction of the stress genes tested was seen in mouse liver 10 days after osmotic mini
548 pump implantation; moreover, HDI itself did not lead to stress gene induction when assayed after 11 days (Fig.
549 S2B). Thus, the 11-day time period after HDI that we selected for evaluation of STARR-seq enhancer activity is
550 sufficient for recovery from any initial stress-related responses to HDI or osmotic minipump implantation.
551
552 Next, we examined the set of 4,126 robust active reporters identified by HDI-STARR-seq (Table S2) to identify
553 reporters whose functional enhancer activity, assayed by HDI-STARR-seq, is sex-biased and/or responsive to
554 cGH infusion. Overall, a total of 1,397 of the 4,126 robust active reporters exhibited enhancer activity that was
555 either male-biased (n=310), female-biased (n=541), cGH-induced (398) and/or cGH-repressed (n=419), as
556 indicated by a > 2-fold change in normalized STARR-seq RNA activity between conditions, averaged over
557 biological replicates (Table S7A, column AZ). Together, these represent 840 distinct DAR regions. The sex-bias
558 of enhancer activity was a key determinant of the response of STARR-seq reporter activity to cGH infusion.
559 Thus, 87 of 310 male-biased enhancers responded to cGH infusion by repression of enhancer activity, whereas
560 only 1 male-biased enhancer showed an increase in enhancer activity. Moreover, enhancer activity was
561 induced by cGH infusion for 179 of the 541 female-biased enhancers, whereas only 4 of the female-biased
562 enhancers showed a decrease in enhancer activity (Fig. 4A). This effect of cGH on STARR-seq enhancer activity
563 is consistent with the well-established feminization of liver chromatin accessibility, and of gene expression, in
564 cGH-infused male mouse liver [33, 55] and lends strong support to the proposal that the episomal STARR-seq
565 reporters recapitulate the physiological responses of the endogenous transcriptional and epigenetic mediators
566 of GH action regulating liver sex differences.
567
568 Further insight was obtained by plotting the fold change in STARR-seq enhancer activity following cGH infusion
569 (cGH-male liver vs. male liver activity log2 ratio) versus the sex bias in STARR-seq enhancer activity (female
570 liver vs. male liver activity log2 ratio) for the full set of 725 STARR-seq reporters mapping to 360 male-biased

571 and cGH-repressed DARs (Fig. 4B), and also for the full set of 717 STARR-seq reporters mapping to 252 female-
572 biased and cGH-induced DARs (Fig. 4C). In both cases there was a highly significant positive correlation
573 between the response to cGH infusion and liver sex bias of enhancer activity (Table S7B). Correlations
574 obtained for other DAR sets whose chromatin accessibility was not significantly responsive to both sex and
575 cGH treatment were weaker and less significant (Fig. S7A, Fig. S7B). Importantly, separation of DARs that
576 showed sex-bias but not cGH responsiveness into autosomal (Fig. 4D) and sex chromosome DARs (Fig. 4E)
577 revealed stronger and more highly significant positive correlations for the autosomal DARs (p = 5.6E-31 vs p =
578 3.5E-03), consistent with the regulation of sex-biased DARs on sex chromosome by distinct, GH-independent
579 mechanisms.
580
581 **Enrichment of regulated active enhancers for DARs with matched response.** Enrichment analysis validated
582 the close association between the sex bias and cGH regulation of HDI-STARR-seq activity and the sex-bias and
583 cGH regulation of endogenous mouse liver chromatin accessibility (Table S7C). Thus, male-biased and cGH-
584 repressed DARs were significantly enriched for male-biased and cGH-repressed STARR-seq activity, when
585 compared to female-biased and cGH-induced DARs (ES=8.3, p=3.7E-08, Fisher's exact test: male-biased and
586 cGH-repressed STARR-seq activity at 41 of 725 STARR-seq-active reporters mapping to male-biased and cGH-
587 repressed ATAC regions, versus only 5 of 694 STARR-seq-active reporters mapping to female-biased and cGH-
588 induced ATAC regions). Moreover, of the STARR-seq active reporters mapping to a male-biased, cGH-
589 repressed ATAC region showing sex-biased and cGH-regulated STARR-seq activity, 41 showed male-biased and
590 cGH-repressed activity and 21 showed female-biased and cGH-induced activity. This compares to a
591 background set comprised of STARR-seq active reporters mapping to a female-biased, cGH-repressed ATAC
592 region that show sex-biased and cGH-regulated STARR-seq activity, of which 5 show male-biased and cGH-
593 repressed activity and 33 show female-biased and cGH-induced activity; ES=12.9, p=1.53E-07 (Fisher's exact
594 test) (Table S7C). Other, related enrichments are shown in Table S7C.
595
596 Comparison of active sex-biased and cGH-regulated DARs to their non-regulated sex-biased and cGH-regulated
597 DAR counterparts revealed no differences in the proportion of linked non-sex-biased genes (Fig. S9A) or in the
598 distance to linked genes or the significance of the linkages, both for male-biased and cGH-repressed DARs (Fig.
599 S9B; DARs in Pick 1 group) and for female-biased and cGH-induced DARs (Fig. S9C; DARs in Pick 2 group). Thus,
600 the metrics of gene linkages do not differ between regulated and non-regulated enhancers. Many of the sex-
601 biased and cGH-regulated DARs were linked to both sex-biased genes and non-sex-biased genes (Table S8,
602 clumns BU-BZ, CA-DB). However, the linkages to non-sex-biased genes were characterized by significantly
603 longer distances and significantly lower reliability (lower linkage score, lower linkage significance than the
604 linkages to sex biased genes (Fig. S9B, Fig, S9C).

605 **Epigenetic features enriched at active sex-biased enhancers.** We investigated the enrichment of chromatin
606 states and epigenetic marks, previously determined for adult male and female mouse liver [48], in the sets of
607 active functional regulated enhancers to identify epigenetic features characteristic active vs inactive sex-
608 biased DARs (Fig. 5). For these analyses, we compared the chromatin states of the sex-biased and cGH-
609 regulated reporters across each ATAC region to those of a background set comprised of 317 robust active but
610 non-regulated ATAC regions (Table S9A, column AI). Four sets of DARs were investigated: 41 regulated male-
611 biased and cGH-repressed DARs, 82 non-regulated male-biased and cGH-repressed DARs, 43 regulated female-
612 biased and cGH-induced DARs, and 53 non-regulated female-biased and cGH-induced DARs (Table S9A,
613 columns T-W). Fig. 5A shows that the active and regulated male-biased and female-biased enhancer DARs
614 identified by HDI-STARR-seq are significantly enriched for distinct subsets of enhancer chromatin states in
615 mouse liver. Thus, the set of 41 regulated male-biased and cGH-repressed DARs is significantly enriched for
616 chromatin state E6, but not for chromatin state E7, in male liver, whereas the set of 43 regulated female-
617 biased DARs is significantly enriched for chromatin state E7 but not for chromatin state E6 in female liver (see
618 Table S10A). States E6 and E7 are both characterized by open chromatin and the active histone mark H3K27ac
619 but differ in that E6 is more highly enriched for the enhancer mark H3K4me1, whereas E7 has high probability
620 for the promoter mark H3K4me3. The enrichment of these sex-specifically regulated DAR sets for chromatin
621 regions in distinct epigenetic states in mouse liver supports the proposal that the HDI-transfected plasmid

622 DNA sequences acquire epigenetic states similar to those found at the corresponding endogenous DNA
623 regions in mouse liver.

624 Consistent with the above findings, the sets of regulated DAR enhancers were also enriched for genomic
625 regions with activating and repressive histone marks (Fig. 5B) characterized previously in adult liver [48].
626 Strikingly, the regulated male-biased, cGH-repressed DARs showed stronger enrichments for histone marks
627 associated with active enhancers (ES = 8.6 for H3K27ac, ES = 14.2 for H3K4me1). The non-regulated male-
628 biased, cGH-repressed DARs showed weaker enrichments for these histone marks (ES = 3.27 for H3K27ac, ES =
629 5.25 for H3K4me1) when compared to the background set comprised of 317 non-regulated enhancers. We
630 also observed significant depletion of H3K27me3 repressive histone marks (ES = 0 to 0.16), but no significant
631 enrichment or depletion of H3K9me3 (repressive mark) or H3K36me3 (transcribed region mark). Notably, a
632 weak enrichment of H3K27me3 marks in male liver was correlated with regulated female-biased, cGH-induced
633 DARs (ES = 2.75, p=2.96E-3), indicating repression of female-specific enhancers in male liver. In contrast, the
634 regulated female-biased, cGH-induced DARs showed stronger enrichments for histone marks associated with
635 active enhancers (ES = 3.65 for H3K27ac, ES = 27.94 for H3K4me1) and promoters (ES = 4.50 for H3K4me3)
636 (Fig.3B, bottom; Table S10A). These differences imply the differential molecular mechanism underlying sex
637 specificity in mouse livers.

638 **Active enhancers identified by HDI-STARR-seq: mapping to disease relevant sex-biased genes and liver**
639 **chromatin states**. The 1,397 robustly active sex-biased and/or cGH-responsive reporters mapped to 840 ATAC
640 regions. Sex-biased genes predicted to be regulated by these ATAC regions, determined by multiOmic gene
641 linkage analysis, are shown in Table S9C. Of the 840 ATAC regions, 165 were male-biased, cGH-repressed DARs
642 that mapped to one or more male-biased genes, and 134 were female-biased, cGH-induced DARs that mapped
643 to one or more female-biased genes. The male-biased gene targets included several MASLD-enabling genes
644 [31], including *Bcl6*, *Cyp4a12a*, *LNC-LFAR1* and *Nox4*; and the female-biased gene targets included several
645 MASLD-protecting genes [31], including *Hao2*, *Fmo2*, *Sall1* and *Trim24*.
646
647 The set of 165 male-biased ATAC regions was comprised of the following chromatin states in adult male liver:
648 126 were in an enhancer state (76%), 17 in a promoter state (10%), 6 in a bivalent state, 12 in a transcribed
649 state, and 4 in an inactive state. The corresponding distribution of chromatin states for the 168 ATAC regions
650 in adult female liver was as follows: 101 were in an enhancer state (61%), 11 in a promoter state (6.5%), 11 in
651 a bivalent state, 21 in a transcribed state, and 21 in an inactive state. Thus, these male-biased ATAC regions
652 exhibit a shift away from enhancer and promoter states in male liver to an increased frequency of bivalent,
653 transcribed and inactive states in female liver. The set of 134 female-biased ATAC regions was comprised of
654 the following chromatin states in adult female liver: 98 were in an enhancer state (73%), 23 in a promoter
655 state (17%), 10 in a bivalent state, 2 in a transcribed state, and 1 in an inactive state. The corresponding
656 distribution of chromatin states for the 134 ATAC regions in adult male liver was as follows: 70 were in an
657 enhancer state (52%), 11 in a promoter state (8%), 23 in a bivalent state, 7 in a transcribed state, and 23 in an
658 inactive state. Thus, these female-biased ATAC regions exhibit a shift away from enhancer and promoter
659 states in female liver to an increased frequency of bivalent, transcribed and inactive states in male liver.

660 **Enrichment of GH-regulated liver TFs at active sex-biased enhancers.** We investigated the enrichment of the
661 regulated DAR enhancers for binding in mouse liver by STAT5 and Bcl6, two essential, GH-regulated liver TFs
662 implicated in GH-regulated sex-biased liver gene expression [22, 48] and in liver metabolic disease [1, 6, 21,
663 56]. Specifically, we examined the enrichment of sex-biased binding sites for each TF in the sets of regulated
664 versus non-regulated sex-biased DARs as compared to a background set comprised of 317 non-regulated
665 enhancers (Fig. 5C). Male-biased STAT5 binding sites identified by ChIP-seq were significantly enriched in the
666 set of regulated male-biased and cGH-repressed DARs but not in the non-regulated male-biased and cGH-
667 repressed DARs or in the set of female-biased and cGH-induced DARs. In contrast, female-biased STAT5
668 binding sites were significantly at both the regulated and the non-regulated female-biased and cGH-induced
669 DARs, but not in the set of male-biased and cGH-repressed DARs. These enrichments reflect, at least in part,
670 sex-dependent and GH-regulated differences in chromatin accessibility, insofar as there were no differences in

671 the frequency of STAT5 motif occurrences for the regulated versus non-regulated sex-biased robust active
672 enhancers for each DAR set (Fig.5D, Pick1-Pick4 sets; Table S13D).

