HDI-STARR-seq Identifies Functional GH-regulated Sex-Biased Hepatocyte 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

Author Disclosure summary:
The authors declare no financial interests

Grant support: NIH DK121998 (to DJW)

Keywords:
Liver sex differences; hepatocyte chromatin accessibility; HDI-STARR-seq; MASLD; epigenomics

*Corresponding author:
Dr. David J Waxman
Dept. of Biology, Boston University
Boston MA 02215 USA
Email: djw@bu.edu
ORCID: **https://orcid.org/0000-0001-7982-9206**

Abstract

Growth hormone (GH) controls sexual dimorphism in hepatocyte gene expression programs governing lipid metabolism, bile acid synthesis and xenobiotic processing, which contribute to sex differences in metabolic dysfunction-associated steatotic liver disease (MASLD) risk. Despite extensive study of GH-regulated sex differences in gene transcription, the functional *cis*-regulatory hepatocyte enhancers that orchestrate these sex-dependent metabolic programs remain largely unknown. Here, we integrated single-nucleus multiomic profiling of hepatocyte chromatin accessibility with *in vivo* functional enhancer assays to identify and validate GH-responsive, sex-biased hepatocyte enhancers in intact mouse liver. We constructed a tiled HDI-STARR-seq library of 23,912 reporters spanning 1,839 liver ATAC regions and delivered it to liver by hydrodynamic injection, enabling functional assessment of enhancer activity *in vivo* across distinct biological conditions. Reporters representing 840 ATAC regions showed sex-biased and/or GH-regulated enhancer activity, in many cases mirroring the regulation of their chromatin accessibility in hepatocytes, validating these sites as functional, physiologically regulated enhancers. The regulated enhancers were enriched for activating histone marks (H3K4me1, H3K27ac), for binding sites for the GH-activated transcriptional regulator STAT5, and for the STAT5-dependent, sex-specific repressors BCL6 and CUX2. Further, *de novo* motif analysis identified binding sites for HNF4A and for several novel factors specifically enriched at the regulated enhancers. Sex-biased and GH-regulated functional enhancers were linked to both MASLD-enabling and MASLD-protective genes, suggesting that GH-dependent chromatin remodeling at these loci contributes to sex-differential metabolic disease susceptibility. This integrated *in vivo* approach defines a validated set of GH-regulated hepatocyte enhancers through which chromatin accessibility and transcription factor binding drive sexual dimorphism in hepatic metabolism and sex-specific MASLD risk.

Introduction

Growth hormone (GH)–regulated sex differences in hepatic gene expression underlie sex differences in lipid and drug metabolism, xenobiotic clearance and susceptibility to/progression of metabolic liver disease, as seen in both humans and in mouse models (1). GH, acting via its sex steroid-determined, sex-differential temporal patterns of pituitary secretion (2,3), controls the sexually dimorphic expression of several hundred hepatocyte-expressed genes (4) and is a key endocrine regulator of sex-dependent liver function (5,6). GH activates intracellular signaling cascades that lead to distinct, sex-specific expression of cytochrome P450 (*Cyp*) and other metabolic genes, including genes active in lipid metabolism (7,8). Many of these downstream transcriptional outputs are mediated by the TF STAT5 (9-11). Upon binding to GH receptor at the hepatocyte cell surface, GH induces conformational changes leading to activation of JAK2, a GH receptor-linked tyrosine kinase that phosphorylates STAT5, promoting its dimerization (12). Activated STAT5 dimers translocate to the nucleus, where they bind TTCNNNGAA consensus motifs (13) and increase transcription of linked target genes (11). GH-STAT5 signaling plays a pivotal role in diverse physiological processes, including body growth (10,14), cell cycle regulation (15) and hepatic metabolism of lipids, bile acids, steroids and drugs (10,16,17). Dysregulation of GH-activated STAT5 signaling has been implicated in liver disease and its progression from steatosis (7,17,18) to fibrosis and hepatocellular carcinoma (8,19).

Two liver STAT5 isoforms, encoded by the genes *Stat5a* and *Stat5b*, mediate the actions of GH in regulating the transcription of sex-biased genes, with active liver STAT5b protein much more abundant than STAT5a (20). Hepatocyte-specific deletion of *Stat5a* and *Stat5b* in adult male liver (21,22) decreases the hepatic expression of many male-biased genes while de-repressing female-biased genes, consistent with earlier findings in global *Stat5b*-null male mice (9). STAT5b can enhance the transcriptional activity of HNF4A, particularly toward male-biased liver target genes (23,24), many of which are co-dependent on HNF4A for sex-biased expression (25). The transcriptional repressor BCL6, which shows male-biased hepatic expression, is regulated by GH and STAT5 (26) and has a DNA recognition motif that partially overlaps with the STAT5 consensus site, which enables BCL6 to inhibit a subset of liver STAT5 transcriptional responses (13,27). A second GH/STAT5-regulated transcriptional repressor, CUX2, shows female-specific hepatic expression (28) and enforces liver sex differences by repressing many male-biased genes and activating a subset of female-biased genes in female liver (29,30). Together, these findings define a GH-responsive transcriptional network in which STAT5 and HNF4A are major transcriptional activators whose actions are modulated by the sex-biased repressors BCL6 and CUX2 to shape sex-dependent liver gene expression programs.

DNA accessibility plays a crucial role in modulating sex-dependent gene expression by controlling chromatin access for STAT5, BCL6, CUX2 and other liver TFs. Genome-wide DNase-seq and chromatin-state mapping have identified a few thousand sex-biased mouse liver DNase-hypersensitive sites that strongly co-localize with enhancer marks and predict sex-biased gene expression, implicating differential chromatin accessibility as a primary determinant of sex-dependent transcriptional programs (31,32). Plasma GH pulsatility (33) and natural genetic variation (34) both impact the sex-bias of these accessible chromatin regions, dynamically turning accessibility on or off in specific genomic regions in a sex-dependent manner, thereby switching nearby sex-biased genes between male-like and female-like expression states. Further, 3D genome organization studies demonstrate that sex-biased distal enhancers physically contact target gene promoters *via* chromatin looping interactions; and loss of cohesin, which mediates long-range DNA looping, selectively suppresses expression of male-biased genes with distal sex-biased enhancers (35), firmly linking sex-biased chromatin accessibility to sex-biased liver gene expression output. Despite these advances, a critical gap remains: prior epigenomic studies infer regulatory function from correlations between accessibility, histone marks and gene expression, leaving key questions unresolved. First, which of the thousands of sex-biased accessible chromatin regions act as functional, *bona fide* enhancers *in vivo*? Second, does GH signaling and its sex-specific temporal secretion pattern directly modulate liver enhancer activity in the intact endogenous hormonal and metabolic context? Third, are these putative regulatory chromatin sequences linked to genes controlling disease-relevant metabolic pathways implicated in sex-differential MASLD susceptibility?

Liver function is highly dependent on its specialized sinusoidal architecture and cellular heterogeneity, encompassing multiple cell types with distinct phenotypes and functions, including hepatocytes, liver sinusoidal endothelial cells, immune cells and hepatic stellate cells (36-38). Single-cell and spatial transcriptomics studies have revealed pronounced zonation of hepatocyte and non-parenchymal cell gene expression across the liver lobule, as well as substantial intercellular heterogeneity in metabolic pathways and disease-relevant programs (39). Critically, recent single-nucleus (sn)RNA-sequencing analysis established that sex-biased liver gene expression is largely restricted to hepatocytes and is not a feature of liver nonparenchymal cells, with 93% of sex-biased genes showing expression specifically in periportal and pericentral hepatocyte populations (4). Many sex-biased transcripts exhibited sex-dependent zonation within the liver lobule, highlighting an intricate interconnection between liver sexual dimorphism and metabolic zonation at the single-cell level (4). However, most prior work on sex differences in liver gene expression and sex-biased chromatin accessibility has been performed in bulk liver, limiting the ability to distinguish hepatocyte-intrinsic effects from those mediated by liver non-parenchymal cells (NPCs).

To better understand how hepatocyte-specific, zone-resolved gene regulation contributes to liver sex-biased function and metabolic disease risk, and how GH influences these programs, cell-type-specific and liver lobule zone-resolved multiomic approaches were recently employed to jointly profile chromatin accessibility and gene expression using 10x Genomics Multiome ATAC + GEX technology (40). ATAC-seq peaks and RNA expression were simultaneously profiled from the same individual nuclei, enabling discovery of differentially accessible chromatin regions (DARs) as candidates for functional sex-biased and GH-responsive enhancers linked to target gene expression. Here, we build on those findings and present an integrated functional genomics strategy that combines single-nucleus Multiome technology with HDI-STARR-seq (41) to functionally assay and validate sex-biased, GH-responsive enhancers in intact mouse liver. Guided by hepatocyte-resolved chromatin accessibility and gene expression data, we selected ~1,000 sex-biased and GH-regulated hepatocyte differentially accessible regions (DARs), many linked to sex-biased genes involved in lipid homeostasis and implicated in GH action and its sex-dependent effects on metabolic dysfunction-associated steatotic liver disease (MASLD) (1), together with several hundred non-regulated control ATAC-seq regions. These 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, under endogenous hormonal and metabolic environments that are absent from conventional cell culture–based reporter assays. This experimental design enabled us to identify physiologically relevant enhancers whose sex-biased, functional enhancer activity mirrors the sex-bias and hormonal regulation of accessibility in hepatocyte chromatin, including enhancers linked to MASLD-enabling and MASLD-protective sex-biased genes, and to discover epigenetic signatures and TF motifs enriched at DARs showing sex-biased, GH-responsive enhancer activity. Together, our findings move beyond correlative epigenomic maps to define a validated set of hepatocyte enhancers through which GH-regulated chromatin accessibility drives sex-dependent function linked to sexual dimorphism in liver metabolism and sex-differential risk of metabolic liver disease.

Methods

Animal procedures. All animal experimentation was conducted in accord with accepted standards of humane animal care and in compliance with procedures approved by the Boston University Institutional Animal Care and Use Committee (protocol # PROTO201800698) and ARRIVE 2.0 Essential 10 guidelines (42), 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). Reporter plasmids and STARR-seq plasmid libraries were transfected into livers of male and female mice at 7-9 weeks of age by hydrodynamic injection (HDI) on day 0, as described (41). For the cGH-treated male mouse group, mice were implanted on day 1 with an Alzet osmotic minipump set to deliver recombinant rat GH at 25 ng/g of BW/h. For the male vehicle control group, the minipump was loaded with vehicle (30 mM $NaHCO_3$) and

154 implanted on day 1. Ten days later, on day 11 after HDI injection, livers were harvested and stored at -80°C.
155 Recombinanat rat GH depleted of bacterial lipopolysaccharide (≤ 7.7 to 15.5 EU/mg) (source: Dr. Arieh Gertler,
156 Protein Laboratories Rehovot, Ltd., Rehovot, Israel) was dissolved in 30 mM $NaHCO_3$ and 0.15 M NaCl containing
157 0.1 mg/ml rat albumin (Sigma-Aldrich, catalog no. A6414) (43). Mice were euthanized under CO_2 anesthesia by
158 cervical dislocation at 8-10 wk of age and livers were collected at a consistent time of day (between 10:30 AM
159 and 11:30 AM) to minimize the impact of circadian rhythms on gene expression (44). Livers were excised and
160 flash frozen in liquid N_2 then stored at -80°C.
161
162 **In vivo transfection by HDI and validation by luciferase assay.** Naked plasmid DNA was delivered to mouse liver
163 by HDI via the tail vein, as described (41). Briefly, 30 μg of pooled STARR-seq plasmid was mixed with 3 μg of
164 plasmid pGYL18 encoding Renilla luciferase and delivered by rapid tail vein injection (HDI) into male and female
165 mice on day 0. The effectiveness of HDI and of plasmid delivery to the liver was verified for each mouse by Renilla
166 luciferase assay, as described (41). Liver extracts were prepared at the same time for each mouse group to
167 minimize batch effects and assayed in technical triplicates. The absolute Renilla activity of each liver was used to
168 assess the success of HDI delivery.
169
170 **Single-cell transcriptional and chromatin accessibility profiling of mouse liver.** Single cell-based mouse liver
171 gene expression and chromatin accessibility profiles were determined for each biological group using 10X
172 Genomics MultiOme technology, which implements both snRNA-seq and snATAC-seq analysis in the same
173 individual nuclei. The 10X Genomics MultiOmic datasets were analyzed as described (40) using Cell Ranger ARC
174 (v.2.0.0) for initial processing, followed by analysis using Signac (v1.1.0) (45) and Seurat (4.1.0) (46). A total of
175 127,957 ATAC-seq peaks were identified across all samples and biological groups using the MACS2 peak caller
176 within the Signac package (implemented using the CallPeaks function), which gives ATAC-seq sequence read
177 density data for each biological condition and in each cell type. The CreateChromatinAssay function within Signac
178 was applied to aggregate the genomic ranges of all peaks from the individual sets of liver samples to generate a
179 'peaks' assay.
180
181 **Differentially accessible regions (DARs) and differentially expressed genes.** Differential chromatin accessibility,
182 and differential gene expression, between biological conditions (male, female and cGH-treated male) was
183 computed (40) using 10x Genomics Loupe Browser 6.0 (https://www.10xgenomics.com/), which uses the exact
184 negative binomial test to test for differences in mean ATAC-seq sequence read density (cut sites per cell) in a
185 given ATAC region, or differences in gene expression between groups of cells. Differential ATAC-seq peaks
186 between cell types were initially identified using the Loupe Browser's integrated Global Distinguishing function
187 with a differential expression threshold of Log2 (fold change) >1 at FDR (Benjamini-Hochberg) < 0.05. ATAC-seq
188 peaks found in 3% or fewer of nuclei, in the case of non-hepatocyte clusters, or in fewer than 5% of nuclei in the
189 case of hepatocytes were removed. Cell type-specific DARs were identified by excluding peaks found to be cell
190 type-specific DARs for 2 or more clusters. We used the Loupe Browser's integrated Locally Distinguishing function
191 to calculate condition-specific DARs across five different pairwise comparisons: Male vs. Female-F1, Male vs.
192 Female-F2, Male vs. cGH-treated male (M-cGH), Female-F1 vs. M-cGH, and Female-F2 vs. M-cGH, with a
193 significance threshold of adjusted P-value (Bonferroni-corrected FDR) <0.05, except as noted (and where Female-
194 F1 and Female-F2 represent two independent pools of nuclei from control female mouse livers). We used the
195 same approach, based on the Loupe Browser's Locally Distinguishing function, described previously (4), to
196 identify differentially expressed genes for the same five comparisons.
197
198 **ATAC-seq peak to target gene feature linkages.** ATAC-seq peak-to-gene correlations were used to associate sex-
199 biased DARs with their inferred gene targets using feature linkage analysis (40). Briefly, we used the peak-to-gene
200 linkage method embedded in 10x Genomics Cell Ranger ARC (v2.0.1) to link ATAC-seq peaks to specific genes
201 based on Pearson correlations between ATAC-seq peaks and genes, which gives a direct and interpretable
202 measure of the linkage strength (linkage score, values ranging from -1 to 1). We used the set of 127,957 ATAC-
203 seq peaks and 75,798 genes, including some 52,000 mouse lncRNA genes represented by the custom GTF file
204 used in our snRNA-seq analysis (47), 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 we analyzed here is shown in Table S8 and Table S4B. Individual linkages were characterized by the ATAC peak to gene TSS distance, the linkage score (Pearson correlation) and the linkage significance (-log(p-value)).

ATAC regions selected 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 regions also belong to the TAD (topologically associating domain)-based group and one 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).

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. 612 of these DARs were both sex-biased and cGH-regulated: 360 were male-biased and cGH-repressed, and were linked to 235 male-biased, cGH-repressed genes (Pick1, Table S2B) but were not linked to any genes that were male-biased but not cGH-regulated, or genes that were female-biased genes regardless of GH responsiveness; and 252 were female-biased and cGH-activated DARs linked to 166 female-biased, cGH-induced genes (Pick2, Table S2C) but were not linked to other sex-biased genes (Table S4B, columns W, AD, AK, AR; Table S8). 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 (47), likely due to their large distance (> 80 kb) from the nearest gene (Table S4E, columns J-K; Pick6 group, Table S2A, Table S4B). These DARs showed HDI-STARR-seq activities similar to those of the sex-biased and cGH-responsive DARs linked to sex-biased genes, 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 GH-regulated chromatin accessibility displays a sex bias and direction of cGH response opposite from that expected for enhancers that regulate expression of female-biased genes, suggesting these DARs may repress their linked female-specific genes. Similarly, we selected 9 female-biased and cGH-activated 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).

STARR-seq reporter design and synthesis. The 1,839 ATAC regions selected for inclusion in the STARR-seq library were tiled to generate a total of 23,950 STARR-seq reporter sequences, of which only 23,912 reporters were included in the final design (Table S5) owing to issues relating to the liftOver between mouse genome mm10 and mouse genome mm9. Each ATAC region was extended by 70 bp on each end and was tiled using a series of 180 nt reporter sequences, arranged to include 125 nt overlap and a 55 nt shift between neighboring reporters (Fig. 2C). This tiling design minimizes the risk of gaps in reporter coverage of an ATAC region that can arise due to a loss of reporter sequence representation during cloning. Each reporter was synthesized as a 210 nt oligonucleotide using SurePrint HiFi oligo synthesis (Agilent, Inc) and is comprised of the 180 nt reporter sequence flanked on each end by a 15 nt adapter sequence. The pool of 23,912 reporters was then cloned into pPromALB_STARR-seq (plasmid pTYC10; Addgene # 220289) (41), as described below. Each 15 nt adaptor sequence serves as a PCR seed and facilitates the generation of the 210 base pair double-stranded DNA library required for STARR-seq library preparation. Illumina primer sequence #1 is 5'-ACGCTCTTCCGATCT-3', and Illumina primer sequence #2 is 5'-AGATCGGAAGAGCAC-3'. Sequences of ATAC regions and their mouse genomic coordinates are shown in Table S5.

256 **STARR-seq plasmid library construction.** Each reporter sequence was PCR amplified from the SurePrint HiFi oligo
257 synthesis product using specific PCR primers (Table S1) as described below. We set up 6 individual 25 μl PCR
258 reactions (10 ng of oligo pool, 12.5 μl of 2x KAPA HiFi HotStart master mix (Cat No. KK2601, Roche), 2.5 μl of 5
259 μM forward primer, 2.5 μl of 5 μM reverse primer per 25 ul reaction) using PCR primers ON8352 and ON8353
260 (Table S1). PCR reactions were set to generate an HA-adaptor-ligated DNA library under low amplification
261 conditions using the following cycling conditions: 98°C for 45 s, followed by 10 cycles of 98°C for 15 s, 65°C for 30
262 s, 72°C for 45 s, and final elongation step for 1 min at 72°C. The HA-adaptor-ligated DNA was purified in two
263 rounds of 1x SPRI bead selection and then cloned into pPromALB_STARR-seq vector (Fig. S1). Next, we used a 2:1
264 molar ratio of insert to 125 ng SalI/AgeI-digested pPromALB_STARR-seq plasmid in two 10 μl In-Fusion® HD
265 Cloning Plus reactions (TaKaRa, cat. #638910) to assemble STARR-seq reporter plasmid libraries via the homology
266 arms. The reaction products were pooled and precipitated using sodium acetate/ethanol, resuspended in 10 μl
267 EB buffer (10 mM Tris HCl, pH 8.5) and stored at -80°C prior to use (48). Two 2.5 μl aliquots of the precipitated
268 material were transformed separately into 20 μl of electrocompetent MegaX DH10B bacterial cells (ThermoFisher
269 Scientific, cat. # EC0113) using Gene Pulser Xcell Total System (Bio-Rad) and Gene Pulser Electroporation Cuvettes
270 (0.1 cm gap, Bio-Rad; cat No. 1652089) under these conditions: 2 kV, 25 μF, 200 ohms. LB
271 bacterial culture medium (1 L) was used to amplify the STARR-seq plasmid library, which was then purified
272 using PureLink™ Endotoxin-Free Maxi Plasmid Purification Kit (ThermoFisher Scientific, cat. #A31231) to obtain ~
273 2 mg of the final STARR-seq plasmid library, STARR-TYC8.
274
275 **Preparation of plasmid, transfected DNA and transcribed RNA reporter Illumina sequencing libraries.** The
276 following methods, adapted from published STARR-seq protocols (48) and detailed previously (41), were used in
277 a nested PCR amplification protocol to prepare 3 types of Illumina sequencing libraries used to determine the
278 transcriptional enhancer activity of each ATAC region-associated reporter sequence (putative regulatory region)
279 to be assayed by HDI-STARR-seq. 1) A plasmid library was prepared from the same batch of purified plasmid DNA
280 injected into each mouse by HDI. Plasmid DNA was used directly as a PCR template in a nested PCR reaction for
281 library construction (primers ON8381, ON8382; Table S1); 2) a DNA library was prepared from DNA extracted
282 from flash frozen liver tissue collected 10 days after in vivo transfection by HDI (primers ON8381, ON8382); and
283 3) an RNA library was prepared from the transcribed RNA (reporter sequences) extracted from each HDI-
284 transfected liver (primers ON8379, ON8380). STARR-seq reporter plasmid DNA remaining in the liver 10 days
285 after HDI was extracted using a commercial QIAprep spin miniprep kit (QIAgen, cat. # 27106). The liver tissue (200
286 mg frozen liver) was initially homogenized in 750 ul of QIAgen Buffer P1 using a rubber-tipped micropestle and
287 passed through a 20G needle three times to homogenize large tissue pieces. The extracted plasmid DNA was
288 bound to a QIAprep spin column, eluted in EB buffer and used as a PCR template in a subsequent nested PCR
289 reaction using primers ON8381 and ON8382. To extract transcribed reporter RNA, 120 μg of total liver RNA was
290 isolated from 60-100 mg liver tissue from each mouse by TRIzol guanidinium thiocyanate-phenol-chloroform
291 extraction followed by polyA selection using oligo(dT) beads (New England Biolabs, cat. #E7490L). The transcribed
292 RNA was digested with Turbo DNase I (Ambion, cat. # AM2238) at 37°C for 30 min, reverse transcribed using
293 primer ON8731 for STARR-seq RNA first strand synthesis and Superscript III first-strand synthesis supermix
294 (Invitrogen, cat. # 18080-400) and treated with 10 μg RNase A per ml (Fermentas, cat. # EN0531) for 1 h at 37°C.
295 The cDNA obtained was purified using 1.8x SPRI beads and then used as a PCR template in the nested PCR
296 reaction.
297
298 The nested PCR reaction requires one set of primers specifically targeting STARR-seq reporter DNA or RNA
299 followed by amplification using generalized Illumina dual index primers. Primers ON8381 and ON8382 (Table S1)
300 were designed to amplify the STARR-seq plasmid DNA libraries by specifically targeting the unspliced intron
301 incorporated into the original plasmid sequence and present in the plasmid and DNA libraries but spliced out of
302 the transcribed reporter RNA sequences. Primers ON8379 and ON8380, which straddle the splice junction, were
303 designed to specifically amplify only the transcribed and spliced RNA molecules (41). The second, nested PCR
304 reaction was performed for all three library types (plasmid, DNA and RNA) using NEBNext® Multiplex Oligos for
305 Illumina (NEB cat. # E7600S) to further amplify the material and attach Illumina i5 and i7 barcoded sequences
306 used for Illumina multiplex sequencing, performed at Novogene, Inc. (Sacramento, CA).

