Title
A Human-Evolved Gli2 Enhancer is Required for Normal Male Aggression

Author
Norman, Andrew Robert

Publication Date
2020

Peer reviewed|Thesis/dissertation

A Human-Evolved Gli2 Enhancer is Required for Normal Male Aggression

by
Andrew Robert Norman

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Developmental and Stem Cell Biology

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:
Jeremy Reiter
F07C889D1B164B3...

Jeremy Reiter
Chair

DocuSigned by:
Nadav Ahituv
DocuSigned by:
Brian Black
DocuSigned by:
Pao-Tien Chuang
9CF8ADC5E474459...

Nadav Ahituv
Brian Black
Pao-Tien Chuang

Committee Members

Acknowledgements

Wow, that was quite a journey. As I sit here now, and reflect on where I've been, how far I've come and all the people I've known it's almost too much to take in. So many have helped me, given advice and constructive criticism along the way. Here I will attempt to name all of them and thank them, but undoubted my memory is not up to the task and I will fail. Please forgive me, in advance, if the help and support you have given me slips my mind. I'm on the home stretch here and running on fumes!

There's the original Reiter Lab crew, of course: Laura Yee with her quiet snark, but with kindness and a deep knowledge of all things SF (her hometown); Bill Dowdle, the life of the party and the man who's got some "theories"; Colin Dinsmore, a beautiful mind and wit, someone the world will hear from someday; Nicole Santos, who was always looking out for me. Then, of course, there's Elle Roberson, from quiet wallflower to firebrand professor-in-waiting, she's come a long way.

And almost back that far, I'm recalling all the wonderful people in the Ahituv lab who helped me out, laughed with me and tolerated me in their weekly lab meetings. Walter Eckalbar, Naveneet Matharu, Aaron Hardin, and Jingjing Zhao. Hane Ryu was a close collaborator with me on this project and the creator of the HARS library. None of this would have possible without her. And of course, I owe a huge debt of gratitude to Taka Inoue. Not only is he one of the most generous people I've ever met in science, but also one of the hardest working and most productive. There are many aspects of my project that would not have happened without his help, and I think many people in his lab would say the same thing.

This project has been extremely collaborative and spanning across disciplines. I must thank and give credit to Kirsty Jamieson, the 4C guru; David Erle and all the fine people in his

RNA sequencing core; Junli Zhang at the Gladstone Gene Targeting Core; certainly, Jeff Simms and Iris Lo at the Gladstone Behavioral Phenotyping, for all their amazing advice and help with the mouse behavioral assays, without which I would have been completely lost; and while she's not officially a core, Lauren Byrnes is a bioinformaticist badass and a single-cell wrangler of the highest order, and I will be forever grateful for her help and collaboration.

Reiter Lab the Next Generation is taking things into the future, and they're a great bunch: Semil Choksi, Melissa Truong, Chang Xie, Ying Yang, Benedict Tsai, Gabi Canales, Gabe Loeb, Galo Garcia and Dhivya Kumar. I really need to give a heartfelt thank you to E Yu, the caretaker of our mouse colony and all-around sharp dresser. She's an amazing person, always there with a smile to ask you how things are going. Without a doubt all science in our lab would grind to a halt without her and the devotion she gives to our mice.

Of course, I must give a special thank you to my committee: Pao-Tien Chuang, I always appreciate your kindness and generosity. When we've met to talk about science you've always been free with your time and spirit. Brian Black, truly you are the most energetic and enthusiastic scientist I've ever met. The feedback and engagement you've given me in our meetings has been second-to-none. It's very rare that someone will grant you their focus attention in the way that you have to me. Nadav Ahituv, this project really belongs as much to you as it does to me and my lab. You've been unconditional in your collaboration and help, with advice and with tools. Thank you!

And of course, there is the Dude, the man himself, Jeremy F. Reiter. He and I go back a ways. I can honestly say I've never met a more intelligent human being, an unconventional man who always sees the good in people. In this cynical, pessimistic age he is the antidote. Yeah, he's got a few crazy hypotheses. He may or may not believe that scRNAseq is a hoax and TSNE plots are randomly generated. But remember, Einstein thought quantum mechanics was wack and Newton was an alchemist who dabbled in physics and calculus. I think Jeremey is in

good company. To him I owe a tremendous debt of gratitude. My PhD would not have been possible without his support, materiel and emotional, and his scientific insight and confidence in me. I will always be thankful and grateful for that, and the many chances he has taken on me.

I will end this on a more personal note. I need to thank my parents. Guess what? I'm finally done with school, I promise. Mom, I know the last few years have been unbelievably rough, but once this fog lifts and the world returns to normal I believe golden years are ahead. Thank you for all the love and support. Papa, I miss you every day and wish you could be here with us now.

ARN

San Francisco, CA

August 28, 2020

A Human-Evolved *Gli2* Enhancer is Required for Normal Male Aggression

Andrew R. Norman

Abstract

Human accelerated regions (HARS), which are genome sequences that have evolved at an accelerated rate in the human lineage, have previously been shown to include developmental enhancers. I have used a massively parallel enhancer reporter assay (MPRA) to test a library of human and chimp HAR orthologs in a cell model of mammalian testes, in order to find HARS which have evolved in human and may underly unique human reproductive traits. Through luciferase assay, 4C-seq, transcriptional analysis and CRISPR/Cas9 deletion I have validated a candidate and shown that it is an enhancer of the hedgehog signaling pathway effector *Gli2* in the testis cell model. Furthermore, I have shown that the enhancer is required for normal *Gli2* expression in Leydig cells, as well as normal testis vascular development and normal inter-male social behavior.

Table of Contents

List of Figures

List of Tables

Table **Page**

Introduction

The testis is a complex organ consisting of tissues derived from multiple lineages. At day 11.5 in the mouse, pre-Sertoli cells migrate from the coelomic epithelium to the developing gonad and aggregate with primordial germ cells. These aggregates interact with nearby interstitial and endothelial cells to form the sex chords, which will eventually become seminiferous tubules, the sperm-producing structures of the testis. Further differentiation of the interstitial cells produces Leydig cells, the principle source of testosterone in males (*1*).

This complexity has made it difficult to faithfully recapitulate testes development *in vitro* for the purposes of study or sperm production, although partial success has been achieved with germ-cell co-culture systems (*2*) and testes *ex-vivo* culture (*3*, *4*). For studies or applications pertaining to a singles testes cell type, recapitulation of all testes cell types is not necessary. In these instances, immortalized testes cell lines derived from the various testes cell types may be preferable. To this end, Hofmann and colleagues derived several SV40 t-antigen immortalized testes cell lines from an adult mouse testis and characterized them based on morphology and limited gene expression (*5*).

When I began my enhancer study, described below, I sought a testis model that would be practical to use in my experiments and obtained the cell line (GC-1) characterized by Hofmann and colleagues. During my work with the line, I began to suspect that instead of being a cell type most similar to mouse spermatocytes, they were in fact most similar to mouse testis somatic cells. I tested this hypothesis with a combination of single-cell RNA sequencing and immunostaining and found it to be well-supported. I present these data in Chapter 1.

The eukaryotic genome is vast, and most of it does not code for proteins. This apparent lack of function has led some to label the non-coding genome as “junk” (*6*), but an enormous body of work has shown that it contains critical regulatory sequences, including enhancers (*7*). Enhancer elements are critical for the regulation of developmental processes (*8*, *9*) and often provide the mechanism of evolution (*10*, *11*). Demonstrating a causative link between enhancer molecular evolution and phenotypic evolution is a significant challenge, given the relatively weak interaction enhancers have with genes and the multitude of enhancers that can interact with a single gene. However, clever comparison of sequence conservation across multiple species, suggesting sequence importance, with unexpected mutation in a single species, indicating species-specific molecular evolution, has yielded datasets of putative gene regulatory sequences that may underly the evolution of species-specific traits (*12*–*15*). One such class of sequences, found to have experienced accelerated molecular evolution in the human lineage (“Human Accelerated Regions”, or “HARs”), includes enhancers (*16*–*18*). HARs are promising candidates for enhancers specific to the evolution of human form and behavior.

Humans are unique among the great apes in terms neuroanatomy, locomotion, reproductive biology and other traits (*19*). HARs have been previously shown to contain enhancers of genes involved in neural function and disease (*17*, *18*), but to my knowledge they have not been directly linked to any non-neuronal human-specific traits. Differences in human and great ape mating strategies are well documented and associated with male reproductive traits such as sperm production and motility (*20*, *21*), suggesting human-lineage evolution of the male reproductive system. However, the genetic basis of this evolution has been relatively unexplored. Thus, I undertook a search for HARs critical to the evolution of unique human reproductive biology.

To select strong, critical enhancers from a full set of 714 HARs (*13*, *22*) I employed a high-throughput enhancer discovery method, massively parallel reporter assay (MPRA) (*23*), in

the GC-1 mouse testis somatic cell line (*5*). Use of the MPRA reduced the search space to a few strong candidates, one of which (2xHAR.238) is adjacent to *GLI2*, an effector of the Hedgehog signaling pathway (*24*). Proper hedgehog pathway signaling is required for normal testes development, genital development, sperm production and androgen synthesis (*25*–*27*). For this reason, and since 2xHAR.238 has been previously identified as an embryonic hindbrain enhancer in transgenic mice (*16*), I chose to narrowed my study to this candidate. To demonstrate that 2xHAR.238 is an enhancer of *Gli2*, I deleted it in the mouse testes cell line and found that it is required for normal expression of *Gli2*. I confirmed the enhancer-looping interaction of the HAR with *Gli2* using a 4C chromatin interaction assay. Consistent with its role as an enhancer of *Gli2*, I found that *Gli2* expression was reduced in the 2xHAR.238-deleted cell line. Finally, to test the requirement of 2xHAR.238 *in vivo*, I deleted it in mice and found that it is necessary normal testis vascular development and for normal inter-male mouse behavior.

This study identifies a single enhancer required for a discrete mammalian behavior. The requirement of the 2xHAR.238 sequence for normal social behavior, and the fact that it has evolved within the human lineage, has important implications for the mechanisms with which evolution has shaped our species.

GC-1 Cells are Derived from Testis Somatic cells

Hofmann et al. (*5*) reported a cell line they called “GC-1 spg” and claimed that it is most similar to a spermatocyte, the male meiotic germ cell, based on morphology and expression of a single spermatocyte-specific marker (*28*). From my experience working with the line I had reason to suspect that GC-1 cells are not representative germ cells. Using modern transcriptome analysis and histological data, I present evidence here that GC-1 cells are, in fact, most similar to mouse testis somatic cells.

If GC-1 cells are derived from spermatocytes, germline cells, then these cells should express markers specific to the germline. DDX4 is a germline-specific RNA helicase, and a classic marker of the germline in both Drosophila and mammals (*29*). In an immunofluorescence assay, GC-1 cells did not express DDX4, similar to the fibroblast cell line NIH/3T3, while the luminal germ cells of a testis tubule had robust expression (Fig. 2.1a-c). SCP3 is a component of the meiotic synaptonemal complex and a marker of spermatocytes (*30*). GC-1 cells did not express SCP3, again similar to 3T3 cells, while the spermatocytes of the mouse testis (identified by large nuclei, adjacent to the condensing post-meiotic spermatids) had robust expression (Fig. 2.1d-f). Finally, WT1 is a transcription factor that marks Sertoli cells and some testis interstitial cell precursors (*31*), and marked the Sertoli cells of mouse testis in my hands (Fig. 2.1g). GC-1 cells did show expression of WT1, while 3T3 did not (Fig. 2.1h-l). Taken together, these results indicate that GC-1 cells are more similar to testis somatic cells than to germ cells.

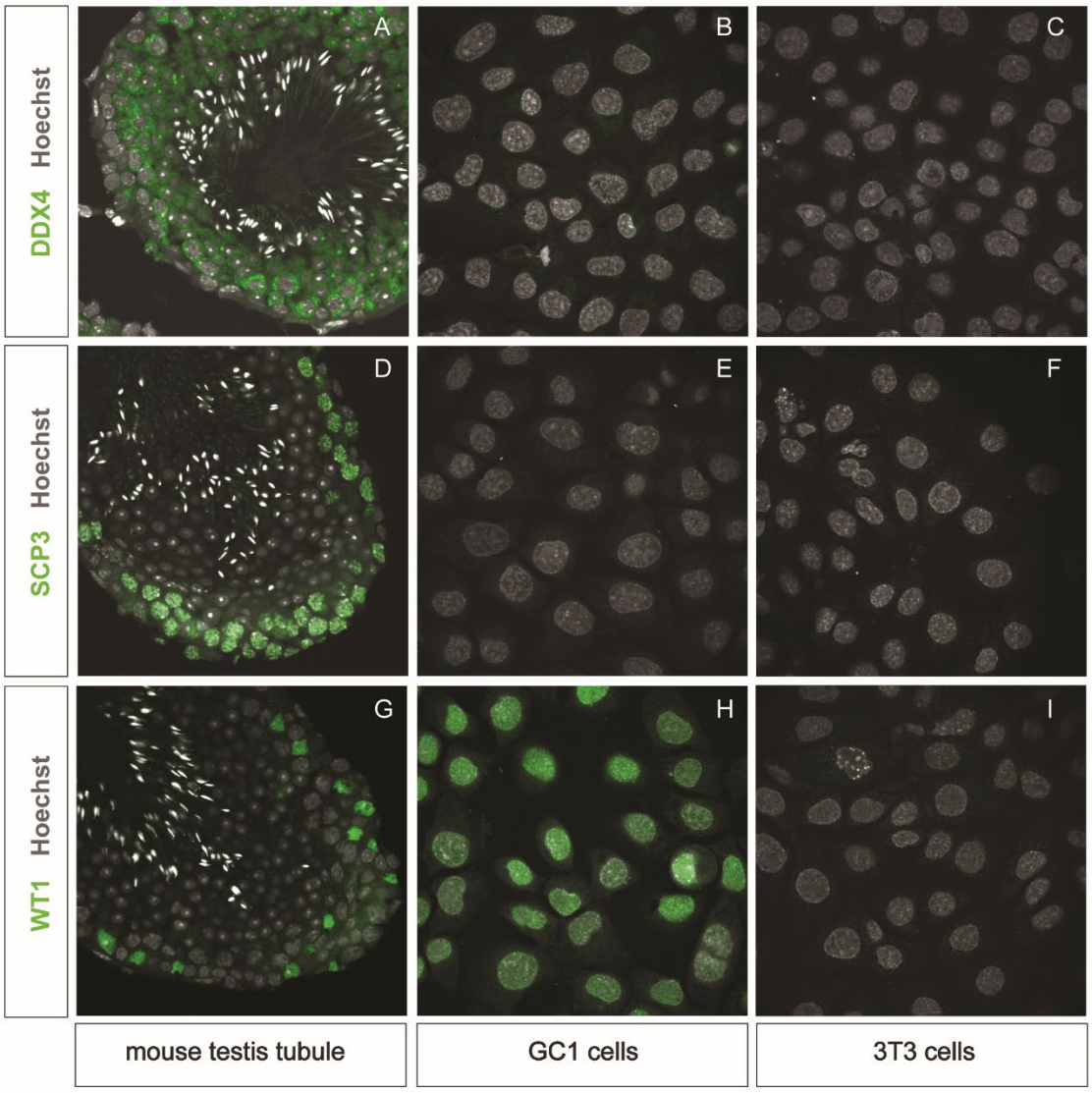

Figure 1.1. Expression of testis markers in GC-1 cells is similar to that of testis somatic cells. Hoechst stain marks nuclear DNA. (A-C) DDX4 expression in mouse testes, GC-1 cells and 3T3 cells. Note that expression is absent in both the GC-1 and 3T3 cell lines, but can be seen clearly in the testis germline-derived cells (cells proximal to the lumen). (D-F) Same cell types but marking SCP3 expression. Note again that expression is absent from GC-1 and 3T3 cells. (G-I) Same cells types but marking WT1. Note that expression can clearly be seen in the Sertoli cells in the testis (G) and the GC-1 cells (H), but it is lacking in the 3T3 cells (I).