673 A moderate enrichment of male mouse liver Bcl6 binding (ES = 4.25, p=1.41E-3) was found at the set of
674 regulated female-biased, cGH-induced DARs (Fig. 5C), consistent with the role of this transcriptional repressor
675 in down regulating female-biased gene expression in male liver [22, 48]. Here, too, no differences in Bcl6 motif
676 frequency were observed for the regulated versus non-regulated sets of cGH-responsive sex-biased robust
677 active enhancers for each DAR set (Pick1 and Pick2 sets; Fig. 5D, Table S13D). In contrast, Bcl6 motifs were
678 significantly enriched at the set of female-biased DARs stringently unresponsive to cGH infusion (Pick3_F DAR
679 set), where the regulated DARs were highly enriched for BCL6 motifs compared to the non-regulated DARs (ES
680 = 7.3, p=0.0007 by Fisher's Exact test; Fig. 5D). However, few regions in regulated Pick3 (female-biased DARs
681 or male-biased DARs with no responsive to GH) overlapped with Bcl6 binding sites (Table S10B). This finding
682 implicates Bcl6, a transcriptional repressor showing male-biased expression [8], as a determinant of the sex-
683 biased regulation of functional enhancer activity for these female-biased DARs. Thus, increased Bcl6 binding at
684 these DARs in male compared to female liver is proposed to repress functional enhancer activity, resulting in
685 the observed sex-biased regulation of enhancer activity. The non-regulated DARs in this set do not show a sex
686 bias in enhancer activity owing to their relative deficiency in Bcl6 motifs and hence Bcl6-dependent repression
687 in male liver. Finally, motif scans using FIMO analysis for Cux2, a female-specific transcriptional repressor [57],
688 showed significantly greater motif frequency in the regulated sets of male-biased DARs from Pick1 (male-
689 biased and cGH-repressed DARs) and Pick3_M (male-biased but not cGH-repressed DARs) when compared to
690 the non-regulated DAR sets (Fig. 5D). The binding of Cux2 to this motif in female liver at the regulated but not
691 at the non-regulated DARs may help explain the observed differences in the sex-dependence of functional
692 enhancer activity, given that Cux2 protein is essentially undetectable in male mouse liver [58]. The motif for
693 Hnf6 (Onecut1), which is highly similar to that of Cux2 [59], also showed significant enrichment at the
694 regulated compared to the non-regulated male-biased DAR sets (Fig. 5D, Table S10A). These findings support
695 the hypothesis that the specific array of GH-responsive transcription factors bound to the regulated vs non-
696 regulated male-biased DARs is a key determinant of enhancer activity manifested in HDI-STARR-seq.

697 **Motif enrichment and discovery in sex-biased and cGH-regulated enhancers.** We sought to discover motifs
698 for other, novel TFs that may discriminate between regulated and non-regulated sex-biased and cGH-
699 responsive DARs, both for the set of 41 regulated male-biased, cGH-repressed DARs, and for the set of 43
700 regulated female-biased, cGH-induced DARs. We used MEME-CHIP to discover *de novo* motifs enriched in
701 each DAR set compared to a suitable background set of sequences, and then identified the discovered motifs
702 by comparison to a database of established mouse TF motifs [51] (Fig. 6, Fig. 7). Three distinct background sets
703 of ATAC sequences were used in these analyses: robust active ATAC regions that showed the same sex bias
704 and cGH responsiveness as the test set of ATAC regions, but whose STARR-seq enhancer activity was not
705 regulated; a larger set comprised of 317 robust active but non-regulated ATAC regions; and a set of 65
706 stringently inactive ATAC regions. Top enriched motifs identified for the set of 41 regulated male-biased, cGH-
707 repressed DARs versus a background set comprised of 82 non-regulated male-biased, cGH-repressed DARs
708 included Esrra/Nr2f2 and Hnf4a (Fig. 6). These same motifs were also top motifs in at least one other
709 background set comparison (Table S13B). Top enriched motifs identified for the set of 43 regulated female-
710 biased, cGH-induced DARs versus a background set comprised of 53 non-regulated female-biased, cGH-
711 induced DARs included Sox7, Crx_3485.1, Zbtb7b_primary (Fig. 7). Other top motifs enriched in the set of 43
712 regulated female-biased, cGH-induced DARs when using other background sets included Sp4/Klf7/Smad3,
713 Ascl2 and Tcfe2a (Table S13B). Finally, we discovered de novo motifs significantly enriched in the set of 41
714 regulated male-biased, cGH-suppressed DARs when compared to a background set comprised of 43 regulated
715 female-biased, cGH-induced DARs (Fig. 6, Table S13B), and vice versa (Fig. 7, Table S13B). Motifs for
716 Mafb/Sox8 and Hnf4a/Isl2 were most significantly enriched in the set of 41 regulated male-biased, cGH-
717 suppressed DARs, and motifs for Nkx2-5 and Tcfap2e/Tcfap2c were most significantly enriched in the set of 43
718 regulated female-biased, cGH-induced DARs (Fig. 6, Table S13B).

719 Results shown in Table S13D aligned with the *de novo* motif discovery when higher frequency of the motifs
720 identified from foreground sets was found in the foreground sets. Male-biased, cGH-suppressed DARs with a

721 matching pattern of STARR-seq activity showed 29-49% occurrence of the *de novo* motifs found in the same
722 regulated DAR enhancers, while the non-regulated male-biased, cGH-suppressed DARs showed 15-27%
723 occurrence (Fig. 6A, Table S13D). Also, female-biased, cGH-induced DARs with matching STARR-seq activity
724 pattern showed 40-77% occurrence while the non-regulated female-biased, cGH-induced DARs showed 23-
725 34% occurrence (Fig. 7A, Table S13D). Higher frequencies of Hnf6 and Cux2 motifs were found in the regulated
726 enhancers that opened to a greater extent in male liver, whereas a higher frequency of Bcl6 motifs was found
727 T the regulated enhancers which opened more in female liver.
728
729 **TAD-based view of active enhancers.** We assayed a total of 749 ATAC regions across 10 TAD regions (Table S3,
730 Table S4), as summarized in Table S3A and detailed in Table S4D. Control TAD regions, which are devoid of sex-
731 biased (SB) genes, generally contained few regulated reporters, as expected. In contrast, TAD_2885, on chr19,
732 with its 111 ATAC regions and 47 linked genes, of which 28 are sex-biased, presents an interesting example.
733 TAD_2885 encompasses the *Cyp2c* gene locus and includes many female-biased genes and female-biased
734 regulatory regions concentrated at the 5' end within a female-biased inta-TAD [54], suggesting genuine
735 regulatory activity supported by tandem reporters with similar biases. Additionally, a strong cGH-suppressed,
736 male-biased enhancer near *Cyp2c70* is present at the end of TAD_2885, outside the female-biased intra-TAD
737 region. TAD_4231 contains male-biased genes with overlapping male-biased regulatory reporters, indicating
738 consistent regulation, while TAD_4827 is characterized by female-biased genes, overlapping with female-
739 biased, cGH-induced regulatory reporters that bind STAT5 weakly, and with chromatin accessibility differences
740 in ATAC regions showing sex-specific patterns.

741 **HDI-STARR-seq activity in NPCs**. Hepatocytes are preferentially transfected by HDI as compared to other liver
742 cell types [60], raising the question of whether in liver NPC enhancer activity can be assayed HDI-STARR-seq.
743 To address this question, we used HDI-STARR-seq to assay enhancer activity of 63 ATAC regions, encompassing
744 972 reporters, whose chromatin is preferentially accessible in either hepatic stellate cells, liver endothelial
745 cells, immune cells, or cholangiocytes (Table S11). Results were compared to those we obtained for a similar
746 number of ATAC regions preferentially accessible in hepatocytes. Overall, we observed a significant 1.8-fold
747 depletion ($p < 0.00001$, Fisher's exact test) of robust active enhancers in NPCs when compared to hepatocytes.
748 The depletion of robust active enhancers reached 3.6-fold in hepatic stellate cells and 1.8-fold in liver
749 endothelial cells, when compared to hepatocytes. Cholangiocytes, which share many characteristics with
750 hepatocytes [26, 27], showed a 1.73-fold enrichment of robust active enhancers compared to hepatocytes,
751 but did not reach statistical significance ($p = 0.2$), likely due to the limited number of cholangiocyte-enriched
752 DARs available for analysis.

753 **Discussion**
754 By integrating single nucleus multiOmics with functional in vivo enhancer profiling, we provide a
755 comprehensive validation of GH-regulated, sex-biased enhancers regulating hepatic sexual dimorphism. Our
756 HDI-STARR-seq analysis identified and validated 1,397 sex-biased and/or GH-regulated reporters representing
757 840 differentially accessible chromatin regions (DARs), revealing the mechanistic basis of hormone-regulated
758 transcriptional control. These findings establish several key principles. First, sex-biased chromatin accessibility
759 strongly predicts functional enhancer activity, as indicated by the strong positive correlations between
760 functional enhancer activity and chromatin accessibility across sex and hormonal conditions. Thus, male-
761 biased, cGH-suppressed DARs were significantly correlated with a matching pattern of STARR-seq enhancer
762 activity (R^2=0.32, p<2.2E-16) and female-biased, cGH-induced DARs showed similar concordance (R^2=0.49,
763 p<2.2E-16). Second, regulated enhancers exhibit distinct chromatin states and epigenetic signatures. Thus,
764 male-biased enhancers were 14-fold enriched for H3K4me1 and 9-fold for H3K27ac histone marks (chromatin
765 state E6), while female-biased enhancers were enriched for both of these enhancer marks and for the
766 promoter mark H3K4me3 (chromatin state E7), indicating sex-specific regulatory logic. Third, binding patterns
767 for key GH-regulated transcription factors differentiate functional from non-functional enhancers, with STAT5
768 binding enriched 5-10 fold at regulated enhancers and BCL6, CUX2, and HNF4α motifs showing sex-specific
769 enrichment patterns. Importantly, a subset of these validated enhancers link to sex-biased genes controlling
770 lipid metabolism (*Elovl3*, *Cyp4a12a*), cholesterol homeostasis (*Csad*), and pathways relevant to MASLD
771 susceptibility (*Bcl6, Nox4, LNC-LFRA1*) or protection (*Hao2, Fmo2, Trim24, Sall1*), providing mechanistic

772 insights into sex differences in metabolic disease susceptibility. Importantly, our HDI-STARR-seq approach
773 overcomes critical limitations of previous studies by assaying enhancer function in intact liver under
774 physiological hormonal conditions, where transfected reporters undergo chromatinization [35] and respond to
775 endogenous regulatory factors. These results provide mechanistic insight into how GH-regulated chromatin
776 accessibility translates into sex-dependent enhancer function and downstream metabolic phenotypes.
777
778 **Synthetic STARR-seq libraries.** Here, we prepared our HDI-STARR-seq reporter library using synthetic
779 oligonucleotide arrays (SurePrint HiFi Oligo Pools, Agilent) for MPRA [61, 62] encompassing selected sex-
780 biased open chromatin regions, which allowed us to incorporate the targeted genomic regions with superior
781 uniformity (Fig. 2A), which helps minimize PCR amplification bias during the initial STARR-seq library
782 preparation and during steps of reporter extraction and sequencing library preparation. We did observe a bi-
783 phasic correlation between library representation and probe region of GC content (Fig. S4B, Fig. S4C), which
784 may be inherent in the sequences to be assayed; however, independently processed plasmid libraries
785 separately prepared exhibited a strong correlation of read counts within 23,912 reporter regions (Fig. S4E).
786 These findings suggest the reporter representation is more dependent to their own sequence.

787 We used short chemically synthesized DNA oligos to tile open chromatin regions of interest. Short oligos allow
788 us to tile and finely map the enhancer activity across open chromatin regions, although longer fragments
789 (>500 bp) may be more useful to uncover synergic effects in longer regions. However, the fragment length has
790 been studied as a factor to affect the outcome of reporter activity [63, 64]. Library preparation from sheared
791 DNA requires accurate size selection when ideal fragment lengths within 300 bp are suggested to avoid biases
792 [43, 65]. As enhancer size increased, a modest but significant increase in the activity of both negative controls
793 and positive controls were observed, but only modestly increased in dynamic range was observed [63]. Library
794 with smaller fragments showed higher correlation as compared to libraries with larger fragments while the
795 extending sequences affect TF binding [64]. The incorporation of 15 bp of homology arm on each end of
796 synthetic oligos requires no adaptor ligation and minimizes the further introduction of PCR bias. The
797 synthesized libraries are subject to synthesis-based length restrictions, and short synthesized fragments may
798 have sequence-specific biases owing to factors such as repetitiveness, single-stranded DNA structures and
799 mappability issues [65]. Despite the fact that the synthesis of single oligo is limited by the current technology
800 in terms of the oligo length [66], it is, nonetheless, possible to assay a longer region at once in a single reporter
801 by ligating individual oligos for synergic enhancers [64, 67]. Thus, a synthesized DNA source allows us to start
802 with an oligo pool of more evenly distributed. The variation may be inherent in the intrinsic sequence, and
803 result from a series of PCR steps even though minimized PCR cycles and multiple PCR reactions were applied.

804 **Active enhancers identified by STARR-seq.** HDI-STARR-seq is a flexible and scalable assay used to identify
805 condition-specific enhancers in a tissue-specific and cell-type-selective manner within an intact, complex
806 organ in vivo. HDI-STARR-seq integrates liver-specific plasmid library delivery with the coupling of putative
807 enhancer sequences to a liver-specific Alb minimal promoter. It supports large-scale multiplexing of functional
808 enhancer activity assays within a single mouse. Here, we leveraged 10x Genomics MulitOmic sn-ATAC/RNA-
809 seq technology to assay those hepatic sex-specific genes and DARs, and to identify regulatory DNA elements in
810 sex specificity and/or cGH regulation in liver. Analyses of 10x Genomics MultiOmic single nucleus ATAC-/RNA-
811 seq datasets showed most (76.5-90.7%) of the GH-regulated, sex-biased DARs and differential expressed
812 genes were almost exclusively limited to hepatocyte clusters [31], consistent with our finding of lower HDI-
813 STARR-seq enhancer activity in the non-hepatocytes ATAC regions (Fig. S8). The HDI-STARR-seq active
814 enhancers were highly enriched for endogenous liver epigenetic signatures of active enhancers and deficient
815 in repressive histone marks in mouse liver (Fig. 5), indicating they retain sufficient sequence information to
816 recreate the essential chromatin state required for functional activity in vivo. These findings support that
817 chromatin structure plays a crucial role in modulating sex-dependent gene expression by influencing
818 chromatin accessibility.
819
820 **Regulated STARR-seq enhancers in sex-biased DARs.** The active enhancers identified here were enriched for
821 regulated enhancer features in untreated male and female liver, indicating HDI-transfected plasmid DNA
822 sequences acquire epigenetic states similar to the corresponding endogenous liver DNA regions (Fig. 5A).