307
308 **Enhancer identification in TAD regions with sex-biased genes.** TADs are fundamental units that
309 compartmentalize genomic features and help constrain interactions of cis-regulatory sequences to their within-
310 TAD genic targets. We examined 6 TAD regions showing a strong tendency toward sex-bias that contain genes
311 encoding important liver metabolic functions, such as cytochromes P450 (CYPs), solute carrier (SLC) family
312 members, cytosolic sulfotransferases (SULTs), UDP glucuronosyltransferases (UGTs), and 3β-hydroxysteroid
313 dehydrogenases (Hsd3b). As controls, we examined ATAC regions from 4 TADs that do not contain any sex-biased
314 genes or sex-biased gene linkages (TAD_203, TAD_1519, TAD_4080, TAD_4241). ATAC regions within the selected
315 TADs were incorporated in the STARR-seq library (Table S3A).
316
317 **Liver cell type-specific ATAC regions.** The liver MultiOme dataset was used to profile open chromatin regions for
318 each liver cell type or subtype: hepatocytes (including periportal, midlobular and pericentral hepatocytes), liver
319 endothelial cells, hepatic stellate cells (HSC), immune cells, Kupffer cells and cholangiocytes. For each liver cell
320 type (cell cluster), we selected 20 ATAC regions showing the highest specificity for that cell type and defined
321 them as cell-type specific ATAC regions based on the following criteria for the ATAC-seq signal: log2 (fold change)
322 > 1 using the Loupe Browser's integrated Global Distinguishing function, FDR (Benjamini-Hochberg) < 0.05, and
323 the ATAC-seq signal detected in > 3% of nuclei (> 5% for hepatocytes). ATAC regions that met the criteria for cell
324 type specificity for 2 or more liver cell clusters were excluded. Very few liver cell type-specific ATAC regions were
325 identified for midlobular hepatocytes and cholangiocytes (Table S3B, Table S4C).
326
327 **Sequencing analysis of STARR-seq libraries.** Paired-end sequence reads, each 150 base pairs long, were
328 generated for the STARR-seq plasmid, DNA, and RNA libraries. A custom pipeline was developed to handle the
329 raw sequencing data (41). This pipeline takes raw FASTQ files as input and produces various quality control
330 metrics, such as FASTQC reports (version 0.0.13.2 of FASTXToolkit), identification and removal of contaminating
331 adapter sequences (using Trim_galore version 0.4.2), and determination of insert-size length distributions for
332 paired-end read sequencing using Picard (version 1.123). The reads were aligned to the mouse genome (version
333 mm9) using Bowtie2 (version 2.2.6) (49). Subsequently, all identified peak sets were filtered to eliminate peaks
334 overlapping ENCODE blacklisted regions (50). An average of 9.5 million fragments were sequenced for the RNA
335 libraries and 27.1 million fragments for the DNA and plasmid libraries (Table S6). Almost all (96-99%) of the
336 plasmid, DNA and RNA library reads mapped to the 23,912 predesigned reporters for all samples except for
337 library G202M16, a negative control for HDI in the absence of a STARR-seq reporter library. Five RNA sequencing
338 samples were excluded from the downstream analysis (G202M1, G202M2, G202M6, G202M10, G202M12) due to
339 low Pearson's correlations with other RNA libraries from the same biological condition (Table S6). Similarly, DNA
340 sequencing sample G202M18 was removed due to its low Pearson's correlation with the other 3 DNA libraries
341 and the 2 plasmid libraries (Table S6). All Raw fastq sequencing files are available at GEO
342 (https://www.ncbi.nlm.nih.gov/geo/) under accession number **GSE315771**.
343
344 **STARR-seq library characterization.** STARR-seq plasmid sequencing reads, and liver-extracted STARR-seq DNA
345 library sequencing reads, were mapped to the set of 23,912 reporters. Since adjacent reporter sequences overlap
346 by 125 bp of sequence, we used Bedtools (v2.31.0) to set > 95% overlap (>171/180 nt) as the threshold for
347 assigning a sequence read to a given reporter to avoid counting partial overlaps from neighboring reporters.
348 Sequence reads obtained for each plasmid library mapped with high frequency (98.8%) to the 23,912 designated
349 reporters (Table S6), indicating very low non-specific PCR amplification and minimal sequencing errors. The two
350 technical replicate plasmid libraries were highly correlated based on normalized counts across the 23,912
351 reporter regions (R^2=0.984, Fig. S2A). Almost all (>98%) of the inserts were 180 nt long, showing that intact oligos
352 were cloned into the library. The median number of normalized plasmid sequence reads in each reporter region
353 was 322, and 81.9% of the 23,912 reporters were represented within 6.4x of the median frequency (Fig. S2B). The
354 GC content of the 23,912 reporters included in our library design ranged from 19-87%, with the highest frequency
355 observed for reporters at ~45-50% GC content (Fig. S2C). The PCR-amplified plasmid and liver-extracted DNA
356 libraries average sequence read counts both peaked at ~55-70% GC content (Fig. S2D), consistent with the known
357 bias against both GC-rich and AT-rich in sequencing libraries prepared using a PCR-amplification step (51),

358 suggesting that differences in reporter representation across the library is largely determined by the sequence of
359 each reporter.
360
361 **Qualification of STARR-seq library reporters for downstream analysis.** A total of 19,136 STARR-seq reporters
362 (Table S7A, column AU) were qualified for inclusion in downstream analysis based on the following stringent
363 threshold for robust representation in the cloned and bacterially amplified STARR-seq plasmid library, STARR-
364 TYC8: > 40 mapped Illumina sequence read pairs in each of the 3 liver-extracted STARR-seq DNA libraries and in
365 both individual plasmid libraries (Table S7A; Fig. S3B, first 2 columns). 4,776 other reporters did not meet this
366 threshold for consistent STARR-seq library representation and were excluded from downstream analysis. A very
367 similar list of qualified reporters was obtained when the threshold of > 40 mapped Illumina sequence read pairs
368 was normalized to the total sequence depth of each plasmid or DNA library.
369
370 **Determination of HDI-STARR-seq enhancer activity.** Illumina sequencing libraries prepared from RNA extracted
371 from each HDI-STARR-seq-transfected liver ('RNA library') were analyzed using Bedtools to count the number of
372 RNA sequence reads that overlapped each of the 19,136 qualified reporters. Qualified reporters with > 30
373 mapped Illumina RNA sequence read pairs in all 3 biological replicates livers from untreated male mice, in both
374 biological replicate livers from cGH-treated male mice, <u>or</u> in both biological replicates from untreated female
375 mice were deemed active enhancers. A total of 11,996 (63%) of the 19,136 qualified STARR-seq reporters were
376 thus determined to be active, with values ranging from n=7,099-8,922 for each biological condition (Fig. S3E). The
377 remaining 7,140 reporters (37%) were inactive under all three biological conditions. Further, a total of 4,126 of
378 the 19,136 qualified STARR-seq reporters met these criteria for active enhancers under all 3 biological conditions
379 and were designated robust active enhancers (Table S7, column AT). The robust active reporters mapped to
380 1,292 ATAC regions (Fig. S3F) were somewhat more highly represented in the STARR-seq plasmid library than the
381 qualified reporters that showed activity under only 1 or 2 of the 3 biological conditions (Fig. S4A), suggesting a
382 higher false positive rate for the latter sets of non-robust STARR-seq enhancers owing to noise associated with
383 genomic regions with lower representation in the STARR-seq plasmid library. We therefore limited our analysis of
384 differences in enhancer activity between biological conditions (see below) to the set of 4,126 robust active
385 reporters. 982 of the 19,136 qualified reporters were designated stringently inactive reporters based on these
386 criteria: ≤ 30 mapped Illumina sequence read pairs in each RNA library across all 3 biological conditions (i.e., in all
387 7 RNA libraries included in our analysis; Table S7, column AT). Since many of these inactive reporters were poorly
388 represented in the STARR-seq library, we subdivided the 982 stringently inactive reporters into two sets based on
389 mean plasmid library read counts: stringently inactive-High (top 400 reporters), which were used as a background
390 set for downstream enrichment analyses; and stringently inactive-Low (bottom 582 reporters), which were
391 excluded from the analysis (Fig. S4A). A further 6,188 qualified reporters were designated 'Other' and were
392 excluded from further analysis. The 'Other' designation was assigned based on inconsistent STARR-seq enhancer
393 activity across biological replicates and may relate to low representation of these reporters in the plasmid library
394 (Fig. S4A). The lists of stringently inactive and robust active reporters were highly similar, or nearly identical,
395 respectively when the thresholds for minimum numbers of raw mapped read counts stated above were
396 normalized for the total sequence depth of each RNA library (e.g., 4,125 of the 4,126 robust active reporters were
397 retained when using those alternative criteria).
398
399 **Relationship between sex bias and cGH responsiveness of STARR-seq enhancer activity.** A linear regression
400 model (52) was implemented in R Studio to assess the relationship between response to cGH treatment (log2
401 ratio of STARR-seq enhancer activity in cGH-treated male liver to STARR-seq enhancer activity in male liver) and
402 sex bias (log2 ratio of STARR-seq enhancer activity in female liver to STARR-seq enhancer activity in male liver).
403
404 **Regulated active STARR-seq enhancers.** Robust active enhancers with > 2-fold differences between biological
405 conditions in their STARR-seq reporter activity (defined as mean RNA library sequence reads for a given HDI-
406 STARR-seq reporter, averaged across biological replicates, normalized for total mapped sequence reads) were
407 designated regulated active enhancers (Table S7A). Regulated active enhancers whose pattern of HDI-STARR-seq
408 enhancer activity regulation matched their pattern of ATAC-seq activity regulation were designated positively

acting regulated enhancers. One example is the set of DARs whose ATAC-seq activity was both male-biased and cGH-suppressed (Pick1 DAR set, Table S2) and whose RNA reporter(s) showed >2-fold increase in STARR-seq enhancer activity in male compared to female liver as well as a >2-fold decrease in HDI-STARR-seq enhancer activity in cGH-treated male liver compared to control male liver. Another example is the set of DARs whose ATAC-seq activity was male-biased but not cGH-regulated (Pick3_M DAR set, Table S2) and whose reporter(s) showed >2-fold increase in STARR-seq enhancer activity in male compared to female liver but did not show >2-fold decrease following cGH infusion. In contrast, male-biased and cGH-suppressed DARs whose STARR-seq reporter activity showed < 2-fold male-bias and < 2-fold response to cGH infusion were non-regulated enhancer DARs. We also identified a set of male-biased, cGH-repressed DARs linked to female-biased genes (putative repressors; Pick4 DAR set, Table S2), whose reporters showed >2-fold increase in STARR-seq enhancer activity in male compared to female liver and also >2-fold decrease in STARR-seq enhancer activity in cGH-treated male compared to control male liver, and correspondingly for female-biased, cGH-activated DARs linked to male-biased genes (Pick4 and Pick5 DAR sets; Table S2). Finally, 452 robust active enhancers whose reporters showed no regulated STARR-seq enhancer activity in any comparison (Table S9A, column AI) were identified as non-regulated DARs.

Chromatin state analysis of STARR-seq regions in liver. Chromatin state maps, previously defined for adult male mouse liver, and separately, for adult female mouse liver (32) using ChromHMM (53), were used to determine the enrichment of individual chromatin states in various sets ATAC regions. These chromatin states are based on a 14-state model (states E1-E14) and were previously used to characterize sex differences in chromatin state and their relationships to sex-biased gene expression (32). Bedtools overlap was used to assign the chromatin state determined for each sex to each ATAC region. An ATAC region that overlapped two or more different chromatin states was assigned to the state with the greatest overlap in bp. The distribution of chromatin state enrichments (overall chromatin landscape) was determined separately for male and for female liver for each set of functionally active enhancer region DARs in comparison to a background set comprised of 317 DARs whose reporters showed no regulated STARR-seq enhancer activity in any comparison (i.e., 452 DARs mentioned above, after excluding 82 non-regulated DARs from the Pick1 DAR set and 53 non-regulated DARs from the Pick2 DAR set) (Table S9A column AJ).

Motif analysis. *De novo* motif discovery was carried out using MEME-CHIP (54) (https://meme-suite.org) run in the classic mode to discover motifs enriched relative to a random model based on frequencies of each nt in the input sequences (Table S13A). MEME-CHIP was also run in the discriminative (non-classic) mode comparing an input set of sequences to a background set of sequences, as specified in the text, using default settings for both modes (Table S13B). For each set of regions analyzed (Table S9A, columns U-X) sequences were input using the coordinates for each region shown in Table S9A. For example, STARR-seq analysis of a set of 360 male-biased and cGH-repressed DARs identified 41 functionally active regulated ATAC regions (i.e., regulated Pick1). 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 (Table S9A, column AK), 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; and 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-activated DARs (Table S9A, columns V, X). The *de novo* discovered motifs were compared with published mouse TF motifs determined by protein-binding microarrays and obtained from the UniPROBE motif database (55). 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) (56,57) (Table S13C, columns AP-BM).

460 **Results**
461
462 **Candidate sex-biased and GH-responsive regulatory regions identified by snATAC-seq + snRNA-seq analysis.** We
463 implemented 10X Genomics MultiOme technology (snATAC-seq in combination with snRNA-seq) to identify open
464 chromatin regions that show sex-dependent and/or GH-regulated accessibility in mouse liver nulcei from male,
465 female, and cGH-infused male mice. 46,188 nuclei across the 3 biological conditions passed quality control filters
466 for both the snATAC-seq and snRNA-seq modalities. ATAC-seq data quality was confirmed by TSS enrichment
467 scores of 5.5–6.4. Five major liver cell types were resolved by UMAP and identified using established marker
468 genes (Fig. 1A). 127,957 ATAC regions were identified across all liver cells and characterized regarding sex-bias
469 and the impact of cGH infusion on chromatin accessibility. We linked individual ATAC regions to predicted gene
470 targets using correlations between chromatin accessibility and gene expression; these 'co-expression'
471 relationships indicate a shared regulatory mechanism and infer enhancer–gene targeting relationships. For
472 example, a feamle-biased, cGH-activated DAR located 46.1 kb downstream of the *Cux2* TSS (ATAC_39124)
473 showed linkages to 4 nearby genes (Fig. 1B, Fig. 1C): two positive linkages, to female-biased genes (*Cux2*,
474 *lnc4593*), indicate positive regulation, and one negative linkage, to *P2rx4*, indicates negative regulation by the
475 ATAC region (Fig. 1D). The linkage to *Cux2* showed the highest correlation (r = 0.41) and significance (–log(*P)* FDR
476 = 70.6).
477
478 **Selection of DARs for HDI-STARR-seq library.** We selected ATAC regions for HDI-STARR-seq analysis, focusing on
479 genomic regions showing cGH-regulated sex differences in accessibility in hepatocytes, the primary liver cell type
480 showing gene expression and chromatin accessibility sex differences (4,40). Three major types of DAR sets were
481 chosen: 1) 360 DARs that were male-biased and cGH-repressed and linked to correspondingly regulated male-
482 biased genes (Fig. 2B, Table S2A, 'Pick1'), and 252 DARs that were female-biased and cGH-activated and linked to
483 correspondingly regulated female-biased genes (Fig. 2B, Table S2A, 'Pick2'); 2) 102 male-biased and 186 female-
484 biased DARs whose chromatin accessibility was unresponsive to cGH, subsets on which were on chrY and chrX,
485 respectively (Table S2C, 'Pick3'); 3) 749 ATAC regions from 10 TAD regions (58) encompassing sex-biased genes
486 important for hepatic lipid metabolism and fatty liver disease, including cytochromes P450 (CYPs), solute carrier
487 (SLC) family members, sulfotransferases (SULTs), UDP glucuronosyltransferases (UGTs), and 3β-hydroxysteroid
488 dehydrogenases (HSD3B family) (Table S3A). Other ATAC regions evaluated include putative repressor sequences
489 (Pick4, Pick5), which are male-biased or female-biased DARs whose regulatory orientation relative to their linked
490 genes suggests repressive rather than activating functions, and sex/cGH-regulated ATAC regions not linked to any
491 genes (Pick6), which may represent long-range enhancers, given their large distances (> 80 kb) from nearest
492 genes (Table S2, Table S4E). These DAR sets all showed similar ranges of HDI-STARR-seq enhancer activity (Fig.
493 S5). Additionally, 124 ATAC regions were included to determine the liver cell type-specificity of HDI-STARR-seq
494 (Table S3B). Overall, we profiled 1,839 ATAC regions (mean 634 nt), tiled with overlapping reporters spaced 55 bp
495 apart to finely map active sub-sequences (Fig. 2C).
496
497 **Active enhancers identified by HDI-STARR-seq.** A pool of 23,912 unique single-stranded oligonucleotides (Table
498 S5, Table S7A) tiled across the 1,839 ATAC regions was cloned in bulk into the STARR-seq reporter vector
499 pPromALB_STARRseq (41). The library was high-quality, with 19,136 (80%) of the reporters passing stringent
500 thresholds for robust representation (see Methods). We used HDI to transfect the STARR-seq reporter library into
501 mouse liver and assay the enhancer activity of each ATAC region under 3 biological conditions: male, female, and
502 10-day cGH-infused male liver. Livers were harvested after 11 days and transfection efficiency was assessed with
503 a co-transfected Renilla luciferase reporter. Low Renilla activity livers, indicating poor transfection, were excluded
504 from downstream analysis (Fig. 3A, asterisks). DNA reporter sequences were recovered from the livers without
505 bias (Fig. S6). RNA reads from the 7 livers showing robust HDI transfection (n=3 males, n=2 females, n=2 cGH-
506 treated males) were mapped to the 23,912 reporters with high efficiency (>98%, Table S6) and general overall
507 consistency within groups (Fig. 3B). In total, 11,996 of the 19,136 reporters (63%) showed enhancer activity under
508 one or more biological conditions and were designated *active enhancers* (Fig. S3E; Table S7A, column AU); 4,126
509 reporters (21.6%), representing 1,292 ATAC regions, were active under all 3 conditions and are termed *robust*
510 *active enhancers* (Fig. S3F; Table S7A, column AT).

Impact of cGH infusion on functional activity of active sex-biased enhancers. Subcutaneous infusion of GH in male mice using osmotic minipumps (cGH infusion) mimics the female plasma GH pattern and overrides the male pulsatile pattern of pituitary GH release, substantially feminizing sex-biased gene expression in mouse liver within 7–10 days (43,59). We used cGH infusion to identify ATAC regions whose enhancer activity responds to changes in temporal plasma GH patterns. cGH infusion for 10 days feminized the liver, confirmed by increased expression of female-biased genes (*A1bg*, *Cux2*, *Sult3a1*) and decreased male-biased gene expression (complete repression, *Elovl3*; partial repression, *Cyp7b1*) (Fig. 3C). Liver stress gene expression, which in a prior study was elevated for up to 4 days following surgery to implant osmotic minipumps (43), was not was observed in response to HDI or 10 days after minipump implantation (Fig. S7). Therefore, the 10-day post-HDI interval is sufficient for recovery from any stress-related responses.

We examined the set of 4,126 robust active reporters to identify those showing sex-dependent or cGH-responsive activity. Of these, 1,397 reporters exhibited HDI-STARR-seq enhancer activity that was either male-biased (n=310), female-biased (n=541), cGH-induced (n=398), and/or cGH-repressed (n=419) (Table S7A, column AW). Concordant effects across neighboring reporters support true local regulation (Fig. 3D). The 1,397 reporters mapped to 840 distinct ATAC DAR regions, including 168 male-biased DARs whose accessibility was repressed by cGH, and 137 female-biased DARs whose accessibility was increased by cGH. Further, our analysis indicated that the sex-bias of enhancer activity was the key determinant of cGH responsiveness. Thus, 87 male-biased reporters showed decreased enhancer activity following cGH infusion, versus only 1 showed an increase (Fig. 4A). Moreover, among the female-biased reporters, cGH infusion increased enhancer activity for 179 reporters, but decreased it for only 4 (Fig. 4A). This pattern mirrors the well-established feminization of liver chromatin accessibility and gene expression following cGH infusion (43,59), supporting the conclusion that HDI-delivered STARR-seq reporters recapitulate physiological GH-regulated enhancer responses.

Plotting the fold-change in enhancer activity following cGH infusion versus the sex-bias in enhancer activity revealed strong positive correlations for HDI-STARR-seq reporters mapping to male-biased/cGH-repressed DARs (360 DARs, n=725 reporters) and to female-biased/cGH-activated DARs (252 DARs, n=693 reporters), respectively (Fig. 4B, Fig. 4C). Non-regulated DAR sets showed weaker, less significant correlations (Table S7B, Fig. S8). Separating sex-biased but cGH-unresponsive reporters into autosomal and sex chromosome subsets revealed stronger correlations for autosomal reporters ($p = 5.6E–31$ vs $p = 3.5E–03$) (Fig. S8C-S8D), consistent with sex chromosome DARs being regulated by GH-independent mechanisms (40). Importantly, analysis of the reporter datasets averaged at the ATAC region level revealed an even stronger correlation between enhancer activity sex bias and cGH-responsiveness (Fig. 4D).

Enrichment of regulated active enhancers for DARs with matched response. Enrichment analysis confirmed the close association between the sex bias and cGH responsiveness of HDI-STARR-seq reporter activity and the sex-bias and cGH regulation of endogenous mouse liver chromatin accessibility (Table S7C). Male-biased and cGH-repressed DARs were significantly enriched for male-biased, cGH-repressed enhancer activity compared with female-biased and cGH-activated DARs (ES = 8.3, $p = 3.7E-08$, Fisher's exact test). Enrichments for other regulated DAR sets appear in Table S7C. The sex-bias in chromatin accessibility in liver did not differ significantly between the regulated and non-regulated DAR sets (Table S7E). Furthermore, in male liver, no significant difference in chromatin accessibility was found for regulated vs. non-regulated male-biased and cGH-repressed DARs, but in female liver, the regulated female-biased and cGH-activated DARs showed significantly greater accessibility than the non-regulated DAR set (Table S7E). Gene linkage metrics did not differ between the regulated and non-regulated sets of sex-biased, cGH-responsive HDI-STARR-seq enhancers, which showed no differences in the proportion of non-sex-biased genes, distance to linked genes, or linkage significance for either DAR type (Fig. S9, Table S12). Finally, many sex-biased and cGH-regulated DARs were linked to both sex-biased and non-sex-biased genes (Table S8), but the linkages to non-sex-biased genes were characterized by significantly longer distances and lower reliability (lower score and significance) than the linkages to sex-biased genes (Fig. S9).

562 **Epigenetic features enriched at active sex-biased enhancers.** We investigated the enrichment of adult male and
563 adult female mouse liver chromatin states and histone marks (32) in the sets of active, functional sex-biased and
564 cGH-regulated enhancers to identify epigenetic features associated with the regulation of HDI-STARR-seq
565 enhancer activity in a sex-dependent and cGH-responsive manner (Fig. 5). We compared the frequency of each of
566 14 chromatin states, defined separately for male and female mouse liver (32), across the ATAC regions
567 comprising each sex-biased, cGH-regulated DAR set with that in a background set of 317 robust active but non-
568 regulated ATAC regions (Table S9A, column AJ). Four DAR sets were compared: 41 male-biased and cGH-
569 repressed DARs with regulated activity, and 82 male-biased and cGH-repressed DARs with non-regulated activity;
570 43 female-biased and cGH-activated DARs with regulated activity, and 53 non-regulated female-biased and cGH-
571 activated DARs (Table S9A, columns U–X). We observed clear differences in chromatin state enrichments
572 between the regulated and non-regulated enhancer DAR sets, with distinct results for the male-biased and the
573 female-biased enhancer DARs. Thus, the 41 regulated male-biased and cGH-repressed DARs were significantly
574 enriched for chromatin state E6, but not for chromatin state E7, in male liver, whereas the 43 regulated female-
575 biased DARs were significantly enriched for chromatin state E7 but not for chromatin state E6 in female liver (Fig.
576 5A, Table S10A). States E6 and E7 are both characterized by open chromatin and the active enhancer mark
577 H3K27ac; E6 is more highly enriched for the active enhancer mark H3K4me1 while E7 has high probability for the
578 promoter mark H3K4me3 (Fig. 5A; ChromHMM emission probabilities, where DHS indicates open chromatin).
579 Further, in female but not male liver, the set of 41 regulated male-biased and cGH-repressed DARs, but not their
580 non-regulated counterparts (n=82), was significantly enriched for chromatin state E9. State E9 is characterized by
581 the same two active enhancer marks as state E6 – H3K4me1 and H3K27ac – but is devoid of open chromatin.
582 Thus, the male bias and cGH regulation of HDI-STARR-seq enhancer activity of these DARs is significantly
583 associated with the presence of active enhancer marks at the corresponding endogenous mouse genomic
584 sequences in both sexes, but with open chromatin present in male but not female liver. The unique patterns of
585 enrichment of each sex-specifically regulated DAR set for distinct chromatin states in mouse liver support the
586 proposal that the HDI-transfected plasmid DNA sequences acquire histone marks and epigenetic states similar to
587 those found at the corresponding endogenous DNA regions in mouse liver.
588
589 Consistent with the above findings, the regulated active enhancer DARs were also enriched for genomic regions
590 with specific activating histone marks. The regulated male-biased, cGH-repressed enhancer DARs showed strong
591 enrichments for genomic regions characterized by active enhancer histone marks in male liver (ES = 8.6 for
592 H3K27ac, ES = 14.2 for H3K4me1) but not female liver, with ~2.7-fold lower enrichments seen for the
593 corresponding non-regulated DARs (Fig. 5B; Table S10A). In contrast, regulated but not non-regulated female-
594 biased, cGH-activated DARs showed strong enrichments for genomic regions characterized by marks for active
595 enhancers (ES = 3.7 for H3K27ac, ES = 27.9 for H3K4me1) and marks for promoters (ES = 4.5 for H3K4me3) in
596 female but not male mouse liver. The regulated male-biased DAR regions were significantly depleted of
597 H3K27me3 repressive histone marks in male liver (ES = 0 to 0.16) (Fig. 5B, *red asterisks*) but showed no significant
598 enrichment or depletion of H3K9me3 or H3K36me3 (repressive and transcribed region marks, respectively; Table
599 S10). In contrast, the set of regulated female-biased, cGH-activated DARs showed weak but significant
600 enrichment for H3K27me3 marks in male liver (ES = 2.75, p = 2.96E-3), which parallels the documented functional
601 repression by H3K27me3 marks of female-biased enhancers in male liver (60).
602
603 **Mapping to HDI-STARR-seq active enhancers to disease relevant sex-biased genes**. We used multiOmic gene
604 linkage analysis to identify sex-biased genes predicted to be regulated by the 840 DARs targeted by the 1,397 sex-
605 biased and/or cGH-responsive HDI-STARR-seq reporters (Table S9C). 305 of the 840 DARs were linked to genes
606 with corresponding regulation. Thus, 168 male-biased, cGH-repressed DARs were linked to male-biased, cGH-
607 repressed genes, and 137 female-biased, cGH-activated DARs were linked to female-biased, cGH-induced genes
608 (Fig. 4A, *bottom*). Of the other DARs, 37 were sex-biased DARs potentially involved in repressive interactions (31
609 male-biased, cGH-repressed DARs and 6 female-biased DARs linked to sex-opposite genes), 18 were sex-biased
610 DARs (of which 6 are on ChrX, 6 on ChrY) showing GH-independent sex-biased regulation, and 416 were sex-
611 biased DARs not linked to any genes (n=261) or not linked to any sex-biased genes (n=155) (Table S13D, column
612 K). The male-biased, cGH-repressed DAR gene linkages included MASLD-enabling genes (e.g., *Bcl6*, *Lnc-Lfar1* and