To confirm the identity of the GC-1 cells, I chose to compare the GC-1 transcriptome with the transcriptomes of known mouse testis cell-types, as reported in a recent study using single-cell RNA sequencing (*32*) by Green and colleagues. First, I re-built the data analysis pipeline provided by Green and colleagues. In order to confirm the integrity of the reconstructed pipeline, I performed the following operations in the same manner they reported: clustering of the whole-testis dataset, principle-component analysis of the dataset and identification of testis cell-types based upon markers used in (*32*). My results matched those reported (Fig. 1.2a). To determine protein-coding transcript read counts of GC-1 cells, I performed bulk RNA-sequencing using an Illumina HiSeq instrument. Single-cell transcriptome data is shallow and sparse, so it is not appropriate to directly compare single-cell RNA sequencing reads to bulk RNA sequencing reads. However, the rank-order of gene-expression in a cell should not be affected by depth of sequencing. For this reason, I used a Spearman Correlation to compare the rank-order of gene expression in the bulk GC-1 dataset to the rank-order of gene expression reported in each of the testis single-cell transcriptomes (Fig. 1.2b-c). When all germ cell-types and somatic cell-types comparisons were pooled, GC-1 cells were found to have a significantly higher correlation with testis somatic cell-types (Fig. 1.2d).

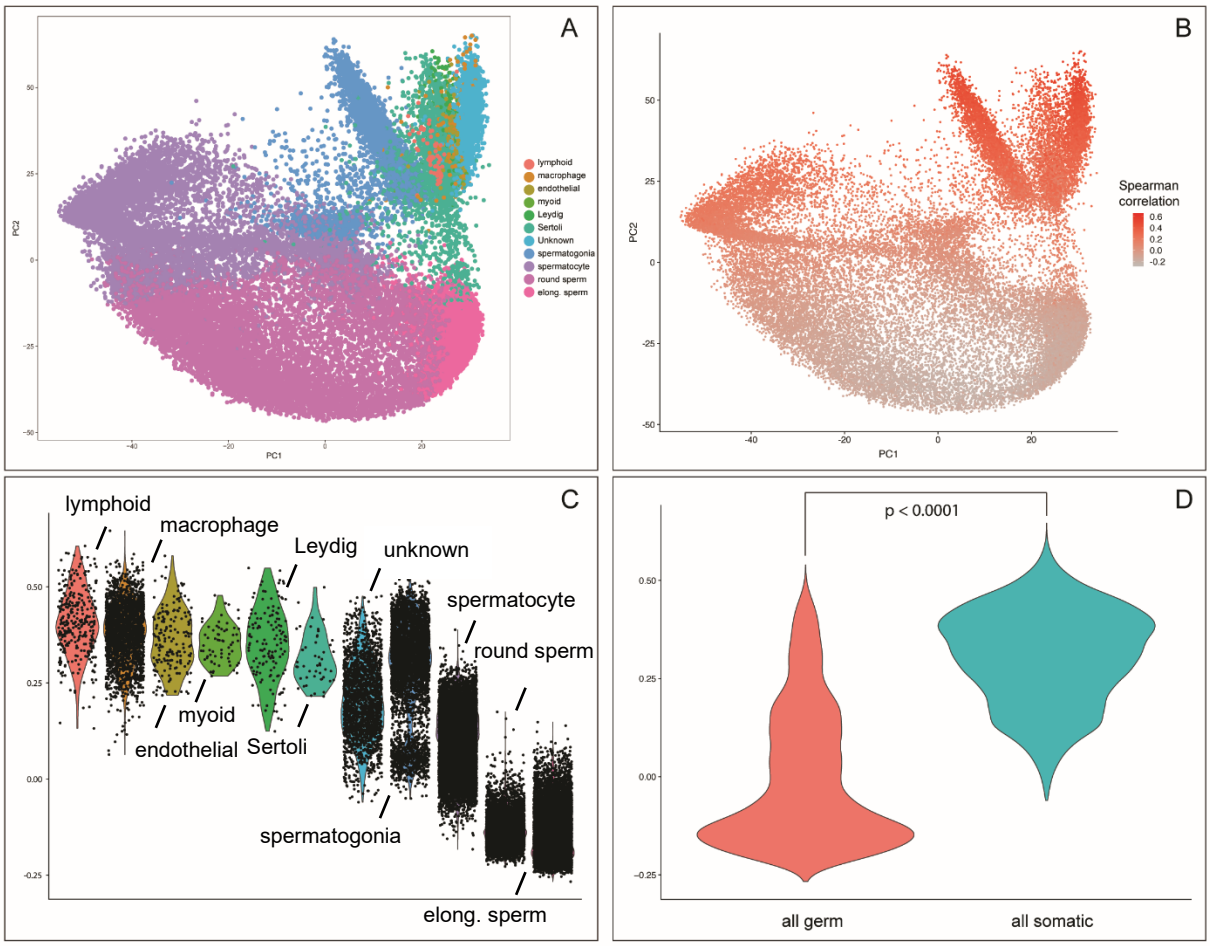

Figure 1.2. GC-1 cells are transcriptionally more similar to testis somatic cells than to testis germ cells. (A) Re-analysis of data from Green and colleagues confirmed the accuracy of the analysis pipeline for Principle Component Analysis. (B) Spearman correlation was used to compare the transcriptomes of the cell types reported by Green and colleagues with the transcriptome of GC-1 cells determined by bulk-sequencing. Intensity of color is proportional to correlation of transcriptomes. (C) The Spearman correlations for each cell type in the single-cell dataset are shown individually. Black dots are individual cell transcriptomes. (D) The Spearman correlations for all germline-derived cell types is compared to all somatic-derived cell types. P-value from Student's t-test.

2xHAR.238 is a *Gli2* Enhancer Required for Normal Inter-male Behavior

Many human accelerated regions (HARS) are gene enhancers (*16*), and changes in enhancers can underlie evolution, particularly of human-specific traits (*12*, *33*–*35*) (Fig 2.1a). To test the possibility that HAR evolution has contributed to the evolution of human male reproduction, I assessed the enhancer function of 714 HARs described within the human (hg19) and chimpanzee (panTro2) genomes (*13*, *22*) using a massively parallel reporter assay (MPRA) (*17*). Briefly, human and chimp orthologs of HARS were packaged into a single lentivirus library which was used to infect the mouse testes somatic cell line GC-1. Human versus chimp relative expression of each HAR was determined by RNA sequencing of the infected cells (Fig 2.1b; see Methods). Of the HARS that fell within my cutoff for a significant FDR, 18 displayed a human-evolved increase in expression greater than Log2F = 0.2, while 20 displayed a human-evolved decrease of at least that magnitude (Fig. 2.1c).

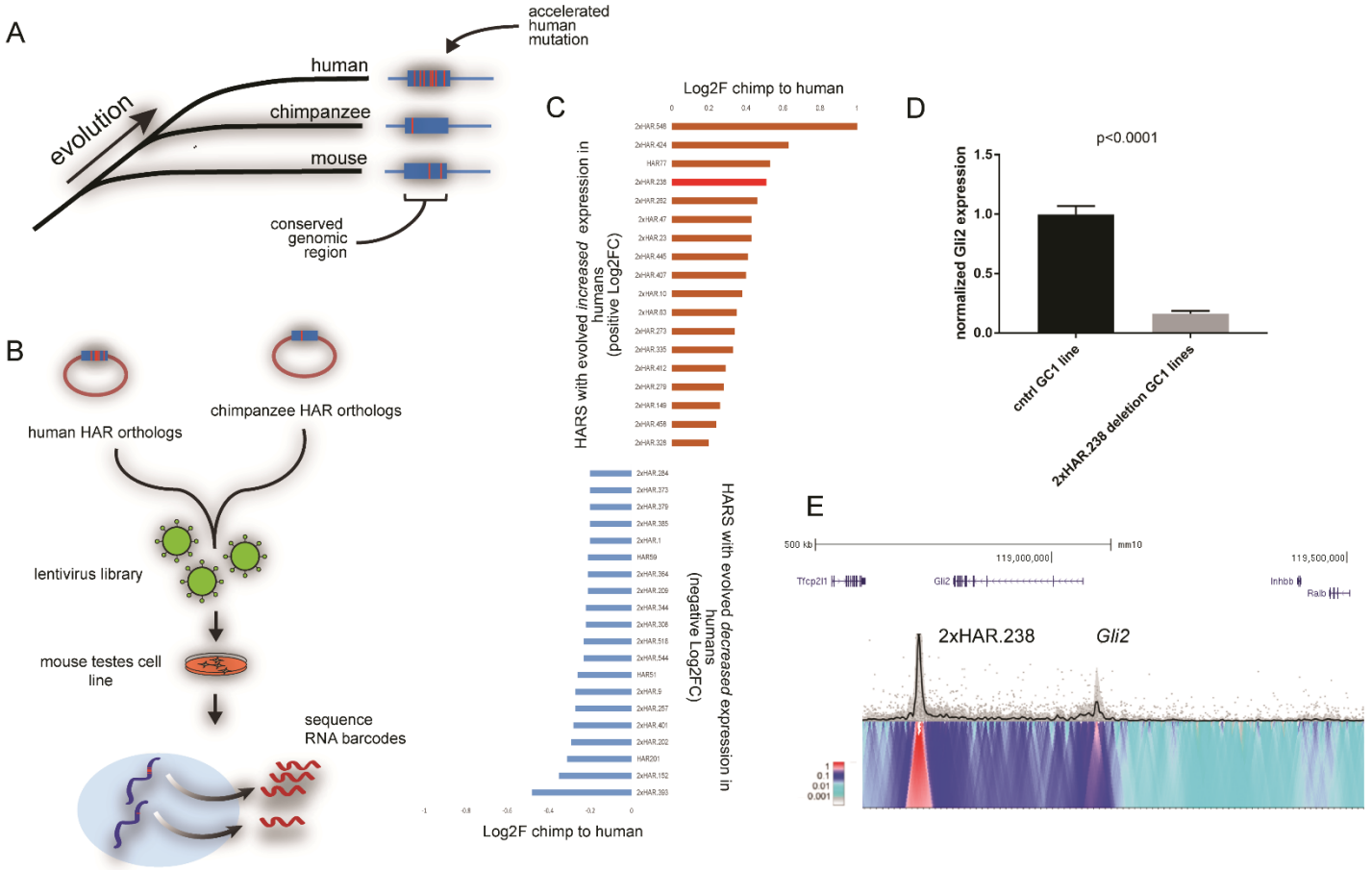

Figure 2.1. A human accelerated region is an enhancer of *Gli2* in testis. (A) Human accelerated regions are defined as sequences highly conserved in evolution but having evolved at an accelerated rate in the human lineage. (B) A library containing all human and chimpanzee HAR orthologs was cloned into a lentiviral library, which was used to infect GC-1 testis cells for the MPRA. (C) All HARS found in the MPRA with an FDR < 0.05 and a fold-change in enhancer activity >0.2 between the human and chimp orthologs are shown. The red bar marks 2xHAR.238. (D) GC-1 lines in which 2xHAR.238 was deleted, as compared to wildtype control lines, had significantly lower expression of *Gli2* by qRT-PCR. Error bars are S.E.M. P-value from Student's t-test. (E) 4C-seq anchored at 2xHAR.238 found a strong interaction with the promoter region of *Gli2* in GC-1 cells.

Among the HARS showing significantly different enhancer activity between orthologs, genes proximal to several are involved in testes function (Table 1.1). I confirmed the differential enhancer activity of these orthologs using a luciferase assay (Fig 2.2). 2xHAR.238 has been previously shown to be an enhancer expressed during embryonic hindbrain development (*16*) and is located ~200 kilobases from the transcription start-site of *Gli2*, a critical effector of the Hedgehog signaling pathway (*24*). Given the importance of the hedgehog signaling pathway for male reproductive function (*26*, *36*), I set out to study the function of 2xHAR.238 in greater detail. Gene enhancers bind transcription factors to control the transcription of genes (*37*). Once bound to transcription activators, enhancers are brought into proximity of gene promoters through the looping of chromatin. To test whether 2xHAR.238 interacts with the *Gli2* promoter in the context of testes gene regulation, I performed chromosome chromatin confirmation capture with sequencing (4C-seq) on mouse GC-1 cells and found that 2xHAR.238 strongly interacts with the *Gli2* promoter region (Fig 2.1e). Thus, 2xHAR.238 interacts with the promoter of *Gli2* in a manner that is consistent with it being an enhancer of *Gli2* in testes.

I hypothesized that if 2xHAR.238 is an enhancer of *Gli2* in testes somatic cells, then disruption of 2xHAR.238 will reduce *Gli2* expression. Using CRISPR/Cas9, I deleted the 2xHAR.238 locus in GC-1 cells and measured *Gli2* expression. Expression of *Gli2* in the 2xHAR.238-deleted cells lines was significantly reduced (Fig 2.1d), indicating that the 2xHAR.238 sequence is necessary for normal *Gli2* expression in GC-1 cells, a testis somatic cell line.