823 Enriched regulated active enhancers in the differentially accessible regions (DARs) with matched responses
824 showed a close association between sex-biased and cGH-regulated chromatin accessibility in mouse liver and
825 the sex bias and cGH-regulation of HDI-STARR-seq activity of the HDI-transfected reporters (Table S7).
826 Regulated male-biased, cGH-repressed DARs showed strong enrichment for active histone marks (Fig. 5B,
827 Table S10A), while non-regulated DARs showed weaker enrichment (Table S10A). Conversely, regulated
828 female-biased, cGH-induced DARs were significantly enriched for active enhancers and promoters, highlighting
829 different molecular mechanisms underlying sex specificity in mouse liver. Although female-biased, cGH-
830 induced regulated DARs showed a more distinct promoter feature, indicated by H3K4me3 histone marks, the
831 distance between the regulated DARs and potential target genes predicted by 10x datasets showed no
832 significant difference; thus, the active female-based DARs are *not* closer in proximity to their sex-biased gene
833 target promoters and transcription start sites. Furthermore, no significant differences in linkage significance
834 were found between regulated DARs and sex-biased and cGH-regulated genes (Fig. S9), suggesting other
835 factors, such as histone modifications or embedded motifs, are more critical for enhancer function. Indeed,
836 motif analysis identified enriched transcription factor binding sites, particularly STAT5, in regulated sex-biased
837 and cGH-regulated DAR enhancers, suggesting direct regulation by STAT5 in male-biased, cGH-repressed
838 DARs.
839
840 **Comparison of STARR-seq activity in ATAC regions showing the highest difference between hepatic and non-**
841 **hepatic cells.** We examine the STARR-seq activity from the selected cell-type specific ATAC regions, and we
842 investigated whether HDI-STARR-seq using a minimal *Albumin* promoter and hydrodynamic injection,
843 specifically designed for hepatocytes, may show high activity for NPC ATAC regions showing the highest
844 specificity toward other cell types. We found enhancer activity of robust reporters in the hepatic cells, as
845 calculated by the ratio of RNA reads divided by DNA reads, was significantly greater than that in the non-
846 hepatic cells (Fig. S8C). Non-hepatic ATAC regions were less active, possibly due to the lack of cell-type specific
847 TF and chromatin modulators in hepatocytes. This finding aligns with the specific transfection method we used
848 –HDI—which shows specificity for hepatocytes, and the effect of incorporation of an *Albumin* minimal
849 promoter on STARR-seq reporter activity. Additionally, we excluded the possibility of lack of active reporter or
850 plasmid representing those cell-type specific ATAC region delivered by HDI (Fig. S8A, Fig. S8B). Of the 23,912
851 predesigned reporters, specific subsets of 11,996 active reporters and 4,126 robust active reporters are
852 categorized under hepatic (including "Hepa", "PC" and "PP") and non-hepatic (including "Endo", "HSC", "Imm"
853 and "cholangio") ATAC regions (Fig. S8A). For the hepatic ATAC regions, 1049 reporters were designed, of
854 which 648 were designated active and 253 showed robust activity; for the non-hepatic ATAC regions, 972
855 reporters were designed, of which 594 were active and 146 inactive. Thus, no enrichment was found in the set
856 of active reporters in hepatocytes as compared to the set of active reporters in non-hepatocytes, indicating
857 the transfection was independent of the sequences represented.

858 **Sex-specificity regulation in sex and autosomal chromosome.** We assayed 288 sex-biased but not cGH
859 responsive DARs, of which 102 are male-specific, non cGH responsive DARs (i.e., ATAC regions with more
860 hypersensitivity in both male liver and cGH-treated liver than that in female liver) and 186 female-specific, not
861 cGH-responsive (i.e., more ATAC activity in female liver than that in male and cGH-treated liver). Note that 61
862 DARs localized to the sex chromosomes highly exhibiting sex-biased hypersensitivity and they are GH-
863 independent DARs. 20 of the 102 male-specific DARs are on chromosome Y, including 18 of the top 20 when
864 ranked by sex-bias. 41 of the 186 are on chromosome X, including 19 of the top 20 when ranked by sex-bias.
865 The 102 male-specific, not cGH responsive DARs were linked to 50 genes, whereas 9 genes are male-specific,
866 cGH-repressed genes and 23 genes are male-specific, but not cGH-responsive genes. In contrast, the 186
867 female-specific, non-cGH responsive DARs were linked to 96 genes, whereas 25 genes are female-specific,
868 cGH-induced genes and 37 genes are female-specific but NOT cGH responsive. Increasing studies have shown
869 that sex differences contributed by sex chromosomes associates with disease susceptibility and progression
870 [68]. By assaying these sex chromosomal DARs and top sex-biased DARs, we can identify how sex-biased DARs
871 on sex chromosome modulate sex-biased genes.

872 **Conclusions.** The HDI-STARR-seq platform implemented here enabled us to characterize in liver in vivo a
873 subset of hepatocyte DARs, discovered by sn-based multiOmics, as functional sex-biased enhancers linked to

874 sex-biased genes implicated in susceptibility or protection from MASLD, bridging the gap between genomic
875 correlation and causal mechanistic insights. HDI-STARR-seq is shown to be a flexible and scalable assay that
876 identifies condition-regulated enhancers in a tissue-specific and cell type-selective manner. By delivering large,
877 multiplexed libraries of putative enhancer sequences, coupled to a liver-specific *Albumin* minimal promoter,
878 HDI-STARR-seq enabled functional enhancer activity profiling within livers of individual mice. Importantly, the
879 sex-biased enhancers identified are highly enriched for endogenous liver epigenetic signatures, suggesting
880 they contain sufficient sequence information to recapitulate essential chromatin states required for functional
881 activity in vivo. Further, our discovery of distinct GH-regulated epigenetic states and transcription factor
882 architectures underlying active male-biased versus female-biased hepatocyte enhancers reveals that sexual
883 dimorphism emerges from nuanced chromatin-level regulation responsive to hormonal cues. This mechanistic
884 understanding has broad relevance for understanding how disruptions—in the form of genetic variants,
885 epigenetic dysregulation, or altered hormone signaling—may precipitate sex-specific disease pathogenesis.
886 Future studies applying this methodology to disease models will accelerate discovery of disease-associated
887 regulatory variants and inform precision medicine approaches that account for sex differences in metabolic
888 outcomes.
889
890 **Figure legends**
891
892 **Fig 1. Combined transcriptional and chromatin profile from 10x Genomics MultiOmic datasets.** (A) UMAP of
893 mouse liver cells based on transcriptomic and chromatin accessibility profiles of 34,364 single nuclei obtained
894 using 10X Genomic MultiOme technology integrated from four samples representing three biological
895 conditions: male liver, female liver and cGH-treated male liver. (B) Example of global links from a female-
896 biased, cGH-induced DARs (named as "ATAC_39124" located at chr5: 122,003,446-122,004,010 in mm10;
897 labelled as "*") identified by single-nucleus 10x Genomics MultiOmic sequencing. Left, top: predicted
898 regulatory relationship from the DAR to other genomic elements including linked genes or ATAC regions. Solid
899 lines indicate potential regulation between ATAC_39124 and 7 target genes, including 6 PCGs and 1 lncRNA,
900 based on the "co-expression" of the accessibility of the DAR and the expression of its target gene across
901 hepatocyte population. Bottom: differential accessible regions identified within three biological conditions.
902 Right: histogram graphs showed 4 of 7 the linked genes (bold) with significant correlation as their -log10(FDR)
903 are greater than 5. The highest correlation of 0.30 was found in the association of ATAC_39124 and Cux2
904 showing a high correlation of female-biased and cGH-induced expression (see the histogram graphs).
905
906 **Fig 2. Design of the synthetic STARR-seq approach.** (A) In vivo STARR-seq library prepared from synthetic
907 oligos. Left, middle: Reporters containing a window of 180 nt and 15 nt of adaptor sequences (shown in green
908 and blue) at two ends tile overs extended ATAC regions to be assayed in STARR-seq. Reporters were designed
909 for 55 bp shifting between two neighboring reporters with a 125 bp overlapping region. The adaptor
910 sequences were designed for PCR reaction, where the primers anneal and incorporate homology arms for the
911 HiFi cloning step. Right: PCR amplification followed by HiFi cloning to attach homology arms to ssDNA oligos.
912 (B) Selection of sex-biased and/or cGH-regulated DARs for synthetic STARR-seq library. Four categories of sex-
913 biased and/or cGH-regulated DARs were defined based on their response of ATAC activity and the linked
914 genes. (C) Number of active DARs with at least 1 robust active reporter regions in each category. The number
915 in the parentheses represents the number of robust active reporter regions from the set of 4,126 reporters,
916 defined later in Fig. S5F. For example, 250 of the 360 male-biased, cGH-repressed DARs are active whereas
917 724 robust active reporters are present in the subset of 4,126 reporters.
918
919 **Fig. 3. Characterization of in vivo STARR-seq reporter library.** (A) Validation of reporter plasmid transfection
920 using luciferase activity. pGYL18 plasmid encoding Renilla luciferase were co-transfected with pooled STARR-
921 seq reporters and delivered to liver cells using hydrodynamic tail vein injection. Five livers with lower
922 luciferase activity (labelled by "*") were removed from the downstream analyses. (B) Clustering analysis of
923 normalized mapped counts in ATAC regions from biological replicates for 2 plasmid libraries, 3 DNA and 7 RNA
924 reporter libraries. DNA and RNA libraries were extracted from HDI livers in the biological conditions (Male
925 labelled as "M", Female labelled as "F", cGH-treated male labelled as "cGH_M"). As the heatmap showed, RNA
926 libraries are separated from plasmid and DNA libraries, and clustered by their biological conditions, supporting

927 that RNA activity are functionally biased. (C-D) Library complexity of reporters extracted from transfected
928 mouse livers. (C) Frequency distribution of averaged normalized read counts within 23,912 reporters from 2
929 plasmid and 3 DNA libraries. 81.9% of reporters are within the range of 6.4x median. (D) Strong correlation of
930 sequence reads mapping to individual STARR-seq reporters for plasmid libraries compared to DNA libraries
931 extracted from HDI-transfected livers, indicating that STARR-seq reporter plasmid uptake and retention
932 following HDI is independent of reporter sequence.
933
934 **Fig. 4. Responsiveness of STARR-seq activity in robust active reporters.** (A) Venn diagram showing the
935 number of reporters that are sex-biased or cGH regulated. (B) Correlation between the response of male liver
936 STARR-seq reporter activity to cGH infusion and the sex bias of STARR-seq reporter activity for the n=725
937 robust active reporters (subset of 4,126 such reporters; Table S2A) that are at male-biased, cGH-repressed
938 DARs. 41 of the male-biased and cGH-repressed DARs have at least one male-biased reporter and at least one
939 cGH-repressed reporter (>2-fold difference in reporter activity between conditions; datapoints within blue
940 box), while 82 male-biased and cGH-repressed DARs have no regulated reporters (subset of datapoints in
941 green box). See Table S7B for correlation data. (C) Correlation as in panel B for n=693 robust active reporters
942 that are at female-biased, cGH-induced DARs. 43 female-biased and cGH-induced DARs have at least one
943 female-biased reporter and at least one cGH-induced reporter (datapoints within red box), while 53 female-
944 biased and cGH-induced DARs have no regulated reporters (subset of datapoints in green box). (D) Correlation
945 as in panel B for n=577 autosomal robust active reporters that are at sex-biased DARs that are not cGH-
946 responsive (Pick3_M and Pick3_F in Table S2). (E) Correlation as in panel B for the n=79 robust active reporters
947 on chrX or chrY that are at DARs that are sex-biased but not cGH-responsive. Blue lines in panels B-E represent
948 the best fit linear regression models, with the 95% confidence interval show in gray.
949
950 **Fig. 5. HDI-STARR-seq regulated enhancers: enrichment for epigenetic features and TF binding in mouse**
951 **liver**. Enrichment scores for the overlap between regulated and non-regulated enhancer DARs identified by
952 HDI-STARR-seq and the sets of epigenetic and TF binding features indicated in each panel. Enrichments in A, B,
953 and C were determined for 4 active enhancers sets: 41 regulated male-biased, cGH-repressed DARs (*blue*), 82
954 non-regulated male-biased, cGH-repressed DARs (*black*), 43 regulated female-biased, cGH-induced DARs (*pink*)
955 and 53 non-regulated female-biased, cGH-induced DARs (*gray*) (see Fig. 4). Enrichments were computed
956 relative to a background set of 317 nonregulated genomic regions, which is a subset of 4,126 robust active
957 reporters showing no regulated activity when comparing sexes or the response to cGH-infusion. Shown are
958 enrichments for: (A) occurrence of the chromatin states exhibiting significance, determined for male or female
959 liver by ChromHMM analysis, with the emission probabilities describing each state shown in the inset; (B)
960 Histone H3 chromatin marks identified by ChIP-seq in male or female mouse liver; (C) TF binding sites for the 2
961 indicated TFs. The significance of enrichment of each active enhancer set was determined by Fisher's exact
962 test: *, (-log10P) > 3; Two **, (-log10P) > 5; Three ***, (-log10P) > 7; Full dataset is shown in Table S10A.
963
964 **Fig. 6, Fig. 7. Motif analysis on regulated DARs**. De novo motifs and the compared motifs in a mouse motif
965 database (UniProbe) identified by MEME-ChIP for a variety of combination of foreground and background
966 sets.
967
968
969 **Supplemental Figure legends:**
970
971 **Fig. S1. Nucleus isolation.** Representative images showing intact and well-separated nuclei were isolated from
972 historical livers from male and female mice.
973
974 **Fig. S2. Gene expression of sex-biased genes and stress genes in vehicle male mice, female mice and 10 day**
975 **cGH-treated male mice as indicated.** (A) Mouse study design for assaying sex-biased and/or cGH-regulated
976 open chromatin regions using in vivo STARR-seq delivered to mouse liver by HDI. A mixture of STARR-seq
977 plasmids, along with an internal control plasmid encoding Renilla luciferase, was administered to mouse livers
978 on day 0 via HDI. Following this, osmotic pumps were implanted on day 1 for cGH treatment or vehicle control,
979 specifically for male livers. Livers were then harvested on day 11 for analysis. (B) The effect of continuous GH