613 *Nox4)*; and the female-biased, cGH-activated DAR linkages included MASLD-protecting genes (e.g., *Hao2*, *Fmo2*,
614 and *Trim24)*. Furthermore, the male-biased, cGH-repressed DARs were primarily in an enhancer state in male
615 liver (n=128 of 168, 76%); the remaining 40 DARs were in promoter (n=18), bivalent (n=6), transcribed (n=12), or
616 inactive (n=4) states. These ATAC regions shifted away from enhancer and promoter states in male liver to
617 increasingly populate bivalent, transcribed and inactive states in female liver (female liver: 102 enhancer (61%),
618 12 promoter, 11 bivalent, 21 transcribed, 22 inactive) (Table S9C). Similarly, the 137 female-biased DARs showed
619 reciprocal shifts from enhancer and promoter states in female liver to bivalent, transcribed and inactive states in
620 male liver (female: 100 enhancer (73%), 23 promoter (17%), 11 bivalent, 2 transcribed, 1 inactive; male: 71
621 enhancer (52%), 11 promoter (8%), 24 bivalent, 7 transcribed, 24 inactive) (Table S9C).
622
623 **Enrichment of GH-dependent liver TFs at regulated vs. non-regulated sex-biased enhancers.** We investigated
624 the regulated sex-biased and cGH-responsive enhancer DARs to determine whether the corresponding genomic
625 sequences show enriched binding by STAT5 and BCL6, two essential, GH-regulated liver TFs implicated in GH-
626 regulated sex-biased liver gene expression (13,32) and in liver metabolic disease (1,5,16,27). We examined the
627 enrichment of sex-biased binding by each TF in regulated versus non-regulated sex-biased DARs compared with a
628 background set of 317 non-regulated enhancers (Fig. 6A). Male-biased STAT5 binding (liver ChIP-seq) (13) was
629 significantly enriched at both regulated and non-regulated male-biased, cGH-repressed DARs. Similarly, female-
630 biased STAT5 binding was enriched at both regulated and non-regulated female-biased, cGH-activated DARs, but
631 was depleted at male-biased, cGH-repressed DARs. Regulated and non-regulated sex-biased, cGH-responsive
632 ATAC sites showed no differences in STAT5 motif frequency (Fig. 6B, Table S13D), suggesting that other features,
633 including histone marks, chromatin states and other bound TFs, are central to the regulation of sex-biased
634 enhancer activity.
635
636 One such factor may be BCL6, whose binding to male mouse liver chromatin was enriched (ES = 4.25, p=1.41E-3)
637 at the set of female-biased, cGH-activated DARs (Fig. 6A), consistent with BCL6 repressing many female-biased
638 genes in male liver (13,32). BCL6 motif frequency was higher in the female-biased than the male-biased DAR sets,
639 but with equal frequency between the regulated and non-regulated sets of sex-biased, cGH-responsive active
640 enhancers (Pick1 and Pick2 sets, Fig. 6B; Table S13D). Of note, BCL6 motifs were significantly enriched at the
641 regulated versus the non-regulated female-biased DARs stringently unresponsive to cGH infusion (Pick3_F DAR
642 set; ES = 7.3, p=0.0007 (Fisher's Exact test); Fig. 6B). This finding implicates BCL6, a transcriptional repressor
643 showing male-biased expression in liver (26), in controlling sex-biased enhancer activity for these female-biased
644 DARs, whose chromatin accessibility is unresponsive to cGH infusion. The absence of sex-biased enhancer activity
645 at these 55 female-biased DARs can therefore be explained by their deficiency in BCL6 motifs and BCL6-
646 dependent repression in male liver.
647
648 Motif scans for CUX2, a female-specific liver TF that represses many male-biased genes in female liver (30),
649 revealed a higher motif frequency in the regulated male-biased versus the regulated female-biased DAR sets
650 (37% vs. 12%, for Pick1 vs. Pick2; 42% vs. 2%, for Pick3_Male vs. Pick3_Female; Fig. 6B). Additionally, CUX2 motifs
651 were significantly enriched in the regulated versus the non-regulated male-biased DAR sets, both those from
652 Pick1 (male-biased DARs suppressed by cGH) and from Pick3_M (male-biased DARs not suppressed by cGH) (ES =
653 2.8 and 4.2, respectively; Fig. 6B). In view of the near absence of CUX2 protein in male mouse liver (28), these
654 results implicate CUX2 in distinguishing regulated from non-regulated male-biased, cGH-repressed DAR enhancer
655 activity. The motif for HNF6 (Onecut1), which is very similar to that of CUX2 (29), also showed significant
656 enrichment at the regulated versus the non-regulated male-biased DAR sets (Fig. 6B, Table S13D). Together,
657 these findings support the conclusion that the regulated and non-regulated male-biased DARs introduced by HDI
658 bind different subsets of TFs, which are key determinants of the sex bias in enhancer activity detected by HDI-
659 STARR-seq.
660
661 **Motif enrichment in sex-biased and cGH-regulated enhancers.** We used MEME-CHIP to discover *de novo* motifs
662 for TFs that distinguish regulated from non-regulated sex-biased, cGH-responsive DARs in regulated male-biased,
663 cGH-repressed (n=41) and female-biased, cGH-activated DARs (n=43). The *de novo* motifs identified were

matched to a database of established mouse TF motifs (55). Three background sets were used: 1) sex-/cGH-matched ATAC regions with non-regulated enhancer activity; 2) 317 non-regulated robust active ATAC regions; and 3) 65 stringently inactive ATAC regions. Top enriched *de novo* motifs for 41 regulated male-biased, cGH-repressed DARs versus 82 non-regulated male-biased, cGH-repressed DARs were most similar to HNF4A and ESRRA/NR2F2 (Fig. S10). This was confirmed using alternative backgrounds (Table S13B); secondary matches included MAFB and LHX8, in addition to several motifs with no known TF matches. Top enriched motifs for 43 regulated female-biased, cGH-activated DARs versus 53 non-regulated female-biased, cGH-activated DARs best matched SOX7 and CRX_3485.1 (Fig. S11). Secondary matches included SP4/KLF7, NKX2-5, ASCl2 and RFX7 (RFXDC2) (Table S13B).

The *de novo*-discovered motifs were scanned across sex-biased and/or cGH-responsive DARs for frequency differences between regulated and non-regulated enhancer activity DAR sets (Table S13C, Table S13D). Results confirmed significant enrichment of the *de novo* HNF4A-like motif in regulated versus non-regulated male-biased, cGH-repressed DARs (56% vs 15% frequency; ES = 7.45, Fisher's exact test, $p < 0.0001$) (Table S13D). Similarly, the *de novo* RFX7-like motif was significantly enriched in regulated versus non-regulated female-biased, cGH-activated DARs (35% vs. 9% frequency; ES = 5.1, Fisher's exact test, $p = 0.0046$; Table S13D).

TAD-based view of active enhancers. We used HDI-STARR-seq to assay 749 ATAC regions across 10 TAD regions (Table S3A, Table S4D). Control TAD regions, devoid of sex-biased genes, generally contained few regulated reporters, as expected. In contrast, the organization of TAD_2885 (chr19; 111 ATAC regions, 47 linked genes, of which 28 are sex-biased) is consistent with domain-level co-localization. This TAD contains the *Cyp2c* gene locus and includes female-biased genes and regulatory regions concentrated at the 5' end and within a female-biased intra-TAD (35,58), suggesting genuine regulatory activity with tandem reporters showing concordant sex-bias patterns. Additionally, a cGH-suppressed male-biased enhancer near *Cyp2c70* at the 3' end of TAD_2885, outside the female-biased intra-TAD region, potentially regulates male-biased genes via CTCF/cohesion-mediated loops ~1 MB to the 5' end (35). TAD_4231 contains male-biased genes with overlapping correspondingly regulated reporters, confirming sex-bias consistency of the measured HDI-STARR-seq enhancer activity. TAD_4827 is characterized by female-biased genes and female-biased, cGH-activated reporters showing weak STAT5 binding, plus chromatin accessibility differences in ATAC regions showing sex-specific patterns.

HDI-STARR-seq activity in NPCs. Plasmid DNA introduced by HDI is reported to preferentially transfect hepatocytes versus other liver cell types (61), raising the question of whether NPC enhancer activity can be assayed by HDI-STARR-seq. To address this, we assayed HDI-STARR-seq enhancer activity of 63 ATAC regions (972 reporters) whose chromatin is preferentially accessible in hepatic stellate cells, liver endothelial cells, immune cells, or cholangiocytes (Table S11). Results were compared to a similar number of hepatocyte-enriched ATAC regions. Overall, robust active reporters from ATAC regions with NPC specificity were 1.8-fold depleted ($p < 0.00001$, Fisher's exact test) versus hepatocyte ATAC region reporters (Fig. S12A). Depletion was 3.6-fold for hepatic stellate cell-specific reporters ($p < 0.0001$) and 1.8-fold for liver endothelial cell reporters ($p=0.0003$) (Table S11). Cholangiocytes, which share many hepatocyte characteristics (36,37), showed a 1.73-fold enrichment of robust active enhancers versus hepatocytes, but this did not reach statistical significance ($p = 0.2$), probably due to the small number of cholangiocyte-enriched DARs available for analysis. HDI-STARR-seq enhancer activity was significantly higher for hepatocyte ATAC regions than for NPC ATAC regions (Fig. S12C), consistent with the hepatocyte preference of HDI.

Discussion

This study establishes a functional genomics framework that moves beyond correlative chromatin maps to identify and then directly validate in intact mouse liver hundreds of sex-biased, GH-responsive hepatocyte enhancers driving liver sexual dimorphic gene expression, including expression of genes associated with sex differential, GH-regulated MASLD susceptibility. By integrating hepatocyte-resolved single-nucleus multiOmics for

both male and female mice (40) with functional enhancer profiling in intact liver using HDI-STARR-seq (41), we experimentally validated sex-biased, GH-regulated reporters representing 840 enhancer DARs, including 168 male-biased and cGH-repressed DARs and 137 female-biased and cGH-activated DARs. Critically, our *in vivo* approach overcomes critical limitations of prior studies by assaying enhancer function in intact liver under physiological conditions, where transfected episomal reporters undergo chromatinization (41,62,63) and respond to endogenous regulatory factors within their native hormonal and metabolic environment. Together, our findings establish three major mechanistic principles, discussed below, that advance our understanding of how GH-regulated chromatin accessibility drives sex-dependent hepatocyte gene expression linked to metabolic disease risk.

First, sex-biased hepatocyte chromatin accessibility is a strong but imperfect predictor of functional sex-biased enhancer activity. While 63% of DNA-qualified reporters (11,996 of 19,136) showed enhancer activity under at least one biological condition, a smaller fraction exhibited regulated enhancer activity in response to sex or GH signaling. This is consistent with our tiling design, in which each ATAC region was covered by multiple overlapping 180-nt reporters spaced 55 bp apart. Functional enhancer activity is often restricted to discrete sequence elements within accessible chromatin regions, and thus only a subset of reporters tiling a regulated DAR would be expected to show regulated activity. Indeed, Many DARs gave concordant regulatory patterns across neighboring reporters, supporting the presence of localized functional elements embedded within broader accessible chromatin regions. Moreover, analysis at the ATAC region (DAR) level revealed a highly significant correlation between sex-bias and cGH-responsiveness of enhancer activity (Fig. 4D). Furthermore, we identified 41 male-biased, cGH-repressed DARs harboring regulated reporters showing matched male-biased, cGH-repressed enhancer activity, and similarly for 43 female-biased, cGH-activated DARs with matched female-biased, cGH-induced enhancer activity, demonstrating that concordance between chromatin state and regulation of enhancer activity is non-random. However, these matching patterns were observed for only a subset of the sex-biased and cGH-regulated hepatocyte DARs, indicating that open, accessible chromatin is a necessary but not sufficient condition for enhancer function. This finding reconciles earlier genome-wide DNase-seq studies, which identified thousands of sex-biased DHS sites in bulk liver tissue that co-localize with or are found nearby sex-biased genes (31,32), with the reality that only a subset of accessible regions function as bona fide enhancers *in vivo* under physiological conditions. The discordance between chromatin accessibility and enhancer function underscores the importance of direct functional validation and indicates that additional determinants, such as specific TF binding motifs, cooperative TF occupancy, histone modifications, and three-dimensional chromatin architecture, are required to convert an accessible chromatin region into an active enhancer, in accord with results from massively parallel reporter assays in other biological systems (64-66). Importantly, our tiling approach, while enabling fine mapping of functional sequences within DARs, may lead to false negatives when critical enhancer elements span reporter boundaries or require longer sequence contexts for full activity. Nevertheless, our study extends these general principles to sex-biased hepatocyte enhancers assayed in intact mouse liver under conditions where episomal reporters become chromatinized and respond to endogenous hormonal cues.

Second, functional sex-biased, GH-responsive enhancers are distinguished from non-functional sex-biased DARs by distinct epigenetic signatures that predict responsiveness to hormonal regulation. Among male-biased, cGH-repressed DARs, those showing correspondingly regulated HDI-STARR-seq enhancer activity were highly enriched for genomic sequences characterized by chromatin state E6 with its dual enhancer histone marks in male mouse liver (H3K4me1, 14.2-fold; H3K27ac, 8.6-fold) compared to a background set of 317 robust active but non-regulated genomic regions. Similarly, female-biased, cGH-activated DARs showing correspondingly regulated enhancer activity were strongly enriched for genomic sequences with enhancer marks in female mouse liver (H3K4me1, 27.9-fold; H3K27ac, 3.7-fold) but in addition showed a unique 4.5-fold enrichment for both the promoter mark H3K4me3 (chromatin state E7) and RNA polymerase-II-bound genomic regions, suggesting that these enhancers — which are not closer to their linked gene targets than male-biased enhancers — adopt promoter-like regulatory architectures. Strikingly, regulated male-biased enhancer DARs were depleted of H3K27me3 repressive histone marks in male liver, while regulated female-biased enhancer DARs were enriched for H3K27me3 marks in male liver, consistent with the H3K27me3-mediated functional repression of female-

766 biased genes established in male liver (60). Taken together, these findings provide strong support for the
767 proposal that episomal HDI-STARR-seq reporters delivered to mouse liver acquire endogenous hepatocyte-like
768 epigenetic states in a sequence-dependent manner, enabling functional enhancer activity when the appropriate
769 activating marks are present and repressive marks are absent. Further, our findings indicate that enhancer
770 sequences 180 bp in length (Fig. 2C), when combined with the minimal *Alb* promoter incorporated in our HDI-
771 STARR-seq reporter plasmids, encode sufficient information to recruit hepatocyte chromatin modifiers, and
772 perhaps could be employed to study the impact on sex-biased enhancer function and sex-differential disease
773 susceptibility when these epigenetic states are disrupted by genetic variants (34) or environmental exposures
774 (67).
775
776 Third, binding site architecture for GH-regulated TFs discriminates functional, regulated enhancers from non-
777 functional DARs with similar chromatin accessibility. Sex-biased STAT5 binding to liver chromatin was significantly
778 enriched at the correspondingly sex-biased and cGH-responsive DAR sequences but with no large differences
779 between regulated and non-regulated DAR enhancers. This may, in part, reflect the direct role that STAT5 plays in
780 sex-biased chromatin opening *per se* for a subset of male-biased open chromatin regions (33). In contrast, motifs
781 for the female-specific transcriptional repressor CUX2 (30) were enriched at regulated compared to non-
782 regulated male-biased DARs, while motifs for the male-biased transcriptional repressor BCL6 (13,26,27) were
783 enriched at regulated versus non-regulated female-biased DARs unresponsive to cGH. These patterns suggest a
784 model in which STAT5 binding at male-biased DARs is necessary but not sufficient for regulated enhancer activity.
785 Hormonal responsiveness of DAR enhancer activity would be conferred by additional determinants, such as CUX2
786 motifs that enable female-specific repression regulated by cGH, and BCL6 motifs that enable male-specific
787 repression of female-biased enhancers. *De novo* motif discovery supported this model, identifying HNF4A-like
788 motifs enriched at regulated versus non-regulated male-biased DARs and RFX7-like motifs enriched at regulated
789 female-biased DARs. HNF4A is a liver-enriched transcription factor that cooperates with STAT5 to activate male-
790 biased genes (25) and its motif enrichment at functional enhancers suggests it acts as a co-determinant of GH-
791 responsive enhancer activity. Motif enrichment for RFX7 at regulated female-biased enhancers nominates this
792 tumor suppressor (68) as a novel candidate for female-biased hepatic gene regulation. These findings further
793 refine the current model of GH-regulated liver sexual dimorphism: rather than STAT5 and STAT5-regulated
794 repressors (BCL6, CUX2) acting in simple opposition, our data suggest a layered regulatory code in which
795 combinatorial motif architecture—STAT5 plus cooperating activators (HNF4A) and/or sex-biased repressors
796 (BCL6, CUX2)—determines which accessible chromatin regions convert GH input signals into functional,
797 hormone-responsive enhancer output.
798
799 Integration of functional enhancer datasets with sex-biased genes revealed mechanistic links between sex-biased,
800 GH-regulated enhancers and sex-differential MASLD susceptibility. Of 840 functional enhancer DARs showing sex-
801 biased or cGH-responsive enhancer activity, 168 were male-biased, cGH-repressed DARs linked to
802 correspondingly regulated male-biased genes and 137 were female-biased, cGH-activated DARs linked to
803 correspondingly regulated female-biased genes. Critically, a subset of the male-biased enhancer DARs linked to
804 known MASLD-enabling genes, including *Bcl6* (male-biased transcriptional repressor that enhances MASH
805 progression by repressing PPARA-dependent hepatoprotective fatty acid oxidation genes (5,27,69)), *Nox4*
806 (NADPH oxidase that prevents oxidative damage in early stages of disease (70) but become maladaptive and
807 mediates oxidative stress, inflammation, and fibrosis as disease advances (71)), and the lncRNA lipid metabolism
808 regulator *Lnc-Lfar1* (72). Similarly, female-biased enhancer DARs were linked to MASLD-protective genes, notably
809 *Hao2* (hydroxyacid oxidase promoting fatty acid oxidation linked to tumor suppression (73,74)), *Fmo2* (flavin
810 monooxygenase with anti-inflammatory roles that inhibits SREBP1 translocation, blocking steatosis (75)) and
811 *Trim24* (epigenetic regulator that represses steatosis, inflammation and fibrosis (76,77)). Notably, chromatin
812 state analysis showed that the male-biased enhancers shift from active enhancer and promoter states in male
813 liver to increasingly populate inactive, bivalent, or transcribed states in female liver, consistent with epigenetic
814 silencing of male-biased disease-promoting pathways in females. Conversely, female-biased enhancers tend to
815 shift from active states in female liver to more repressed states in male liver, consistent with loss of protective
816 epigenetic modifications in males. These findings provide a mechanistic framework for understanding the higher

817 MASLD prevalence in males and post-menopausal females (1) and suggest therapeutic strategies targeting GH
818 signaling or enhancer epigenetic states to modulate sex-biased metabolic disease risk. Moreover, identification of
819 functional enhancers linked to *Bcl6,* a male-biased, cGH-repressed regulator balancing immunity and metabolism
820 and whose disruption drives fatty liver specifically in males (5,27,69), suggests that targeting BCL6-regulated
821 enhancer networks may offer sex-specific therapeutic opportunities.
822
823 The HDI-STARR-seq platform circumvents limitations of cell culture-based massively parallel reporter assays
824 (MPRAs) and enables massively parallel functional assays of enhancer activity within hepatocytes *in situ*, where
825 episomal reporters apparently undergo chromatinization, recruit endogenous TFs and, as we establish here,
826 respond to plasma GH pattern-dependent hormonal signaling. Unlike traditional transgenic enhancer-reporter
827 assays, which are limited to single enhancers and require extensive mouse breeding, HDI-STARR-seq enabled the
828 parallel testing of 1,839 ATAC regions (23,912 tiled reporters) in individual mice, dramatically accelerating
829 enhancer discovery and characterization. The 11-day interval between HDI and liver tissue collection ensured
830 recovery from the initial stress associated with HDI (61,78), as verified by absence of stress gene induction, while
831 retaining robust plasmid expression and sufficient time for cGH-driven chromatin remodeling and epigenetic
832 reprogramming. Importantly, our finding that the robust active set of HDI-transfected regulated enhancers shows
833 strong, significant enrichment for genomic regions whose endogenous liver loci have active histone marks and
834 are in active enhancer chromatin states with depletion of repressive histone marks provides strong support for
835 the proposal that the transfected episomal DNA retains sufficient *cis*-regulatory sequence information to recruit
836 endogenous TFs and their associated transcriptional regulatory complexes and chromatin modifiers to recreate
837 native-like chromatin environments. Of note, a potential limitation of HDI-STARR-seq is the preferential
838 transfection of hepatocytes by HDI (78), which is supported by our finding of an up to 3.6-fold depletion of robust
839 active enhancers in NPC–specific ATAC regions. However, this likely has limited impact on our conclusions, given
840 that the large majority of sex-biased, cGH-regulated DARs and genes are almost exclusively limited to hepatocyte
841 clusters (4,40). Future adaptations of HDI-STARR-seq using cell-type–targeted viral delivery methods or lipid
842 nanoparticles may extend this method to NPCs.
843
844 Our findings also have important implications for understanding hormone-regulated gene expression and for
845 developing sex-specific therapeutic strategies. The principle that chromatin accessibility is necessary but
846 insufficient for enhancer function raises the possibility that genome-wide association studies identifying disease-
847 associated SNPs in accessible chromatin may substantially overestimate the fraction of variants that directly alter
848 enhancer function. Functional validation using HDI-STARR-seq or similar in vivo assays will be essential to
849 prioritize variants for mechanistic follow-up. Further, our discovery that specific TF motifs, such as HNF4A,
850 discriminate regulated from non-regulated enhancers identifies these factors as promising candidates for future
851 mechanistic investigation. Moreover, our finding that regulated female-biased enhancer DARs are specifically
852 enriched for repressive H3K27me3 marks in male liver suggests a mechanism for epigenetic suppression of
853 female-typical enhancer activity in males and the possibility that interventions blocking H3K27me3 deposition or
854 activating H3K27me3 demethylases at these enhancers in males could confer metabolic protection. Finally,
855 extending HDI-STARR-seq to widely studied liver disease models, including high-fat diet, methionine-choline–
856 deficient diet and genetic obesity models (79), may identify enhancers whose activity is dysregulated by
857 metabolic stress, revealing how environmental and genetic risk factors interact with sex-biased regulatory
858 programs.
859
860 Study limitations – We recognize several limitations. First, our 180-nt reporter design with 55 bp overlapping
861 reporters, while enabling fine mapping within ATAC regions, may fragment longer enhancers and miss
862 cooperative interactions between distal elements. Future studies may investigate the extent to which hepatocyte
863 enhancer activity is enhanced by extending sequence context using full-length ATAC regions (mean length, 634
864 bp) as single reporters. Second, by coupling all reporters to the liver-specific *Albumin* minimal promoter (41) we
865 may bias detection toward enhancers compatible with this promoter architecture; testing reporters with
866 alternative minimal liver promoters could reveal promoter-specific enhancer–promoter compatibility issues or
867 biases. Third, the multiOme peak-to-gene linkage analysis we employed predicts enhancer–gene relationships

868 based on correlations, which do not prove causality; definitive validation will require CRISPR-mediated deletion of
869 individual enhancers *in vivo* and measurement of target gene expression, e.g., using AAV-delivered Cas9 and
870 multiplexed gRNA libraries. Fourth, as HDI-STARR-seq interrogates enhancer activity on episomal plasmids, our
871 data do not establish necessity or sufficiency at native chromosomal loci. Finally, our finding that 37% of DNA-
872 qualified STARR-seq reporters showed no enhancer activity in any of three biological conditions suggests either
873 that many accessible chromatin regions are not enhancers, or that they require specific TFs, co-factors, metabolic
874 states or promoter architectures not represented under our experimental conditions; testing reporters under
875 diverse physiological perturbations (fasting, high-fat diet, xenobiotic exposures) may activate currently silent
876 enhancers and reveal context-dependent regulatory logic.
877
878 In summary, this study establishes HDI-STARR-seq as a powerful *in vivo* platform for systematically validating
879 hormone-regulated enhancers and defines a functionally validated enhancer atlas linking GH-regulated
880 chromatin accessibility to hepatocyte expression of many sex-biased genes, including genes impacting MASLD
881 susceptibility. Our findings that only a minority of sex-biased DARs function as regulated enhancers, that
882 functional enhancers are distinguished by combinatorial epigenetic and TF codes, and that these enhancers are
883 linked to disease-relevant metabolic pathways, provide a mechanistic framework for understanding how
884 hormonal cues—encoded in pulsatile versus continuous pituitary GH secretion patterns—are decoded by the
885 genome into sex-specific transcriptional programs and sex-differential disease outcomes. Together, these insights
886 establish a foundation for future studies aimed at elucidating genetic and environmental perturbations that
887 disrupt sex-biased enhancers and uncover novel therapeutic targets within enhancer-regulated pathways
888 contributing to sex differences in liver metabolism and disease susceptibility.
889
890 **Acknowledgments:** The authors thank Drs. Maxim Pyatkov and Kritika Karri for advice and collaborative
891 assistance and Drs. Christine Goldfarb and Ravi Sonkar for assistance with tail vein injections.
892
893 **Data availability:** All Raw fastq sequencing files are available at GEO (https://www.ncbi.nlm.nih.gov/geo/) under
894 accession number GSE315771. Other data generated and analyzed during this study are included in this published
895 article or in the data repositories listed in References.
896
897 **Author contributions:** This study was jointly designed by TYC and DJW. TYC carried out all wet lab work, including
898 HDI-STARR-seq library preparation, HDI-STARR-seq mouse studies, and Illumina sequencing library preparation,
899 validation and initial quality control analysis. TYC carried out primary analysis of the HDI-SATARR-seq data, and
900 TYC and DJW jointly performed secondary data analysis, figure preparation and preparation of Supplementary
901 tables and other Supplementary files for publication. TYC wrote a first draft of the manuscript, which was revised,
902 edited and finalized by DJW. DJW supervised the overall project.
903
904
905 **References**
906
907 1. Kineman RD, Del Rio-Moreno M, Waxman DJ. Liver-specific actions of GH and IGF1 that protect against
908 MASLD. Nat Rev Endocrinol. 2025;21(2):105-117.
909 2. Brie B, Ramirez MC, De Winne C, Lopez Vicchi F, Villarruel L, Sorianello E, Catalano P, Ornstein AM, Becu-
910 Villalobos D. Brain Control of Sexually Dimorphic Liver Function and Disease: The Endocrine Connection.
911 Cell Mol Neurobiol. 2019;39(2):169-180.
912 3. Matz-Soja M, Berg T, Kietzmann T. Sex-related variations in liver homeostasis and disease: From zonation
913 dynamics to clinical implications. J Hepatol. 2026;84(1):181-193.
914 4. Goldfarb CN, Karri K, Pyatkov M, Waxman DJ. Interplay Between GH-regulated, Sex-biased Liver
915 Transcriptome and Hepatic Zonation Revealed by Single-Nucleus RNA Sequencing. Endocrinology.
916 2022;163(7).
917 5. Waxman DJ, Kineman RD. Sex matters in liver fat regulation. Science. 2022;378(6617):252-253.