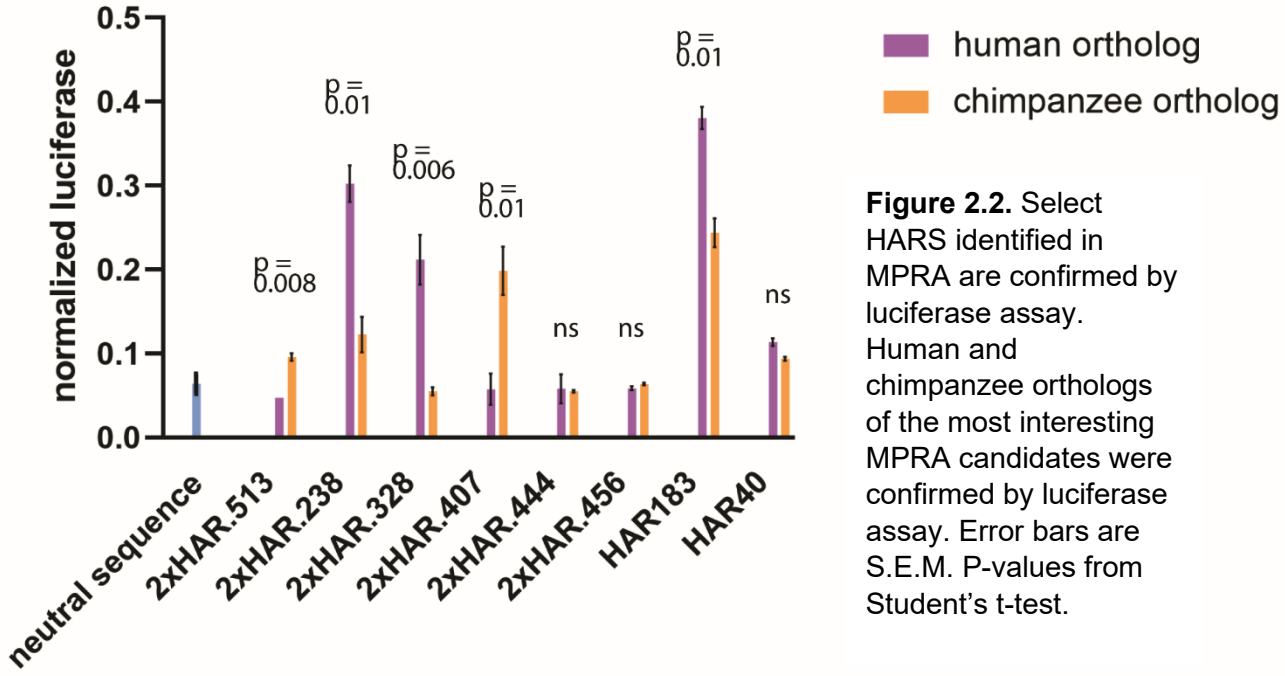

Figure 2.2. Select HARS identified in MPRA are confirmed by luciferase assay. Human and chimpanzee orthologs of the most interesting MPRA candidates were confirmed by luciferase assay. Error bars are S.E.M. P-values from Student's t-test.

Table 1.1. Several HARS Flank Genes Implicated in Testis Biology.

Examples of human accelerated regions with increased enhancer activity in human ortholog	Immediately flanking gene	Relationship to testes function
2xHAR.238	*Gli2 (21)*	Critical Shh transcription factor
2xHAR.444	*Gpr37 (15)*	Deletion seems to perturb Hh signalling in testes
HAR183	*Rabep1 (17)*	Required for ciliary, and potentially flagella, trafficking
2xHAR.513	*Cdk4 (18)*	Required for normal progression through spermatogenesis
HAR40	*Zbtb16 (16)*	Loss leads to progressive failure of spermatogenesis
2xHAR.328	*Nr2f2 (19)*	Required for normal androgen expression in Leydig cells
2xHAR.456	*Pax6 (20)*	Dynamic expression during spermatogenesis

The abrogation of *Gli2* expression upon deletion of 2xHAR.238 in the GC-1 cells suggested that the locus might be required for normal testes development or function. To test this hypothesis, I deleted 2xHAR.238 using CRISPR/Cas9 in mice. *2xHAR.238*$^{-/-}$ mice were viable and exhibited no gross defects in body or testes morphology (data not shown).

During perinatal testis development, Desert hedgehog (DHH) is required for the maturation of Leydig cells, the source of androgens in males (*27*, *36*). In the perinatal period, testosterone is synthesized by fetal Leydig cells to generate an early surge of androgens, sometimes referred to as a “mini-puberty”. This early production of androgens is required for brain masculinization (*38*, *39*). To investigate the possibility that the 2xHAR.238 enhancer is required for the normal expression of *Gli2* in fetal Leydig cells, I assayed the expression of *Gli2* RNA in peri-natal testes of *2xHAR.238*$^{-/-}$ and wildtype mice by in situ hybridization. Simultaneously I assayed expression of CYP17A1, a marker of Leydig cells, by immunofluorescence. *Gli2* expression was reduced in the Leydig cells of *2xHAR.238*$^{-/-}$ perinatal testes as compared to wildtype control (compare Figs 2.3a and 2.3b), indicating that 2xHAR.238 is required for normally timed *Gli2* expression in fetal Leydig cells *in vivo*. These expression differences are quantified in Fig 2.3c and Fig 2.3d, showing that *Gli2* RNA expression is significantly reduced in *2xHAR.238* $^{-/-}$ testes as compared to wildtype controls.

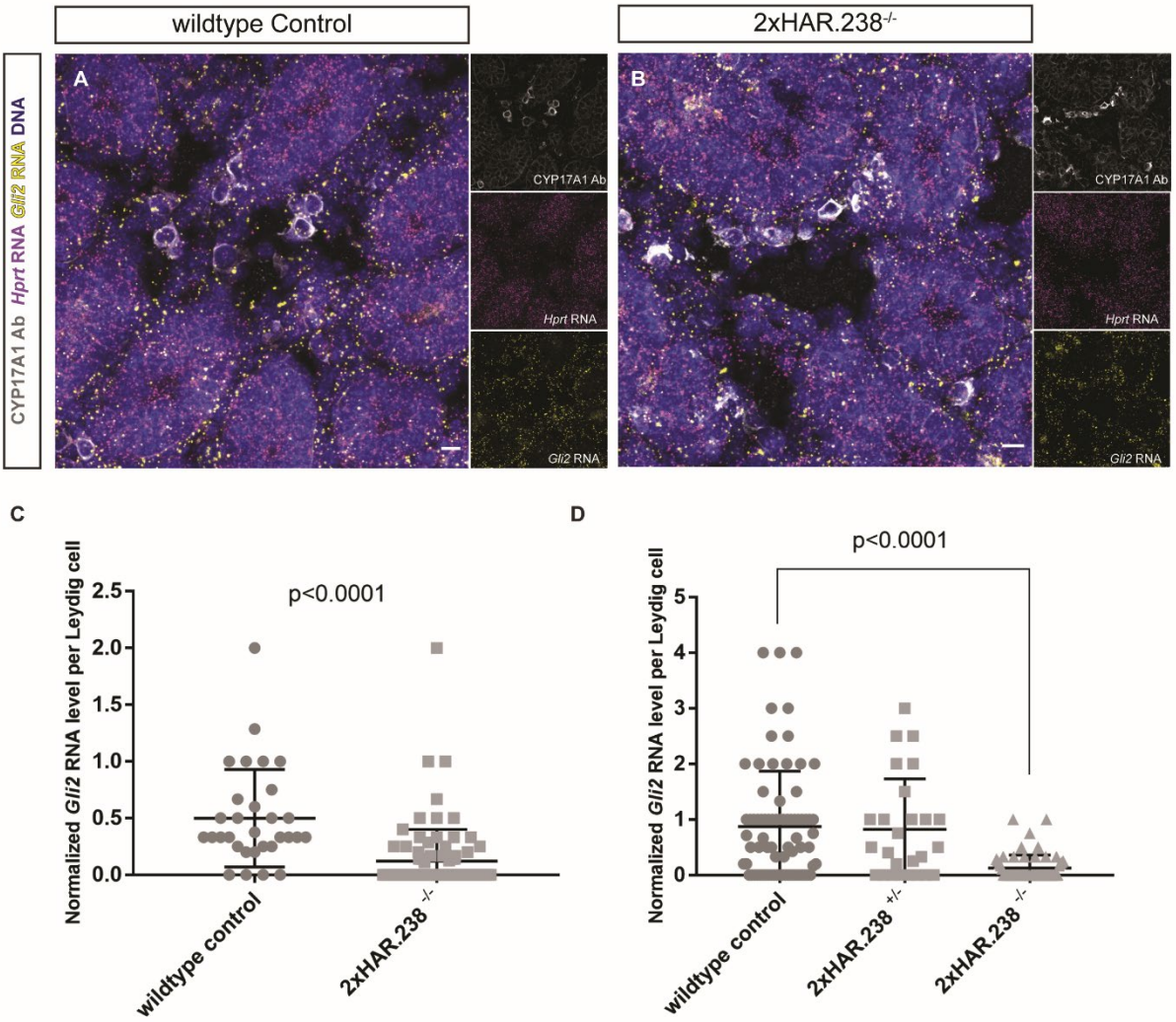

Figure 2.3. 2xHAR.238 is required for the normal expression of *Gli2* in perinatal mouse Leydig cells. RNAscope in situ hybridizations were performed on sections of perinatal mouse testes. (A) is wildtype control and (B) is 2xHAR.238 $^{-/-}$. CYP17A1 antibody signal (white) denotes the Leydig cells. RNA is visualized as puncta. *Gli2* RNA is yellow and *Hprt* RNA is magenta. (C) *Gli2* RNA signal is normalized to *Hprt* housekeeping RNA signal in P0 mouse testes. Individual dots represent single Leydig cells. (n = 2 testes). (D) Same analysis as (C), but testes from P6 mouse (n = 3 testes). Error bars are S.E.M. P-values from Welch's t-test.

Normal perinatal testosterone is required to masculinize the mammalian hypothalamus and enable male-typical behaviors (*40*). The activity of androgens, produced by the testes and circulating in the developing brain during a discrete time window in the perinatal period, has been shown to be critical for masculinizing behavior in male mammals (*41–43*). If the *2xHAR.238*$^{-/-}$ mice had perturbed perinatal testosterone, I reasoned that this could affect inter-male behavior. To test this hypothesis, I utilized the rodent Resident Intruder assay (see Methods for behavioral assay descriptions). Briefly, *2xHAR.238*$^{-/-}$ mice or wildtype controls ("residents") were placed in a cage for one day to establish a home territory. On days 2, 3 and 4 the residents were challenged with an intruder male and responses were recorded (Fig 2.4f). The interactions of the resident and intruder males were scored for specific behaviors including mounting, sniffing, licking, circling, tail grabs, and pushing by the intruder male. The sum of these behaviors by an intruder in a single session is expressed as an "intruder dominance score". Interestingly, intruders in the cages of *2xHAR.238* $^{-/-}$ males had a significantly higher Intruder dominance score than those intruding on wildtype controls on the 4th day, indicating that the mutant males were more likely to be dominated by an intruding male in their cage (Fig 2.4g). To further assess the inter-male behavior of *2xHAR.238* $^{-/-}$ males, experimental mice were tested using the standard Social Approach assay. In this test, a second male mouse is placed in a cage adjoining that of the subject mouse with free airflow, and the subject mouse can choose to investigate the novel male or retreat (Fig 2.4a). Consistent with altered inter-male behavior, *2xHAR.238*$^{-/-}$ mice were significantly more likely than wildtype controls to socially interact with the novel male, both in terms of number of interaction bouts and the mean interaction time per bout (Fig 2.4b-e). Importantly, subject mice did not display any significant defects in locomotion or olfaction (Fig 2.5).

In addition to control of inter-male behavior, testosterone acts upon the hypothalamus to program male-typical mating behavior (*40*). To assess the interest of *2xHAR.238*$^{-/-}$ mice in

sexually receptive females, subject mice were placed in a cage with a receptive female and observed on 3 consecutive days. *2xHAR.238*$^{-/-}$ males engaged in significantly more following bouts across all 3 days than wildtype controls, but did not spend more time following females overall, indicating that mutant mice spent less time following females in each bout (Fig 2.6). This may indicate a reduced persistence of sexual interest in receptive females as compared to wildtype control males.

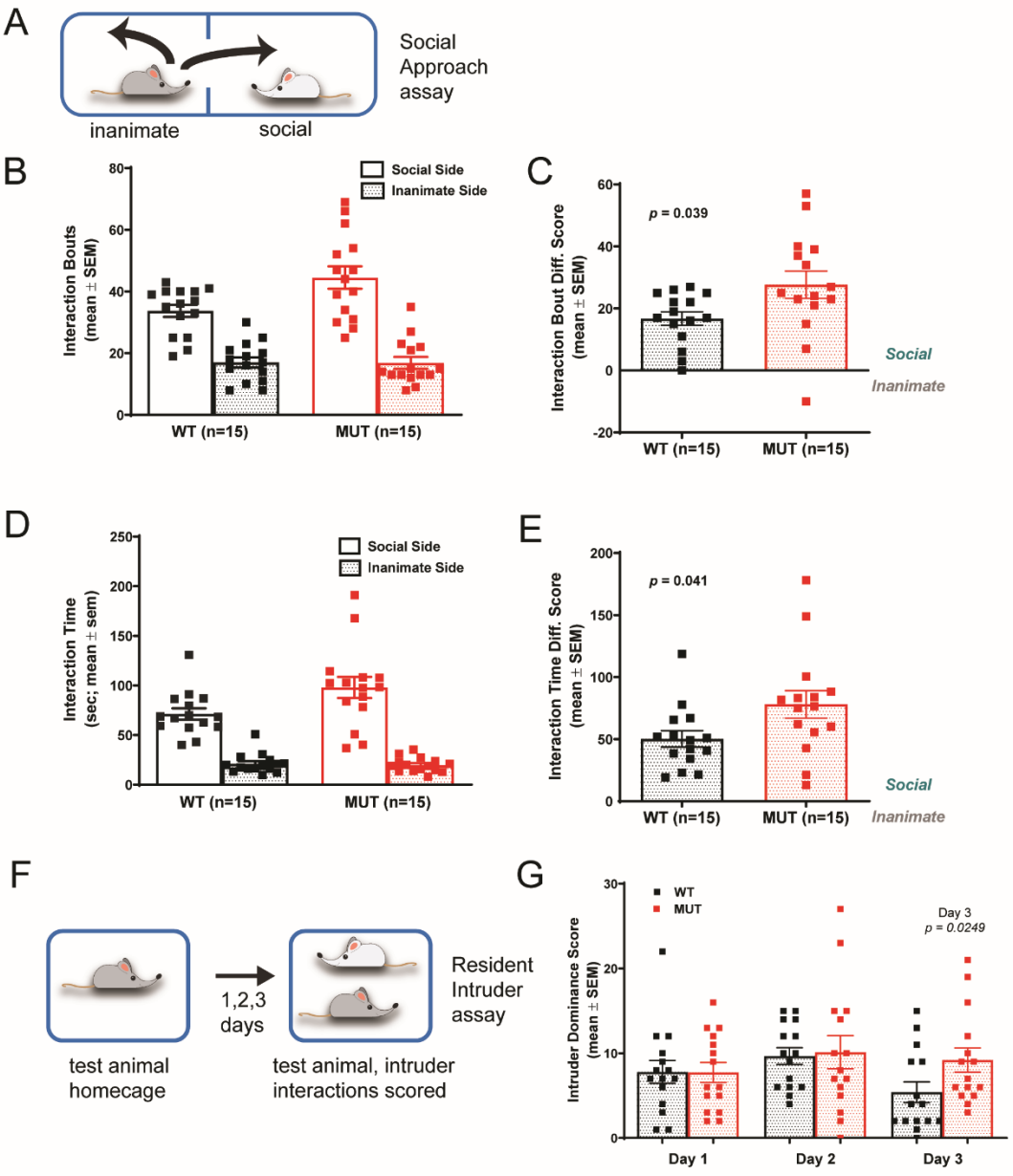

Figure 2.4. 2xHAR.238 is required for normal inter-male social behavior. (A) In the Social Approach assay, the subject mouse can choose to remain in the "inanimate" chamber or move to the "social chamber", where a stimulus mouse is confined. (B,C) *2xHAR.238*$^{-/-}$ mice chose to engage in significantly more interaction bouts with the stimulus mouse as compared to wildtype controls. (D, E) Mutant mice also chose to spend significantly more time overall interacting with the stimulus mouse. (F) In the Resident Intruder assay, the subject mouse establishes a home cage and is challenged with an intruder male. Mounting, sniffing, licking, circling, tail grabs, and pushing by the intruding male is recorded. (G) On Day 3, male mice intruding on *2xHAR.238*$^{-/-}$ mice engaged in significantly more of these behaviors than those intruding on wildtype control males. P-values for Social Approach from Welch's t-test. P-value for Resident Intruder assay from Mann Whitney test.