980 treatment on the RNA expression of sex-biased genes. RT-qPCR revealed that female-biased genes (A1bg, Cux2
981 and Sult3a1) were up-regulated after 10 days of cGH treatment in male livers. Male-biased genes (Cyp7b1 and
982 Elovl3) were partially or fully repressed after cGH treatment. Statistical tests by Kolmogorov-Smirnov test (N=6
983 per group), followed by multiple comparison with false discovery rate (desired FDR = 1%) for multiple
984 comparison correction. Data shown are mean +/- SEM values, with "male livers with vehicle pump" mean
985 value set = 1 for normalization. (C) No stress genes were induced after 10 days of pump implementation and 7
986 days of HDI administration. Statistical tests by Kolmogorov-Smirnov test (N=3-5 per group), followed by
987 multiple comparison with false discovery rate (desired FDR = 1%) for multiple comparison correction. Data
988 shown are mean +/- SEM values, with "male livers w/o HDI administration" mean value set = 1 for
989 normalization.
990
991 **Fig. S3. Synthetic STARR-seq preparation.** (A) Pooled inserts, started from synthetic ssDNA oligos, were
992 amplified by PCR to attach sequences of homology arms at ends, were expected to be 307 bp long. (B) Crude
993 PCR products with 10 and 21 PCR cycles. Gel electrophoresis showed the size of the PCR fragments after 10
994 amplification cycles (Lane 3) and 21 amplification cycles (Lane 5) matched with the expected length. The
995 fragments with 10 PCR cycles were proceeded for plasmid library preparation. Lanes 1 and 6 are 100 bp NEB
996 marker (New England Biolabs, #N3231S). Lane 3 represents the amplified HA-adapted DNA fragments with
997 minimized PCR cycles, to be used in the library preparation and the lower molecular bands of unused primers
998 and primer dimer will be removed using SPRI beads prior to the subsequential cloning steps. Lane 5 represents
999 the overamplification of STARR-seq inserts. Lanes 2 and 4 are irrelevant.
1000
1001 **Fig. S4. Reporter representation with respect to their GC content.** (A) The frequency distribution showing the
1002 GC content of 23,912 individual synthetic reporters (within the 180 nt sequence). Y-axis, number of library
1003 reporters at each % GC that were included in the design for the set of 23,912 reporters. (B-C) Averaged read
1004 counts of 23,912 reporters from 5 plasmid/DNA libraries (including 2 plasmid (B) and 3 extracted DNA (C)
1005 libraries) were sorted by GC content, then grouped into 500 bins. Data show biphasic correlation of reporter
1006 representation as judged by read counts in PCR-amplified libraries versus GC content. (D) Scatterplot showed
1007 the alignment of normalized read counts in two technical replicates (G202M1 and G202M2) of the plasmid
1008 libraries amplified separately.
1009

1010 **Fig S5. Active reporters in the STARR-seq assay across three biological conditions.** (A) 900 selected sex-biased
1011 DARs were tiled with multiple reporters to provide a profile at a finer resolution. Read count was calculated in
1012 a reporter basis with >95% overlaps using the coordinates of 23,912 reporters. Numbers in the parenthesis
1013 indicates the number of designated reporters in the selected categories (e.g., 4,949 reporters are designed for
1014 360 male-specific, cGH-repressed DARs). (B) A summary of the number of active reporters in each of the three
1015 conditions based on the normalized read counts in DNA/plasmid libraries (n=5) or RNA libraries (n=2-3). Table
1016 showed the distributions of 19,136 DNA qualified reporter regions (i.e., meeting the threshold of 40
1017 normalized reads per 10 million mapped reads for each of the DNA/plasmid libraries) for RNA regions that
1018 were qualified in either 0, 1, or 2 of the 2 livers analyzed in untreated livers (male or female, as indicated), or
1019 in 0, 1, 2, or 3 of the 3 livers analyzed in cGH-treated livers (cGH_M), with 'qualified' indicating the liver met
1020 the threshold of 30 normalized reads per 10 million mapped reads for STARR-seq RNA libraries recovered from
1021 the HDI-transfected livers. (C) Histogram showed average RNA activity of 19,136 DNA qualified reporter
1022 regions in cGH-M (n=2), male (n=3) and female (n=2) mice, grouped into 150 bins. (D) Table showed the
1023 number of reporter regions using two DNA and RNA qualification thresholds within 6 different categories (Fig.
1024 2, Table S2). (E-F) Venn diagrams showing the 11,996 active reporters from the union of three active sets in
1025 three biological conditions (8,922 active reporters in cGH-treated male livers, 7,099 active reporters in male
1026 livers and 7,748 active reporters in female livers), and the number of ATAC regions represented by each
1027 subset. 4,126 reporters are robustly active commonly in three biological conditions in terms of RNA
1028 qualification and DNA qualification, and comprised of 1,446 unique ATAC regions, ideally for condition-specific
1029 comparison. Note: Designated reporters within an ATAC region can have different combination of
1030 qualification. 2,008 cGH-active only reporters come from 1,568 ATAC regions assayed in the STARR-seq library.
1031 Of the 1,568 ATAC regions, 1,446 ATAC regions have at least one robust active reporter in the same region, 58

1032 regions have at least one active reporter qualified in cGH and female conditions, 36 regions have at least one
1033 active reporter qualified in cGH and male conditions, and 28 active reporters have at least one active reporter
1034 qualified only in cGH male liver.

1035

1036 **Fig S6. Active reporters and stringently inactive reporters in the three biological conditions.** Plasmid reads in
1037 sets of active reporters (see Venn diagram in Fig. S5E) STARR-seq reporter regions. 4,126 robust active
1038 reporters whose recovered liver RNA met the threshold for qualification in all 2 replicate livers (cGH-treated
1039 male or female) or in all 3 replicate livers (male) have the highest representation in the original STARR-seq
1040 plasmid library used for HDI. The 982 stringently inactive reporters are defined as the regions consistently
1041 exhibiting low activity (less than 30 normalized reads) in RNA libraries of all biological replicates. We defined
1042 the top 400 stringently inactive reporters with higher plasmid representation (labeled as "str_inactive_H" in
1043 Fig. S6A) and the other 582 stringently inactive reporters with lower plasmid reads (labeled as "str_inactive_L"
1044 in Fig. S6A). The set of 582 stringently inactive reporters will be excluded from downstream comparison to
1045 avoid artificial effect of the low representation in the library. (B) Bar plot showed the number of ATAC regions
1046 represented by robust active reporters only, stringently inactive reporters only, and mixed of robust active and
1047 stringently inactive reporters. (C) Boxplots showed the ATAC Cut sites (ATAC intensity*100) in hepatocytes of
1048 the three subsets in panel (B) from the four pooled biological samples. In conclusion, the stringent inactive set
1049 with higher plasmid representation is a comparable background set in terms of plasmid representation and
1050 chromatin accessibility in hepatocytes for downstream enrichment analyses.
1051

1052 **Fig. S7 Characterization of regulated reporters.** (A) Correlation of cGH treatment and sex specificity in term of
1053 the ratio of STARR-seq reporter activity for the n=220 subset of 4,126 active reporters (Fig. 3) that are male-
1054 biased, not cGH-responsive DARs (Pick3_M in Table S2). (B) Correlation of cGH treatment and sex specificity in
1055 term of the ratio of STARR-seq reporter activity for the n=436 subset of 4,126 active reporters (Fig. 3) that are
1056 female-biased, not cGH-responsive DARs Blue lines in panel A through D represent the fitting linear regression
1057 models and the gray area around the lines represents a confidence interval of 95%. See Table S7B for
1058 correlation data. (C) Boxplot showed comparable plasmid representation of the subsets of regulated reporters
1059 and non-regulated reporters in each category.
1060

1061 **Fig. S8 HDI-STARR-seq activity in cell-type specific ATAC regions.** (A) Incorporation of reporters representing
1062 cell-type specific ATAC regions in the sets of 11,996 active reporters and 4,126 robust reporters extracted
1063 from HDI-administrated livers. Percentages were calculated by dividing the number of active reporters (11,996
1064 active reporters or 4,126 robust active reporters) by the 23,912 predesigned reporters in each cell type
1065 category. No enrichment (Fisher exact test) of active reporters representing hepatic ATAC regions was
1066 observed compared to non-hepatic ATAC regions, suggesting that the uptake of plasmid reporters with
1067 sequences from non-hepatic ATAC regions is independent of those from hepatic ATAC regions in the STARR-
1068 seq assay. (B) Boxplots showing the distribution of DNA reads and RNA reads of the robust active reporters in
1069 each category. (C) Boxplots showing the distribution of enhancer activity of the robust active reporters in each
1070 category. Enhancer activity in hepatic ATAC regions was significantly higher than in non-hepatic ATAC regions,
1071 as determined by the Kolmogorov-Smirnov test.
1072

1073 **Fig. S9. Gene linkage data for regulated vs non-regulated active enhancer DARs.** Metrics of individual
1074 linkages between ATAC regions and genes grouped based on their STARR-seq activity and gene response, as
1075 derived from Table S4B and Table S9. (A) Numbers shown are the number of genes that sex-biased or not sex-
1076 biased and are linked to regulated regions or to non-regulated regions, as indicated. There is no significant
1077 difference between regulated and non-regulated sex-biased and cGH-responsive DARs regarding the
1078 proportion of linked sex-biased genes vs linked non-sex-biased genes, as seen for both Pick1 DARs (red; male-
1079 specific and cGH-repressed ATAC regions) and Pick2 DARs (green; female-biased and cGH-induced ATAC
1080 regions) (also see Table S12). (B, C) Boxplots showed the distribution of sex-biased regulated enhancers
1081 (salmon) or non-regulated regions (blue) with linked genes which are sex-biased (male biased and cGH-
1082 suppressed genes for regulated male biased and cGH-suppressed DARs (i.e., 41 regulated regions); female
1083 biased and cGH-induced genes for regulated female biased and cGH-induced DARs (i.e., 43 regulated regions),

not sex-biased, or any genes including sex-biased and not sex-biased (as denoted "all") in the subpanels. Statistical analyses were performed by applying the Kolmogorov-Smirnov test to the distribution of each of metric (i.e., score, significance, distance) between the linkage to sex-biased genes and the linkage to non sex-biased genes, considering the regulated regions and non-regulated regions as a group: '***' labeled in the figure indicates p-values are < 0.001. Moreover, there are no significant differences (p-values >0.05, not drawn in the figure) between regulated and non-regulated sex-biased and cGH-responsive DARs regarding the distributions of linkage scores, linkage distances (in nt), and significance of linkages (-log(pvalue)) in the groups regardless of their linkage to sex-biased genes or non-sex-biased genes.

Fig. S10. HDI-STARR-seq reporter activity for autosomal compared to sex chromosomes.
Boxplots showing the distributions of enhancer activity (female/male liver ratio, and cGH-infused male/male liver ratio, as indicated) for each mouse chromosomes for all robust active reporters from Pick3_F and Pick3_M. The data shown represents 436 and 220 robust active STARR-seq probes in Pick3_F and Pick3_M, respectively. No difference in STARR-seq reporter activity was seen for sex chromosomes as compared to autosomal chromosomes, by judged from Kolmogorov-Smirnov test. Almost all STARR-seq reporters assaying sex chromosome are in Pick3_F and Pick3_M, except 1 reporter from Pick6.

Fig. S11. HDI-STARR-seq reporter activity for ATAC regions that are far from genes.
Boxplots (top) show the distribution of STARR-seq activity in three biological conditions for individual robust active reporters in each of the 6 indicated DAR Pick categories. Results suggest that DAR regions in Pick #6, which are distant from genes, are active (bottom) as the fraction of the number of robust active reporters in the predesignated reporters in Pick1 and Pick6 are similar. Based on enrichment analysis using Fisher's exact test, there is no enrichment of robust active reporters in Pick6 as compared to robust active reporters in Pick1.