918 6. Waxman DJ, Holloway MG. Sex differences in the expression of hepatic drug metabolizing enzymes. Mol
919 Pharmacol. 2009;76(2):215-228.
920 7. Berryman DE, List EO. Growth Hormone's Effect on Adipose Tissue: Quality versus Quantity. Int J Mol Sci.
921 2017;18(8).
922 8. Mueller KM, Kornfeld JW, Friedbichler K, Blaas L, Egger G, Esterbauer H, Hasselblatt P, Schlederer M,
923 Haindl S, Wagner KU, Engblom D, Haemmerle G, Kratky D, Sexl V, Kenner L, Kozlov AV, Terracciano L,
924 Zechner R, Schuetz G, Casanova E, Pospisilik JA, Heim MH, Moriggl R. Impairment of hepatic growth
925 hormone and glucocorticoid receptor signaling causes steatosis and hepatocellular carcinoma in mice.
926 Hepatology. 2011;54(4):1398-1409.
927 9. Clodfelter KH, Holloway MG, Hodor P, Park SH, Ray WJ, Waxman DJ. Sex-dependent liver gene expression
928 is extensive and largely dependent upon signal transducer and activator of transcription 5b (STAT5b):
929 STAT5b-dependent activation of male genes and repression of female genes revealed by microarray
930 analysis. Mol Endocrinol. 2006;20(6):1333-1351.
931 10. Baik M, Yu JH, Hennighausen L. Growth hormone-STAT5 regulation of growth, hepatocellular carcinoma,
932 and liver metabolism. Ann N Y Acad Sci. 2011;1229:29-37.
933 11. Waters MJ, Brooks AJ. Growth hormone and cell growth. Endocr Dev. 2012;23:86-95.
934 12. Waters MJ. The growth hormone receptor. Growth Horm IGF Res. 2016;28:6-10.
935 13. Zhang Y, Laz EV, Waxman DJ. Dynamic, sex-differential STAT5 and BCL6 binding to sex-biased, growth
936 hormone-regulated genes in adult mouse liver. Mol Cell Biol. 2012;32(4):880-896.
937 14. Rotwein P. Mapping the growth hormone--Stat5b--IGF-I transcriptional circuit. Trends Endocrinol Metab.
938 2012;23(4):186-193.
939 15. Yu JH, Zhu BM, Wickre M, Riedlinger G, Chen W, Hosui A, Robinson GW, Hennighausen L. The
940 transcription factors signal transducer and activator of transcription 5A (STAT5A) and STAT5B negatively
941 regulate cell proliferation through the activation of cyclin-dependent kinase inhibitor 2b (Cdkn2b) and
942 Cdkn1a expression. Hepatology. 2010;52(5):1808-1818.
943 16. Kaltenecker D, Themanns M, Mueller KM, Spirk K, Suske T, Merkel O, Kenner L, Luis A, Kozlov A, Haybaeck
944 J, Muller M, Han X, Moriggl R. Hepatic growth hormone - JAK2 - STAT5 signalling: Metabolic function,
945 non-alcoholic fatty liver disease and hepatocellular carcinoma progression. Cytokine. 2019;124:154569.
946 17. Huang Z, Huang L, Waters MJ, Chen C. Insulin and Growth Hormone Balance: Implications for Obesity.
947 Trends Endocrinol Metab. 2020;31(9):642-654.
948 18. Al-Samerria S, Radovick S. Exploring the Therapeutic Potential of Targeting GH and IGF-1 in the
949 Management of Obesity: Insights from the Interplay between These Hormones and Metabolism. Int J Mol
950 Sci. 2023;24(11).
951 19. Abu El-Makarem MA, Kamel MF, Mohamed AA, Ali HA, Mohamed MR, Mohamed AEM, El-Said AM,
952 Ameen MG, Hassnine AA, Hassan HA. Down-regulation of hepatic expression of GHR/STAT5/IGF-1
953 signaling pathway fosters development and aggressiveness of HCV-related hepatocellular carcinoma:
954 Crosstalk with Snail-1 and type 2 transforming growth factor-beta receptor. PLoS One.
955 2022;17(11):e0277266.
956 20. Choi HK, Waxman DJ. Growth hormone, but not prolactin, maintains, low-level activation of STAT5a and
957 STAT5b in female rat liver. Endocrinology. 1999;140(11):5126-5135.
958 21. Hao P, Waxman DJ. STAT5 Regulation of Sex-Dependent Hepatic CpG Methylation at Distal Regulatory
959 Elements Mapping to Sex-Biased Genes. Mol Cell Biol. 2021;41(2).
960 22. Holloway MG, Cui Y, Laz EV, Hosui A, Hennighausen L, Waxman DJ. Loss of sexually dimorphic liver gene
961 expression upon hepatocyte-specific deletion of Stat5a-Stat5b locus. Endocrinology. 2007;148(5):1977-
962 1986.
963 23. Park SH, Wiwi CA, Waxman DJ. Signalling cross-talk between hepatocyte nuclear factor 4alpha and
964 growth-hormone-activated STAT5b. Biochem J. 2006;397(1):159-168.
965 24. Wiwi CA, Waxman DJ. Role of hepatocyte nuclear factors in transcriptional regulation of male-specific
966 CYP2A2. J Biol Chem. 2005;280(5):3259-3268.

967 25. Holloway MG, Laz EV, Waxman DJ. Codependence of growth hormone-responsive, sexually dimorphic
968 hepatic gene expression on signal transducer and activator of transcription 5b and hepatic nuclear factor
969 4alpha. Mol Endocrinol. 2006;20(3):647-660.
970 26. Meyer RD, Laz EV, Su T, Waxman DJ. Male-specific hepatic Bcl6: growth hormone-induced block of
971 transcription elongation in females and binding to target genes inversely coordinated with STAT5. Mol
972 Endocrinol. 2009;23(11):1914-1926.
973 27. Nikkanen J, Leong YA, Krause WC, Dermadi D, Maschek JA, Van Ry T, Cox JE, Weiss EJ, Gokcumen O,
974 Chawla A, Ingraham HA. An evolutionary trade-off between host immunity and metabolism drives fatty
975 liver in male mice. Science. 2022;378(6617):290-295.
976 28. Laz EV, Holloway MG, Chen CS, Waxman DJ. Characterization of three growth hormone-responsive
977 transcription factors preferentially expressed in adult female liver. Endocrinology. 2007;148(7):3327-
978 3337.
979 29. Conforto TL, Steinhardt GFt, Waxman DJ. Cross Talk Between GH-Regulated Transcription Factors HNF6
980 and CUX2 in Adult Mouse Liver. Mol Endocrinol. 2015;29(9):1286-1302.
981 30. Conforto TL, Zhang Y, Sherman J, Waxman DJ. Impact of CUX2 on the female mouse liver transcriptome:
982 activation of female-biased genes and repression of male-biased genes. Mol Cell Biol. 2012;32(22):4611-
983 4627.
984 31. Ling G, Sugathan A, Mazor T, Fraenkel E, Waxman DJ. Unbiased, genome-wide in vivo mapping of
985 transcriptional regulatory elements reveals sex differences in chromatin structure associated with sex-
986 specific liver gene expression. Mol Cell Biol. 2010;30(23):5531-5544.
987 32. Sugathan A, Waxman DJ. Genome-wide analysis of chromatin states reveals distinct mechanisms of sex-
988 dependent gene regulation in male and female mouse liver. Mol Cell Biol. 2013;33(18):3594-3610.
989 33. Rampersaud A, Connerney J, Waxman DJ. Plasma growth hormone pulses induce male-biased pulsatile
990 chromatin opening and epigenetic regulation in adult mouse liver. Elife. 2023;12.
991 34. Matthews BJ, Melia T, Waxman DJ. Harnessing natural variation to identify cis regulators of sex-biased
992 gene expression in a multi-strain mouse liver model. PLoS Genet. 2021;17(11):e1009588.
993 35. Matthews BJ, Waxman DJ. Impact of 3D genome organization, guided by cohesin and CTCF looping, on
994 sex-biased chromatin interactions and gene expression in mouse liver. Epigenetics Chromatin.
995 2020;13(1):30.
996 36. Cunningham RP, Porat-Shliom N. Liver Zonation - Revisiting Old Questions With New Technologies. Front
997 Physiol. 2021;12:732929.
998 37. Hildebrandt F, Andersson A, Saarenpaa S, Larsson L, Van Hul N, Kanatani S, Masek J, Ellis E, Barragan A,
999 Mollbrink A, Andersson ER, Lundeberg J, Ankarklev J. Spatial Transcriptomics to define transcriptional
1000 patterns of zonation and structural components in the mouse liver. Nat Commun. 2021;12(1):7046.
1001 38. Xiong X, Kuang H, Liu T, Lin JD. A Single-Cell Perspective of the Mammalian Liver in Health and Disease.
1002 Hepatology. 2020;71(4):1467-1473.
1003 39. Ben-Moshe S, Itzkovitz S. Spatial heterogeneity in the mammalian liver. Nat Rev Gastroenterol Hepatol.
1004 2019;16(7):395-410.
1005 40. Karri K, Chang TY, Pyatkov M, Gandhi S, Siggers T, Waxman DJ. Single Nucleus MultiOmics Links Novel
1006 Transcription Factor Motifs to Murine Hepatic Sex Differences in Chromatin Accessibility and Metabolic
1007 Dysfunction-Associated Steatotic Liver Disease. bioRxiv. 2025.
1008 41. Chang TY, Waxman DJ. HDI-STARR-seq: Condition-specific enhancer discovery in mouse liver in vivo. BMC
1009 Genomics. 2024;25(1):1240.
1010 42. Percie du Sert N, Hurst V, Ahluwalia A, Alam S, Avey MT, Baker M, Browne WJ, Clark A, Cuthill IC, Dirnagl
1011 U, Emerson M, Garner P, Holgate ST, Howells DW, Karp NA, Lazic SE, Lidster K, MacCallum CJ, Macleod M,
1012 Pearl EJ, Petersen OH, Rawle F, Reynolds P, Rooney K, Sena ES, Silberberg SD, Steckler T, Wurbel H. The
1013 ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. PLoS Biol. 2020;18(7):e3000410.
1014 43. Lau-Corona D, Suvorov A, Waxman DJ. Feminization of Male Mouse Liver by Persistent Growth Hormone
1015 Stimulation: Activation of Sex-Biased Transcriptional Networks and Dynamic Changes in Chromatin
1016 States. Mol Cell Biol. 2017;37(19).

1017 44. Zhang R, Lahens NF, Ballance HI, Hughes ME, Hogenesch JB. A circadian gene expression atlas in
1018 mammals: implications for biology and medicine. Proc Natl Acad Sci U S A. 2014;111(45):16219-16224.
1019 45. Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. Single-cell chromatin state analysis with Signac. Nat
1020 Methods. 2021;18(11):1333-1341.
1021 46. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, Hao Y, Stoeckius M, Smibert P,
1022 Satija R. Comprehensive Integration of Single-Cell Data. Cell. 2019;177(7):1888-1902 e1821.
1023 47. Karri K, Waxman DJ. Dysregulation of murine long noncoding single-cell transcriptome in nonalcoholic
1024 steatohepatitis and liver fibrosis. RNA. 2023;29(7):977-1006.
1025 48. Neumayr C, Pagani M, Stark A, Arnold CD. STARR-seq and UMI-STARR-seq: Assessing Enhancer Activities
1026 for Genome-Wide-, High-, and Low-Complexity Candidate Libraries. Curr Protoc Mol Biol.
1027 2019;128(1):e105.
1028 49. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA
1029 sequences to the human genome. Genome Biol. 2009;10(3):R25.
1030 50. Mouse EC, Stamatoyannopoulos JA, Snyder M, Hardison R, Ren B, Gingeras T, Gilbert DM, Groudine M,
1031 Bender M, Kaul R, Canfield T, Giste E, Johnson A, Zhang M, Balasundaram G, Byron R, Roach V, Sabo PJ,
1032 Sandstrom R, Stehling AS, Thurman RE, Weissman SM, Cayting P, Hariharan M, Lian J, Cheng Y, Landt SG,
1033 Ma Z, Wold BJ, Dekker J, Crawford GE, Keller CA, Wu W, Morrissey C, Kumar SA, Mishra T, Jain D, Byrska-
1034 Bishop M, Blankenberg D, Lajoie BR, Jain G, Sanyal A, Chen KB, Denas O, Taylor J, Blobel GA, Weiss MJ,
1035 Pimkin M, Deng W, Marinov GK, Williams BA, Fisher-Aylor KI, Desalvo G, Kiralusha A, Trout D, Amrhein H,
1036 Mortazavi A, Edsall L, McCleary D, Kuan S, Shen Y, Yue F, Ye Z, Davis CA, Zaleski C, Jha S, Xue C, Dobin A,
1037 Lin W, Fastuca M, Wang H, Guigo R, Djebali S, Lagarde J, Ryba T, Sasaki T, Malladi VS, Cline MS, Kirkup
1038 VM, Learned K, Rosenbloom KR, Kent WJ, Feingold EA, Good PJ, Pazin M, Lowdon RF, Adams LB. An
1039 encyclopedia of mouse DNA elements (Mouse ENCODE). Genome Biol. 2012;13(8):418.
1040 51. Benjamini Y, Speed TP. Summarizing and correcting the GC content bias in high-throughput sequencing.
1041 Nucleic Acids Res. 2012;40(10):e72.
1042 52. Wilkinson GN, Rogers CE. Symbolic Description of Factorial Models for Analysis of Variance. The Royal
1043 Statistical Society Series C-Applied Statistics. 1973;22(3):392-399.
1044 53. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. Nat Protoc.
1045 2017;12(12):2478-2492.
1046 54. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics.
1047 2011;27(12):1696-1697.
1048 55. Newburger DE, Bulyk ML. UniPROBE: an online database of protein binding microarray data on protein-
1049 DNA interactions. Nucleic Acids Res. 2009;37(Database issue):D77-82.
1050 56. Rauluseviciute I, Riudavets-Puig R, Blanc-Mathieu R, Castro-Mondragon JA, Ferenc K, Kumar V, Lemma
1051 RB, Lucas J, Cheneby J, Baranasic D, Khan A, Fornes O, Gundersen S, Johansen M, Hovig E, Lenhard B,
1052 Sandelin A, Wasserman WW, Parcy F, Mathelier A. JASPAR 2024: 20th anniversary of the open-access
1053 database of transcription factor binding profiles. Nucleic Acids Res. 2024;52(D1):D174-D182.
1054 57. Vorontsov IE, Eliseeva IA, Zinkevich A, Nikonov M, Abramov S, Boytsov A, Kamenets V, Kasianova A,
1055 Kolmykov S, Yevshin IS, Favorov A, Medvedeva YA, Jolma A, Kolpakov F, Makeev VJ, Kulakovskiy IV.
1056 HOCOMOCO in 2024: a rebuild of the curated collection of binding models for human and mouse
1057 transcription factors. Nucleic Acids Res. 2024;52(D1):D154-D163.
1058 58. Matthews BJ, Waxman DJ. Computational prediction of CTCF/cohesin-based intra-TAD loops that insulate
1059 chromatin contacts and gene expression in mouse liver. Elife. 2018;7.
1060 59. Lau-Corona D, Ma H, Vergato C, Sarmento-Cabral A, Del Rio-Moreno M, Kineman RD, Waxman DJ.
1061 Constitutively Active STAT5b Feminizes Mouse Liver Gene Expression. Endocrinology. 2022;163(5).
1062 60. Lau-Corona D, Bae WK, Hennighausen L, Waxman DJ. Sex-biased genetic programs in liver metabolism
1063 and liver fibrosis are controlled by EZH1 and EZH2. PLoS Genet. 2020;16(5):e1008796.
1064 61. Budker VG, Subbotin VM, Budker T, Sebestyen MG, Zhang G, Wolff JA. Mechanism of plasmid delivery by
1065 hydrodynamic tail vein injection. II. Morphological studies. J Gene Med. 2006;8(7):874-888.
1066 62. Riu E, Chen ZY, Xu H, He CY, Kay MA. Histone modifications are associated with the persistence or
1067 silencing of vector-mediated transgene expression in vivo. Mol Ther. 2007;15(7):1348-1355.

1068 63. Zhang C, Liu D. Transcription Factor Binding Site in Promoter Determines the Pattern of Plasmid-Based
1069 Transgene Expression In Vivo. Pharmaceutics. 2024;16(4).
1070 64. Georgakopoulos-Soares I, Deng C, Agarwal V, Chan CSY, Zhao J, Inoue F, Ahituv N. Transcription factor
1071 binding site orientation and order are major drivers of gene regulatory activity. Nat Commun.
1072 2023;14(1):2333.
1073 65. Agarwal V, Inoue F, Schubach M, Penzar D, Martin BK, Dash PM, Keukeleire P, Zhang Z, Sohota A, Zhao J,
1074 Georgakopoulos-Soares I, Noble WS, Yardimci GG, Kulakovskiy IV, Kircher M, Shendure J, Ahituv N.
1075 Massively parallel characterization of transcriptional regulatory elements. Nature. 2025;639(8054):411-
1076 420.
1077 66. Zhang J, Leung AK, Zhu Y, Yao L, Willis A, Pan X, Ozer A, Zhou Z, Siklenka K, Barrera A, Liang J, Tippens ND,
1078 Reddy TE, Lis JT, Yu H. Comprehensive evaluation of diverse massively parallel reporter assays to
1079 functionally characterize human enhancers genome-wide. Genome Biol. 2025;26(1):378.
1080 67. Rampersaud A, Lodato NJ, Shin A, Waxman DJ. Widespread Epigenetic Changes to the Enhancer
1081 Landscape of Mouse Liver Induced by a Specific Xenobiotic Agonist Ligand of the Nuclear Receptor CAR.
1082 Toxicol Sci. 2019;171(2):315-338.
1083 68. Coronel L, Riege K, Schwab K, Forste S, Hackes D, Semerau L, Bernhart SH, Siebert R, Hoffmann S, Fischer
1084 M. Transcription factor RFX7 governs a tumor suppressor network in response to p53 and stress. Nucleic
1085 Acids Res. 2021;49(13):7437-7456.
1086 69. Sommars MA, Ramachandran K, Senagolage MD, Futtner CR, Germain DM, Allred AL, Omura Y, Bederman
1087 IR, Barish GD. Dynamic repression by BCL6 controls the genome-wide liver response to fasting and
1088 steatosis. Elife. 2019;8.
1089 70. Greatorex S, Kaur S, Xirouchaki CE, Goh PK, Wiede F, Genders AJ, Tran M, Jia Y, Raajendiran A, Brown WA,
1090 McLean CA, Sadoshima J, Watt MJ, Tiganis T. Mitochondria- and NOX4-dependent antioxidant defense
1091 mitigates progression to nonalcoholic steatohepatitis in obesity. J Clin Invest. 2023;134(3).
1092 71. Bettaieb A, Jiang JX, Sasaki Y, Chao TI, Kiss Z, Chen X, Tian J, Katsuyama M, Yabe-Nishimura C, Xi Y,
1093 Szyndralewiez C, Schroder K, Shah A, Brandes RP, Haj FG, Torok NJ. Hepatocyte Nicotinamide Adenine
1094 Dinucleotide Phosphate Reduced Oxidase 4 Regulates Stress Signaling, Fibrosis, and Insulin Sensitivity
1095 During Development of Steatohepatitis in Mice. Gastroenterology. 2015;149(2):468-480 e410.
1096 72. Zhang K, Han X, Zhang Z, Zheng L, Hu Z, Yao Q, Cui H, Shu G, Si M, Li C, Shi Z, Chen T, Han Y, Chang Y, Yao
1097 Z, Han T, Hong W. The liver-enriched lnc-LFAR1 promotes liver fibrosis by activating TGFbeta and Notch
1098 pathways. Nat Commun. 2017;8(1):144.
1099 73. Mattu S, Fornari F, Quagliata L, Perra A, Angioni MM, Petrelli A, Menegon S, Morandi A, Chiarugi P,
1100 Ledda-Columbano GM, Gramantieri L, Terracciano L, Giordano S, Columbano A. The metabolic gene
1101 HAO2 is downregulated in hepatocellular carcinoma and predicts metastasis and poor survival. J Hepatol.
1102 2016;64(4):891-898.
1103 74. Li Y, Zhang M, Li X, Wang Y, Wang Y, Li Y, Zhu H, Ding H, Li X, Sun S. Hydroxyacid Oxidase 2 (HAO2) Inhibits
1104 the Tumorigenicity of Hepatocellular Carcinoma and Is Negatively Regulated by miR-615-5p. J Immunol
1105 Res. 2022;2022:5003930.
1106 75. Ke C, Xiao C, Li J, Wu X, Zhang Y, Chen Y, Sheng S, Fu Z, Wang L, Ni C, Zhao J, Shi Y, Wu Y, Zhong Z, Nan J,
1107 Zhu W, Chen J, Wu R, Hu X. FMO2 ameliorates nonalcoholic fatty liver disease by suppressing ER-to-Golgi
1108 transport of SREBP1. Hepatology. 2025;81(1):181-197.
1109 76. Jiang S, Minter LC, Stratton SA, Yang P, Abbas HA, Akdemir ZC, Pant V, Post S, Gagea M, Lee RG, Lozano G,
1110 Barton MC. TRIM24 suppresses development of spontaneous hepatic lipid accumulation and
1111 hepatocellular carcinoma in mice. J Hepatol. 2015;62(2):371-379.
1112 77. Zhang H, Lei S, Zhuo H, Xu Y, Ye Y, Luo Y. TRIM24 Up-Regulates ORM2 to Alleviate Abnormal Lipid
1113 Metabolism, Inflammation, and Oxidative Stress in Mice with Obstructive Sleep Apnea Syndrome and
1114 Metabolic Dysfunction-Associated Steatotic Liver Disease. Am J Pathol. 2024;194(11):2091-2105.
1115 78. Suda T, Gao X, Stolz DB, Liu D. Structural impact of hydrodynamic injection on mouse liver. Gene Ther.
1116 2007;14(2):129-137.
1117 79. Basaranoglu M, Kayacetin S, Yilmaz N, Kayacetin E, Tarcin O, Sonsuz A. Understanding mechanisms of the
1118 pathogenesis of nonalcoholic fatty liver disease. World J Gastroenterol. 2010;16(18):2223-2226.