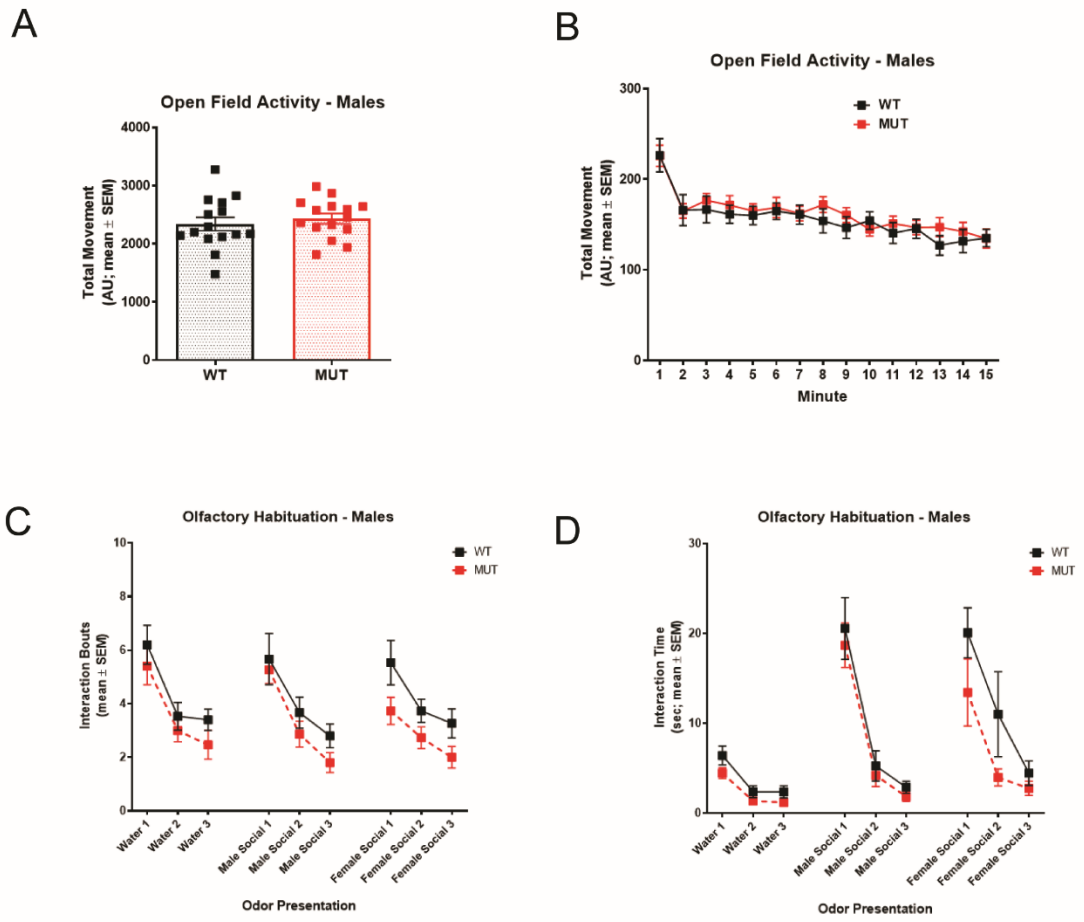

Figure 2.5. Disruption of 2xHAR.238 does not affect normal locomotion or olfaction. (A,B) No difference was observed between *2xHAR.238*$^{-/-}$ male mice and wildtype controls in the Open Field Activity test, which assays locomotion. (C,D) Both *2xHAR.238*$^{-/-}$ and wildtype control males could distinguish between water and mouse bedding odor ("social" odor) and there were no significant differences between genotypes. However, mutant mice may display a trend toward faster habituation to female social odor.

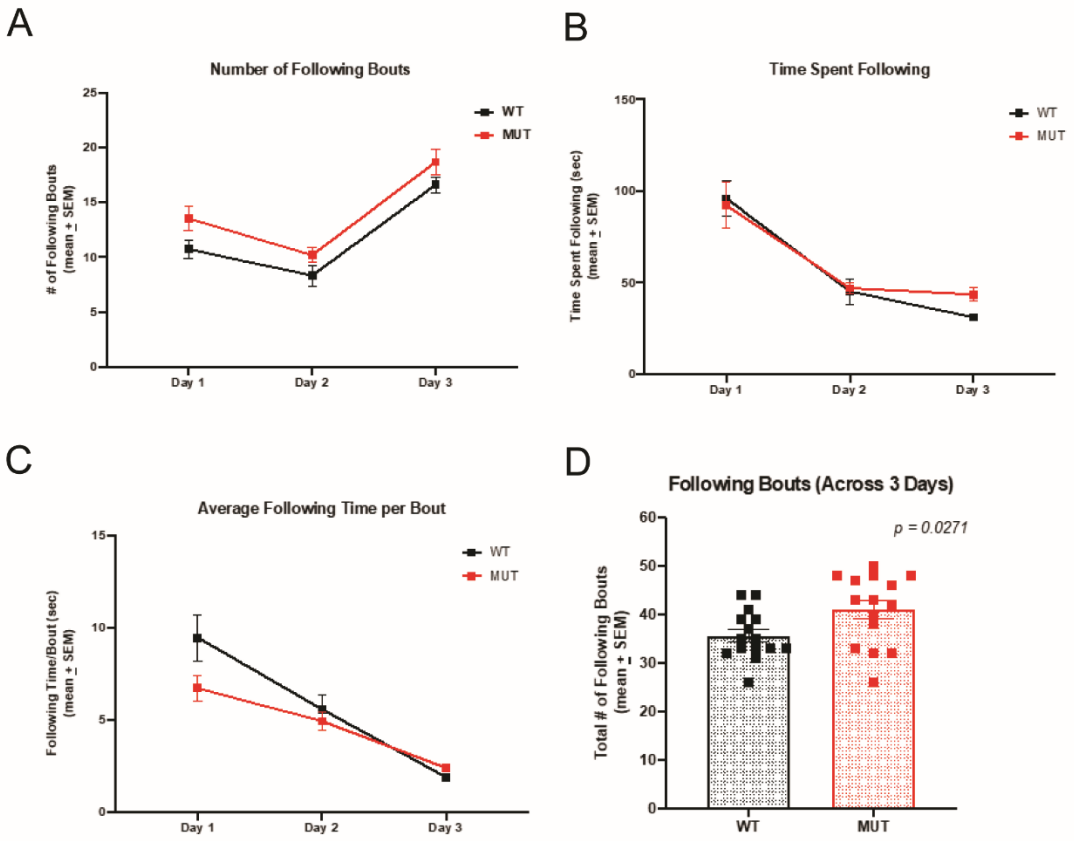

Figure 2.6. 2xHAR.238 may be required for normal persistent male interest in receptive females. (A,C,D) Mutant males engaged in more following bouts of receptive females, but did not spend more overall time following females (B), indicating that these males spent less time following females per bout. P-value from Welch's t-test.

2xHAR.238 is Required for Normal Vascular Development in the Testis

The tubules of the mammalian testis are surrounded by a dense network of capillaries that provide the necessary blood circulation for sperm development (*44*). These endothelial cells, along with the Leydig cells with which they are in close association, make up a significant fraction of the testis interstitial space (Fig. 3.1) I have previously discussed the observed loss of normal *Gli2* expression in the testis of perinatal *2xHAR.238* $^{-/-}$ mice, and normal hedgehog pathway activity critical to normal Leydig cell biology. While no gross defects were observed in the testes of the perinatal mice, aberrant distribution of Leydig cells was seen in the testes of adult *2xHAR.238* $^{-/-}$ mice (Fig. 3.2c,d) as compared to wildtype control mice (Fig. 3.2a,b), indicating spatial disorganization of the adult testis.

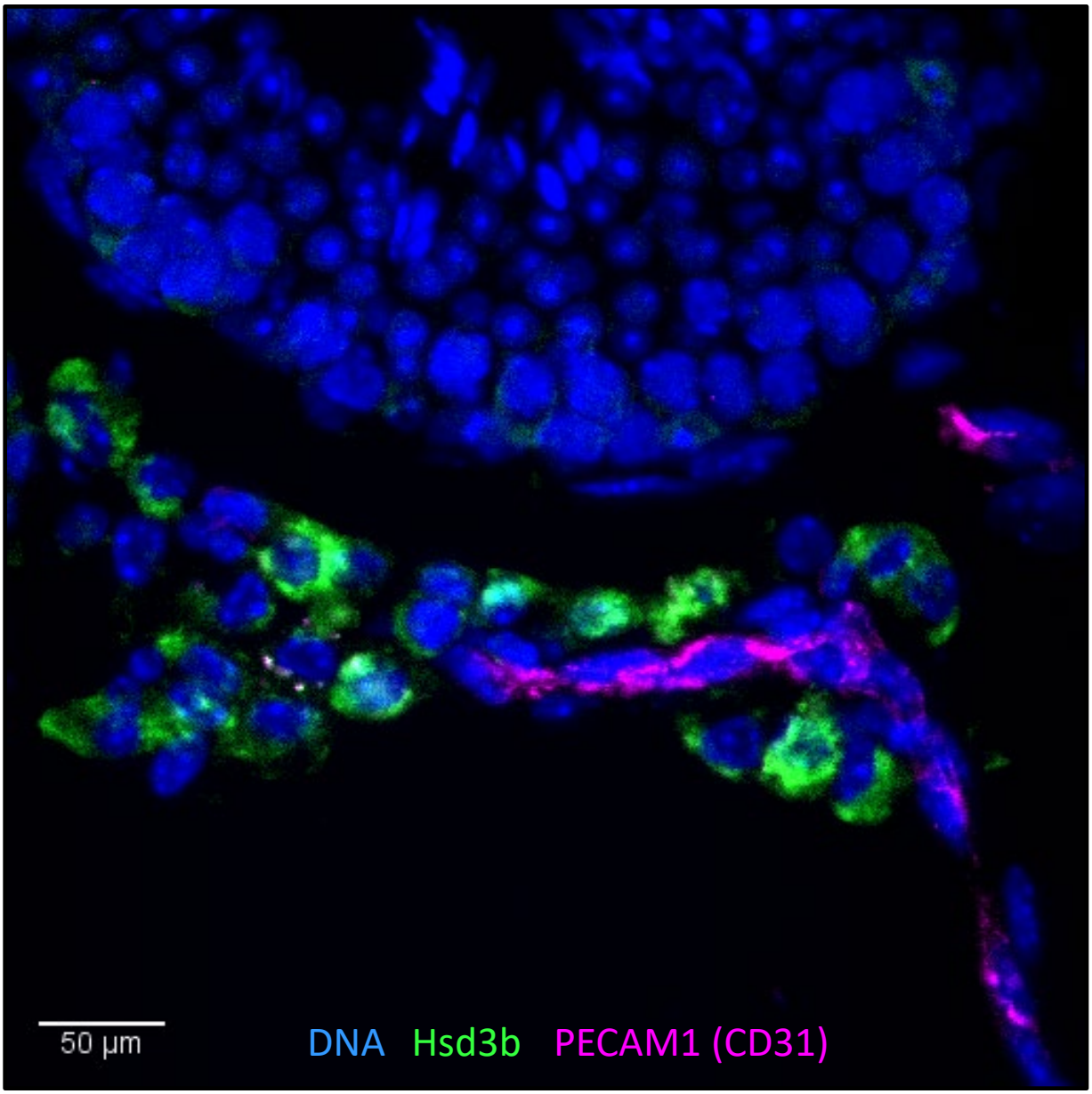

Figure 3.1. Leydig cells are in close association with the interstitial testis vasculature. A 12 µm section of adult mouse testis presented as a maximum projection. Blue is Hoechst stain marking the DNA, green is anti-Hsd3b immunostain marking the Leydig cells and magenta is anti-PECAM1 immunostain marking the endothelial cells of the interstitial capillaries.

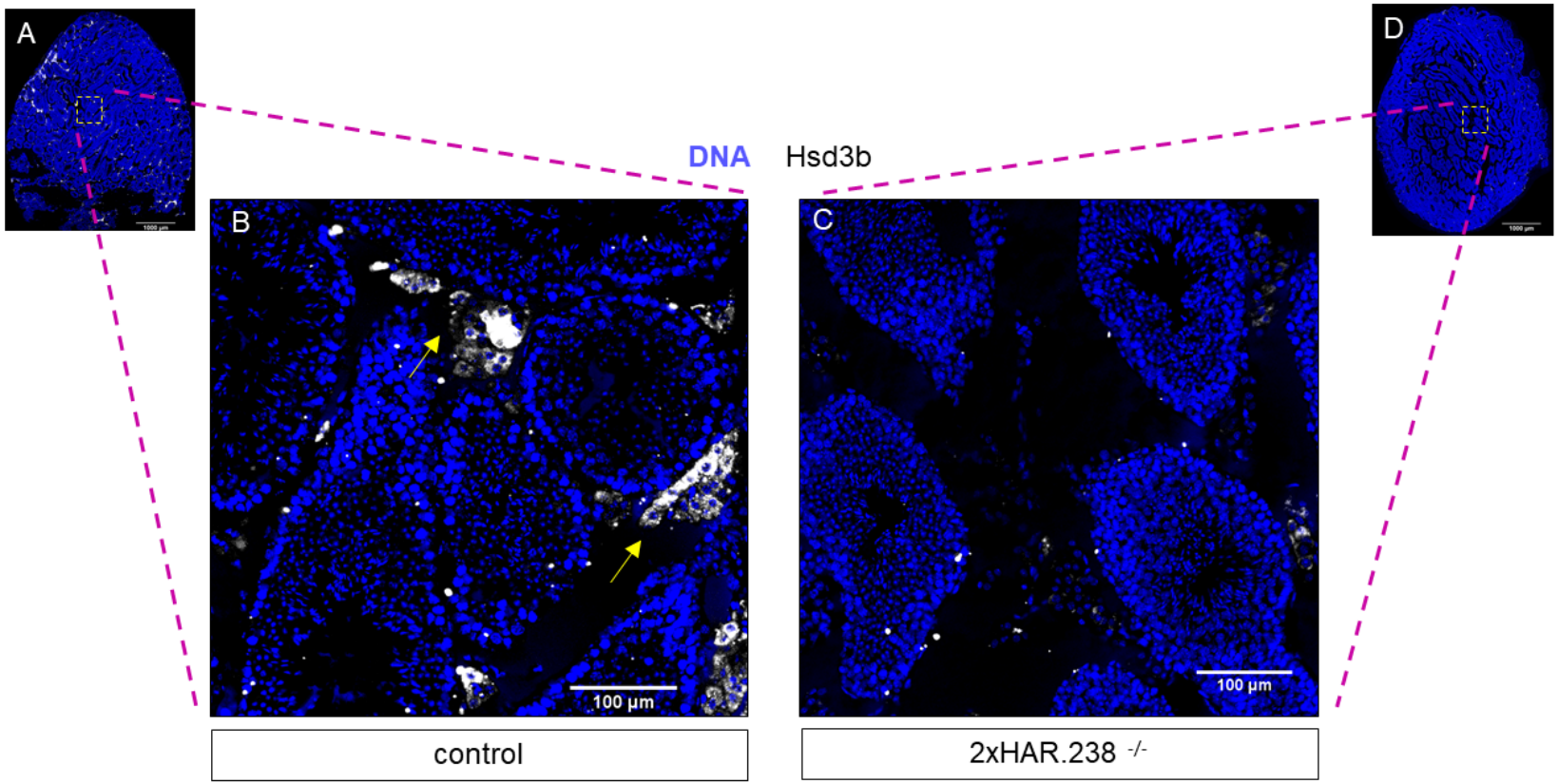

Figure 3.2. Leydig cell distribution appears broadly disturbed in adult 2xHAR.238 $^{-/-}$ testes. (A, D) Whole 12 µm sections of adult mouse testes presented as maximum projections. Blue is Hoechst stain marking DNA. (B, C) Magnified in-sets, white is Hsd3b marking Leydig cells. Yellow arrows indicate large Leydig cell clumps rarely seen in 2xHAR.238 $^{-/-}$ testes.