References

[1] D.J. Waxman, R.D. Kineman, Sex matters in liver fat regulation, Science, 378 (2022) 252-253.
[2] D.J. Waxman, M.G. Holloway, Sex differences in the expression of hepatic drug metabolizing enzymes, Mol Pharmacol, 76 (2009) 215-228.
[3] D.J. Chia, J.J. Young, A.R. Mertens, P. Rotwein, Distinct alterations in chromatin organization of the two IGF-I promoters precede growth hormone-induced activation of IGF-I gene transcription, Mol Endocrinol, 24 (2010) 779-789.
[4] D.E. Berryman, E.O. List, Growth Hormone's Effect on Adipose Tissue: Quality versus Quantity, Int J Mol Sci, 18 (2017).
[5] K.M. Mueller, J.W. Kornfeld, K. Friedbichler, L. Blaas, G. Egger, H. Esterbauer, P. Hasselblatt, M. Schlederer, S. Haindl, K.U. Wagner, D. Engblom, G. Haemmerle, D. Kratky, V. Sexl, L. Kenner, A.V. Kozlov, L. Terracciano, R. Zechner, G. Schuetz, E. Casanova, J.A. Pospisilik, M.H. Heim, R. Moriggl, Impairment of hepatic growth hormone and glucocorticoid receptor signaling causes steatosis and hepatocellular carcinoma in mice, Hepatology, 54 (2011) 1398-1409.
[6] D. Kaltenecker, M. Themanns, K.M. Mueller, K. Spirk, T. Suske, O. Merkel, L. Kenner, A. Luis, A. Kozlov, J. Haybaeck, M. Muller, X. Han, R. Moriggl, Hepatic growth hormone - JAK2 - STAT5 signalling: Metabolic function, non-alcoholic fatty liver disease and hepatocellular carcinoma progression, Cytokine, 124 (2019) 154569.
[7] M. Baik, J.H. Yu, L. Hennighausen, Growth hormone-STAT5 regulation of growth, hepatocellular carcinoma, and liver metabolism, Ann N Y Acad Sci, 1229 (2011) 29-37.
[8] R.D. Meyer, E.V. Laz, T. Su, D.J. Waxman, Male-specific hepatic Bcl6: growth hormone-induced block of transcription elongation in females and binding to target genes inversely coordinated with STAT5, Mol Endocrinol, 23 (2009) 1914-1926.
[9] M.J. Waters, A.J. Brooks, Growth hormone and cell growth, Endocr Dev, 23 (2012) 86-95.
[10] M.J. Waters, The growth hormone receptor, Growth Horm IGF Res, 28 (2016) 6-10.
[11] T. Decker, P. Kovarik, A. Meinke, GAS elements: a few nucleotides with a major impact on cytokine-induced gene expression, J Interferon Cytokine Res, 17 (1997) 121-134.

1137 [12] D.J. Chia, M. Ono, J. Woelfle, M. Schlesinger-Massart, H. Jiang, P. Rotwein, Characterization of distinct
1138 Stat5b binding sites that mediate growth hormone-stimulated IGF-I gene transcription, J Biol Chem, 281
1139 (2006) 3190-3197.
1140 [13] Z. Huang, L. Huang, M.J. Waters, C. Chen, Insulin and Growth Hormone Balance: Implications for Obesity,
1141 Trends Endocrinol Metab, 31 (2020) 642-654.
1142 [14] S. Al-Samerria, S. Radovick, Exploring the Therapeutic Potential of Targeting GH and IGF-1 in the
1143 Management of Obesity: Insights from the Interplay between These Hormones and Metabolism, Int J Mol Sci,
1144 24 (2023).
1145 [15] M.A. Abu El-Makarem, M.F. Kamel, A.A. Mohamed, H.A. Ali, M.R. Mohamed, A.E.M. Mohamed, A.M. El-
1146 Said, M.G. Ameen, A.A. Hassnine, H.A. Hassan, Down-regulation of hepatic expression of GHR/STAT5/IGF-1
1147 signaling pathway fosters development and aggressiveness of HCV-related hepatocellular carcinoma: Crosstalk
1148 with Snail-1 and type 2 transforming growth factor-beta receptor, PLoS One, 17 (2022) e0277266.
1149 [16] P. Hao, D.J. Waxman, STAT5 Regulation of Sex-Dependent Hepatic CpG Methylation at Distal Regulatory
1150 Elements Mapping to Sex-Biased Genes, Mol Cell Biol, 41 (2021).
1151 [17] M.G. Holloway, Y. Cui, E.V. Laz, A. Hosui, L. Hennighausen, D.J. Waxman, Loss of sexually dimorphic liver
1152 gene expression upon hepatocyte-specific deletion of Stat5a-Stat5b locus, Endocrinology, 148 (2007) 1977-
1153 1986.
1154 [18] K.H. Clodfelter, M.G. Holloway, P. Hodor, S.H. Park, W.J. Ray, D.J. Waxman, Sex-dependent liver gene
1155 expression is extensive and largely dependent upon signal transducer and activator of transcription 5b
1156 (STAT5b): STAT5b-dependent activation of male genes and repression of female genes revealed by microarray
1157 analysis, Mol Endocrinol, 20 (2006) 1333-1351.
1158 [19] S.H. Park, C.A. Wiwi, D.J. Waxman, Signalling cross-talk between hepatocyte nuclear factor 4alpha and
1159 growth-hormone-activated STAT5b, Biochem J, 397 (2006) 159-168.
1160 [20] C.A. Wiwi, D.J. Waxman, Role of hepatocyte nuclear factors in transcriptional regulation of male-specific
1161 CYP2A2, J Biol Chem, 280 (2005) 3259-3268.
1162 [21] J. Nikkanen, Y.A. Leong, W.C. Krause, D. Dermadi, J.A. Maschek, T. Van Ry, J.E. Cox, E.J. Weiss, O.
1163 Gokcumen, A. Chawla, H.A. Ingraham, An evolutionary trade-off between host immunity and metabolism
1164 drives fatty liver in male mice, Science, 378 (2022) 290-295.
1165 [22] Y. Zhang, E.V. Laz, D.J. Waxman, Dynamic, sex-differential STAT5 and BCL6 binding to sex-biased, growth
1166 hormone-regulated genes in adult mouse liver, Mol Cell Biol, 32 (2012) 880-896.
1167 [23] M. Endo, Y. Takahashi, Y. Sasaki, T. Saito, T. Kamataki, Novel gender-related regulation of CYP2C12 gene
1168 expression in rats, Mol Endocrinol, 19 (2005) 1181-1190.
1169 [24] E.V. Laz, A. Sugathan, D.J. Waxman, Dynamic in vivo binding of STAT5 to growth hormone-regulated genes
1170 in intact rat liver. Sex-specific binding at low- but not high-affinity STAT5 sites, Mol Endocrinol, 23 (2009) 1242-
1171 1254.
1172 [25] N. Aizarani, A. Saviano, Sagar, L. Mailly, S. Durand, J.S. Herman, P. Pessaux, T.F. Baumert, D. Grun, A
1173 human liver cell atlas reveals heterogeneity and epithelial progenitors, Nature, 572 (2019) 199-204.
1174 [26] F. Hildebrandt, A. Andersson, S. Saarenpaa, L. Larsson, N. Van Hul, S. Kanatani, J. Masek, E. Ellis, A.
1175 Barragan, A. Mollbrink, E.R. Andersson, J. Lundeberg, J. Ankarklev, Spatial Transcriptomics to define
1176 transcriptional patterns of zonation and structural components in the mouse liver, Nat Commun, 12 (2021)
1177 7046.
1178 [27] R.P. Cunningham, N. Porat-Shliom, Liver Zonation - Revisiting Old Questions With New Technologies,
1179 Front Physiol, 12 (2021) 732929.
1180 [28] M.J. McConnell, E. Kostallari, S.H. Ibrahim, Y. Iwakiri, The evolving role of liver sinusoidal endothelial cells
1181 in liver health and disease, Hepatology, 78 (2023) 649-669.
1182 [29] F. Tacke, H.W. Zimmermann, Macrophage heterogeneity in liver injury and fibrosis, J Hepatol, 60 (2014)
1183 1090-1096.
1184 [30] K. Du, J.H. Jun, R.K. Dutta, A.M. Diehl, Plasticity, heterogeneity, and multifunctionality of hepatic stellate
1185 cells in liver pathophysiology, Hepatol Commun, 8 (2024).
1186 [31] K. Karri, T.-Y. Chang, M. Pyatkov, S. Gandhi, T. Siggers, D.J. Waxman, Single nucleus multiomics links novel
1187 transcription factor motifs to murine hepatic sex differences in chromatin accessibility and metabolic
1188 dysfunction-associated steatotic liver disease, bioRxiv, https://doi.org/10.1101/2025.10.31.685903 (2025).

1189 [32] N. Percie du Sert, V. Hurst, A. Ahluwalia, S. Alam, M.T. Avey, M. Baker, W.J. Browne, A. Clark, I.C. Cuthill,
1190 U. Dirnagl, M. Emerson, P. Garner, S.T. Holgate, D.W. Howells, N.A. Karp, S.E. Lazic, K. Lidster, C.J. MacCallum,
1191 M. Macleod, E.J. Pearl, O.H. Petersen, F. Rawle, P. Reynolds, K. Rooney, E.S. Sena, S.D. Silberberg, T. Steckler,
1192 H. Wurbel, The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research, PLoS Biol, 18 (2020)
1193 e3000410.
1194 [33] D. Lau-Corona, A. Suvorov, D.J. Waxman, Feminization of Male Mouse Liver by Persistent Growth
1195 Hormone Stimulation: Activation of Sex-Biased Transcriptional Networks and Dynamic Changes in Chromatin
1196 States, Mol Cell Biol, 37 (2017).
1197 [34] R. Zhang, N.F. Lahens, H.I. Ballance, M.E. Hughes, J.B. Hogenesch, A circadian gene expression atlas in
1198 mammals: implications for biology and medicine, Proc Natl Acad Sci U S A, 111 (2014) 16219-16224.
1199 [35] T.-Y. Chang, D.J. Waxman, HDI-STARR-seq: Condition-specific enhancer discovery in mouse liver in vivo,
1200 submitted for publication, (2024).
1201 [36] N. Habib, I. Avraham-Davidi, A. Basu, T. Burks, K. Shekhar, M. Hofree, S.R. Choudhury, F. Aguet, E.
1202 Gelfand, K. Ardlie, D.A. Weitz, O. Rozenblatt-Rosen, F. Zhang, A. Regev, Massively parallel single-nucleus RNA-
1203 seq with DroNc-seq, Nat Methods, 14 (2017) 955-958.
1204 [37] P. Hu, E. Fabyanic, D.Y. Kwon, S. Tang, Z. Zhou, H. Wu, Dissecting Cell-Type Composition and Activity-
1205 Dependent Transcriptional State in Mammalian Brains by Massively Parallel Single-Nucleus RNA-Seq, Mol Cell,
1206 68 (2017) 1006-1015 e1007.
1207 [38] B. Lacar, S.B. Linker, B.N. Jaeger, S.R. Krishnaswami, J.J. Barron, M.J.E. Kelder, S.L. Parylak, A.C.M. Paquola,
1208 P. Venepally, M. Novotny, C. O'Connor, C. Fitzpatrick, J.A. Erwin, J.Y. Hsu, D. Husband, M.J. McConnell, R.
1209 Lasken, F.H. Gage, Nuclear RNA-seq of single neurons reveals molecular signatures of activation, Nat Commun,
1210 7 (2016) 11022.
1211 [39] T. Stuart, A. Srivastava, S. Madad, C.A. Lareau, R. Satija, Single-cell chromatin state analysis with Signac,
1212 Nat Methods, 18 (2021) 1333-1341.
1213 [40] T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W.M. Mauck, 3rd, Y. Hao, M. Stoeckius, P.
1214 Smibert, R. Satija, Comprehensive Integration of Single-Cell Data, Cell, 177 (2019) 1888-1902 e1821.
1215 [41] C.N. Goldfarb, K. Karri, M. Pyatkov, D.J. Waxman, Interplay Between GH-regulated, Sex-biased Liver
1216 Transcriptome and Hepatic Zonation Revealed by Single-Nucleus RNA Sequencing, Endocrinology, 163 (2022).
1217 [42] K. Karri, D.J. Waxman, Dysregulation of murine long noncoding single-cell transcriptome in nonalcoholic
1218 steatohepatitis and liver fibrosis, RNA, 29 (2023) 977-1006.
1219 [43] C. Neumayr, M. Pagani, A. Stark, C.D. Arnold, STARR-seq and UMI-STARR-seq: Assessing Enhancer
1220 Activities for Genome-Wide-, High-, and Low-Complexity Candidate Libraries, Curr Protoc Mol Biol, 128 (2019)
1221 e105.
1222 [44] B. Langmead, C. Trapnell, M. Pop, S.L. Salzberg, Ultrafast and memory-efficient alignment of short DNA
1223 sequences to the human genome, Genome Biol, 10 (2009) R25.
1224 [45] E.C. Mouse, J.A. Stamatoyannopoulos, M. Snyder, R. Hardison, B. Ren, T. Gingeras, D.M. Gilbert, M.
1225 Groudine, M. Bender, R. Kaul, T. Canfield, E. Giste, A. Johnson, M. Zhang, G. Balasundaram, R. Byron, V. Roach,
1226 P.J. Sabo, R. Sandstrom, A.S. Stehling, R.E. Thurman, S.M. Weissman, P. Cayting, M. Hariharan, J. Lian, Y.
1227 Cheng, S.G. Landt, Z. Ma, B.J. Wold, J. Dekker, G.E. Crawford, C.A. Keller, W. Wu, C. Morrissey, S.A. Kumar, T.
1228 Mishra, D. Jain, M. Byrska-Bishop, D. Blankenberg, B.R. Lajoie, G. Jain, A. Sanyal, K.B. Chen, O. Denas, J. Taylor,
1229 G.A. Blobel, M.J. Weiss, M. Pimkin, W. Deng, G.K. Marinov, B.A. Williams, K.I. Fisher-Aylor, G. Desalvo, A.
1230 Kiralusha, D. Trout, H. Amrhein, A. Mortazavi, L. Edsall, D. McCleary, S. Kuan, Y. Shen, F. Yue, Z. Ye, C.A. Davis,
1231 C. Zaleski, S. Jha, C. Xue, A. Dobin, W. Lin, M. Fastuca, H. Wang, R. Guigo, S. Djebali, J. Lagarde, T. Ryba, T.
1232 Sasaki, V.S. Malladi, M.S. Cline, V.M. Kirkup, K. Learned, K.R. Rosenbloom, W.J. Kent, E.A. Feingold, P.J. Good,
1233 M. Pazin, R.F. Lowdon, L.B. Adams, An encyclopedia of mouse DNA elements (Mouse ENCODE), Genome Biol,
1234 13 (2012) 418.
1235 [46] Y. Benjamini, T.P. Speed, Summarizing and correcting the GC content bias in high-throughput sequencing,
1236 Nucleic Acids Res, 40 (2012) e72.
1237 [47] G.N. Wilkinson, C.E. Rogers, Symbolic Description of Factorial Models for Analysis of Variance, The Royal
1238 Statistical Society Series C-Applied Statistics, 22 (1973) 392-399.
1239 [48] A. Sugathan, D.J. Waxman, Genome-wide analysis of chromatin states reveals distinct mechanisms of sex-
1240 dependent gene regulation in male and female mouse liver, Mol Cell Biol, 33 (2013) 3594-3610.