1119
1120 **Figure legends**
1121
1122 **Fig 1. Joint transcriptional and chromatin accessibility profiling of mouse liver. (A)** UMAP of 46,188 mouse liver
1123 nuclei based on 10X Genomic MultiOmic joint transcriptomic + chromatin accessibility profiles harmonized across
1124 three biological conditions: male liver, female liver and cGH-treated male liver. Cell types were identified using
1125 established marker genes. **(B)** snATAC-seq linkage analysis for female-biased, cGH-activated ATAC_39124 (red
1126 arrow, mouse mm10 genomic coordinates indicated), revealing inferred short and long distance positive
1127 regulatory linkages (blue arcs) to female-biased genes *Cux2* and *lnc4593*, and negative regulatory linkages (red
1128 arcs) to male-biased gene *P2rx4*. Region within yellow box is expanded in panel C. **(C)** Closer view of ATAC_39124
1129 region with linkages (*top*) and normalized ATAC peaks for each biological condition (*bottom*). Select sex-biased
1130 and cGH-responsive ATAC peaks are numbered, including negative regulation to male-biased, cGH-repressed
1131 ATAC_39115. **(D)** Bar graphs showing ATAC activity (cut sites/cell) for region ATAC_39124 (*top*), and RNA
1132 expression levels for the 4 indicated genes in female, male and in cGH-infused male liver (*bottom*), whose
1133 linkages to ATAC_39124 are all significant at -log10(FDR) > 5 (Table S4B).
1134
1135 **Fig 2. Synthetic STARR-seq reporter library. (A)** Sex-biased and/or cGH-regulated DARs selected for library (red
1136 box) based on the sex bias (SB) and cGH-responsiveness of their ATAC activity and that of their linked genes
1137 (Table S2). DARs selected were derived from the full sets of ATAC-seq peaks indicated at top. **(B)** Number of
1138 active DARs from each set in panel A with at least 1 robust active reporter. Numbers in parentheses indicate the
1139 number of robust active reporters from the full set of 4,126 robust active reporters. For example, 250 of the 360
1140 male-biased, cGH-repressed DARs include a total of 725 robust active reporters from the full set of 4,126. **(C)**
1141 Design of synthetic STARR-seq library reporters based on liver ATAC region plus 70 nt of flanking sequences. *Left*,
1142 Schematic showing a female-biased ATAC region activated by cGH infusion and a series of tiled reporters. *Middle,*
1143 Each reporter is comprised of a genomic sequence flanked by 15 nt of adaptor sequences (olive, blue).
1144 Neighboring reporters are shifted by 55 bp and overlap by 125 bp. *Right*, adaptor sequences for PCR, where the
1145 primers anneal and incorporate homology arms for the HiFi cloning step. PCR amplification is followed by HiFi
1146 cloning to attach homology arms to single stranded DNA oligos (see Methods).
1147
1148 **Fig. 3. In vivo evaluation of HDI-STARR-seq library transfection and cGH response. (A)** Validation of reporter
1149 plasmid transfection based on Renilla luciferase reporter activity in each of 12 HDI-transfected mouse livers. Five
1150 livers with lower luciferase activity (asterisks) were excluded from downstream analyses. **(B)** Clustering of
1151 normalized mapped read counts in ATAC regions from biological replicates for 2 plasmid libraries, 3 DNA libraries
1152 and 7 RNA (enhancer activity) reporter libraries, based on Pearson's correlation. DNA and RNA libraries were
1153 extracted from HDI livers for each biological condition: M, male; F, female; and cGH-treated male. **(C)** Impact of
1154 cGH infusion on RNA levels of the indicated sex-biased genes. Significance: Kolmogorov-Smirnov test, followed by
1155 multiple comparison with FDR for multiple comparison correction. Data are mean +/- SEM (n=6 livers/group),
1156 with male livers (vehicle pump) mean value = 1. **(D)** UCSC Browser screen shots with examples of sex-biased,
1157 cGH-regulated ATAC regions from Pick1 set (*top*) and Pick2 set (*bottom*) whose HDI-STARR-seq enhancer activity
1158 shows matched regulation for cGH response (Repressed, green; Induced, red) and sex (male-biased, blue;
1159 Female-biased, light red) across multiple concordant reporters (*lower two tracks*). Boundaries of the ATAC region
1160 identified in the current study and of the overlapping sex-biased DNase-hypersensitive sites (DHS) identified in
1161 bulk liver ('72K DHS set') (31) are shown. Female-biased ATAC regions shown within *Cux2* intron are detailed in
1162 Fig. 1C.
1163
1164 **Fig. 4. Relationship between sex bias and cGH-responsiveness of HDI-STARR-seq enhancer activity. (A)** Venn
1165 diagram showing overlaps between the 1,397 HDI-STARR-seq reporters whose enhancer activity is sex-biased
1166 and/or cGH-regulated (Table S7A, columns AW-BB). These reporters map to 840 DARs, including 229 male-biased
1167 (MB) DARs repressed by cGH and 161 female-biased (FB) DARs activate by cGH, whose distributions of gene
1168 linkages to sex-biased and other genes are indicated in the table. **(B, C)** Correlation between sex bias of reporter
1169 activity and response of reporter activity to cGH infusion in male liver for the set of n=725 robust active reporters

1170 (subset of 4,126 such reporters; Table S7A) that map to male-biased, cGH-repressed DARs (panel B, Pick1 DARs)
1171 and for n=693 robust active reporters that map to female-biased, cGH-activated DARs (panel C, Pick2 DARs). Blue
1172 and Red boxes, reporters whose enhancer activity is regulated by both sex and cGH. Red box: 41 male-biased and
1173 cGH-repressed reporters; red box, 43 female-biased and cGH-activated reporters. See Table S7B for other
1174 correlation data. Blue lines in panels B, C: best fit linear regression models, with the 95% confidence interval
1175 show in gray. **(D)** Mean DAR-level analysis of correlations, as in B and C, applied to 305 DARs: 168 male-biased
1176 and cGH-repressed DARs, and 137 female-biased and cGH-activated DAR. Data are mean values of log2 ratios
1177 calculated across all robust active reporters mapping to each DAR.
1178
1179 **Fig. 5. Epigenetic features enriched at regulated versus non-regulated enhancer DARs**. Enrichment scores (ES)
1180 for the overlap between the indicated sets of sex- and cGH-regulated and non-regulated enhancer DARs
1181 identified by HDI-STARR-seq and chromatin states (panel A) and histone marks (panel B) present in mouse liver at
1182 the corresponding genomic DNA sequences. Enrichments were calculated for four sets of robust active DAR
1183 enhancers: 1) 41 regulated male-biased, cGH-repressed DARs (*dark blue;* also see Fig. 4B, blue box); 2) 82 non-
1184 regulated male-biased, cGH-repressed DARs (*light blue*); 3) 43 regulated female-biased, cGH-activated DARs (*dark*
1185 *pink;* also see Fig. 4C, red box); and 4) 53 non-regulated female-biased, cGH-activated DARs (*light pink*).
1186 Enrichments were computed relative to a background set of 317 non-regulated genomic regions, which is a
1187 subset of 4,126 robust active reporters showing no regulated activity when comparing sexes or the response to
1188 cGH-infusion. **(A)** Enrichments for chromatin states E6 and E9 (both are enhancer-like states; see ChromHMM
1189 emission probabilities at *right*) and state E7, a promoter-like state, with the chromatin states being those
1190 determined in male mouse liver (top row) and those determined in female mouse liver (bottom row). **(B)**
1191 Histone-H3 marks identified in male (*top*) or female mouse liver (*bottom*) by ChIP-seq. Horizontal dashed line,
1192 ES=1. The significance of enrichment was determined by Fisher's exact test: *, (-log10P) > 3; **, (-log10P) > 5;
1193 ***, (-log10P) > 7, with red asterisks indicating significant depletion. Full datasets, including other chromatin
1194 states and other histone marks, are shown in Table S10A.
1195
1196 **Fig. 6. TFs enriched at regulated versus non-regulated enhancer DARs. (A)** Enrichment scores for the overlap
1197 between TF binding sites in mouse liver chromatin and the four indicated sets of sex- and cGH-responsive DARs,
1198 grouped by whether the ATAC sites displayed sex- and cGH-regulated enhancer activity by HDI-STARR-seq, or not
1199 ('non-regulated'). STAT5 ChIP-seq sites were either male-biased (n=1,765 sites) or female-biased (n=1,790 sites),
1200 as determined by ChIP-seq (13). BCL6 sites: n=6,432 ChIP-seq sites determined in male liver; and RNA-
1201 polymerase-II, global liver binding sites (n=31,987). Horizontal red line, ES=1 (no enrichment). The significance of
1202 enrichment was determined by Fisher's exact test: *, p<0.05; **, p<0.01; ***, p<0.001. **(B)** Motif scans for the
1203 indicated TF motifs (last 4 columns) list the percentage of each DAR set that contains the motif. Sex-biased DAR
1204 sets showing a significant enrichment of the TF motif in the regulated compared to the non-regulated sets are
1205 highlighted in green boxes, along with their enrichment scores (ES) and p-values (Fisher's exact test).
1206
1207
1208

Fig. 1

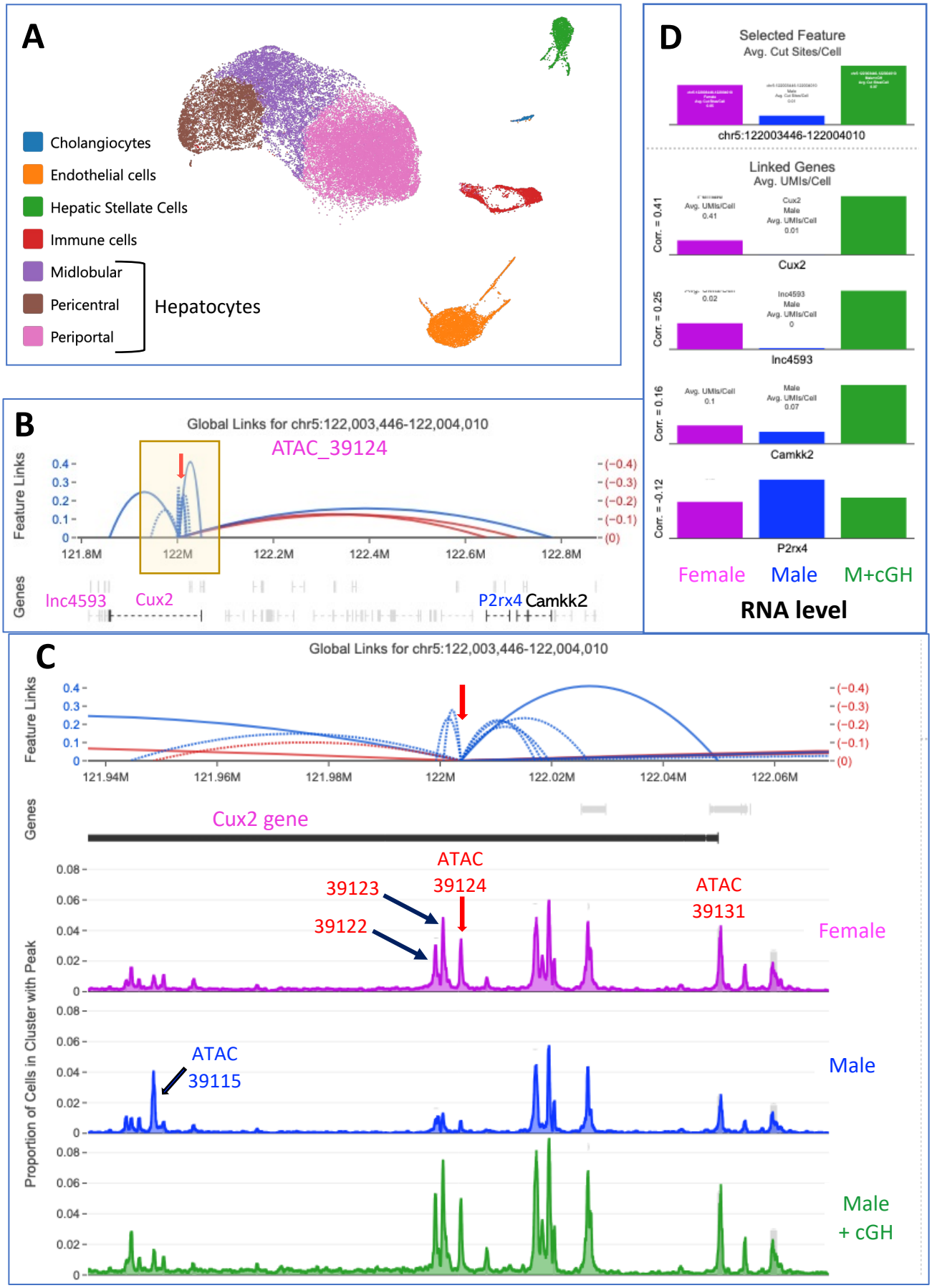

A
Cholangiocytes
Endothelial cells
Hepatic Stellate Cells
Immune cells
Midlobular
Pericentral
Periportal
Hepatocytes
B
Global Links for chr5:122,003,446-122,004,010
ATAC_39124
Feature Links
Genes
Inc4593
Cux2
P2rx4
Camkk2
121.8M
122M
122.2M
122.4M
122.6M
122.8M
C
Global Links for chr5:122,003,446-122,004,010
Feature Links
Genes
Cux2 gene
121.94M
121.96M
121.98M
122M
122.02M
122.04M
122.06M
Proportion of Cells in Cluster with Peak
ATAC
39123
39124
39122
ATAC
39131
Female
ATAC
39115
Male
Male
+ cGH
D
Selected Feature
Avg. Cut Sites/Cell
chr5:122003446-122004010
Linked Genes
Avg. UMIs/Cell
Corr. = 0.41
Cux2
Corr. = 0.25
Inc4593
Corr. = 0.16
Camkk2
Corr. = -0.12
P2rx4
Female
Male
M+cGH
RNA level

Fig. 2

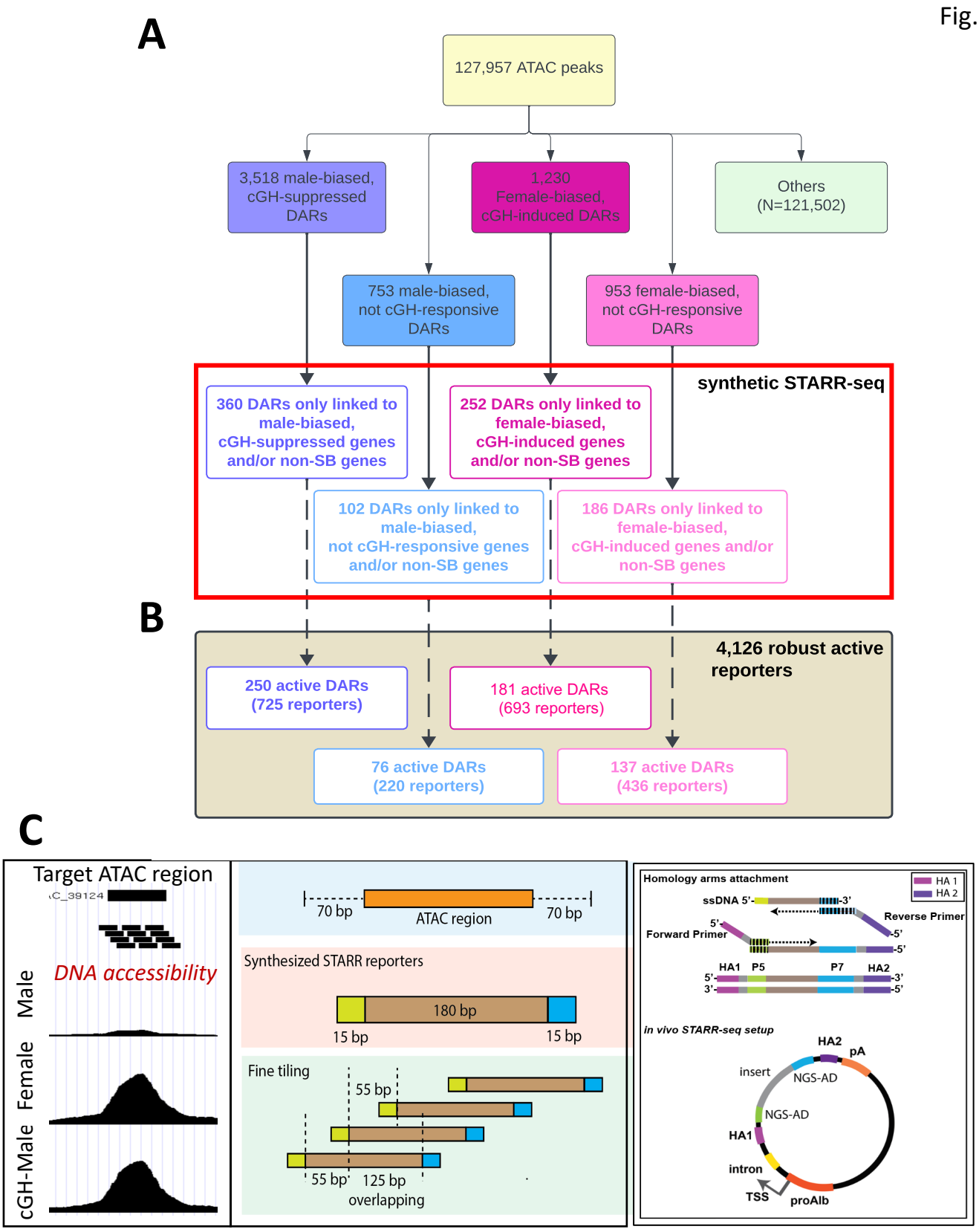

A
127,957 ATAC peaks
3,518 male-biased, cGH-suppressed DARs
1,230 Female-biased, cGH-induced DARs
Others (N=121,502)
753 male-biased, not cGH-responsive DARs
953 female-biased, not cGH-responsive DARs
synthetic STARR-seq
360 DARs only linked to male-biased, cGH-suppressed genes and/or non-SB genes
252 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
B
4,126 robust active reporters
250 active DARs (725 reporters)
181 active DARs (693 reporters)
76 active DARs (220 reporters)
137 active DARs (436 reporters)
C
Target ATAC region
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'
-5'
HA1
P5
P7
HA2
5'-
3'-
-3'
-5'
in vivo STARR-seq setup
HA2
pA
insert
NGS-AD
NGS-AD
HA1
intron
TSS
proAlb

Fig. 3

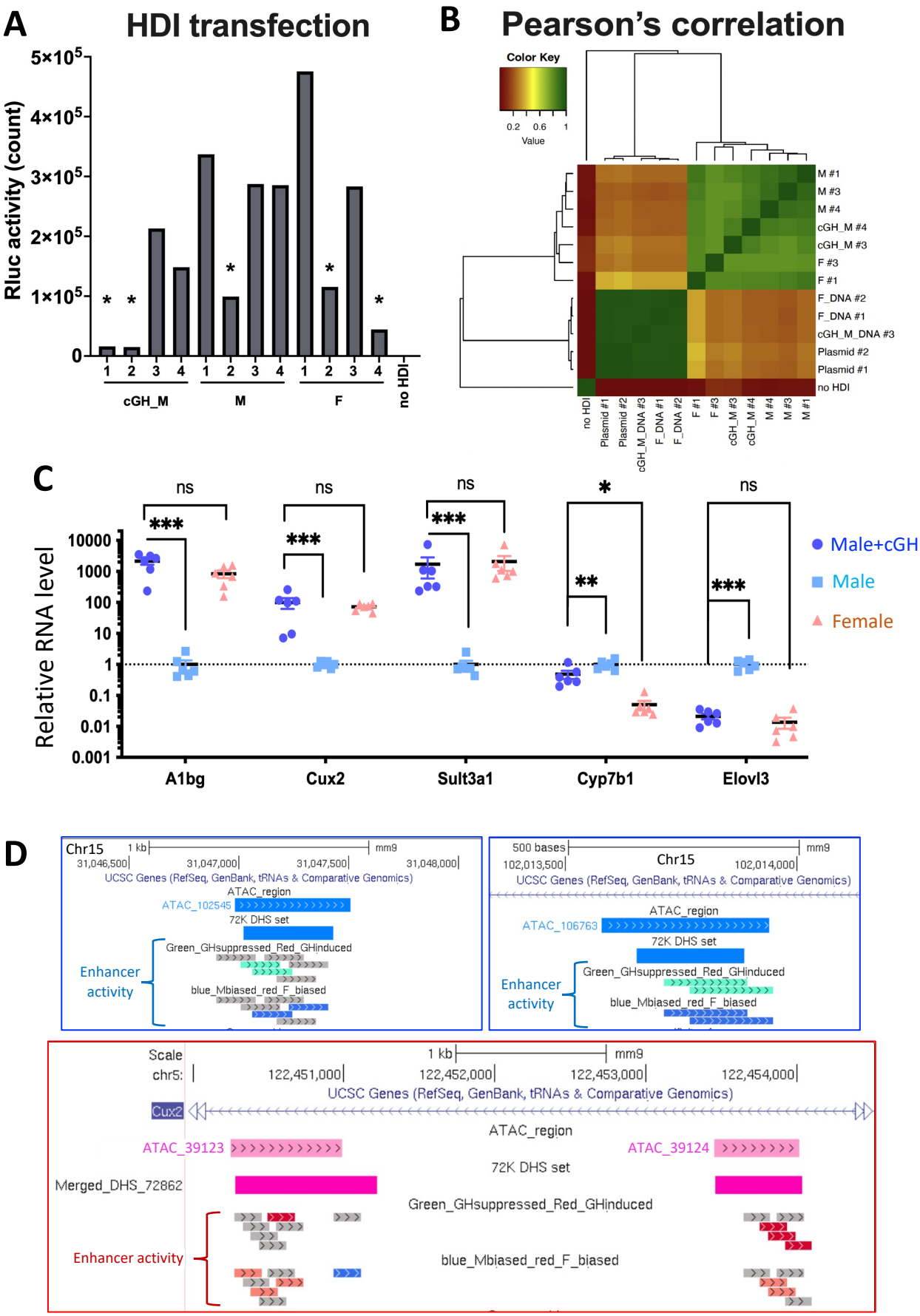
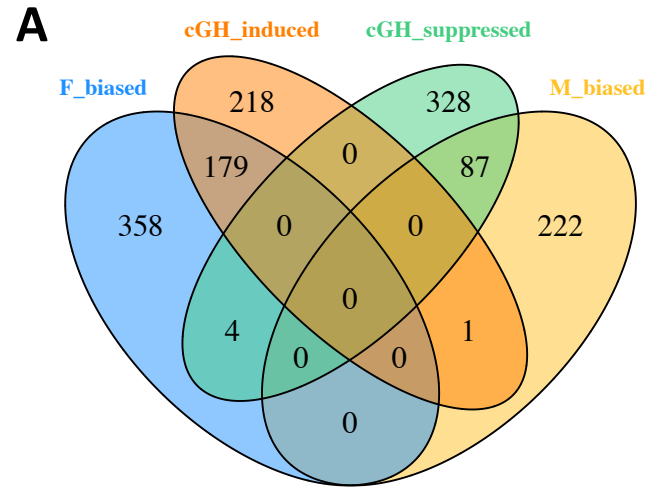

A
HDI transfection
Rluc activity (count)
5×10^5
4×10^5
3×10^5
2×10^5
1×10^5
0
cGH_M
M
F
no HDI
B
Pearson's correlation
Color Key
Value
M #1
M #3
M #4
cGH_M #4
cGH_M #3
F #3
F #1
F_DNA #2
F_DNA #1
cGH_M_DNA #3
Plasmid #2
Plasmid #1
no HDI
C
Relative RNA level
ns
***
**
*
Male+cGH
Male
Female
A1bg
Cux2
Sult3a1
Cyp7b1
Elovl3
D
Chr15
mm9
UCSC Genes (RefSeq, GenBank, tRNAs & Comparative Genomics)
ATAC_region
ATAC_102545
72K DHS set
Green_GHsuppressed_Red_GHinduced
blue_Mbiased_red_F_biased
Enhancer activity
ATAC_106763
Scale
chr5:
Cux2
ATAC_39123
ATAC_39124
Merged_DHS_72862

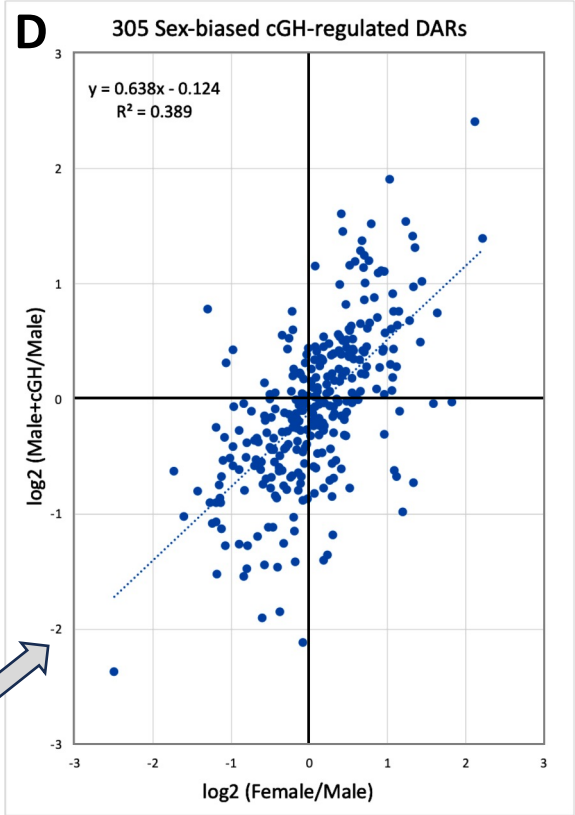
Fig. 4

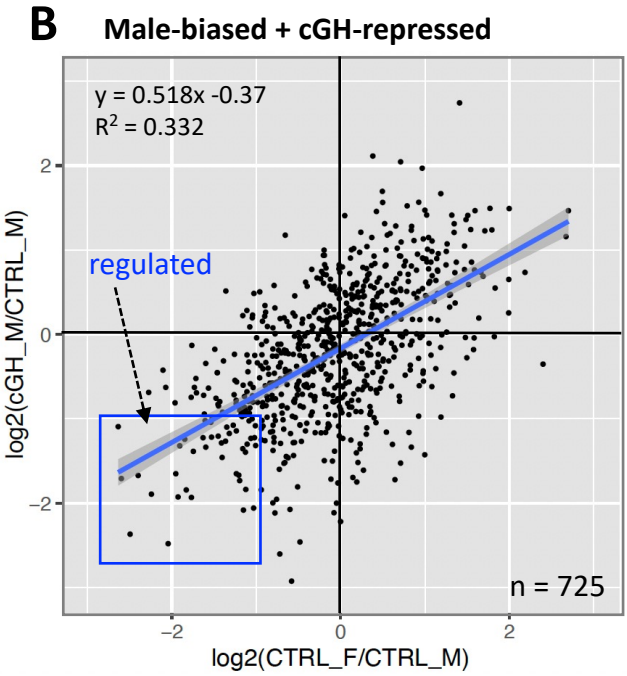


1,397 Sex-biased, cGH-responsive reporters

840 ATAC regions

	Male-biased cGH-repressed	Female-biased, cGH-activated
# DARs -->	229	161
# Linked to MB genes	168	6
# Linked to FB genes	31	137
No sex-biased gene links	1	3
No gene linkages	29	15