Given the close association of the Leydig and endothelial cells, I investigated the possibility of blood vessel disorganization in the *2xHAR.238* $^{-/-}$ testes. A defect in blood vessel could impact sperm production and affect the fitness of the organism. Since it is difficult to ascertain subtle defects of blood vessel branching from thin sections in a reproducible way, I cleared and imaged whole-mount testes and derived thick optical sections (see Methods). The testes were harvested from neonates to improve the optical clarity of the tissue (Fig. 3.3a). Upon immunostaining with anti-PECAM1 to visualize vasculature and whole-mount imaging, it was apparent that the vasculature of *2xHAR.238* $^{-/-}$ testes (Fig. 3.3c,e) was aberrant as compared to age-matched wildtype control (Fig. 3.3b,d). I quantified the cleared sections by tracing observable capillary segments per area in wildtype controls and mutants (see Methods) and found that there were significantly more segments in wildtype testes, indicating that 2xHAR.238 is required for normal capillary development in testes (Fig. 3.3f).

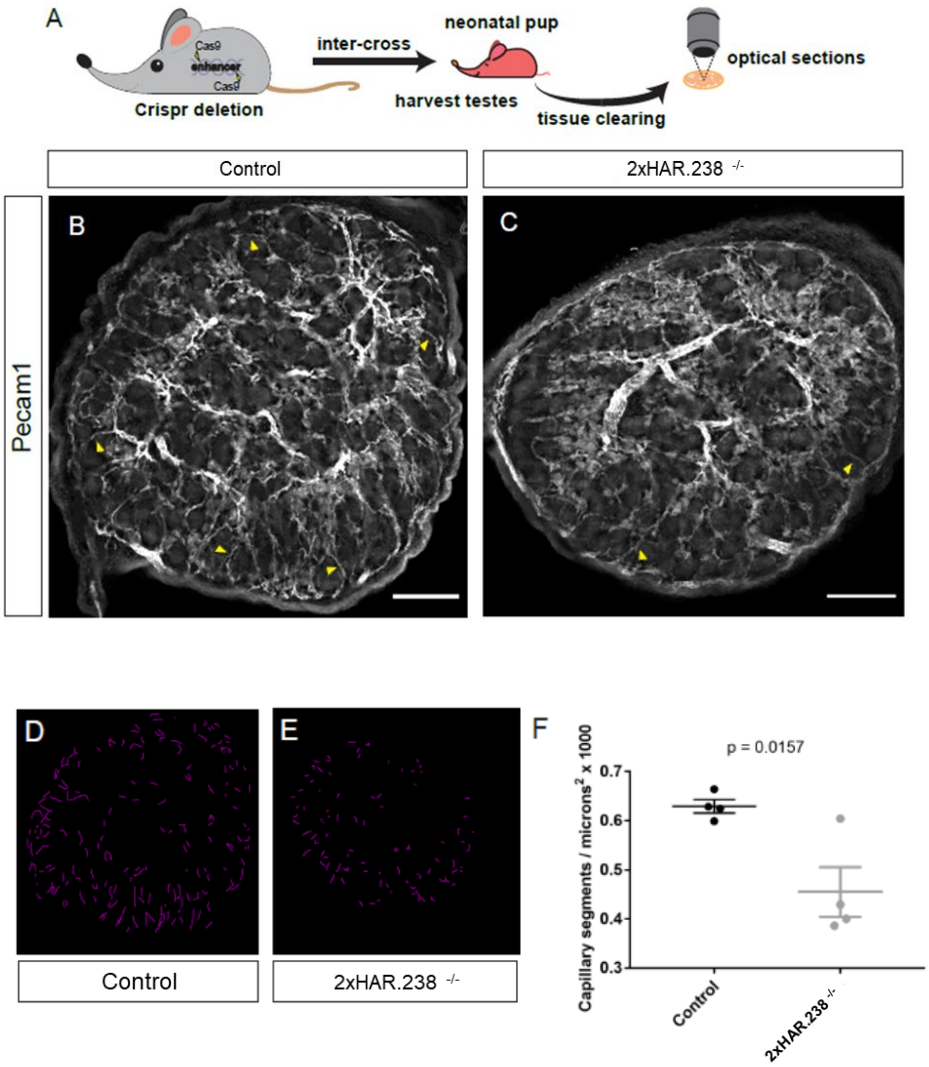

Figure 3.3. 2xHAR.238 is required for normal vascular development in the testis. (A) 2xHAR.238 was deleted in mice using Cripsr/Cas9 and homozygous mutant mice were derived. Testes were harvested from neonatal pups, cleared and optically sectioned using a confocal microscope. (B,C) Single optical sections were taken at consistent depth across samples in cleared wildtype control and 2xHAR.238 $^{-/-}$ mice. (D,C) Observable capillary segments, indicated by anti-PECAM1 immunostaining were hand-traced for each sample. (F) Capillary segments per area were found to be significantly shorter in mutant testes as compared to wildtype controls (Student's t-test). Error bars are S.E.M. Scale bar = 100μm.

Discussion and Future Directions

Here I report the use of an MPRA to discover a set of enhancers, contained within HARS, that have evolved function in testes cells. While conducting these experiments, I discovered that GC-1 cells, previously reported to be spermatocytes (*5*, *45*) have more in common with mouse testis somatic cells and were likely misidentified in the literature. I have shown that a top candidate tested the MPRA assay, 2xHAR.238, is required for normal expression of the hedgehog pathway effector *Gli2* in the GC-1 mouse testis somatic cell line and in perinatal mouse testis Leydig cells. Furthermore, I have demonstrated an interaction between 2xHAR.238 and the *Gli2* promoter-region using 4C-seq. Finally, I have shown that deletion of 2xHAR.238 in mice disrupts normal mouse inter-male. Deletion of the enhancer also disrupts normal testis vascular development.

GC-1 cells were previously reported to be mouse spermatocytes and provide as such by the American Tissue Culture Collection. In my study, through complementary histological and transcriptomic analysis, I have shown that GC-1 cells are more similar to testis somatic cells than to germ cells. The transcriptomic data suggest that GC-1 cells are most similar to Leydig cells or an unknown cell-type near Leydig cells in transcriptional principle component space. However, the strong WT1 expression detected by immunofluorescence is not consistent with adult Leydig cell identity and could suggest an identity more akin to Sertoli cells. WT1 is expressed in the shared, embryonic precursors of Leydig and Sertoli cells (*31*, *46*). Since GC-1 are immortalized cells, it is possible that the line has reverted to a precursor-like state, consistent with both WT1 expression and a transcriptome similar to that of Leydig cells.

In their study identifying the human accelerated regions, Pollard and colleagues (*47*) sought a set of human genome regions that had remained conserved through mammalian

evolution but which had undergone mutational acceleration in the human lineage. This pattern of molecular evolution is consistent with a relaxation of purifying selection, an increase in selection pressure, or both. In this work I have identified an enhancer of *Gli2*, 2xHAR.238, and given that it is a HAR, it follows that it is an enhancer evolved in the human lineage.

I have shown that deletion of 2xHAR.238 reduces *Gli2* expression in the Leydig cells of the testes during the perinatal period, indicating that the enhancer is required for normal *Gli2* expression. Leydig cells are the primary source of androgen production in males (*48*). Previous work has shown that hedgehog pathway activity is required for development of Leydig cells and maturation of Leydig cells is critical for normal androgen production (*26*, *27*, *36*, *49*). It seems plausible, therefore, that disruption of an enhancer required for normal expression of the major hedgehog pathway effector *Gli2* in the testes could affect the normal production of androgens. The activity of androgens, produced by the testes and circulating in the developing brain during the perinatal period, has been shown to be critical for masculinizing behavior in male mammals (*41–43*). Moreover, conditional inactivation of the androgen receptor in the nervous system of mice diminishes stereotypical masculine behavior, such as aggression(*50*). I have demonstrated that deletion of 2xHAR.238 causes an aberrant inter-male behavior in adult mice, characterized by decreased inter-male aggression and increased inter-male sociality. I observed no correlation between 2xHAR.238 genotype and male fertility, as demonstrated by normal fecundity, or between genotype and adult male testosterone levels (data not shown). With these data in mind, I propose that 2xHAR.238 is an enhancer necessary for normal expression of *Gli2* in the perinatal testes. Because of the importance hedgehog pathway activity for normal androgen synthesis, 2xHAR.238 is necessary to establish normal male-specific brain development and adult male-specific behavior (Fig 4.1).

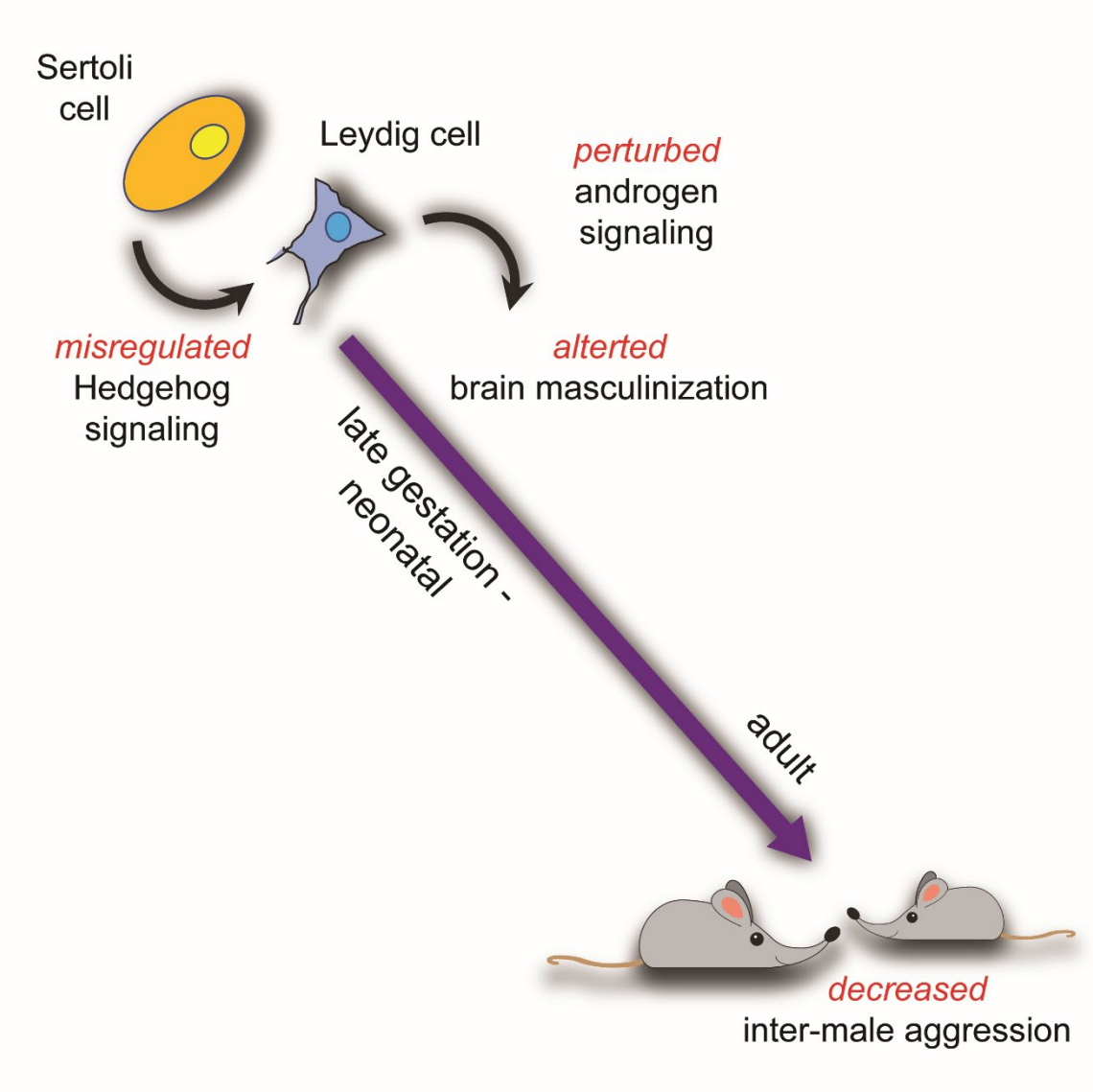

Figure 4.1. A model for the regulation of inter-male behavior by a human-evolved *Gli2* enhancer.

Previous work has established HARs as developmental enhancers (*16*), specifically linked to neurobiology (*17*, *18*), but to my knowledge this is the first example of a single HAR

shown to be required for a normal phenotype. Indeed, it is rare for mutation of single enhancers to generate phenotypes, regardless of their conserved or accelerated status. Perhaps the accelerated-region discovery criteria are well-suited to finding especially potent regions of control, which have been stringently conserved by evolution and modified only when absolutely required by selection pressure. I will be interested to see if other HARS are singularly required for normal phenotypes, as Crispr/Cas9 has enabled facile deletions in the mouse genome. The classic example of a strong enhancer, the mutation of which produces a profound phenotype, is the ZRS, which is required for normal *sonic hedgehog* expression in the developing limb and normal limb outgrowth (*7*). Progressive loss function of the ZRS has been shown to underly the evolution of reduced limbs in snakes (*10*), and reduced expression of the hedgehog receptor *Ptch1*, due to an evolved enhancer, may underly the evolution of bovine limbs (*51*). Here I have presented evidence of a human-evolved enhancer required for the expression of a second member of the hedgehog pathway, *Gli2.* The hedgehog pathway has served repeatedly as the molecular substrate of evolution, and further investigation of putative enhancers of hedgehog pathway members is warranted.