1241 [49] J. Ernst, M. Kellis, Chromatin-state discovery and genome annotation with ChromHMM, Nat Protoc, 12
1242 (2017) 2478-2492.
1243 [50] P. Machanick, T.L. Bailey, MEME-ChIP: motif analysis of large DNA datasets, Bioinformatics, 27 (2011)
1244 1696-1697.
1245 [51] D.E. Newburger, M.L. Bulyk, UniPROBE: an online database of protein binding microarray data on protein-
1246 DNA interactions, Nucleic Acids Res, 37 (2009) D77-82.
1247 [52] I. Rauluseviciute, R. Riudavets-Puig, R. Blanc-Mathieu, J.A. Castro-Mondragon, K. Ferenc, V. Kumar, R.B.
1248 Lemma, J. Lucas, J. Cheneby, D. Baranasic, A. Khan, O. Fornes, S. Gundersen, M. Johansen, E. Hovig, B.
1249 Lenhard, A. Sandelin, W.W. Wasserman, F. Parcy, A. Mathelier, JASPAR 2024: 20th anniversary of the open-
1250 access database of transcription factor binding profiles, Nucleic Acids Res, 52 (2024) D174-D182.
1251 [53] I.E. Vorontsov, I.A. Eliseeva, A. Zinkevich, M. Nikonov, S. Abramov, A. Boytsov, V. Kamenets, A. Kasianova,
1252 S. Kolmykov, I.S. Yevshin, A. Favorov, Y.A. Medvedeva, A. Jolma, F. Kolpakov, V.J. Makeev, I.V. Kulakovskiy,
1253 HOCOMOCO in 2024: a rebuild of the curated collection of binding models for human and mouse transcription
1254 factors, Nucleic Acids Res, 52 (2024) D154-D163.
1255 [54] B.J. Matthews, D.J. Waxman, Computational prediction of CTCF/cohesin-based intra-TAD loops that
1256 insulate chromatin contacts and gene expression in mouse liver, Elife, 7 (2018).
1257 [55] D. Lau-Corona, H. Ma, C. Vergato, A. Sarmento-Cabral, M. Del Rio-Moreno, R.D. Kineman, D.J. Waxman,
1258 Constitutively Active STAT5b Feminizes Mouse Liver Gene Expression, Endocrinology, 163 (2022).
1259 [56] R.D. Kineman, M. Del Rio-Moreno, D.J. Waxman, Liver-specific actions of GH and IGF1 that protect against
1260 MASLD, Nat Rev Endocrinol, 21 (2025) 105-117.
1261 [57] T.L. Conforto, Y. Zhang, J. Sherman, D.J. Waxman, Impact of CUX2 on the female mouse liver
1262 transcriptome: activation of female-biased genes and repression of male-biased genes, Mol Cell Biol, 32 (2012)
1263 4611-4627.
1264 [58] E.V. Laz, M.G. Holloway, C.S. Chen, D.J. Waxman, Characterization of three growth hormone-responsive
1265 transcription factors preferentially expressed in adult female liver, Endocrinology, 148 (2007) 3327-3337.
1266 [59] T.L. Conforto, G.F.t. Steinhardt, D.J. Waxman, Cross Talk Between GH-Regulated Transcription Factors
1267 HNF6 and CUX2 in Adult Mouse Liver, Mol Endocrinol, 29 (2015) 1286-1302.
1268 [60] V.G. Budker, V.M. Subbotin, T. Budker, M.G. Sebestyen, G. Zhang, J.A. Wolff, Mechanism of plasmid
1269 delivery by hydrodynamic tail vein injection. II. Morphological studies, J Gene Med, 8 (2006) 874-888.
1270 [61] J.E. Davis, K.D. Insigne, E.M. Jones, Q.A. Hastings, W.C. Boldridge, S. Kosuri, Dissection of c-AMP Response
1271 Element Architecture by Using Genomic and Episomal Massively Parallel Reporter Assays, Cell Syst, 11 (2020)
1272 75-85 e77.
1273 [62] N. Doni Jayavelu, A. Jajodia, A. Mishra, R.D. Hawkins, Candidate silencer elements for the human and
1274 mouse genomes, Nat Commun, 11 (2020) 1061.
1275 [63] Y. Zheng, N.J. VanDusen, Massively Parallel Reporter Assays for High-Throughput In Vivo Analysis of Cis-
1276 Regulatory Elements, J Cardiovasc Dev Dis, 10 (2023).
1277 [64] J.C. Klein, V. Agarwal, F. Inoue, A. Keith, B. Martin, M. Kircher, N. Ahituv, J. Shendure, A systematic
1278 evaluation of the design and context dependencies of massively parallel reporter assays, Nat Methods, 17
1279 (2020) 1083-1091.
1280 [65] M. Das, A. Hossain, D. Banerjee, C.A. Praul, S. Girirajan, Challenges and considerations for reproducibility
1281 of STARR-seq assays, Genome Res, 33 (2023) 479-495.
1282 [66] B.P. Kuiper, R.C. Prins, S. Billerbeck, Oligo Pools as an Affordable Source of Synthetic DNA for Cost-
1283 Effective Library Construction in Protein- and Metabolic Pathway Engineering, Chembiochem, 23 (2022)
1284 e202100507.
1285 [67] B.N. Akerberg, F. Gu, N.J. VanDusen, X. Zhang, R. Dong, K. Li, B. Zhang, B. Zhou, I. Sethi, Q. Ma, L. Wasson,
1286 T. Wen, J. Liu, K. Dong, F.L. Conlon, J. Zhou, G.C. Yuan, P. Zhou, W.T. Pu, A reference map of murine cardiac
1287 transcription factor chromatin occupancy identifies dynamic and conserved enhancers, Nat Commun, 10
1288 (2019) 4907.
1289 [68] R.R. Voskuhl, A.H. Sawalha, Y. Itoh, Sex chromosome contributions to sex differences in multiple sclerosis
1290 susceptibility and progression, Mult Scler, 24 (2018) 22-31.
1291

Fig. 1

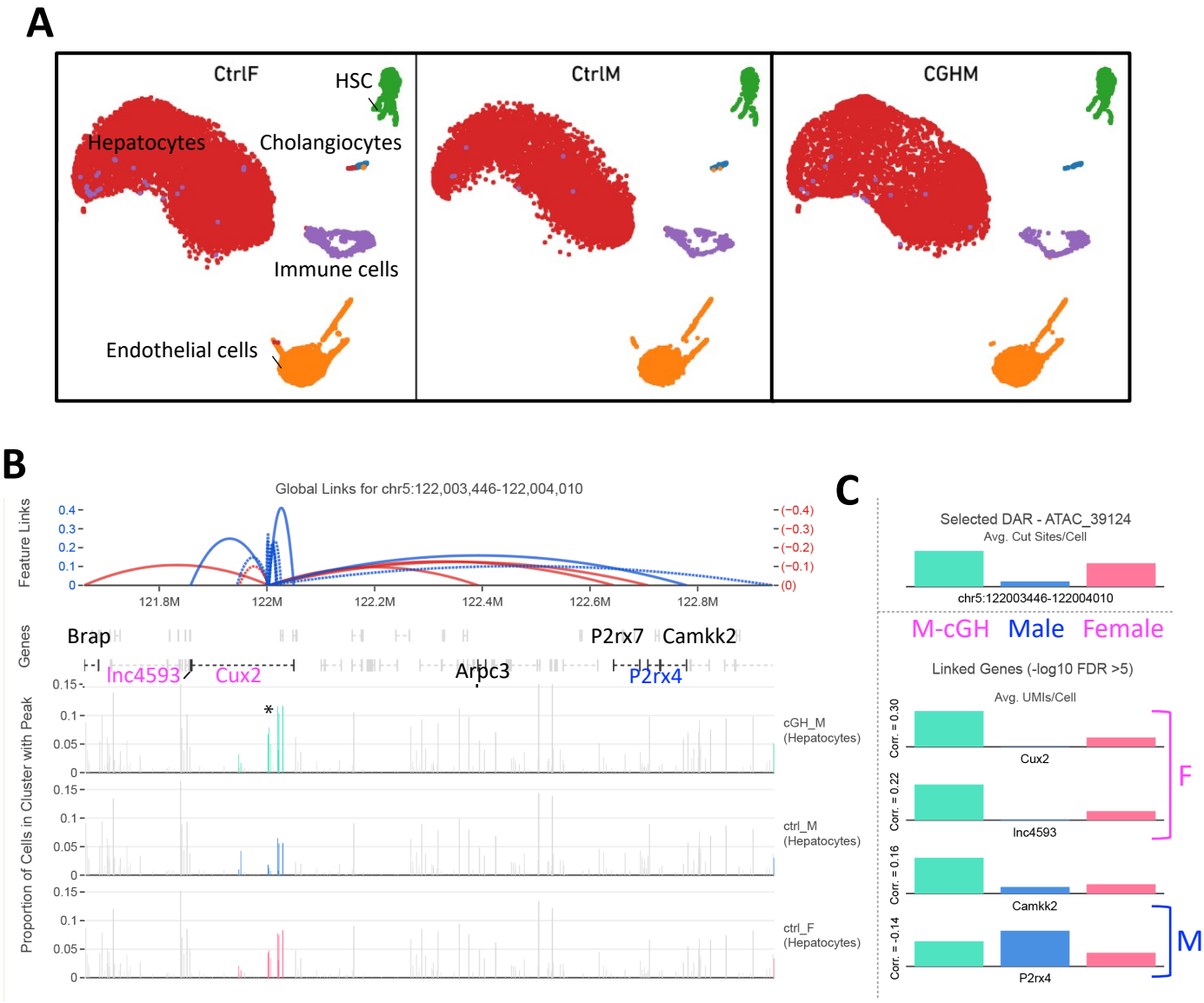

A
CtrlF
HSC
Hepatocytes
Cholangiocytes
Immune cells
Endothelial cells
CtrlM
CGHM
B
Global Links for chr5:122,003,446-122,004,010
Feature Links
Genes
Brap
P2rx7
Camkk2
lnc4593
Cux2
Arpc3
P2rx4
Proportion of Cells in Cluster with Peak
cGH_M (Hepatocytes)
ctrl_M (Hepatocytes)
ctrl_F (Hepatocytes)
121.8M
122M
122.2M
122.4M
122.6M
122.8M
C
Selected DAR - ATAC_39124
Avg. Cut Sites/Cell
chr5:122003446-122004010
M-cGH
Male
Female
Linked Genes (-log10 FDR >5)
Avg. UMIs/Cell
Corr. = 0.30
Cux2
Corr. = 0.22
lnc4593
Corr. = 0.16
Camkk2
Corr. = -0.14
P2rx4
F
M

Fig 2.