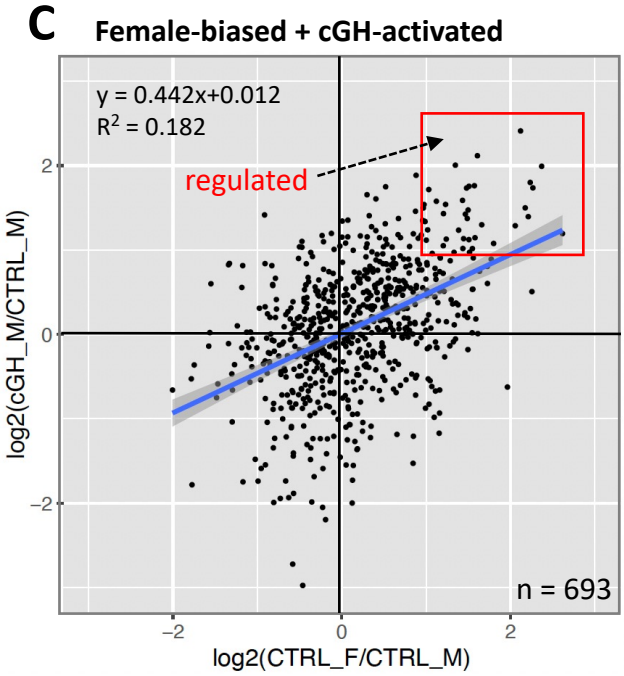


Fig. 5

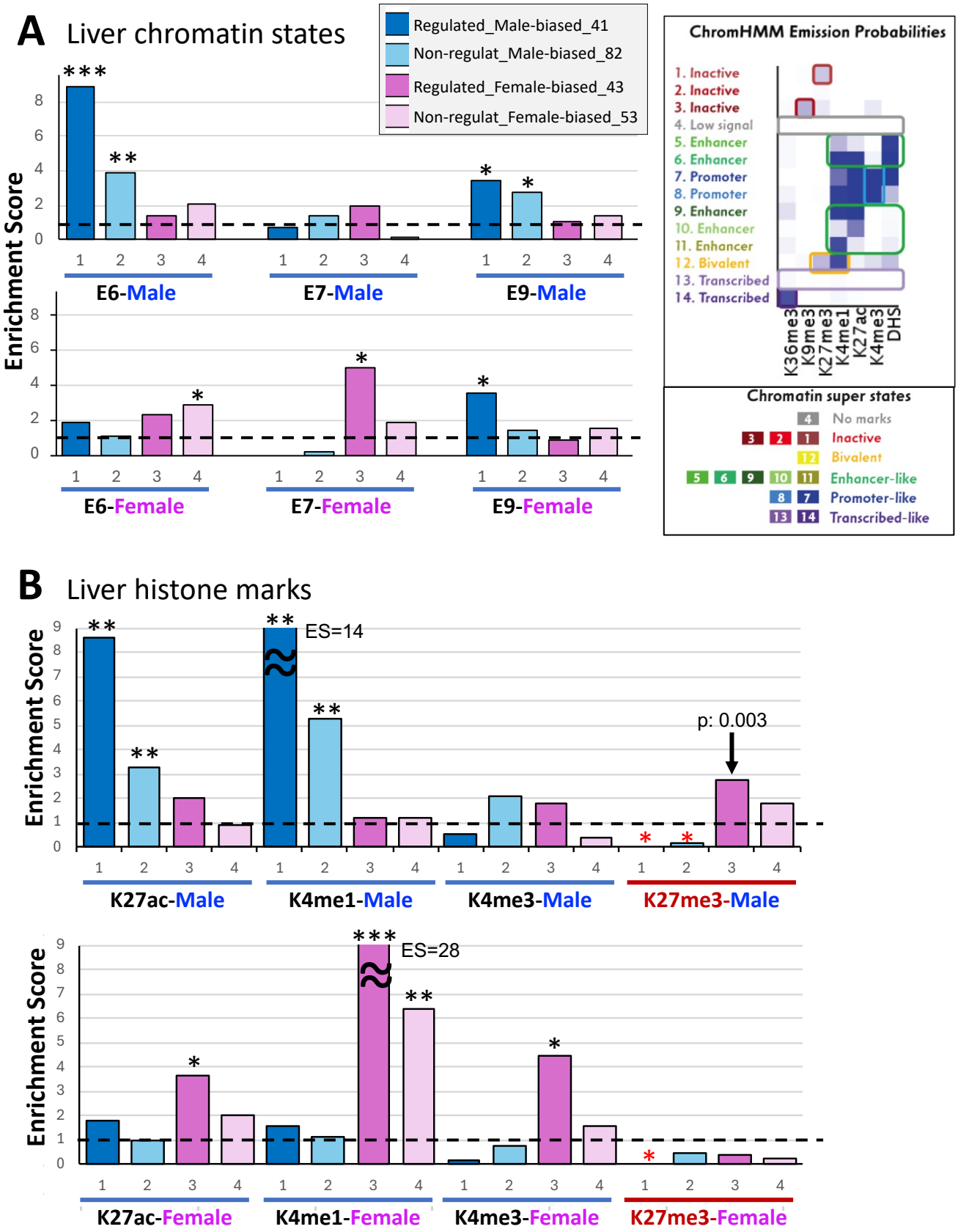
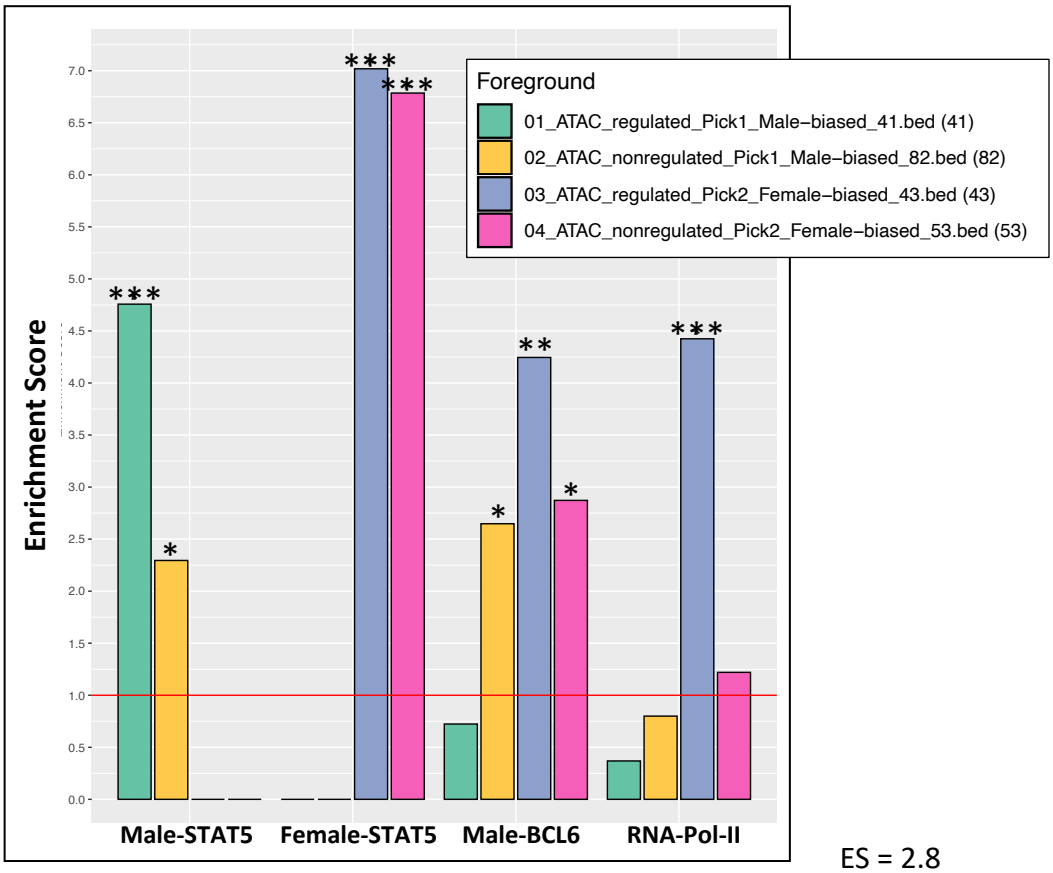

A
Liver chromatin states
Regulated_Male-biased_41
Non-regulat_Male-biased_82
Regulated_Female-biased_43
Non-regulat_Female-biased_53
Enrichment Score
E6-Male
E7-Male
E9-Male
E6-Female
E7-Female
E9-Female
ChromHMM Emission Probabilities
1. Inactive
2. Inactive
3. Inactive
4. Low signal
5. Enhancer
6. 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
B
Liver histone marks
ES=14
p: 0.003
ES=28
K27ac-Male
K4me1-Male
K4me3-Male
K27me3-Male
K27ac-Female
K4me1-Female
K4me3-Female
K27me3-Female

Fig. 6

A



B

Sex-biased DAR set	# of robust active ATAC regions in set	DAR 'Pick' set	STAT5/MA0519.1	BCL6_MOUSE.H11MO.0.A	CUX2_MOUSE.H11MO.0.C	HNF6_MOUSE.H11MO.0.A
Male-biased, cGH repressed_Regulated	41	1	20%	29%	37%	41%
Male-biased, cGH repressed_Non-Reg	82	1	17%	26%	17%	16%
Female-biased, cGH induced_Regulated	43	2	35%	40%	12%	7%
Female-biased, cGH induced_Non-Reg	53	2	34%	42%	11%	9%
Male-biased, NOT cGH responsive_Reg	19	3_Male	16%	21%	42%	53%
Male-biased, NOT cGH responsive_Non-Reg	27	3_Male	22%	26%	15%	4%
Female-biased, NOT cGH responsive_Reg	58	3_Female	16%	36%	2%	3%
Female-biased, NOT cGH responsive_Non-Reg	55	3_Female	13%	9%	7%	4%

ES = 2.8
p=0.024

ES = 3.8
p=0.0033

ES = 7.3
p=0.0007

ES = 4.2
p=0.049

ES = 28.8
p=0.0002

1209 **Supplemental Figure legends:**
1210
1211 **Fig. S1. Synthetic STARR-seq preparation.** (A) Pooled inserts, started from synthetic ssDNA oligos, were amplified
1212 by PCR to attach sequences of homology arms at ends, were expected to be 307 bp long. (B) Crude PCR products
1213 with 10 and 21 PCR cycles. Gel electrophoresis showed the size of the PCR fragments after 10 amplification cycles
1214 (Lane 3) and 21 amplification cycles (Lane 5) matched with the expected length. The fragments with 10 PCR cycles
1215 were proceeded for plasmid library preparation. Lanes 1 and 4 are 100 bp NEB marker (New England Biolabs,
1216 #N3231S). Lane 2 represents the amplified HA-adapted DNA fragments when minimizing the numbers of PCR
1217 cycles, conditions that are used in library preparation. The lower molecular bands indicate unused primers and
1218 primer dimer, to be removed using SPRI beads prior to the subsequential cloning steps. Lane 3 shows the pattern
1219 observed when STARR-seq inserts are over-amplified.
1220
1221 **Fig. S2. Reporters cloned into HDI-STARR-seq library: characterization with respect to GC content.**
1222 (A) Scatterplot showing alignment of normalized read counts in two technical replicates of the plasmid libraries
1223 amplified separately (Illumina sequencing libraries G202M1 and G202M2; see Table S6). (B) Frequency
1224 distribution of averaged normalized read counts within 23,912 reporters from 2 plasmid libraries and 3 DNA
1225 libraries. 81.9% of reporters are within the range of 6.4x median. (C) Frequency distribution showing the GC
1226 content of 23,912 individual synthetic reporters (based on the 180 nt sequence). Y-axis, number of library
1227 reporters at each % GC that were included in the design for the set of 23,912 reporters. (D) Averaged read counts
1228 of 23,912 reporters from 5 plasmid/DNA libraries (including 2 plasmid (D1) and 3 extracted DNA (D2) libraries)
1229 were sorted by GC content, then grouped into 500 bins. Data show biphasic correlation of reporter
1230 representation as judged by read counts in PCR-amplified libraries versus GC content.
1231
1232 **Fig. S3. Active reporters identified in mouse liver by HDI-STARR-seq across three biological conditions.** (A) 900
1233 selected sex-biased DARs were tiled with reporters (numbers in parenthesis, e.g., 4,949 reporters designed for
1234 360 male-specific, cGH-repressed DARs). Read count for each reporter was calculated based on >95% overlap
1235 with the genomic coordinates of 23,912 reporters. (B) The first table shows the number of DNA or plasmid
1236 libraries (out of the 5 such libraries examined) that met the threshold of >= 40 normalized reads per 10 million
1237 mapped reads for each DNA or plasmid library. Thus, 19,136 reporters were so qualified in all 5 plasmid + DNA
1238 libraries, 1,081 other reporters were qualified in only 4 of the 5 libraries, etc. The 3 other tables indicate the
1239 number of active reporters (out of the 19,136 reporters qualified based ion DNA) in each of the three biological
1240 conditions based on the normalized read counts in either both (n=2) or all 3 (n=3) of the liver RNA libraries
1241 analyzed by HDI-STARR-seq for that biological condition. For the RNA libraries, 'qualified' indicates the liver met
1242 the threshold of 30 normalized reads per 10 million mapped reads for STARR-seq RNA libraries recovered from
1243 the HDI-transfected livers. (C) Histograms showing average RNA activity of 19,136 DNA qualified reporter regions
1244 in cGH-M (n=2), male (n=3) and female (n=2) mice, grouped into 150 bins. (D) Table showing the total number of
1245 reporters designed, the number of reporters within the HDI-STARR-seq library that were qualified (DNA qualified)
1246 and the numbers of reporters that were RNA qualified after HDI-STARR-seq for each of the indicated DAR
1247 categories selected for HDI-STARR-seq analysis (Table S2). (E, F) Venn diagrams showing the 11,996 active
1248 reporters from the union of three active sets in three biological conditions (8,922 active reporters in cGH-treated
1249 male livers, 7,099 active reporters in male livers and 7,748 active reporters in female livers), and the number of
1250 ATAC regions represented by each subset. 4,126 reporters are robustly active in common in all three biological
1251 conditions in terms of RNA qualification and DNA qualification and comprise 1,292 unique ATAC regions. See
1252 Table S7A. Note: Although, for example, there are 2,008 active reporters that are uniquely active in cGH-treated
1253 male liver (panel E), these reporters map to a total of 1,063 of the 1,292 ATAC regions represented by at least 1
1254 of the 4,126 robust active reporters. Thus, many of these 1,063 ATAC regions have one or more other individual
1255 reports that are robust active reporters. Consequently, there are only n=24 ATAC regions that harbor reporters
1256 active in only cGH-infused mouse liver.
1257
1258 **Fig. S4. Active reporters and stringently inactive reporters: plasmid library representation, number of ATAC**
1259 **regions, and ATAC intensity values.** (A) Plasmid reads in sets of 4,126 robust active reporter regions whose

recovered liver RNA met the threshold for qualification in both replicate livers (cGH-treated male, female livers) or in all 3 replicate livers (male) have the highest representation in the original STARR-seq plasmid library used for HDI. The set of 982 stringently inactive reporters are defined as the regions consistently exhibiting low activity (less than 30 normalized reads) in RNA libraries of all biological replicates. We defined the top 400 stringently inactive reporters with higher plasmid representation (labeled as "str_inactive_H") and the other 582 stringently inactive reporters with lower plasmid reads (labeled as "str_inactive_L"). The set of 582 stringently inactive reporters was excluded from downstream comparisons to avoid artificial effect of the low representation in the library. (B) Number of ATAC regions represented by robust active reporters only, by a mixed of robust active and by stringently inactive reporters only. (C) Boxplots show the ATAC cut sites intensity distributions (ATAC intensity*100) in hepatocytes for the three subsets in panel B from the four biological samples analyzed by MultiOmics. Result shows that stringent inactive set_H is comparable as a background set for enrichment analyses, in terms of its plasmid representation and chromatin accessibility in hepatocytes.

Fig. S5. HDI-STARR-seq reporter activity for ATAC regions that are far from genes (A) and are located on autosomal vs sex-chromosomes (B). (A): Boxplots (top) show the distribution of STARR-seq activity for all three biological conditions for individual robust active reporters in each of the 6 indicated DAR Pick categories. Results indicate that DAR regions in Pick6, which are distant from genes, are active, and the table at the bottom shows that Pick1 and Pick6 are very similar regarding their frequencies of robust active reporters. 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. (B): 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 represent 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. S6. Liver-extracted HDI-STARR-seq DNA library reads correlate well with HDI-STARR-seq plasmid library reads, indicating no significant reporter sequence-dependent bias in plasmid uptake or retention by hepatocytes following HDI.

Fig. S7. Stress gene expression. No stress genes were induced after 10 days of osmotic minipump implantation or 7 days after HDI. Statistical tests by Kolmogorov-Smirnov test (N=3-5 per group), followed by multiple comparison with FDR for multiple comparison correction. Data are mean +/- SEM values, with mean values for male without HDI set = 1.

Fig. S8. Correlation analysis of regulated sex-based and other reporters. (A) Correlation of cGH treatment and sex specificity in term of the ratio of STARR-seq reporter activity for the n=220 subset of 4,126 robust active reporters that are male-biased, not cGH-responsive DARs (Pick3_M, Table S2). **(B)** Correlation of cGH treatment and sex specificity in term of the ratio of STARR-seq reporter activity for the n=436 subset of 4,126 robust active reporters that are female-biased, not cGH-responsive DARs. Blue lines indicate linear regression models and gray region around each line represents the 95% confidence interval. **(C)** Correlation, as above, for n=577 autosomal robust active reporters that are at sex-biased DARs that are not cGH-responsive (Pick3_M and Pick3_F, Table S2). **(D)** Correlation as above for the n=79 robust active reporters on chrX or chrY that map to DARs that are sex-biased but not cGH-responsive. Blue lines in panels represent the best fit linear regression models, with the 95% confidence interval show in gray. **(E)** Boxplots indicate comparable plasmid library representation of the subsets of regulated and non-regulated reporters in each DAR set ('Pick' group, Table S2). **(F)** TAD-based correlations, as indicated. See Table S7B for detailed correlation data.

Fig. S9. Gene linkage data for regulated vs non-regulated active enhancer DARs. Metrics of individual linkages between ATAC regions and genes grouped based on their HDI-STARR-seq activity and gene response, as derived from Table S4B and Table S9. Results show that linkages to non-sex-biased genes are characterized by

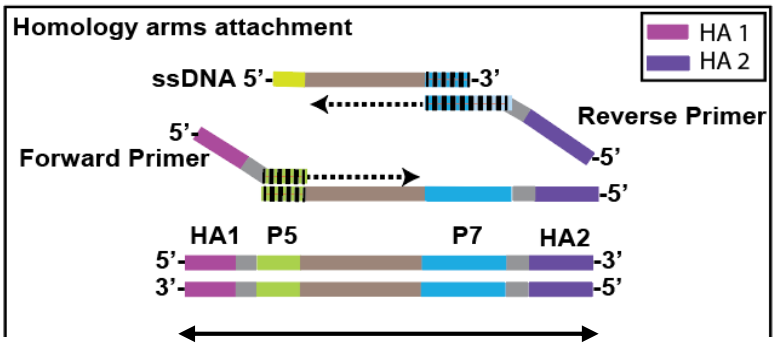
significantly longer distances and significantly lower reliability (lower linkage score, lower linkage significance) than the linkages to sex biased genes; compare *middle* vs *right* subpanels of (A) and of (B). Shown are boxplots displaying the distributions of sex-biased regulated ATAC regions (salmon) or non-regulated regions (blue) with linkages to genes that are sex-biased [in (A), male biased and cGH-suppressed genes linked to regulated male-biased and cGH-suppressed DARs (i.e., 41 regulated regions); and in (B), female-biased and cGH-induced genes linked to regulated female-biased and cGH-activated DARs (i.e., 43 regulated regions)] in the *middle* panels, or that are linked to genes that are not sex-biased (*right* panels), or that are linked to any genes (*left* panels), where 'any genes' includes both sex-biased and not sex-biased genes (denoted as "all" in the subpanels). Statistical analyses were performed by applying the Kolmogorov-Smirnov test to the distributions for each of three metrics [i.e., linkage distance in nts (*top* subpanels), linkage score (*middle* subpanels), and linkage significance (-log(pvalue), *bottom* subpanels)] and are significant at $p < 0.001$ (***) for linkages to sex-biased genes versus linkage to non sex-biased genes, when considering the regulated regions and non-regulated regions as a group. In contrast, there are no significant differences (p-values >0.05, not marked in the figure) between regulated and non-regulated sex-biased and cGH-responsive DARs regarding the distributions of linkage scores, distances and significance, regardless of whether the linkage are to sex-biased genes or to non-sex-biased genes.

Fig. S10, Fig. S11. Motif analysis on regulated DARs. De novo motifs and the compared motifs in a mouse motif database (UniProbe) identified by MEME-ChIP for a variety of combination of foreground and background sets (see text).

Fig. S12. HDI-STARR-seq activity in cell-type specific ATAC regions. (A) Reporters representing liver cell-type specific ATAC regions included in the sets of 11,996 active reporters and 4,126 robust reporters identified in Table S7A. Percentages were calculated by dividing the number of active reporters (11,996 active reporters or 4,126 robust active reporters) by the 23,912 predesigned reporters in each cell type category. (A) Robust active reporters representing hepatic ATAC regions were significantly depleted in NPC ATAC regions. (B) Boxplots showing the distribution of DNA reads and RNA reads of the robust active reporters in each category. (C) Boxplots showing the distribution of enhancer activity for robust active reporters whose ATAC regions are preferentially accessible in the indicated liver cell type. Enhancer activity in hepatic ATAC regions was significantly higher in hepatocyte-specific ATAC regions than in NPC-specific ATAC regions (Kolmogorov-Smirnov test). Also see Table S11.