The requirement of 2xHAR.238 for normal testis vascular development is curious. Testis vascular development may be controlled by the hedgehog pathway as it is in the lung (*52*) and yolk sac (*53*), in which case the enhancer deletion might perturb *Gli2* expression in an unknown cell type involved in vascular formation. An alternative explanation could be that hedgehog signal disruption simply affects Leydig cell-cell interactions and/or Leydig cell distribution which, because of the close association of Leydig cells to capillaries, secondarily disrupts vascular development. In any case, to my knowledge there are no reports of testis vasculature variability among primates, so the evolutionary significance of the phenotype is not clear.

However, it *is* clear that chimpanzee males engage in substantially more non-lethal aggression than hunter-gatherer human males (*54*) and display no apparent paternal care (*55*),

indicating evolution of male aggressive behavior in the human lineage. The behavioral phenotype data presented in my study suggest that the molecular evolution of 2xHAR.238 at least partially underlies this. More broadly, my work suggests that the evolution of male sociality across various species may be linked to androgen signaling, via molecular evolution of the hedgehog pathway.

There are two clear paths forward for this research program: one climbs upstream to identify the signaling pathway(s) that impinge upon 2xHAR.238, and the other flows downstream, to unambiguously determine the contribution of 2xHAR.238 molecular evolution to the evolution of the human phenotype.

I did attempt to knock the human and chimpanzee orthologs of 2xHAR.238 into the mouse genomic locus of fertilized oocytes using CRISPR/Cas9, to generate mice. I hoped to ascertain the effect of the human ortholog on the mouse phenotype and contrast that with the effect of the chimpanzee ortholog. I was not able to derive intact human orthologs in the mouse genome from these experiments. It is worth noting that the knock-in vector was co-injected with only a single CRISPR guide RNA species that cut in only one homology-arm of the knock-in vector. These experiments were done several years ago, and this was based on the best available advice at the time. Interestingly, I was able to derive partial human orthologs, and the human sequence was always on the same end targeted by the co-injected guide RNA. Recent work suggests that co-injection with guides targeting both homology arms may allow for more efficient replacement of endogenous sequence with intact knock-in vectors (Aki Ushiki, personal communication). I believe it is worth repeating the knock-in experiments, but with co-injection of guide RNA species that target both homology arms of the knock-in vector.

What are the signaling pathways upstream of 2xHAR.238, the effectors of which presumably bind to the enhancer and support normal Leydig cell biology? In parallel to the MPRA experiment described above, I assayed several MPRA constructs that included

permutations of all 5 missense mutations chimp > human (Fig. 4.2a, each marked by red “x”), in all possible combinations. The purpose was to test the effect of all chimp > human mutations, alone and in combination, on enhancer strength. The A > G mutation (Fig 4.2a, arrow “2”) alone accounted for most of the change, chimp > human, in 2xHAR.238 enhancer strength (data not shown). Intriguingly, this base pair is invariant in a JASPAR (*56*) transcription factor binding motif for SP1. The transcription factor SP1 has been shown to regulate the expression of *Wt1* (*57*), which is expressed in Leydig cell progenitors and required for Sertoli cells (*31*), and in multiple studies it has been linked to steroidal hormone biogenesis and Leydig cell function (*58*–*60*). Moreover, in the ENCODE Transcription Factor ChIP Peaks track data (*61*), SP1 is only one of four factors with peak enrichment overlapping 2xHAR.238 in human H1-ESCs, out of 341 assayed transcription factors (Fig 4.2c). I believe SP1 is an excellent candidate effector for the functional evolution of human 2xHAR.238.

Finally, when examining the genomes of archaic humans, I noticed that the Denisovan genome is identical to modern humans along 2xHAR.238, except for one base pair (Fig 4.2a,b arrow “1”). At this position it has mutated from the shared ancestral genome, G > T. Since this site is invariant in the JASPAR motif, and if the hypothesized connection between the permutation data and the JASPAR invariance is correct, then I would expect that the Denisovan G > T transition reduces the activity of the enhancer.

I will leave you with this: these mysterious cousins of ours, the Denisovans, are only known by a few bone fragments (*62*), some curiously sophisticated jewelry (*63*), and their genome (*64*). While we know nothing yet of their behavior, any knowledge we do glean will likely come from hard-fought, but worthwhile, experiments in functional genomics.

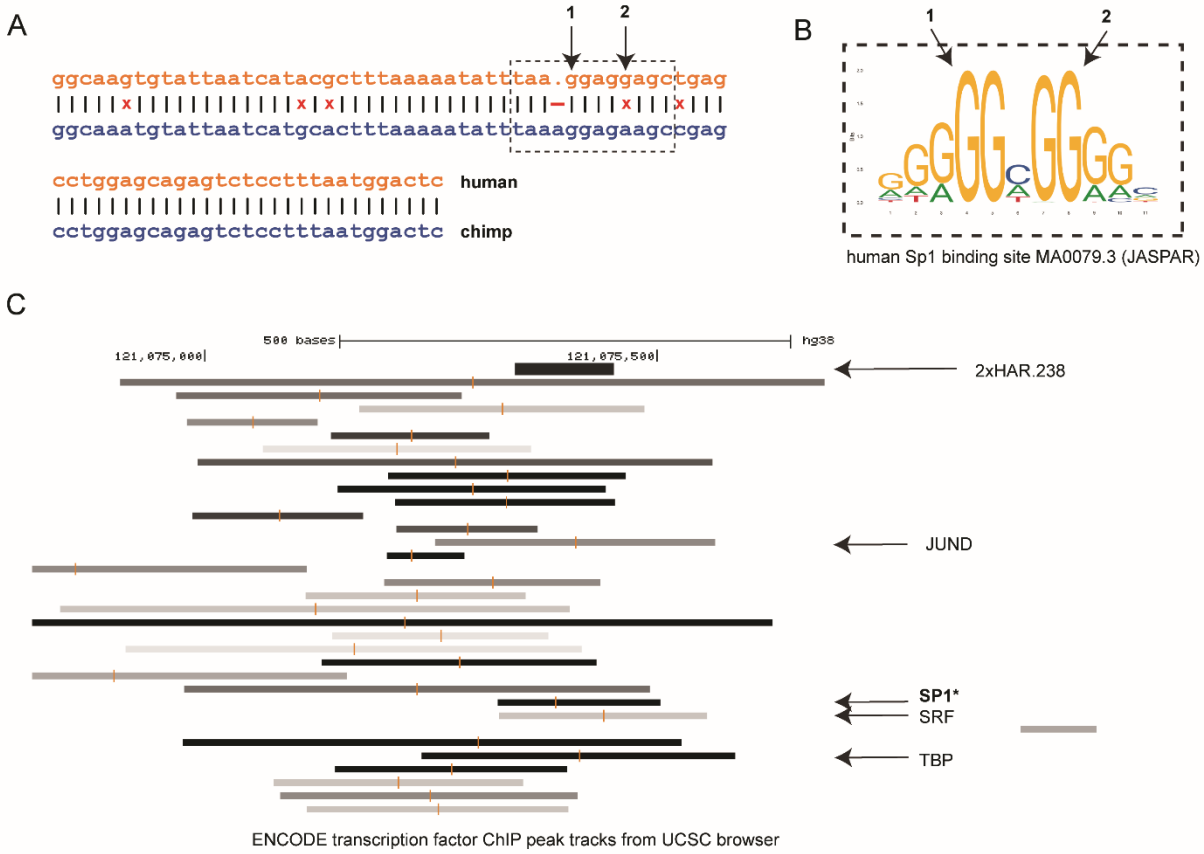

Figure 4.2. Permutation analysis suggests that an SP1 binding site may be restored in the human 2xHAR.238 ortholog. (A) The human (top, orange) and chimpanzee (bottom, blue) genomes are aligned along 2xHAR.238. There are 5 missense mutations from chimp > human, denoted by red "x". The checkered outlines the putative JASPAR SP1 binding site. (B) The JASPAR SP1 binding site MA0079.3 is shown. (C) The UCSC hg38 genome track for the region around 2xHAR.238 is shown. The ENCODE TF ChIP-seq Peaks track is active and set to display all 340 assayed factors found in H1-hESCs for this region. The vertical orange bar marks the point-source of each peak. The darkness of the horizontal bar is proportional to the level of enrichment of the factor. Only 4 factors have peaks centered on 2xHAR.238, and they are noted at the right.

Methods

Immunofluorescence

Testes were harvested from neonatal or adult mice and fixed in 4% PFA overnight, equilibrated in 30% sucrose for an additional night and frozen in OCT (Tissue-Tek). Twelve micron sections were cut from sample blocks and mounted on slides. For samples incubated with anti-DDX4 (Proteintech, Cat# 51042-1-AP), anti-SCP3 (Abcam, Cat# ab97672) and anti-HSD3B (Santa Cruz, Cat# sc-515120) heat-mediated antigen-retrieval was performed prior to blocking (100% ethanol for 3 minutes; 10 mM Sodium citrate, 0.05% Tween 20, pH 6.0 buffer incubation at 95°C for 20 minutes). Samples were blocked in 0.1% Triton X-100, 2% BSA and 1% donkey serum for 1 hour at 4°C, then incubated overnight in blocking solution with anti-HSD3B at 1:100 dilution, anti-PECAM1 (BD Pharmingen, Cat# 550274) at 1:50 dilution, anti-WT1 (Abcam, Cat# ab89901) at 1:100 dilution, anti-DDX4 at 1:500 dilution or anti-SCP3 at 1:250 dilution. The samples were labelled with highly cross-absorbed secondary antibody conjugated to Alexa Fluor 488, 568 or 647 (LifeTech) and 1:1000 Hoechst dye for 45 minutes at RT. Anti-HSD3B is conjugated to Alexa 488 and was only incubated with Hoechst. All samples were mounted in ProLong Diamond for imaging (LifeTech, Cat# P36961).

Bulk RNA sequencing

GC-1 cells were grown to 90% confluency in triplicate, then lysed directly in buffer RLT according to the standard RNeasy kit protocol (Qiagen). Lysate was homogenized with the QIAshredder columns (Qiagen) and processed according to the standard RNeasy kit protocol. RNA integrity was confirmed with a BioAnalyzer (Agilent) and a sequencing library was constructed using the TruSeq PolyA kit (Illumina). Sequencing was performed with an Illumina HiSeq at a length of 50bp. Read alignment was performed using STAR (v2.6.0), with the mouse

GRCm38.88 genome and default settings (*65*). Read counts per coding gene (TPM values) were determined using RSEM (v1.2.25) with the default settings (*66*).

Single-cell Analysis and Bulk Dataset Correlation

Scripts from URL and data from GSE112393 were downloaded and processes as described in (*32*). The genes used for correlation were those determined by Green and colleagues to be markers of the 11 cell types identified in the single-cell dataset (Table S2 in reference). The log(UMI count + 1) for each single cell in the single-cell dataset was then correlated to the log(TPM + 1) of the bulk dataset using Spearman rank correlation. For correlation among the somatic cell types only, the somatic cell dataset was generated using the Somatic_7SomaticCellTypes.R script from (*32*). Marker genes for each cluster were determined by a Likelihood-ratio test, using a 2-fold expression threshold cut-off with the FindAllMarkers function in Seurat (*67*). These marker genes were then used for the Spearman rank correlation as described above.

Whole Mount Testis Clearing, Blood Vessel Imaging and Analysis

Neonatal testes were harvested and fixed as described above but placed in PBS with sodium azide instead of 30% sucrose. Samples were process as described in the iDisco protocol (*68*) with the improvements described for “May 2016” in https://idiscodotinfo.files.wordpress.com/2015/03/idisco-whole-mount-staining-bench-protocol-may-2016.pdf. Samples were immunolabelled with the rat anti-PECAM1 antibody described above at a dilution of 1:100, and the donkey anti-rat Alex 488 antibody described above at a dilution of 1:100. An “n” value of 3 was used, as described in the “May 2016” improvements. After the final methanol dehydration step samples were incubated in ethyl cinnamate (Sigma Cat# 112372-100G) overnight at RT and imaged in ethyl cinnamate. Z-stacks were taken of entire testes with the 10x objective on a Zeiss 800 confocal microscope. A maximum projection

was made of the 3 optical slices in the middle of each sample and this was processed with ImageJ (*69*) as follows: the Enhance Local Contrast tool was run to improve the image contrast, and the Local Thickness tool was used to mask signals too small to be capillaries. The unmasked segments were manually traced with the Overlay Pen. The total area of each testis image was measured, and the number of capillary segments traced was normalized to this value.

Cell culture

Mouse GC-1 testes cells and NIH 3T3 cells were cultured in Dulbecco's Modified Eagle's Medium (Glutamax and HEPES supplementation; Invitrogen Cat# 10564) supplemented with Pen/Strep (10,000 U/mL; Invitrogen Cat# 15140148). Medium was changed every 2 days and cells were passaged upon 90% confluency using Trypsin-EDTA (0.05%; Invitrogen Cat# 25300054). Cells were maintained at 37°C in a 5% CO_2 atmosphere.

MPRA

Library construction, infection and sequencing

Library design/synthesis/cloning and lentivirus library packaging was carried out as described in (*17*). Infection with lentiviral packaged library and DNA/RNA isolation from GC-1 cells and sequencing was conducted in triplicate, and carried out as described in (*17*).

Analysis

RNA and DNA reads were mapped and counted, so that expression of each construct in the library could be normalized to genomic integration. The ratio of RNA reads to DNA reads (ratio score) for each HAR construct was calculated individually for each of the three replicates. Since each HAR construct was represented by multiple fragments, the fragment from each HAR construct with the maximum ratio score was selected to represent the ratio score for that HAR

construct. This was done separately in each library replicate, across both human and chimp HAR construct orthologs. Limma (*70*) was used to determine if the ratio scores observed in one species ortholog across replicates significantly diverged from the ratio score observed for the other. Finally, false discovery rate control was performed, using Benjamini-Hochberg (*71*), to calculate the FDR estimate for the significance of the observed ratio score differences given the number of HARs tested. This value was reported as the negative log of the FDR estimate.

Luciferase Assay

Significant results from the MPRA assay of interest were amplified by PCR and cloned into the pLS-mP-Luc vector (Addgene Cat# 106253) in place of GFP. These included the human and chimpanzee orthologs of 2xHAR.513 (hg19 chr12:59202762-59202994), 2xHAR.238 (hg19 chr2:121549404-121549483), 2xHAR.328 (hg19 chr15:94084865-94085048), 2xHAR.407 (hg19 chr3:140176936-140176987), 2xHAR.444 (hg19 chr7:124454421-124454447) , 2xHAR.456 (hg19 chr11:31993804-31993909), HAR183 (hg19 chr17:6050354-6050576), HAR40 (hg19 chr11:114099775-114100193) and a neutral sequence (hg19 chr2:124107843-124108061) showing no evidence of enhancer activity based on publicly available ENCODE ChIP data (*61*). The renilla control vector used was from (*17*). Lentivirus was generated using 10:1 molar mixtures of firefly-luciferase experimental to renilla-luciferase control vectors for each HAR listed above using a standard protocol outlined in (*72*). NIH 3T3 cells were seeded in a well at 1/10 density and a well of cells was infected with a lentivirus mixture for each tested HAR, in triplicate, 24 hours after seeding. Lucerifase measurements were taken on a Promega Glomax plate reader after sample preparation using the Promega Duel-Luciferase Reporter System (Promega Cat# E1910) 72 hours after lentivirus infection. Enhancer activity was calculated as the ratio of firefly to renilla activity for each construct.