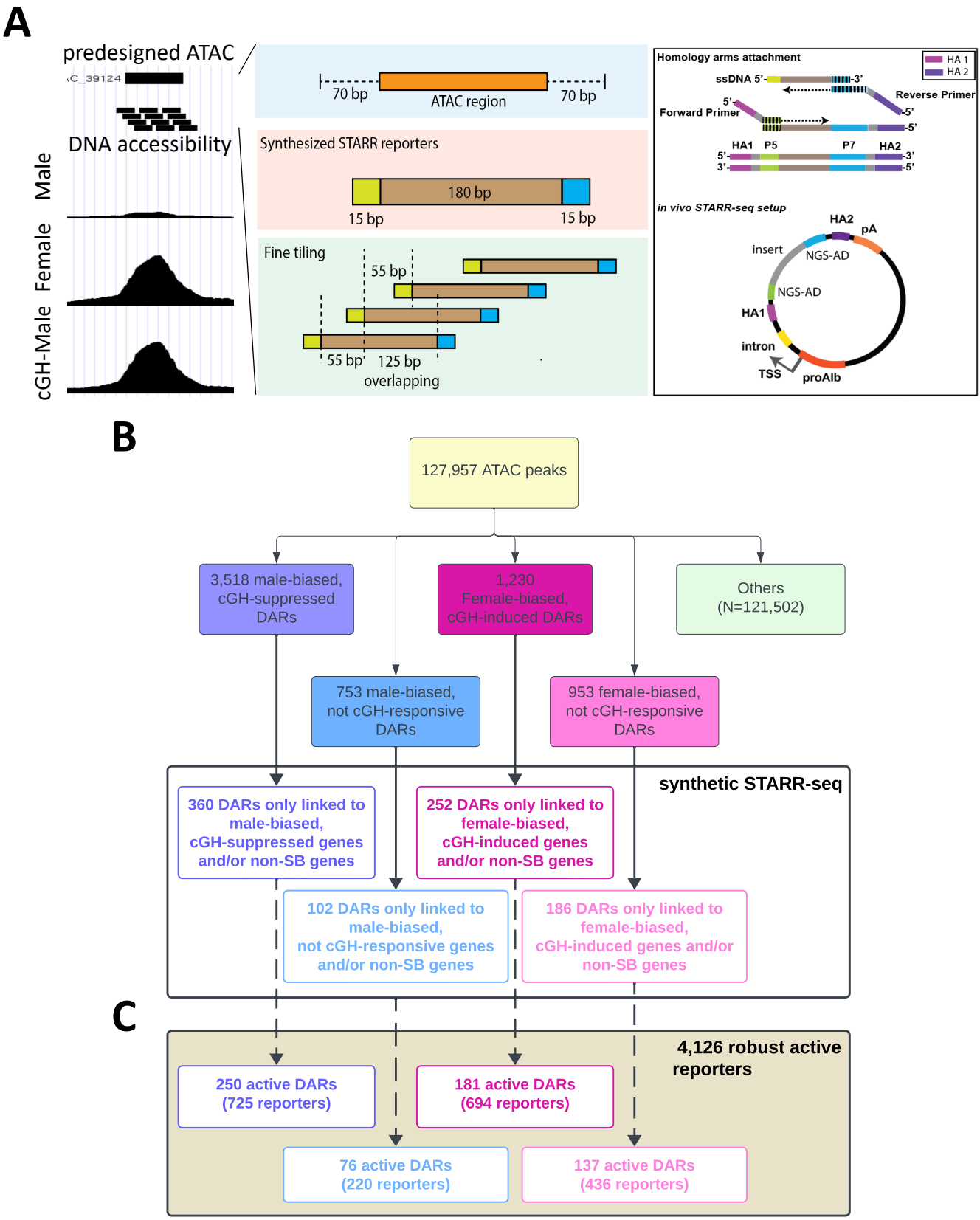

A
preddesigned ATAC
DNA accessibility
Male
Female
cGH-Male
70 bp
ATAC region
70 bp
Synthesized STARR reporters
180 bp
15 bp
15 bp
Fine tiling
55 bp
55 bp
125 bp
overlapping
Homology arms attachment
HA 1
HA 2
ssDNA 5'
3'
Reverse Primer
Forward Primer
5'
-5'
HA1
P5
P7
HA2
in vivo STARR-seq setup
HA2
pA
insert
NGS-AD
NGS-AD
HA1
intron
TSS
proAlb
B
127,957 ATAC peaks
3,518 male-biased, cGH-suppressed DARs
1,230 Female-biased, cGH-induced DARs
Others (N=121,502)
753 male-biased, not cGH-responsive DARs
953 female-biased, not cGH-responsive DARs
synthetic STARR-seq
360 DARs only linked to male-biased, cGH-suppressed genes and/or non-SB genes
252 DARs only linked to female-biased, cGH-induced genes and/or non-SB genes
102 DARs only linked to male-biased, not cGH-responsive genes and/or non-SB genes
186 DARs only linked to female-biased, cGH-induced genes and/or non-SB genes
C
4,126 robust active reporters
250 active DARs (725 reporters)
181 active DARs (694 reporters)
76 active DARs (220 reporters)
137 active DARs (436 reporters)

Fig. 3

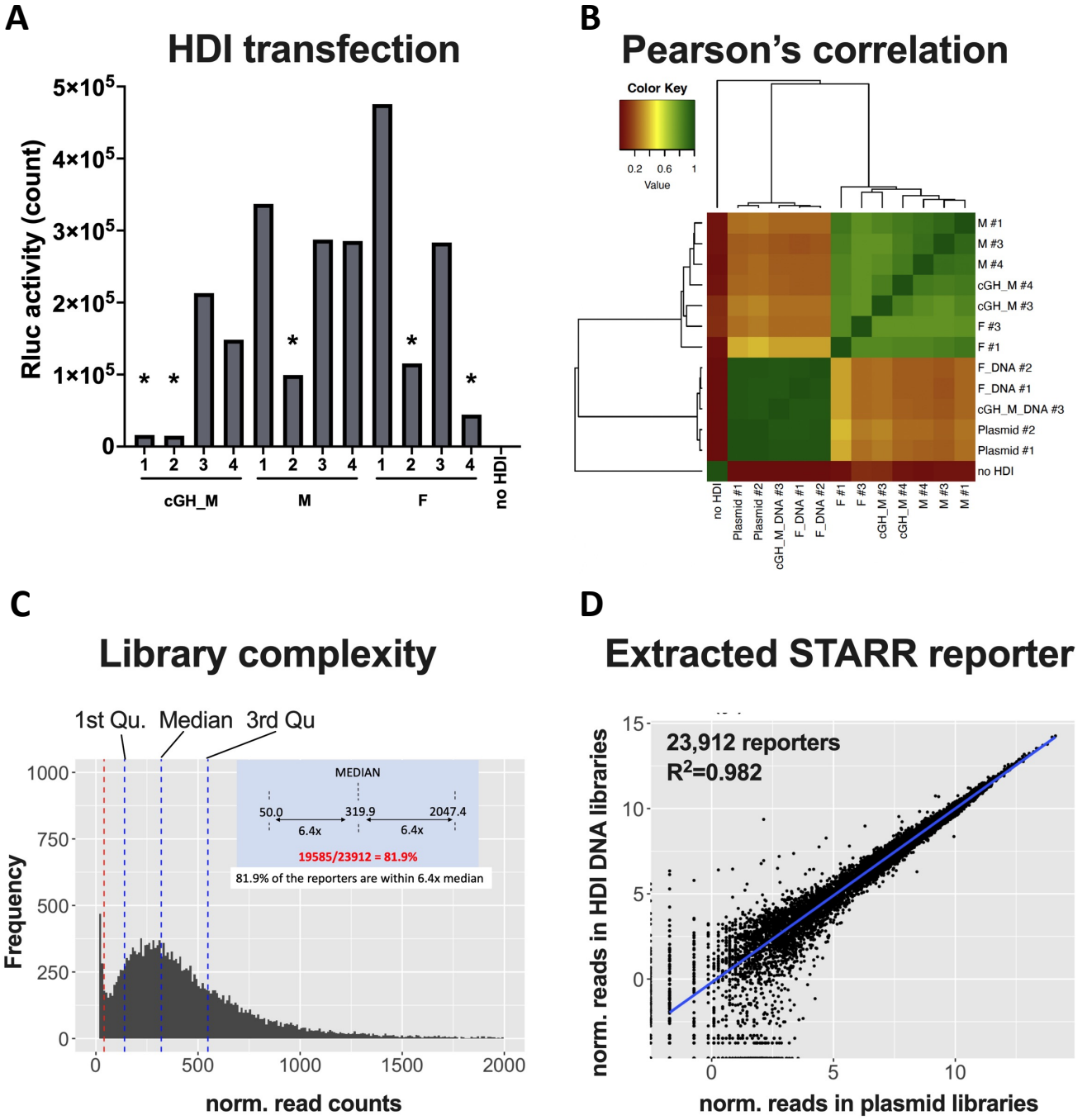

A
HDI transfection
Rluc activity (count)
5×10^5
4×10^5
3×10^5
2×10^5
1×10^5
0
1 2 3 4 1 2 3 4 1 2 3 4
cGH_M
M
F
no HDI
B
Pearson's correlation
Color Key
0.2 0.6 1
Value
M #1
M #3
M #4
cGH_M #4
cGH_M #3
F #3
F #1
F_DNA #2
F_DNA #1
cGH_M_DNA #3
Plasmid #2
Plasmid #1
no HDI
C
Library complexity
1st Qu. Median 3rd Qu
MEDIAN
50.0
319.9
2047.4
6.4x
6.4x
19585/23912 = 81.9%
81.9% of the reporters are within 6.4x median
Frequency
norm. read counts
D
Extracted STARR reporter
23,912 reporters
R2=0.982
norm. reads in HDI DNA libraries
norm. reads in plasmid libraries

Fig. 4

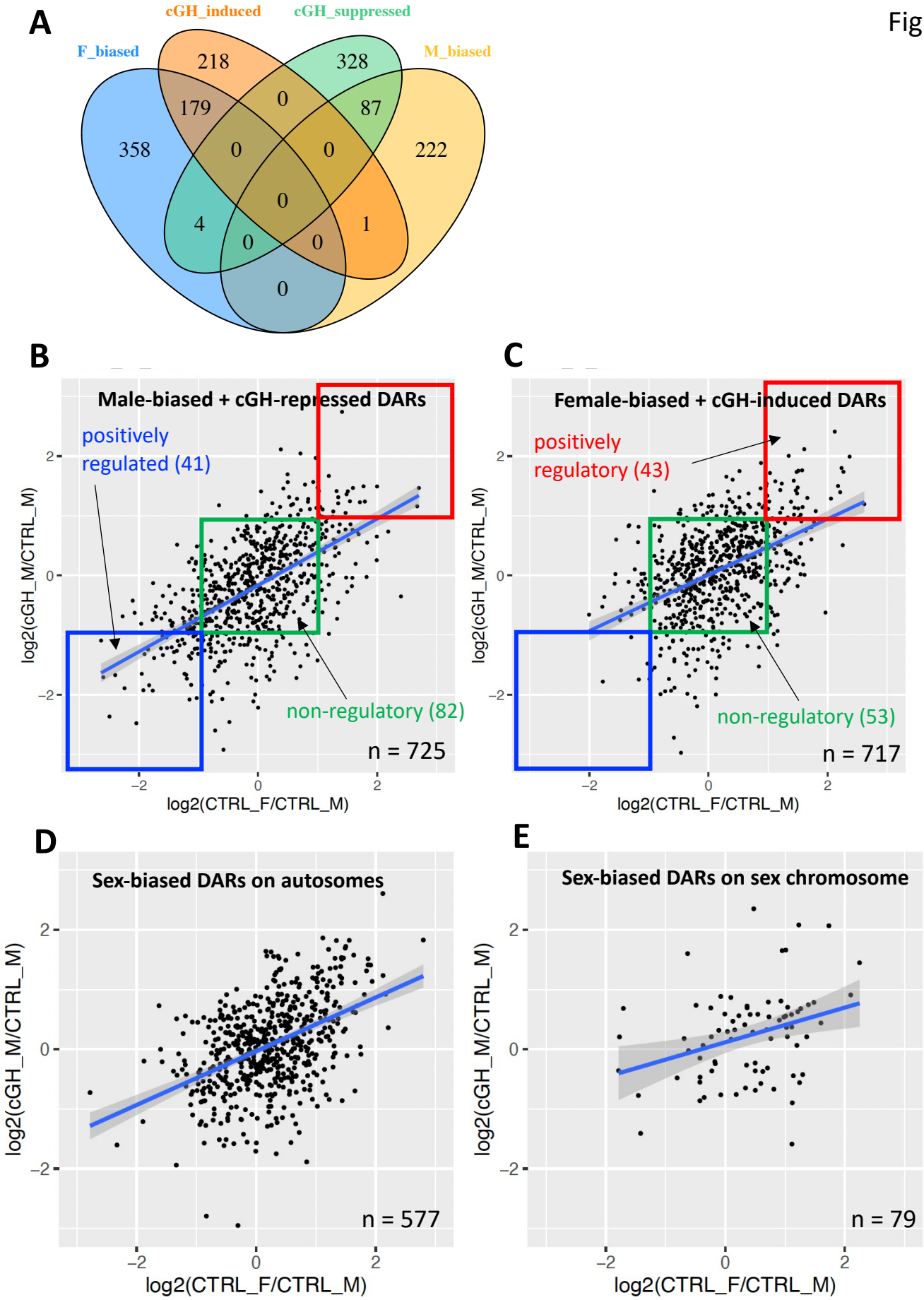
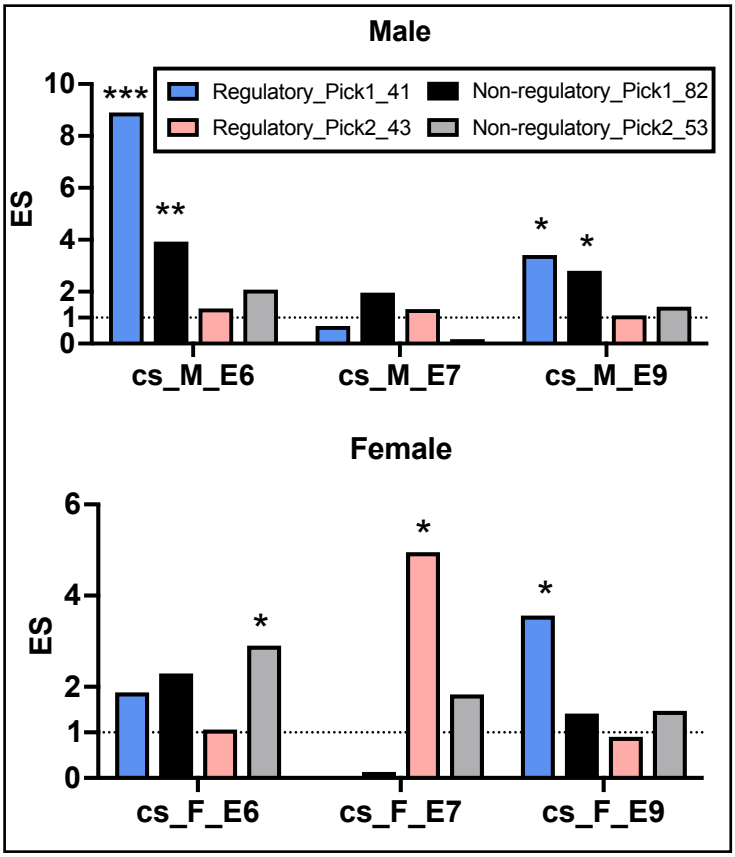
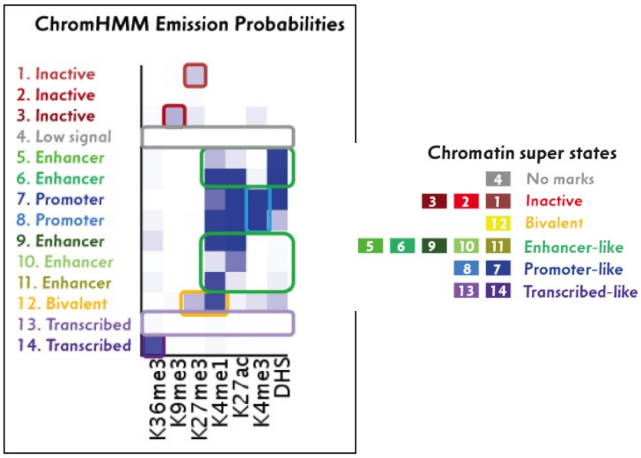