Fig. S1

A



Insert (210 nt) + F/R primer (48/49 nt) =307 bp

B

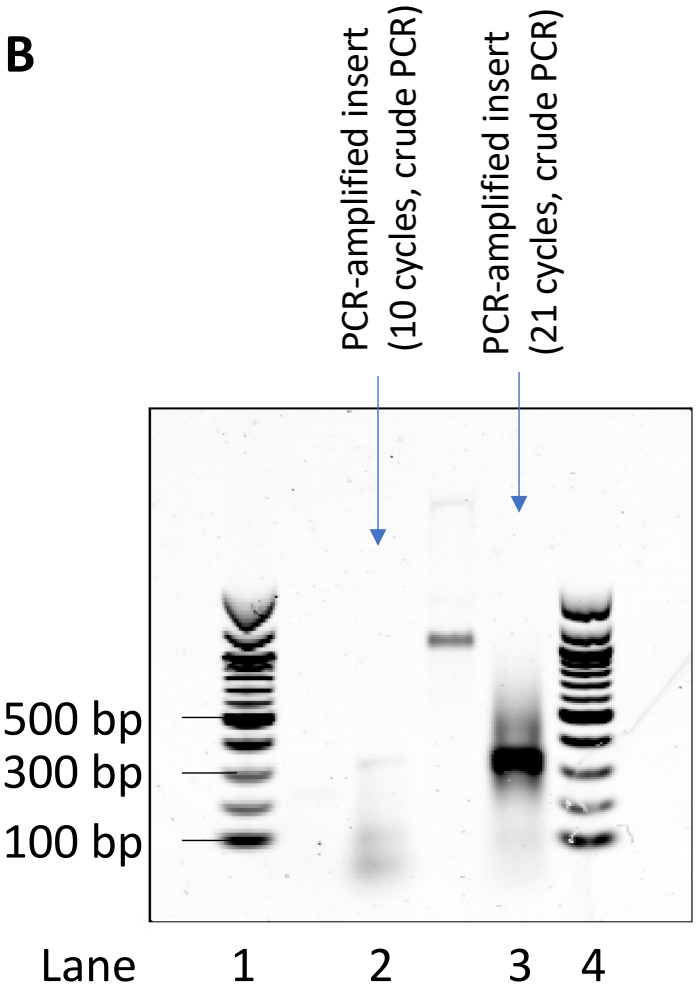


Fig. S2

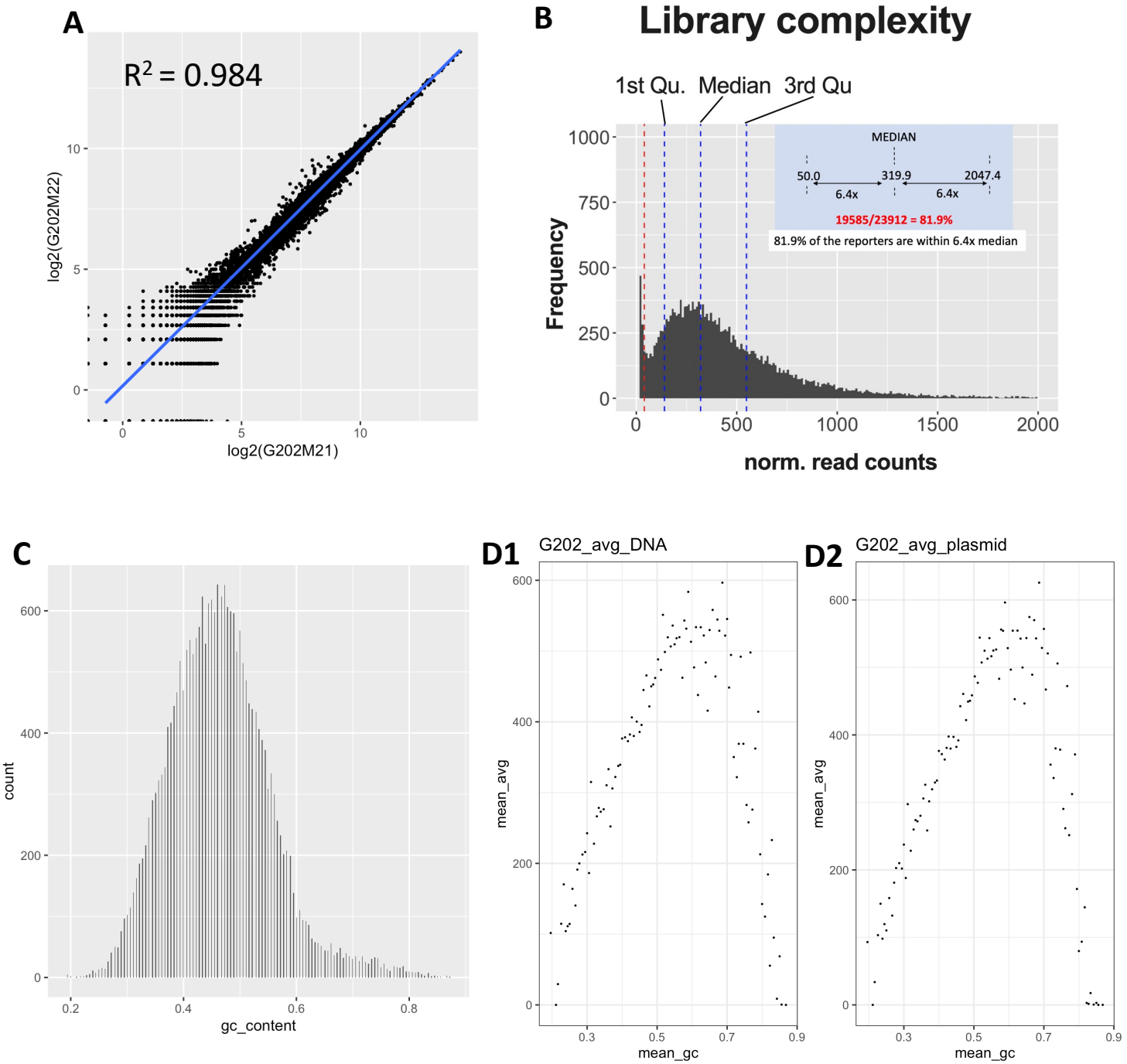
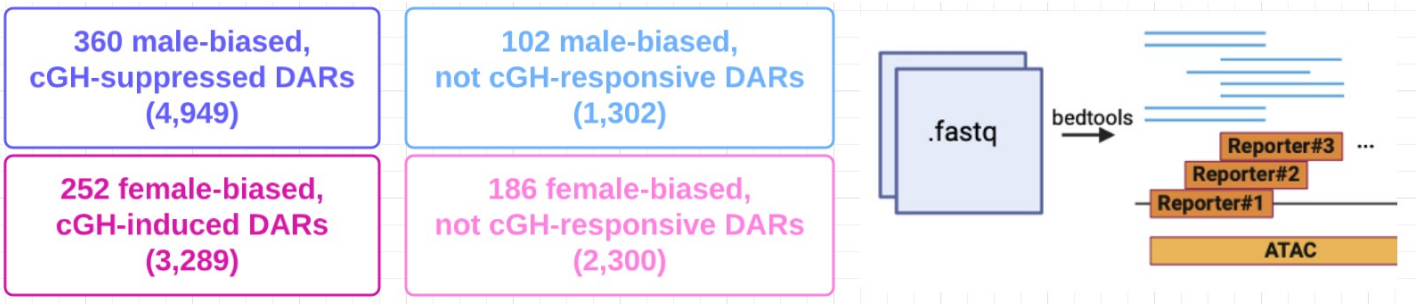

A
R2 = 0.984
log2(G202M22)
log2(G202M21)
B
Library complexity
1st Qu.
Median
3rd Qu
MEDIAN
50.0
319.9
2047.4
6.4x
6.4x
19585/23912 = 81.9%
81.9% of the reporters are within 6.4x median
Frequency
norm. read counts
C
count
gc_content
D1
G202_avg_DNA
mean_avg
mean_gc
D2
G202_avg_plasmid
mean_avg
mean_gc

Fig. S3ABC

A



B

# of qualified DNA/plasmid libraries (norm. read >40)	n_reporter
0	2095
1	739
2	475
3	386
4	1081
5	19136

# of qualified RNA libraries in cGH-M (norm. read >30)	n_reporter
0	4106
1	6108
2	8922

# of qualified RNA libraries in CTRL-M (norm. read >30)	n_reporter
0	3013
1	4287
2	4737
3	7099

# of qualified RNA libraries in CTRL-F (norm. read >30)	n_reporter
0	3181
1	8207
2	7748

C

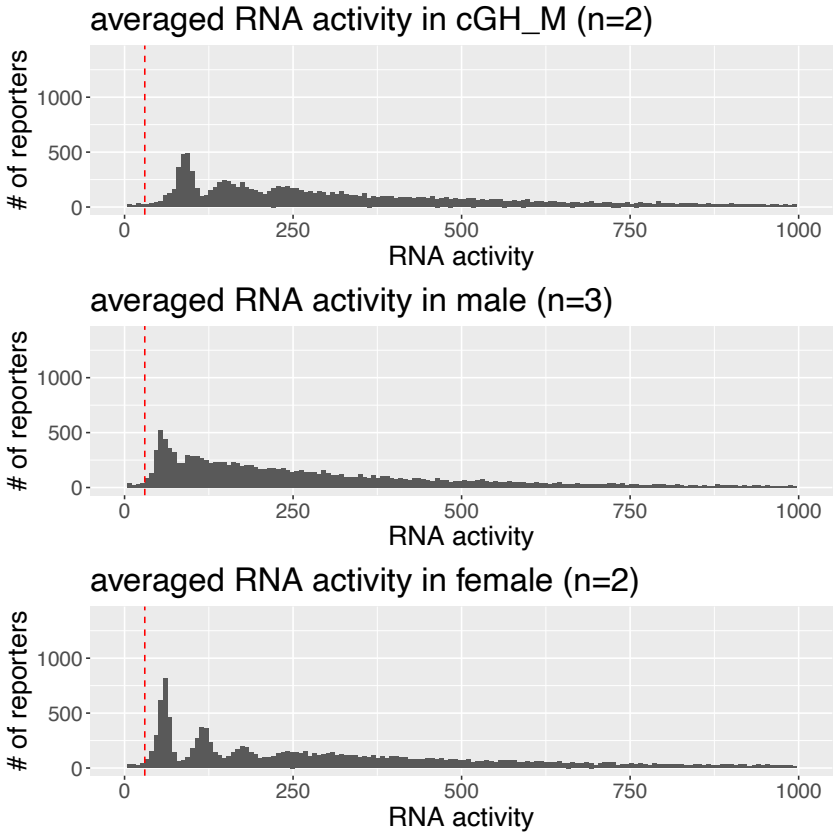


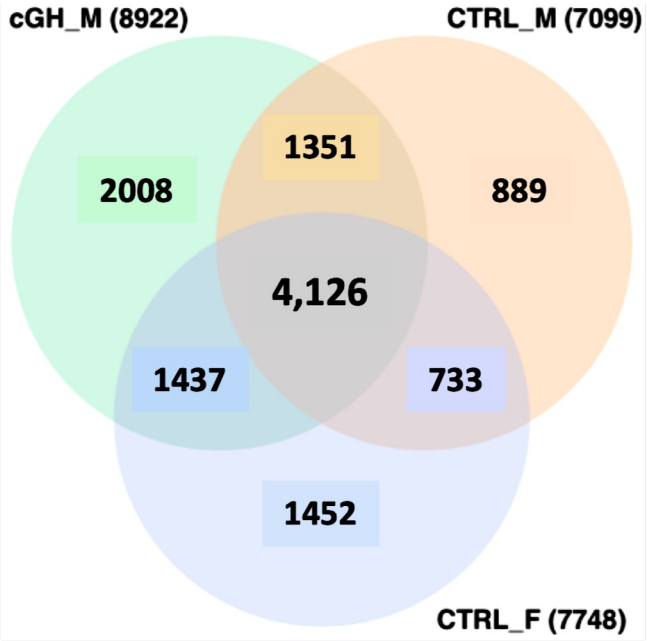
Fig. S3DE

D

Category_Pick	Predesignated reporters (23,950)	DNA qualified reporters (19,136)	RNA qualified reporters (4,126)
Pick1_M_360	4,949	3,990	725
Pick2_F_252	3,289	2,698	693
Pick3_F_186	2,300	1,806	436
Pick3_M_102	1,302	1,074	220
Pick4_M_62	994	886	181
Pick5_F_9	135	131	40
Pick6F	205	194	40
Pick6M	649	590	94
Others	10,127	7,767	1,697

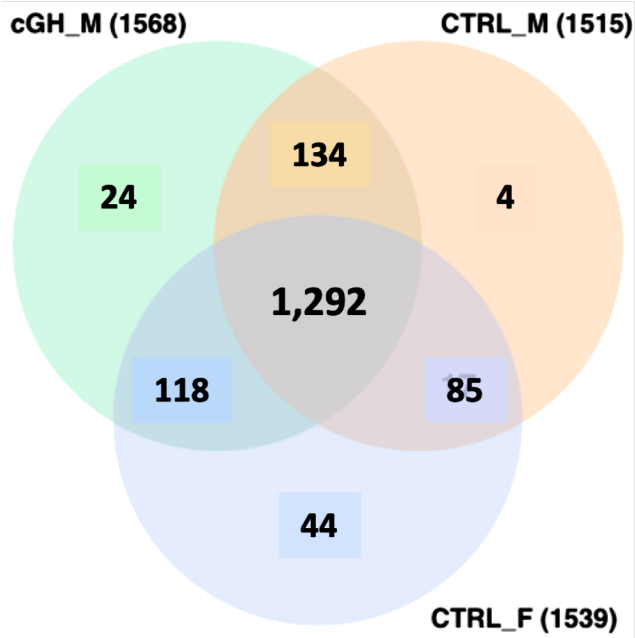
E

11,996 active reporters



F

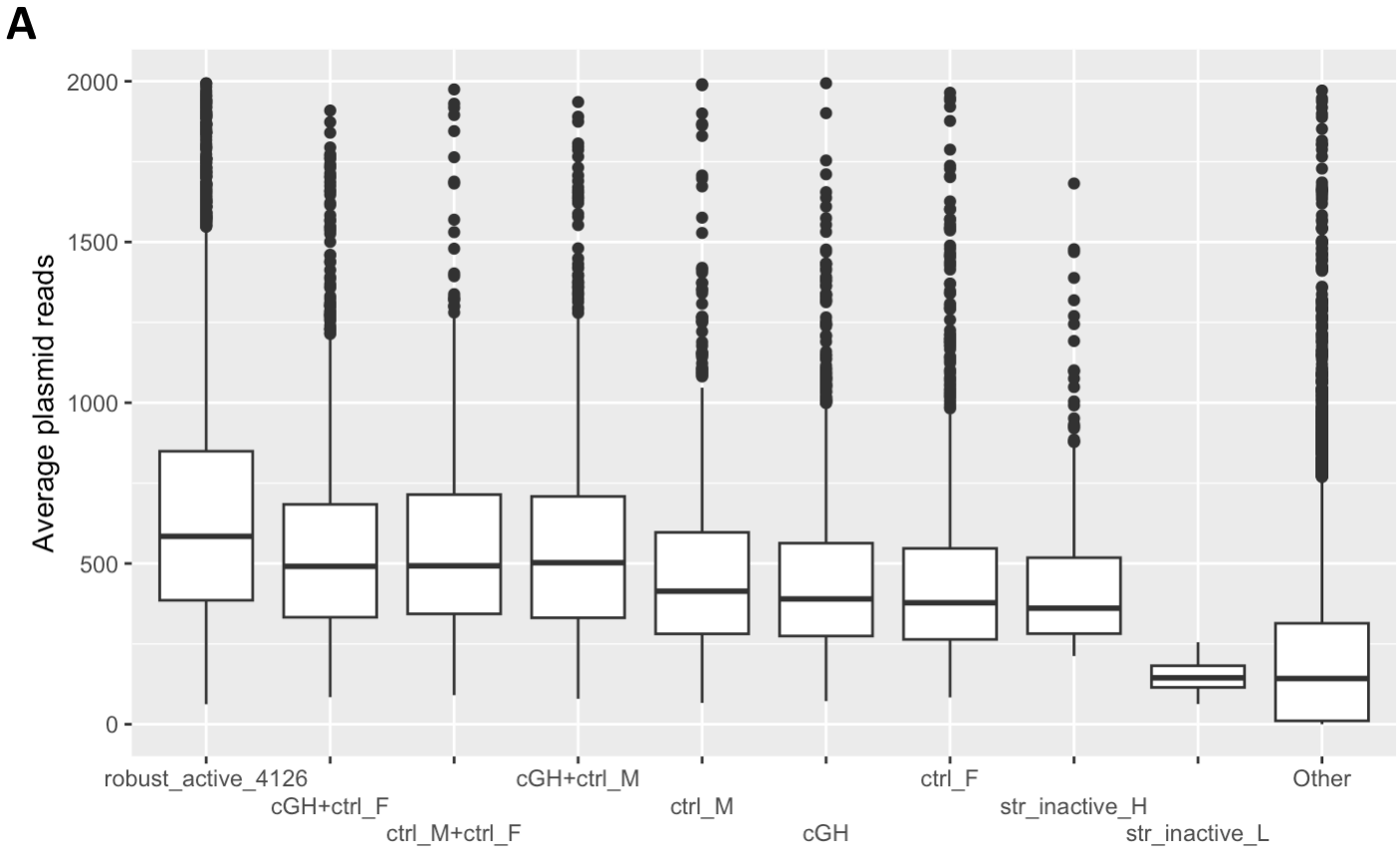
1,619 ATAC regions represented

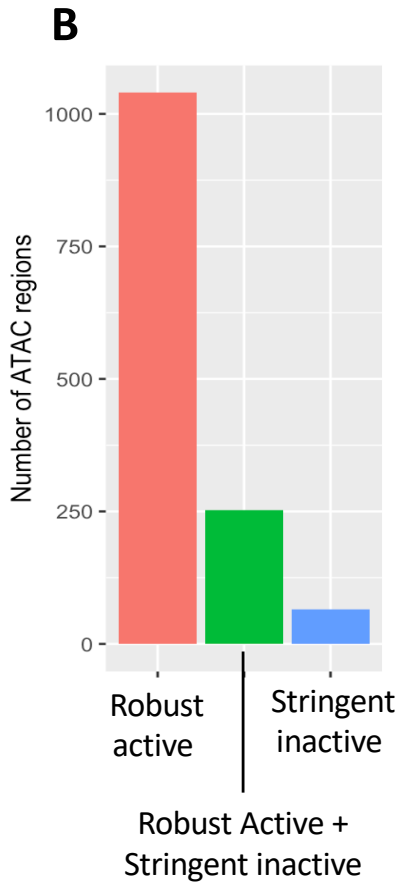


HDI-STARR-seq Reporters:
23,912 Reporters → 19,136 DNA qualified → 11,996 RNA-qualified → **4,126** Robust active reporters

# of corresponding unique ATAC regions:
1,839 ATAC → 1,676 DNA qualified → 1,619 RNA-qualified → **1,292** Robust active Enhancers

Fig. S4


A
Average plasmid reads
2000
1500
1000
500
0
robust_active_4126
cGH+ctrl_F
ctrl_M+ctrl_F
cGH+ctrl_M
ctrl_M
cGH
ctrl_F
str_inactive_H
str_inactive_L
Other


B
Number of ATAC regions
1000
750
500
250
0
Robust active
Stringent inactive
Robust Active + Stringent inactive

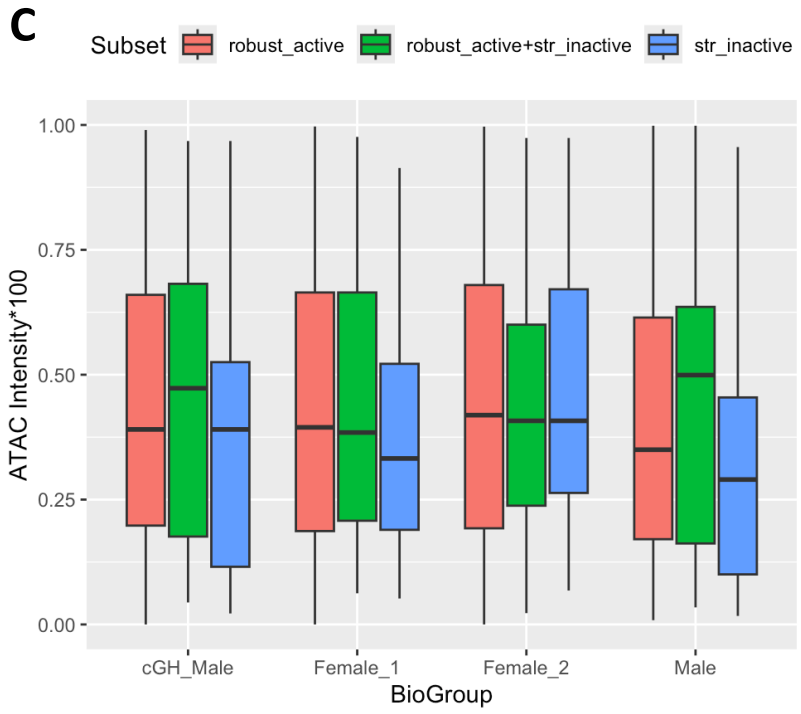
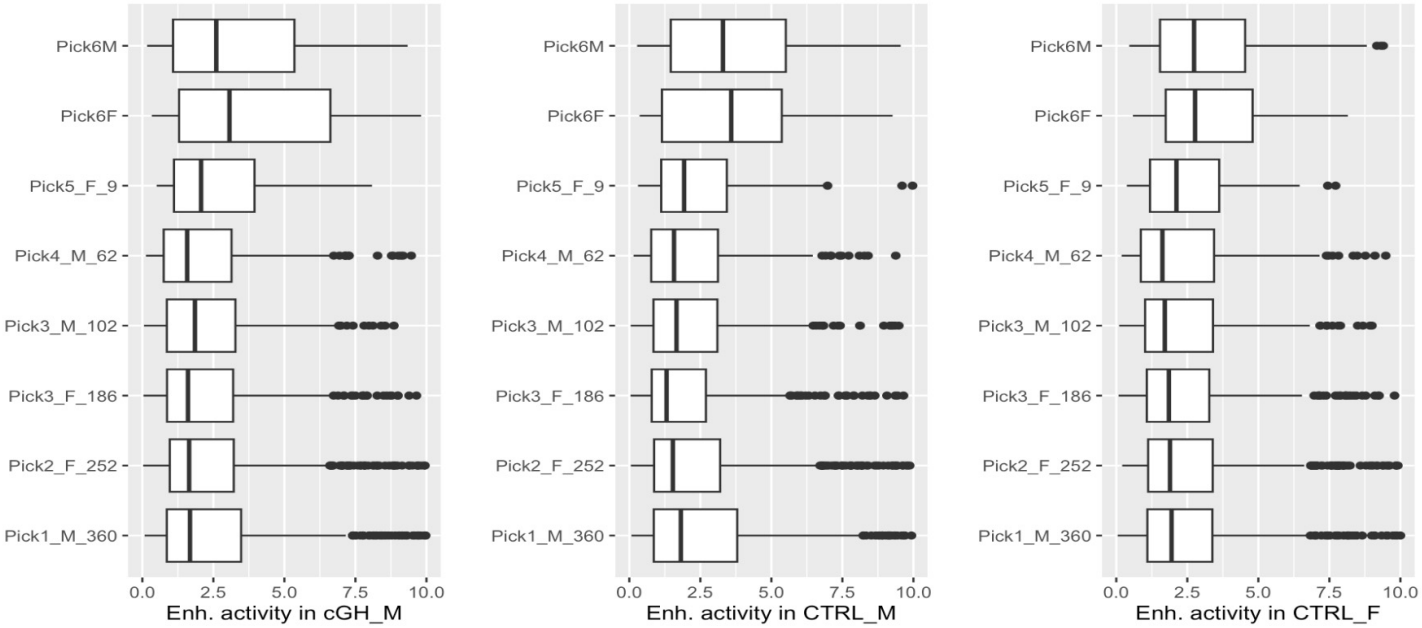

C
Subset
robust_active
robust_active+str_inactive
str_inactive
ATAC Intensity*100
1.00
0.75
0.50
0.25
0.00
cGH_Male
Female_1
Female_2
Male
BioGroup

Fig. S5A

A



	# Male-biased reporters + # female-biased reporters	# Non-regulated reporters (fold-change ≤2)
Pick1_M_360	178	447
Pick6_M_40	19	60
p=0.5 (Fisher's exact test)	ES = 1.26	

Fig. S5B

B

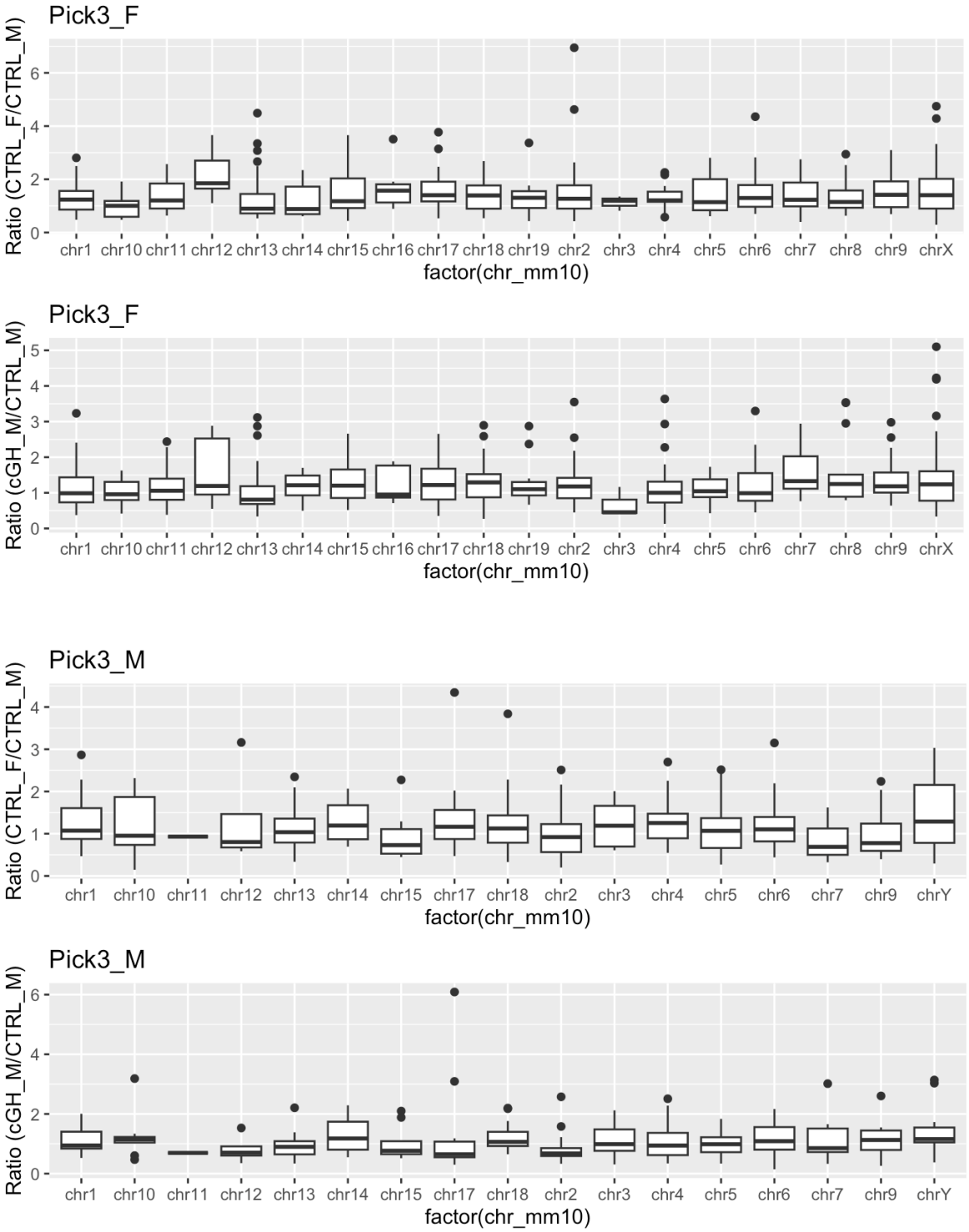

Pick3_F
Ratio (CTRL_F/CTRL_M)
0
2
4
6
chr1 chr10 chr11 chr12 chr13 chr14 chr15 chr16 chr17 chr18 chr19 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9 chrX
factor(chr_mm10)
Pick3_F
Ratio (cGH_M/CTRL_M)
0
1
2
3
4
5
chr1 chr10 chr11 chr12 chr13 chr14 chr15 chr16 chr17 chr18 chr19 chr2 chr3 chr4 chr5 chr6 chr7 chr8 chr9 chrX
factor(chr_mm10)
Pick3_M
Ratio (CTRL_F/CTRL_M)
0
1
2
3
4
chr1 chr10 chr11 chr12 chr13 chr14 chr15 chr17 chr18 chr2 chr3 chr4 chr5 chr6 chr7 chr9 chrY
factor(chr_mm10)
Pick3_M
Ratio (cGH_M/CTRL_M)
0
2
4
6
chr1 chr10 chr11 chr12 chr13 chr14 chr15 chr17 chr18 chr2 chr3 chr4 chr5 chr6 chr7 chr9 chrY
factor(chr_mm10)