Enhancer deletions

GC-1 cells

Crispr/Cas9 expression constructs were made using the PX459 plasmid (Addgene Cat# 62988) as described in (*73*). Crispr small guide RNAs (sgRNAs) were selected using the GPP sgRNA Designer (*74*). A construct targeting the left flank of 2xHAR.238 was made using the sgRNA sequence GCTGCAGAGAGTCCATTAGA (mm10) and another targeting the right flank using the sequence TGTGATTAATACACTTGCCC (mm10). The sgRNA sequences were selected to maximize predicted on-target cutting and minimize predicted off-target cutting. These two constructs were co-transfected in GC-1 mouse testes cells at 80% confluency with Lipofectamine 3000 (Invitrogen Cat# L3000001) using 1.5µL of Lipofectamine and 250 ng of each vector in a well of a 24-well plate. After 2 days these cells were passaged and treated with puromycin (3µg/mL, Thermofisher Cat# A1113802) for an additional 2 days to reduce the number of un-transfected cells in the culture. Cells in the culture were counted and passaged onto 10 cm cell culture plates at a density of 1000 cells per plate. After 2 weeks, colonies were picked, expanded and Sanger sequenced for genotype using amplicons from PCR primers TAAGCCTAGCATTACAAGAAGGTG and CAATACTCAGACTGTTGGTGGAAAC.

Mice

All animal procedures were approved by the Institutional Animal Care and Use Committee at the University of California, San Francisco. Super-ovulated female C57BL/6 mice (4 weeks old) were mated to C57BL/6 stud males. Fertilized zygotes were collected from oviducts. Cas9 protein (30 - 60 ng/ul) and the same 2xHAR.238-flanking sgRNA species (15-30 ng/ul) used in the GC-1 deletion experiment were co-injected into the pronucleus of fertilized zygotes. Injected zygotes were implanted into oviducts of pseudopregnant CD1 female mice.

Pups were genotyped by Sanger sequencing in the same manner as the GC-1 cells. The 2xHAR.238-deleted mouse line was maintained through intercross of 2xHAR.238$^{-/-}$ animals.

qRT-PCR

RNA was isolated from cells using the Qiagen RNeasy Mini Kit (Cat# 74104) and cDNA was prepared using the iScript cDNA Synthesis Kit (Biorad, Cat# 1708890). Assays were done in triplicate using Express SYBR GreenER master mix (Invitrogen) and an ABI 7900HT RT-PCR machine (Applied Biosystems). Expression levels were normalized to the expression level of the housekeeping gene *HMBS* and quantitative changes were calculated using the method described in (*75*). The primers used for *Gli2* were CCGAGGTGGTCATCTACGAG and AACTCCTTCTTCTCCCCGTG. The primers used for *HMBS* were CAGAGAAAGTTCCCCCACCT and GGATGATGGCACTGAACTCC. qRT-PCR primers were designed using the qPCR Primers track on the UCSC genome browser (*76*).

4C-seq

2x10^7 cells were fixed in PBS with 2% formaldehyde for 10 minutes at RT. Crosslinking was quenched with 125 mM Glycine. The cells were centrifuged and resuspended in lysis buffer (10 mM Tris-HCl pH7.5, 10 mM NaCl, 5 mM MgCl2, 100 mM EGTA, 1x complete proteinase inhibitors (Roche, Cat# 11697498001)) and incubated for 5 minutes on a nutator. Cells were then centrifuged and washed with lysis buffer. The cell pellet was stored at -80°C until use. The cell pellet was suspended in 0.5% SDS and incubated at 62°C for 10 min. 10% Triton X-100 was added to the pellet and incubated at 37°C for 15 min. Chromatin was digested with 100 units of *Nla*III (New England Biolabs, Cat# R0125) at 37°C overnight. After heat inactivation of the enzyme, 2,000 units of T4 DNA ligase (New England Biolabs, Cat# M0202) was added and incubated at room temperature for 4 hours. The ligated chromatin was pelleted and resuspended in EB buffer (Qiagen, Cat# 19086). Proteinase K (New England Biolabs, Cat#

P8107) and 10% SDS were added, mixed by inversion and incubated at 55°C for 30 min. 5M NaCl was added and the sample was incubated at 68°C overnight. DNA was purified by SPRISelect beads (B23319) and resuspended in 50 ul EB buffer. DNA was digested with 50 units of *Dpn*III (New England Biolabs, Cat# R0543) at 37°C for 4 hours. Following heat inactivation of the enzyme at 65°C, DNA was purified by SPRISelect beads. A second round of ligation was performed with 400 units of T4 DNA ligase in a 14 ml reaction volume at 16°C overnight. DNA purified using phenol-chloroform and ethanol precipitation. Inverse PCR was performed using NEBNext high-fidelity 2X PCR master mix (New England Biolabs, catalog no. M0541). DNA was purified with SPRISelect beads. Massively parallel sequencing was performed on an Illumina HiSeq4000 with 50 bp single-end reads or Illumina NovaSeq with 150 bp paired-end reads that were subsequently trimmed to 50 bp for data analysis. 4C-seq was carried out using two biological replicates. 4C-seq data was analyzed using the 4C-seq pipeline (*77*). The primers used for the PCR were:

FP for HAR (4C)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTG
TGATGCCCAAGGGCATG

RP for HAR (4C)

CAAGCAGAAGACGGCATACGAGATTCCGTACGGTGACTGGAGTTCAGACGTGTGCTCTTC
CGATCTAAATCACGGCAGATGGCAGG

In situs hybridizations

Testes were harvested from mouse pups from P0 to P6 and fixed in 4% paraformaldehyde in PBS overnight at 4°C. Testes were washed in PBS, incubated in 30% sucrose in PBS overnight at 4°C and embedded in OCT (Tissue-Tek). Twelve micron sections were cut from the embedded blocks and used for subsequent in situ hybridizations with the

RNAscope Multiplex Fluorescent Reagent Kit v2 (ACD, Cat# 323110), as described in (*78*) with the following modifications: slides were washed twice with PBS at the start of the protocol and dehydrated in 100% ethanol for 3 minutes; antigen retrieval was performed for 15 minutes; RNAscope Protease IV was applied to the sections for less than 1 minute at RT, and immediately washed twice in PBS before beginning the hybridization procedure. The in situ probes used were Mm-Gli2 (ACD, Cat# 405771), Mm-Hprt-C2 (ACD, Cat# 312951-C2), Mm-Hsd3b2 (ACD, Cat# 585381-C3). For sections co-stained with antibody, slides were blocked (2% BSA/1% donkey serum in PBS) for 1 hour at RT immediately following RNA scope protocol. Rabbit anti-Cyp17a1 (Proteintech, Cat# 14447-1-AP) was applied at dilution 1:500 for 1 hour at room temperature, followed by application of donkey anti-rabbit 647 secondary antibody (Thermofisher, cat# A-31573) at dilution 1:500 for 45 minutes at RT. Analysis of RNAscope expression level was done by counting RNAscope in situ signal puncta using the Find Maxima function in ImagJ (*69*): channels with probe signal were subjected to Gaussian Blur using ImageJ (radius = 1), puncta identified with Find Maxima (prominence = 20) and maxima within marker-defined Leydig cells counted.

Behavioral tests

Behavioral assays were conducted by the UCSF Gladstone Mouse Phenotyping Core Facility. All observers were blinded to mouse genotype. C57BL/6 used as experimental animals were aged-matched and bred from 2xHAR.238 $^{-/-}$ x 2xHAR.238 $^{-/-}$ and 2xHAR.238 $^{+/+}$ x 2xHAR.238 $^{+/+}$ parents.

Open-Field Test

Mice were placed in a clear chamber and habituated for 1 hour, then allowed to explore for 15 minutes while movement was automatically monitored. These data were used to establish a

baseline of general locomotor activity, exploration, and anxiety levels. Procedure conducted as described in (*79*–*81*).

Olfactory Habituation Test

Mice were placed in a fresh cage and habituated for 30 minutes. Odors (water, then odor from a male cage, then odor from a female cage) were presented to each mouse on a cotton tip applicator. Time that each mouse spent sniffing each applicator was measured and number of separate investigations with each mouse was counted. Trials were done in triplicate. These data can be used to assess ability to discriminate odors and infer interest in certain types of odors. Procedure conducted as described in (*82*, *83*).

Two-Trial Social Approach Assay

Experimental mice were placed in a two-chambered enclosure for 1 hour prior to trial for acclimation. An opening allowed free movement between chambers. A male BALB/c stimulus mouse was placed in a sub-enclosure in one chamber ("social"). During the trials, the number of bouts of investigation and time spent investigating the social or non-social chambers was recorded. Greater or less interest in the social chamber can indicate social abnormalities. Procedure conducted as described in (*82*, *84*, *85*).

Resident Intruder Assay

This assay is conducted over 3 days. The experimental mouse ("resident") is placed in a home cage that is unchanged for one week. On the first day of the assay, a male BALB/c mouse ("intruder") was placed in the resident home cage for 10 minutes. The occurrence, severity and length of each attack by the resident on the intruder was recorded and scored. Any additional aggressive, defensive and/or fearful behaviors by the resident or the intruder was scored. Note that any excessive attack lasting longer than 30 seconds terminated the trial. On days 2 and 3, the trial was repeated with different intruder mice. The intruder dominance score

is the total bouts of following, sniffing, mounting, licking, circling, tail grabs, and pushing during each 10-minute Resident Intruder session. Detail of procedure in (*83*, *86*, *87*).

Sexual Reproductive Behavior Assay

Experimental mice were introduced to the home cage of a receptive female mouse. Sexual reproductive behaviors are scored over 10 minutes, after which the males are returned to their home cages. Male mice are tested with a different female each day, for three days. Sexual reproductive behaviors being scored include bouts and time of: anogenital sniffing, mounting, chasing, lordosis, intromission, and ejaculation. Significant differences in these behaviors between experimental and control animals can indicate important differences in male sexual behavior. Procedure adapted from (*88*).

References

1. P. O'Shaughnessy, in *Knobil and Neill's Physiology of Reproduction* (Elsevier, 2015; https://linkinghub.elsevier.com/retrieve/pii/B9780123971753000144), pp. 567–594.

2. C. A. Easley, B. T. Phillips, M. M. McGuire, J. M. Barringer, H. Valli, B. P. Hermann, C. R. Simerly, A. Rajkovic, T. Miki, K. E. Orwig, G. P. Schatten, Direct differentiation of human pluripotent stem cells into haploid spermatogenic cells. *Cell Rep*. **2**, 440–446 (2012).

3. A. Steinberger, E. Steinberger, W. H. Perloff, MAMMALIAN TESTES IN ORGAN CULTURE. *Exp. Cell Res.* **36**, 19–27 (1964).

4. T. Sato, K. Katagiri, A. Gohbara, K. Inoue, N. Ogonuki, A. Ogura, Y. Kubota, T. Ogawa, In vitro production of functional sperm in cultured neonatal mouse testes. *Nature*. **471**, 504–507 (2011).

5. M. C. Hofmann, S. Narisawa, R. A. Hess, J. L. Millán, Immortalization of germ cells and somatic testicular cells using the SV40 large T antigen. *Exp. Cell Res.* **201**, 417–435 (1992).

6. S. R. Eddy, The C-value paradox, junk DNA and ENCODE. *Curr. Biol.* **22**, R898-899 (2012).

7. L. A. Lettice, S. J. H. Heaney, L. A. Purdie, L. Li, P. de Beer, B. A. Oostra, D. Goode, G. Elgar, R. E. Hill, E. de Graaff, A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum. Mol. Genet.* **12**, 1725–1735 (2003).

8. J.-W. Hong, D. A. Hendrix, M. S. Levine, Shadow enhancers as a source of evolutionary novelty. *Science*. **321**, 1314 (2008).

9. M. Osterwalder, I. Barozzi, V. Tissières, Y. Fukuda-Yuzawa, B. J. Mannion, S. Y. Afzal, E. A. Lee, Y. Zhu, I. Plajzer-Frick, C. S. Pickle, M. Kato, T. H. Garvin, Q. T. Pham, A. N. Harrington, J. A. Akiyama, V. Afzal, J. Lopez-Rios, D. E. Dickel, A. Visel, L. A. Pennacchio, Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature*. **554**, 239–243 (2018).

10. E. Z. Kvon, O. K. Kamneva, U. S. Melo, I. Barozzi, M. Osterwalder, B. J. Mannion, V. Tissières, C. S. Pickle, I. Plajzer-Frick, E. A. Lee, M. Kato, T. H. Garvin, J. A. Akiyama, V. Afzal, J. Lopez-Rios, E. M. Rubin, D. E. Dickel, L. A. Pennacchio, A. Visel, Progressive Loss of Function in a Limb Enhancer during Snake Evolution. *Cell*. **167**, 633-642.e11 (2016).

11. M. Rebeiz, J. E. Pool, V. A. Kassner, C. F. Aquadro, S. B. Carroll, Stepwise Modification of a Modular Enhancer Underlies Adaptation in a Drosophila Population. *Science*. **326**, 1663–1667 (2009).

12. C. Y. McLean, P. L. Reno, A. A. Pollen, A. I. Bassan, T. D. Capellini, C. Guenther, V. B. Indjeian, X. Lim, D. B. Menke, B. T. Schaar, A. M. Wenger, G. Bejerano, D. M. Kingsley, Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature*. **471**, 216–219 (2011).

13. K. S. Pollard, S. R. Salama, B. King, A. D. Kern, T. Dreszer, S. Katzman, A. Siepel, J. S. Pedersen, G. Bejerano, R. Baertsch, K. R. Rosenbloom, J. Kent, D. Haussler, Forces shaping the fastest evolving regions in the human genome. *PLoS Genet.* **2**, e168 (2006).

14. C. P. Bird, B. E. Stranger, M. Liu, D. J. Thomas, C. E. Ingle, C. Beazley, W. Miller, M. E. Hurles, E. T. Dermitzakis, Fast-evolving noncoding sequences in the human genome. *Genome Biol.* **8**, R118 (2007).

15. S. Prabhakar, J. P. Noonan, S. Pääbo, E. M. Rubin, Accelerated evolution of conserved noncoding sequences in humans. *Science*. **314**, 786 (2006).

16. J. A. Capra, G. D. Erwin, G. McKinsey, J. L. R. Rubenstein, K. S. Pollard, Many human accelerated regions are developmental enhancers. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **368**, 20130025 (2013).

17. H. Ryu, F. Inoue, S. Whalen, A. Williams, M. Kircher, B. Martin, B. Alvarado, Md. A. H. Samee, K. Keough, S. Thomas, A. Kriegstein, J. Shendure, A. Pollen, N. Ahituv, K. S. Pollard, "Massively parallel dissection of human accelerated regions in human and chimpanzee neural progenitors" (preprint, Evolutionary Biology, 2018), , doi:10.1101/256313.