A
cGH_induced
cGH_suppressed
F_biased
M_biased
218
328
179
0
87
358
0
0
222
0
4
1
0
0
0
B
Male-biased + cGH-repressed DARs
positively regulated (41)
non-regulatory (82)
n = 725
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)
C
Female-biased + cGH-induced DARs
positively regulatory (43)
non-regulatory (53)
n = 717
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)
D
Sex-biased DARs on autosomes
n = 577
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)
E
Sex-biased DARs on sex chromosome
n = 79
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)

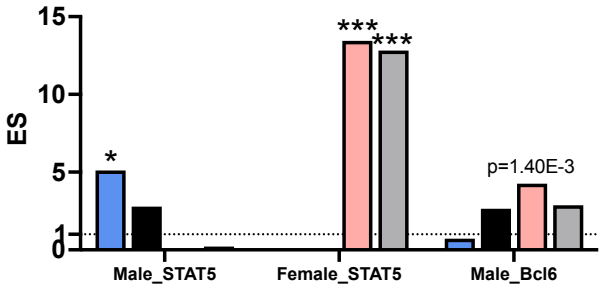
Fig. 5ABC

A

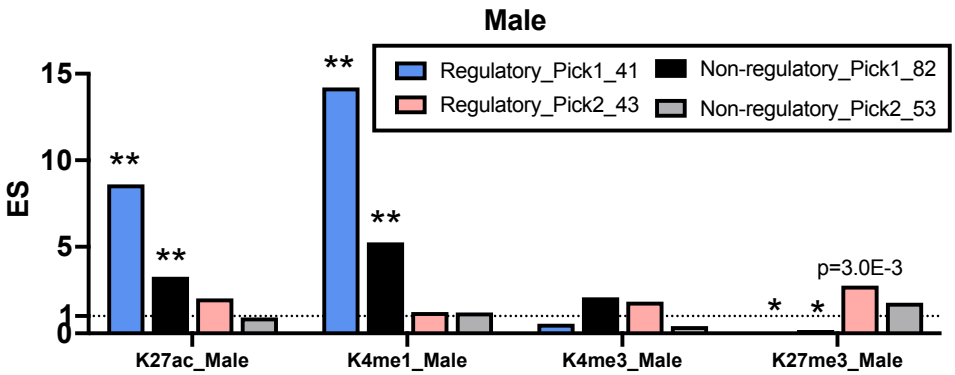

Male
Regulatory_Pick1_41
Non-regulatory_Pick1_82
Regulatory_Pick2_43
Non-regulatory_Pick2_53
ES
cs_M_E6
cs_M_E7
cs_M_E9
Female
cs_F_E6
cs_F_E7
cs_F_E9


ChromHMM Emission Probabilities
1. Inactive
2. Inactive
3. Inactive
4. Low signal
5. Enhancer
6. Enhancer
7. Promoter
8. Promoter
9. Enhancer
10. Enhancer
11. Enhancer
12. Bivalent
13. Transcribed
14. Transcribed
K36me3
K9me3
K27me3
K4me1
K27ac
K4me3
DHS
Chromatin super states
No marks
Inactive
Bivalent
Enhancer-like
Promoter-like
Transcribed-like

C


ES
p=1.40E-3
Male_STAT5
Female_STAT5
Male_Bcl6

B


Male
Regulatory_Pick1_41
Non-regulatory_Pick1_82
Regulatory_Pick2_43
Non-regulatory_Pick2_53
ES
p=3.0E-3
K27ac_Male
K4me1_Male
K4me3_Male
K27me3_Male

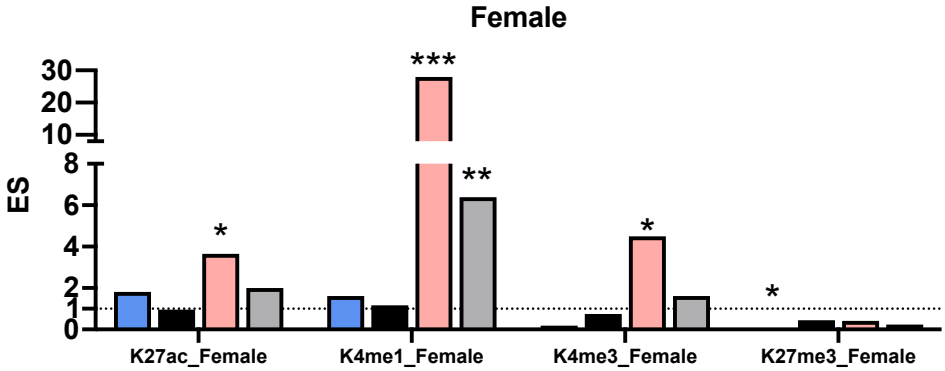

Female
ES
K27ac_Female
K4me1_Female
K4me3_Female
K27me3_Female

Fig. 5D

ES = 2.8
p=0.0237

ES = 3.76
p=0.0033

# of ATAC regions (Table S9A, columns U-AG)	Summary (Pick1 thru Pick6)	STAT5/MA0519.1_full_des	BCL6_MOUSE.H11MO.0.A_full_des	CUX2_MOUSE.H11MO.0.C_full_des	HNF6_MOUSE.H11MO.0.A_full_des
41	regulated_Pick1	20%	29%	37%	41%
82	nonregulated_Pick1	17%	26%	17%	16%
43	regulated_Pick2	35%	40%	12%	7%
53	nonregulated_Pick2	34%	42%	11%	9%
19	regulated_Pick3_M	16%	21%	42%	53%
27	nonregulated_Pick3_M	22%	26%	15%	4%
58	regulated_Pick3_F	16%	36%	2%	3%
55	nonregulated_Pick3_F	13%	9%	7%	4%
6	regulated_Pick4	17%	33%	50%	33%
19	nonregulated_Pick4	20%	35%	20%	20%

ES = 7.3
p=0.0007

ES = 4.18
p=0.0487

ES = 28.8
p=0.0002

Fig. 6

A

foreground: 41 **regulated** male-biased and cGH-suppressed DARs
background: 82 **non-regulated** male-biased and cGH-suppressed DARs

	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs	Distribution
#1	Reverse Complement	STREME	9.4e-008	Esrra_primary (UP00079_1) Nr2f2_primary (UP00009_1) Mtf1_primary (UP00097_1)	Not Centrally Differentially Enriched
#2	Reverse Complement	STREME	4.2e-007	Hnf4a_secondary (UP00066_2) Isgf3g_primary (UP00074_1)	Not Centrally Differentially Enriched
#3	Reverse Complement	STREME	7.5e-006	Zfp187_secondary (UP00082_2)	Not Centrally Differentially Enriched
#4	Reverse Complement	STREME	6.2e-005		Not Centrally Differentially Enriched
#5	Reverse Complement	STREME	9.3e-005		Not Centrally Differentially Enriched

B

foreground:41 **regulated** male-biased and cGH-suppressed DARs
background: 43 **regulated** female-biased and cGH-induced DARs

	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs
#1	Reverse Complement	STREME	1.4e-008	Mafb_primary (UP00045_1) Sox8_secondary (UP00051_2)
#2	Reverse Complement Show 1 More	STREME	2.5e-008	Hnf4a_secondary (UP00066_2) Isl2_3430.1 (UP00170_1) Barhl2_3868.1 (UP00145_1)
#3	Reverse Complement	STREME	3.5e-005	
#4	Reverse Complement	STREME	7.0e-005	Bbx_primary (UP00012_1)
#5	Reverse Complement	STREME	4.1e-004	Hoxd8_2644.1 (UP00168_1) Zbtb3_secondary (UP00031_2) Hoxa7_3750.1 (UP00164_2)

Fig. 6

C

foreground: 41 **regulated** male-biased and cGH-suppressed DARs
background: 317 **non-regulated** ATAC/DARs

	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs
#1	Reverse Complement	STREME	1.5e-009	
#2	Reverse Complement	STREME	7.4e-009	Hnf4a_secondary (UP00066_2) Isl2_3430.1 (UP00170_1)
#3	Reverse Complement	STREME	1.2e-008	
#4	Reverse Complement	STREME	1.3e-007	Lhx8_2247.2 (UP00184_1)
#5	Reverse Complement	STREME	1.6e-007	

D

foreground: 41 **regulated** male-biased and cGH-suppressed DARs
background: 65 **stringently inactive ATAC regions**

	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs
#1	Reverse Complement	STREME	1.5e-008	Esrra_primary (UP00079_1) Nr2f2_primary (UP00009_1) Hoxc9_2367.2 (UP00197_1)
#2	Reverse Complement Show 1 More	STREME	1.1e-006	Sox15_primary (UP00075_1) Sox17_primary (UP00014_1) Sox18_primary (UP00064_1)
#3	Reverse Complement	STREME	7.1e-006	
#4	Reverse Complement	STREME	3.5e-005	Zscan4_secondary (UP00026_2) Foxl1_secondary (UP00061_2)

Fig. 7

A

foreground: 43 **regulated** female-biased and cGH-induced DARs
background: 53 non-**regulated** female-biased and cGH-induced DARs

	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs
#1	Reverse Complement	STREME	1.8e-007	Sox7_secondary (UP00034_2)
#2	Reverse Complement	STREME	2.0e-007	Crx_3485.1 (UP00176_1) Zbtb7b_primary (UP00047_1)
#3	Reverse Complement	STREME	1.2e-006	
#4	Reverse Complement	STREME	5.1e-006	Zbtb3_secondary (UP00031_2)
#5	Reverse Complement	STREME	7.2e-005	

B

foreground:43 **regulated** female-biased and cGH-induced DARs
background: 41 **regulated** male-biased and cGH-suppressed DARs

	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs
#1	Reverse Complement	STREME	3.8e-009	Nkx2-5_3436.1 (UP00249_1)
#2	Reverse Complement	STREME	8.9e-008	
#3	Reverse Complement	STREME	5.3e-006	Tcfap2e_primary (UP00028_1) Tcfap2c_primary (UP00087_1)
#4	Reverse Complement	STREME	1.0e-005	Rfxdc2_secondary (UP00076_2)
#5	Reverse Complement	STREME	1.0e-004	

Fig. 7

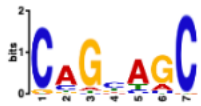
C

foreground: 43 **regulated** female-biased and cGH-induced DARs
background: 317 **non-regulated** ATAC/DARs

	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs
#1	Reverse Complement ⇆ Show 2 More ↧	STREME	1.7e-009	
#2	Reverse Complement ⇆	STREME	2.2e-007	Ascl2_primary (UP00099_1) Pknox2_3077.2 (UP00205_1) Tgif1_2342.2 (UP00122_1)
#3	Reverse Complement ⇆	STREME	1.6e-006	Nkx2-4_3074.1 (UP00107_1) Nkx2-5_3436.1 (UP00249_1) Nkx2-6_3437.1 (UP00147_1)
#4	Reverse Complement ⇆	STREME	5.2e-006	

D

foreground: 43 **regulated** female-biased and cGH-induced DARs
background: 65 str. inactive ATAC regions

	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs
#1	Reverse Complement ⇆ Show 1 More ↧	STREME	9.3e-012	Sp4_secondary (UP00002_2) Klf7_primary (UP00093_1) Smad3_secondary (UP00000_2)
#2	Reverse Complement ⇆	STREME	4.3e-008	
#3	Reverse Complement ⇆	STREME	1.8e-006	Tcfe2a_secondary (UP00046_2)
#4	Reverse Complement ⇆	STREME	6.7e-006	Rara_primary (UP00048_1) Nr2f2_secondary (UP00009_2)
#5	 Reverse Complement ⇆	STREME	7.7e-006	Ascl2_primary (UP00099_1) Osr1_primary (UP00027_1) Osr2_primary (UP00052_1)