Fig. S6

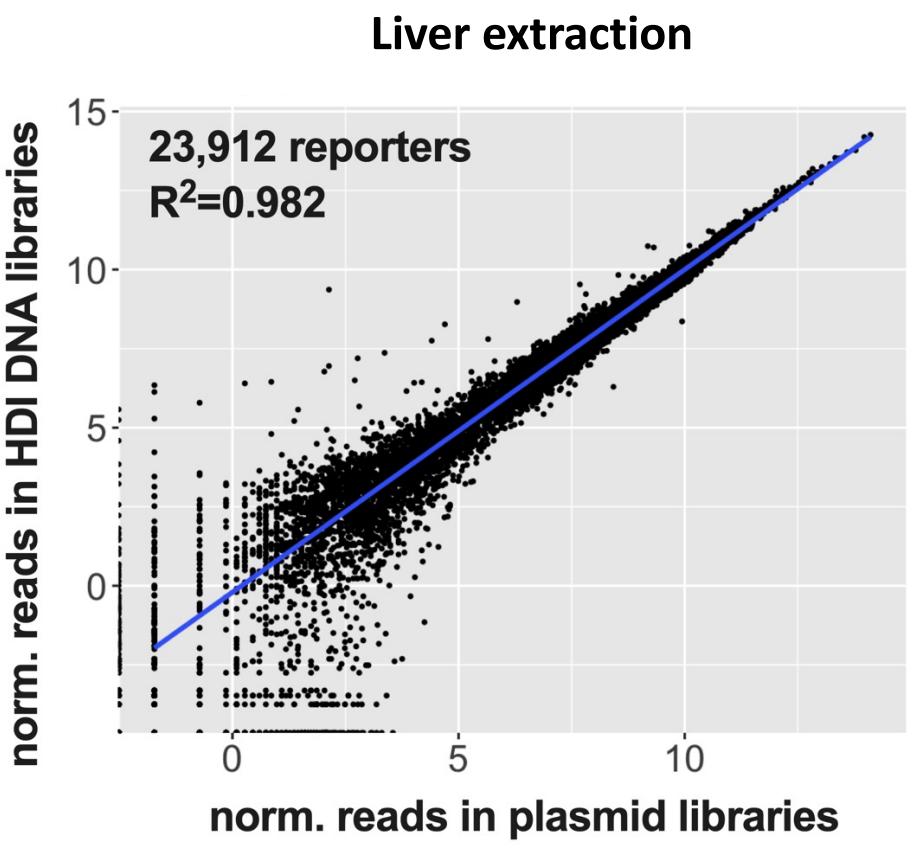

Liver extraction
23,912 reporters
R2=0.982
norm. reads in HDI DNA libraries
norm. reads in plasmid libraries
15
10
5
0
0
5
10

Fig. S7

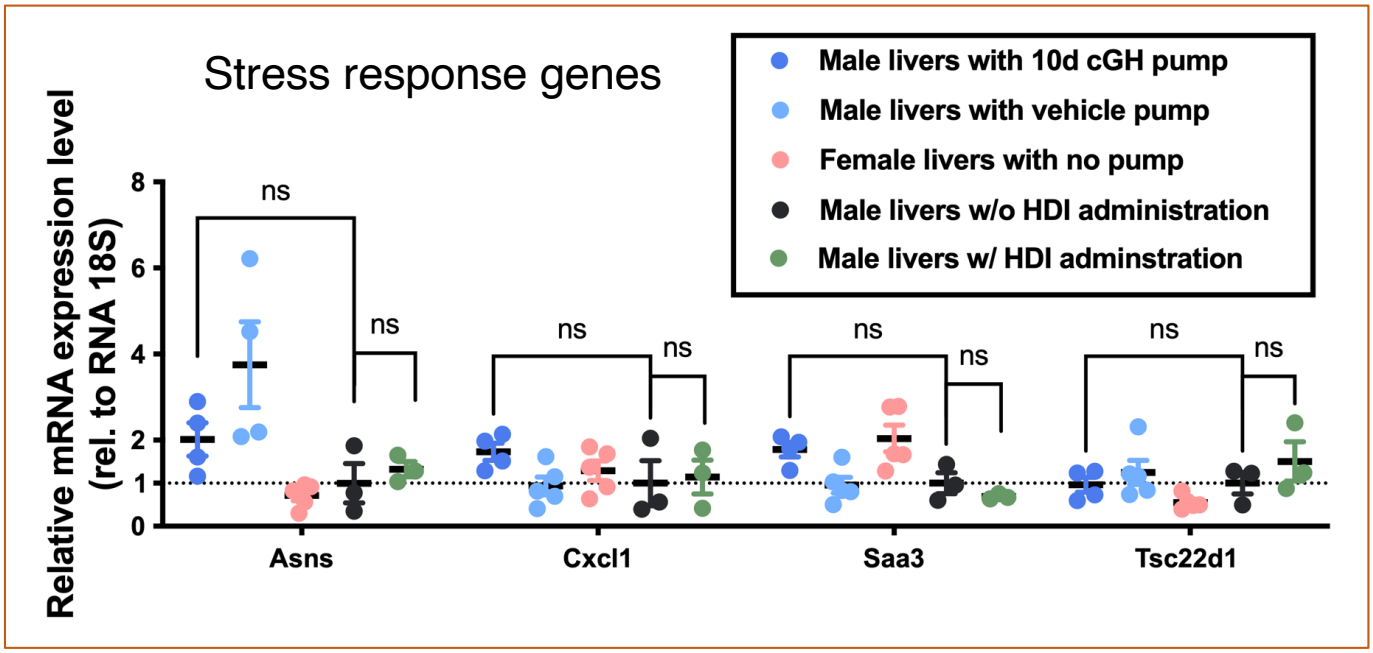

Stress response genes
Male livers with 10d cGH pump
Male livers with vehicle pump
Female livers with no pump
Male livers w/o HDI administration
Male livers w/ HDI adminstration
Relative mRNA expression level
(rel. to RNA 18S)
8
6
4
2
1
0
ns
Asns
Cxcl1
Saa3
Tsc22d1

Fig. S8A-S8E

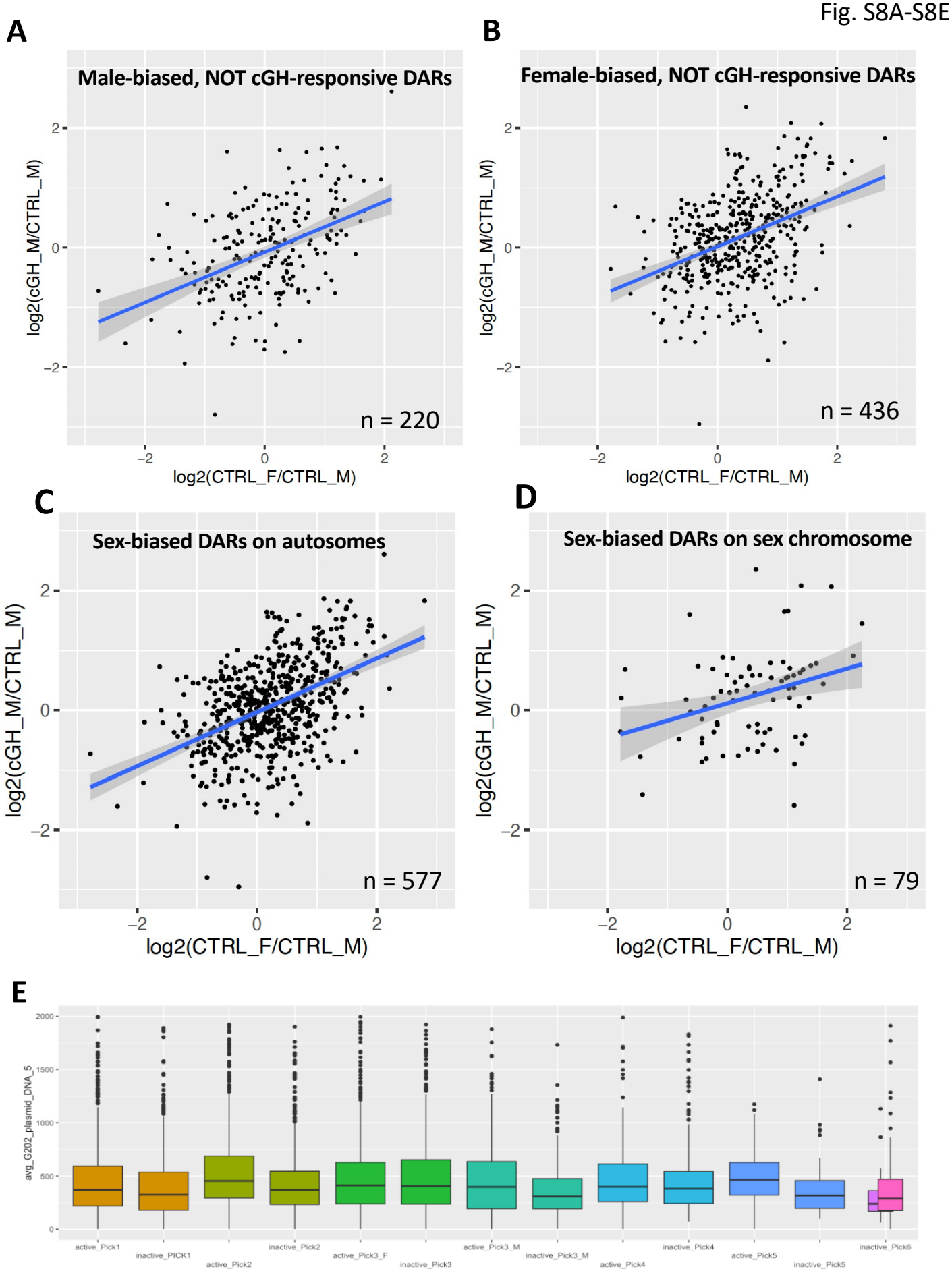

A
Male-biased, NOT cGH-responsive DARs
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)
n = 220
B
Female-biased, NOT cGH-responsive DARs
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)
n = 436
C
Sex-biased DARs on autosomes
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)
n = 577
D
Sex-biased DARs on sex chromosome
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)
n = 79
E
avg_G202_plasmid_DNA_5
active_Pick1
inactive_PICK1
active_Pick2
inactive_Pick2
active_Pick3_F
inactive_Pick3
active_Pick3_M
inactive_Pick3_M
active_Pick4
inactive_Pick4
active_Pick5
inactive_Pick5
inactive_Pick6

Fig. S8F

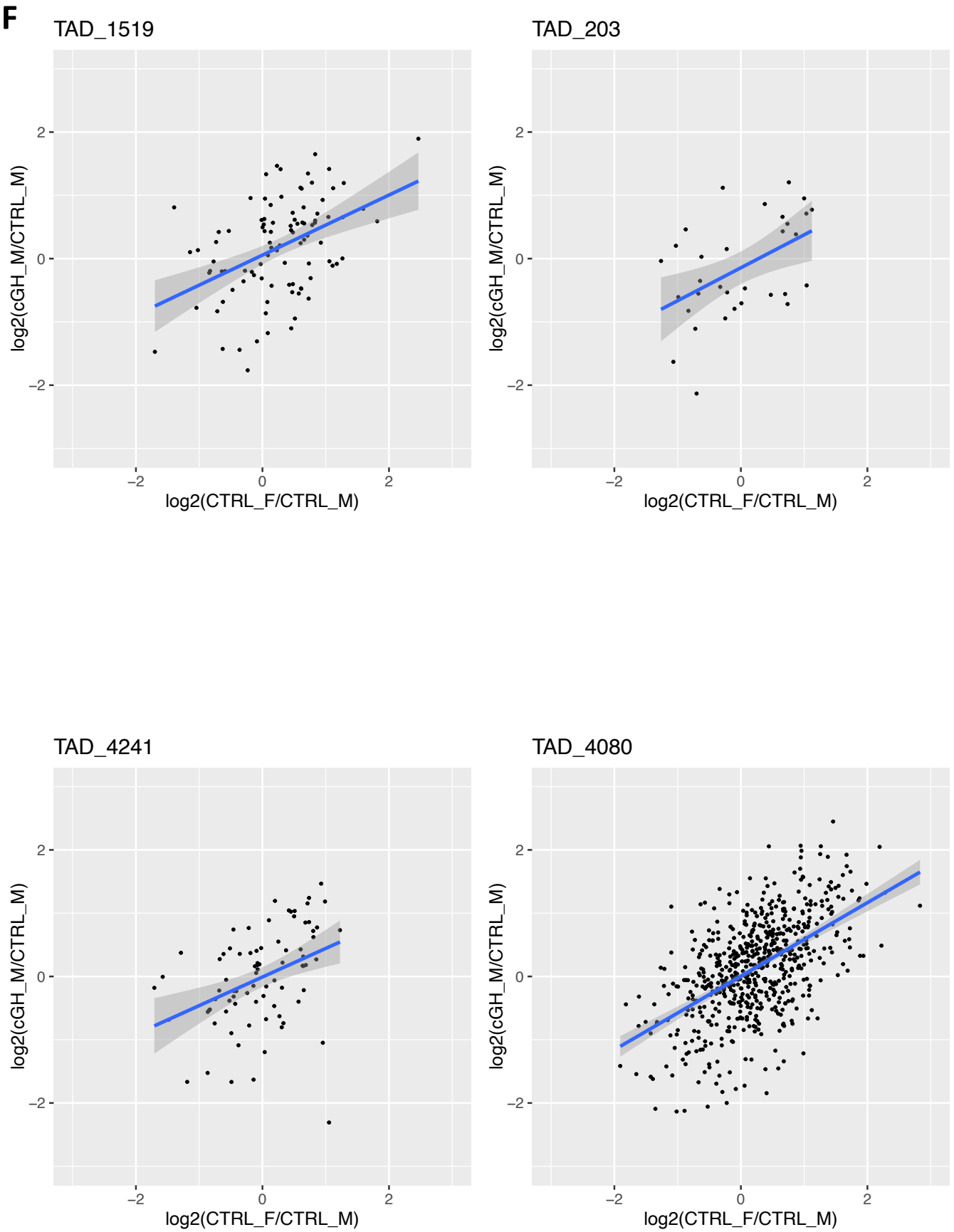
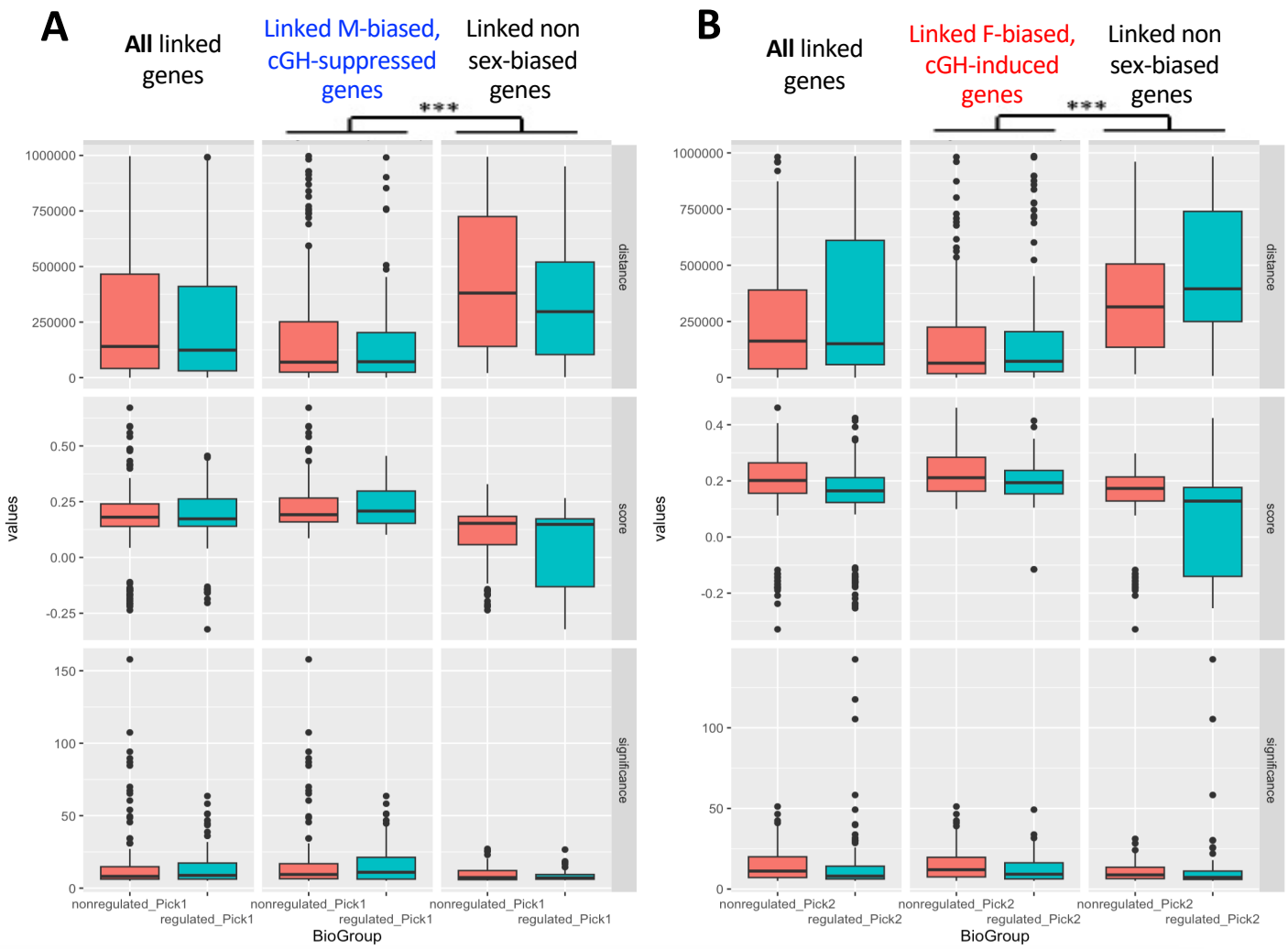

F
TAD_1519
TAD_203
TAD_4241
TAD_4080
log2(cGH_M/CTRL_M)
log2(CTRL_F/CTRL_M)
2
0
−2

Fig. S9



Note: Significance between *middle* and *right* panel datasets (***, shown at the top) applies to all three panels, i.e., *Distance*, *Score*, and *Significance.*

Fig. S10AB

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. S10CD

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. S11AB

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. S11CD

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

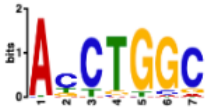
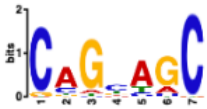
	Motif Found	Discovery/Enrichment Program	E-value	Known or Similar Motifs
#1	Reverse Complement ⇆ Show 1 More ↧	STREME	9.3e-012	Sp4_secondary (UP00002_2) Klf7_primary (UP00093_1) Smad3_secondary (UP00000_2)
#2	Reverse Complement ⇆	STREME	4.3e-008	
#3	 Reverse Complement ⇆	STREME	1.8e-006	Tcfe2a_secondary (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)

Fig. S12

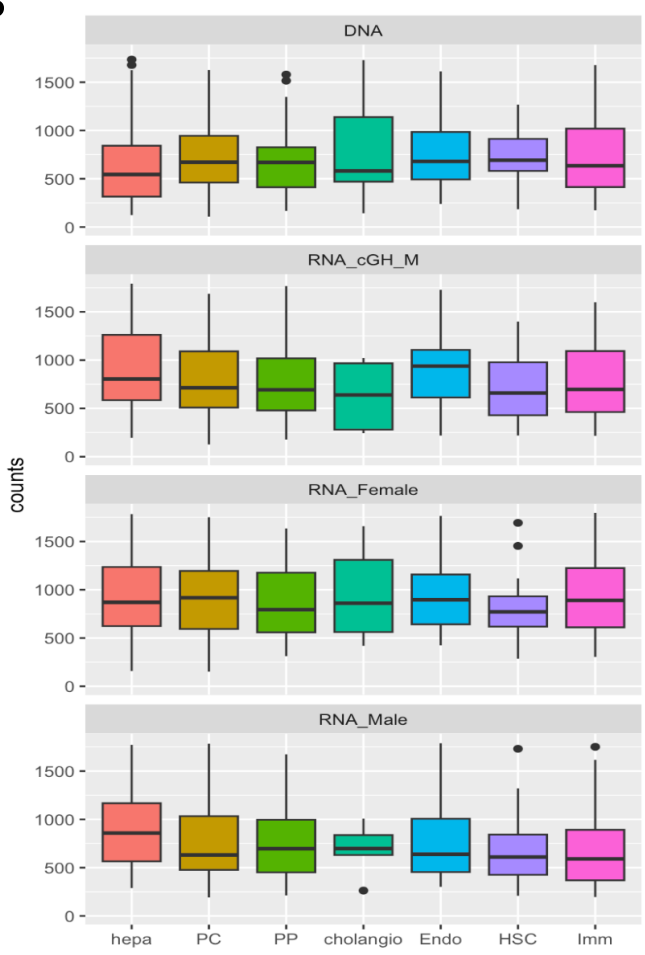
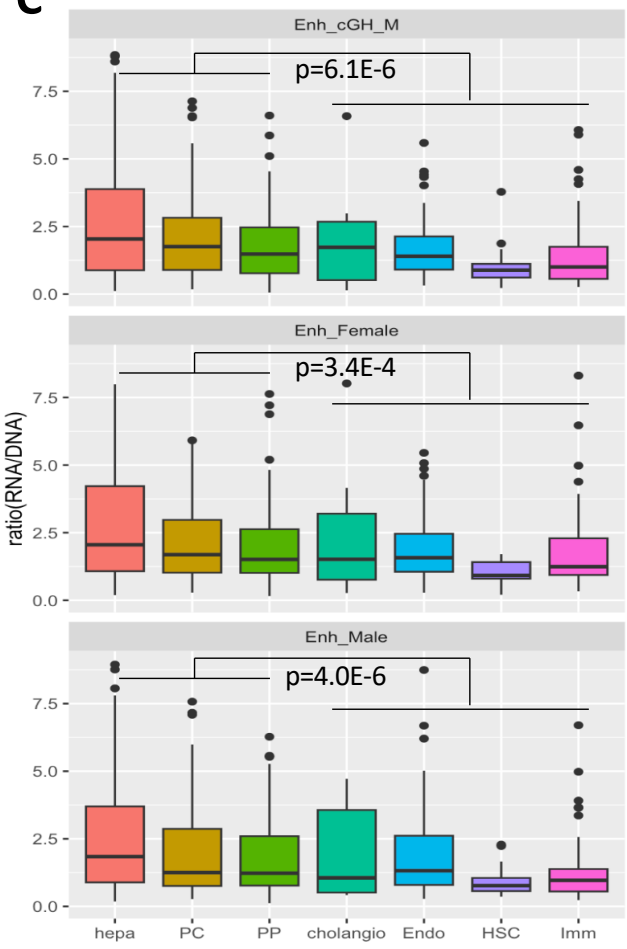
A

		23,912 predesigned		11,996 active reporters			4,126 robust reporters		
Cell type		n_reporter	n_ATAC	n_reporter	%	n_ATAC	n_reporter	%	n_ATAC
Endo	non hepatic	346	20	222	64.2%	20	52	15.0%	16
HSC		299	20	146	48.8%	19	24	8.0%	11
Imm		295	20	207	70.2%	20	59	20.0%	19
cholangio		32	3	19	59.4%	3	11	34.4%	3
hepa	hepatic	391	20	249	63.7%	20	96	24.6%	20
PC		333	20	203	61.0%	19	86	25.8%	17
PP		325	20	196	60.3%	20	71	21.8%	20

	All designed reporters (23,912)	**Active (11,996)**	**Not active**
Endo+HSC+Imm+ cholangio	972	594	378
Hepa+PC+PP	1049	648	401
		ES=	1.03
		Not significant	

(all NPCs)	All designed reporters (23,912)	**Robust active (4,126)**	**Not robust active**
Endo+HSC+Imm+ Cholangio	972	146	826
Hepa+PC+PP	1049	253	796
		ES=	0.56
	(2-tailed)	p< 0.0001	Fisher's exact test

B



C