18. R. N. Doan, B.-I. Bae, B. Cubelos, C. Chang, A. A. Hossain, S. Al-Saad, N. M. Mukaddes, O. Oner, M. Al-Saffar, S. Balkhy, G. G. Gascon, Homozygosity Mapping Consortium for Autism, M. Nieto, C. A. Walsh, Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior. *Cell*. **167**, 341-354.e12 (2016).

19. A. Fuentes, How Humans and Apes Are Different, and Why It Matters. *Journal of Anthropological Research*. **74**, 151–167 (2018).

20. M. J. Anderson, A. F. Dixson, Sperm competition: motility and the midpiece in primates. *Nature*. **416**, 496 (2002).

21. H. Fujii-Hanamoto, K. Matsubayashi, M. Nakano, H. Kusunoki, T. Enomoto, A comparative study on testicular microstructure and relative sperm production in gorillas, chimpanzees, and orangutans. *Am. J. Primatol.* **73**, 570–577 (2011).

22. K. Lindblad-Toh, M. Garber, O. Zuk, M. F. Lin, B. J. Parker, S. Washietl, P. Kheradpour, J. Ernst, G. Jordan, E. Mauceli, L. D. Ward, C. B. Lowe, A. K. Holloway, M. Clamp, S. Gnerre, J. Alföldi, K. Beal, J. Chang, H. Clawson, J. Cuff, F. Di Palma, S. Fitzgerald, P. Flicek, M. Guttman, M. J. Hubisz, D. B. Jaffe, I. Jungreis, W. J. Kent, D. Kostka, M. Lara, A. L. Martins, T. Massingham, I. Moltke, B. J. Raney, M. D. Rasmussen, J. Robinson, A. Stark, A. J. Vilella, J. Wen, X. Xie, M. C. Zody, Broad Institute Sequencing Platform and Whole Genome Assembly

Team, J. Baldwin, T. Bloom, C. W. Chin, D. Heiman, R. Nicol, C. Nusbaum, S. Young, J. Wilkinson, K. C. Worley, C. L. Kovar, D. M. Muzny, R. A. Gibbs, Baylor College of Medicine Human Genome Sequencing Center Sequencing Team, A. Cree, H. H. Dihn, G. Fowler, S. Jhangiani, V. Joshi, S. Lee, L. R. Lewis, L. V. Nazareth, G. Okwuonu, J. Santibanez, W. C. Warren, E. R. Mardis, G. M. Weinstock, R. K. Wilson, Genome Institute at Washington University, K. Delehaunty, D. Dooling, C. Fronik, L. Fulton, B. Fulton, T. Graves, P. Minx, E. Sodergren, E. Birney, E. H. Margulies, J. Herrero, E. D. Green, D. Haussler, A. Siepel, N. Goldman, K. S. Pollard, J. S. Pedersen, E. S. Lander, M. Kellis, A high-resolution map of human evolutionary constraint using 29 mammals. *Nature*. **478**, 476–482 (2011).

23. F. Inoue, N. Ahituv, Decoding enhancers using massively parallel reporter assays. *Genomics*. **106**, 159–164 (2015).

24. C. B. Bai, W. Auerbach, J. S. Lee, D. Stephen, A. L. Joyner, Gli2, but not Gli1, is required for initial Shh signaling and ectopic activation of the Shh pathway. *Development*. **129**, 4753–4761 (2002).

25. F. He, P. Akbari, R. Mo, J. J. Zhang, C.-C. Hui, P. C. Kim, W. A. Farhat, Adult Gli2+/-;Gli3Δ699/+ Male and Female Mice Display a Spectrum of Genital Malformation. *PLoS ONE*. **11**, e0165958 (2016).

26. F. Pierucci-Alves, A. M. Clark, L. D. Russell, A developmental study of the Desert hedgehog-null mouse testis. *Biol. Reprod.* **65**, 1392–1402 (2001).

27. A. M. Clark, K. K. Garland, L. D. Russell, Desert hedgehog (Dhh) gene is required in the mouse testis for formation of adult-type Leydig cells and normal development of peritubular cells and seminiferous tubules. *Biol. Reprod.* **63**, 1825–1838 (2000).

28. E. Goldberg, D. Sberna, T. E. Wheat, G. J. Urbanski, E. Margoliash, Cytochrome c: immunofluorescent localization of the testis-specific form. *Science*. **196**, 1010–1012 (1977).

29. D. H. Castrillon, B. J. Quade, T. Y. Wang, C. Quigley, C. P. Crum, The human VASA gene is specifically expressed in the germ cell lineage. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9585–9590 (2000).

30. L. Yuan, J.-G. Liu, J. Zhao, E. Brundell, B. Daneholt, C. Höög, The Murine SCP3 Gene Is Required for Synaptonemal Complex Assembly, Chromosome Synapsis, and Male Fertility. *Molecular Cell*. **5**, 73–83 (2000).

31. Q. Wen, Y. Wang, J. Tang, C. Y. Cheng, Y.-X. Liu, Sertoli Cell Wt1 Regulates Peritubular Myoid Cell and Fetal Leydig Cell Differentiation during Fetal Testis Development. *PLoS ONE*. **11**, e0167920 (2016).

32. C. D. Green, Q. Ma, G. L. Manske, A. N. Shami, X. Zheng, S. Marini, L. Moritz, C. Sultan, S. J. Gurczynski, B. B. Moore, M. D. Tallquist, J. Z. Li, S. S. Hammoud, A Comprehensive Roadmap of Murine Spermatogenesis Defined by Single-Cell RNA-Seq. *Dev. Cell*. **46**, 651-667.e10 (2018).

33. R. P. Zinzen, J. Cande, M. Ronshaugen, D. Papatsenko, M. Levine, Evolution of the Ventral Midline in Insect Embryos. *Developmental Cell*. **11**, 895–902 (2006).

34. J. Cotney, J. Leng, J. Yin, S. K. Reilly, L. E. DeMare, D. Emera, A. E. Ayoub, P. Rakic, J. P. Noonan, The evolution of lineage-specific regulatory activities in the human embryonic limb. *Cell*. **154**, 185–196 (2013).

35. R. N. Doan, B.-I. Bae, B. Cubelos, C. Chang, A. A. Hossain, S. Al-Saad, N. M. Mukaddes, O. Oner, M. Al-Saffar, S. Balkhy, G. G. Gascon, Homozygosity Mapping Consortium for Autism, M. Nieto, C. A. Walsh, Mutations in Human Accelerated Regions Disrupt Cognition and Social Behavior. *Cell*. **167**, 341-354.e12 (2016).

36. H. H.-C. Yao, W. Whoriskey, B. Capel, Desert Hedgehog/Patched 1 signaling specifies fetal Leydig cell fate in testis organogenesis. *Genes Dev.* **16**, 1433–1440 (2002).

37. S. Schoenfelder, P. Fraser, Long-range enhancer-promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).

38. D. W. Pfaff, Morphological changes in the brains of adult male rats after neonatal castration. *J. Endocrinol.* **36**, 415–416 (1966).

39. J. Clarkson, A. E. Herbison, Hypothalamic control of the male neonatal testosterone surge. *Phil. Trans. R. Soc. B*. **371**, 20150115 (2016).

40. K. Hashikawa, Y. Hashikawa, A. Falkner, D. Lin, The neural circuits of mating and fighting in male mice. *Curr. Opin. Neurobiol.* **38**, 27–37 (2016).

41. J. M. Casto, O. B. Ward, A. Bartke, Play, copulation, anatomy, and testosterone in gonadally intact male rats prenatally exposed to flutamide. *Physiol. Behav.* **79**, 633–641 (2003).

42. C. E. Roselli, J. M. Schrunk, H. L. Stadelman, J. A. Resko, F. Stormshak, The effect of aromatase inhibition on the sexual differentiation of the sheep brain. *Endocrine*. **29**, 501–511 (2006).

43. C. H. Phoenix, R. W. Goy, A. A. Gerall, W. C. Young, Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. *Endocrinology*. **65**, 369–382 (1959).

44. T. Svingen, P. Koopman, Building the mammalian testis: origins, differentiation, and assembly of the component cell populations. *Genes & Development*. **27**, 2409–2426 (2013).

45. Y. Wang, W. Song, S. Li, X. Guan, S. Miao, S. Zong, S. S. Koide, L. Wang, GC-1 mRHBDD1 knockdown spermatogonia cells lose their spermatogenic capacity in mouse seminiferous tubules. *BMC Cell Biol.* **10**, 25 (2009).

46. C. Liu, K. Rodriguez, H. H.-C. Yao, Mapping lineage progression of somatic progenitor cells in the mouse fetal testis. *Development*. **143**, 3700–3710 (2016).

47. K. S. Pollard, S. R. Salama, B. King, A. D. Kern, T. Dreszer, S. Katzman, A. Siepel, J. S. Pedersen, G. Bejerano, R. Baertsch, K. R. Rosenbloom, J. Kent, D. Haussler, Forces shaping the fastest evolving regions in the human genome. *PLoS Genet.* **2**, e168 (2006).

48. Q. Wen, C. Y. Cheng, Y.-X. Liu, Development, function and fate of fetal Leydig cells. *Semin. Cell Dev. Biol.* **59**, 89–98 (2016).

49. X. Li, Z. Wang, Z. Jiang, J. Guo, Y. Zhang, C. Li, J. Chung, J. Folmer, J. Liu, Q. Lian, R. Ge, B. R. Zirkin, H. Chen, Regulation of seminiferous tubule-associated stem Leydig cells in adult rat testes. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2666–2671 (2016).

50. K. Raskin, K. de Gendt, A. Duittoz, P. Liere, G. Verhoeven, F. Tronche, S. Mhaouty-Kodja, Conditional inactivation of androgen receptor gene in the nervous system: effects on male behavioral and neuroendocrine responses. *J. Neurosci.* **29**, 4461–4470 (2009).

51. J. Lopez-Rios, A. Duchesne, D. Speziale, G. Andrey, K. A. Peterson, P. Germann, E. Unal, J. Liu, S. Floriot, S. Barbey, Y. Gallard, M. Müller-Gerbl, A. D. Courtney, C. Klopp, S. Rodriguez, R. Ivanek, C. Beisel, C. Wicking, D. Iber, B. Robert, A. P. McMahon, D. Duboule, R. Zeller, Attenuated sensing of SHH by Ptch1 underlies evolution of bovine limbs. *Nature*. **511**, 46–51 (2014).

52. A. C. White, K. J. Lavine, D. M. Ornitz, FGF9 and SHH regulate mesenchymal Vegfa expression and development of the pulmonary capillary network. *Development*. **134**, 3743–3752 (2007).

53. N. Byrd, S. Becker, P. Maye, R. Narasimhaiah, B. St-Jacques, X. Zhang, J. McMahon, A. McMahon, L. Grabel, Hedgehog is required for murine yolk sac angiogenesis. *Development*. **129**, 361–372 (2002).

54. R. W. Wrangham, M. L. Wilson, M. N. Muller, Comparative rates of violence in chimpanzees and humans. *Primates*. **47**, 14–26 (2006).

55. C. Boesch, J. Lehmann, G. Fickenscher, Kin biased investment in wild chimpanzees. *Behav*. **143**, 931–955 (2006).

56. O. Fornes, J. A. Castro-Mondragon, A. Khan, R. van der Lee, X. Zhang, P. A. Richmond, B. P. Modi, S. Correard, M. Gheorghe, D. Baranašić, W. Santana-Garcia, G. Tan, J. Chèneby, B. Ballester, F. Parcy, A. Sandelin, B. Lenhard, W. W. Wasserman, A. Mathelier, JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, gkz1001 (2019).

57. H. T. Cohen, S. A. Bossone, G. Zhu, G. A. McDonald, V. P. Sukhatme, Sp1 is a critical regulator of the Wilms' tumor-1 gene. *J. Biol. Chem.* **272**, 2901–2913 (1997).

58. P. Roumaud, A. Rwigemera, L. J. Martin, Transcription factors SF1 and cJUN cooperate to activate the Fdx1 promoter in MA-10 Leydig cells. *The Journal of Steroid Biochemistry and Molecular Biology*. **171**, 121–132 (2017).

59. Y. Geng, C.-H. Tsai-Morris, Y. Zhang, M. L. Dufau, The Human Luteinizing Hormone Receptor Gene Promoter: Activation by Sp1 and Sp3 and Inhibitory Regulation. *Biochemical and Biophysical Research Communications*. **263**, 366–371 (1999).

60. Z. Liu, E. R. Simpson, Steroidogenic Factor 1 (SF-1) and SP1 Are Required for Regulation of Bovine CYP11A Gene Expression in Bovine Luteal Cells and Adrenal Y1 Cells. *Molecular Endocrinology*. **11**, 127–137 (1997).

61. K. R. Rosenbloom, C. A. Sloan, V. S. Malladi, T. R. Dreszer, K. Learned, V. M. Kirkup, M. C. Wong, M. Maddren, R. Fang, S. G. Heitner, B. T. Lee, G. P. Barber, R. A. Harte, M. Diekhans, J. C. Long, S. P. Wilder, A. S. Zweig, D. Karolchik, R. M. Kuhn, D. Haussler, W. J. Kent, ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res.* **41**, D56-63 (2013).

62. Z. Jacobs, B. Li, M. V. Shunkov, M. B. Kozlikin, N. S. Bolikhovskaya, A. K. Agadjanian, V. A. Uliyanov, S. K. Vasiliev, K. O'Gorman, A. P. Derevianko, R. G. Roberts, Timing of archaic hominin occupation of Denisova Cave in southern Siberia. *Nature*. **565**, 594–599 (2019).

63. A. P. Derevianko, M. V. Shunkov, P. V. Volkov, A PALEOLITHIC BRACELET FROM DENISOVA CAVE*. *Archaeology, Ethnology and Anthropology of Eurasia*. **34**, 13–25 (2008).

64. D. Reich, R. E. Green, M. Kircher, J. Krause, N. Patterson, E. Y. Durand, B. Viola, A. W. Briggs, U. Stenzel, P. L. F. Johnson, T. Maricic, J. M. Good, T. Marques-Bonet, C. Alkan, Q. Fu, S. Mallick, H. Li, M. Meyer, E. E. Eichler, M. Stoneking, M. Richards, S. Talamo, M. V. Shunkov, A. P. Derevianko, J.-J. Hublin, J. Kelso, M. Slatkin, S. Pääbo, Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature*. **468**, 1053–1060 (2010).

65. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. **29**, 15–21 (2013).

66. B. Li, C. N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. **12**, 323 (2011).

67. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

68. N. Renier, Z. Wu, D. J. Simon, J. Yang, P. Ariel, M. Tessier-Lavigne, iDISCO: A Simple, Rapid Method to Immunolabel Large Tissue Samples for Volume Imaging. *Cell*. **159**, 896–910 (2014).

69. T. Ferreira, W. Rasband, ImageJ user guide. *ImageJ/Fiji*. **1**, 155–161 (2012).

70. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*. **43**, e47–e47 (2015).

71. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. **57**, 289–300 (1995).

72. X. Wang, M. McManus, Lentivirus Production. *JoVE*, 1499 (2009).

73. F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, F. Zhang, Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. **8**, 2281–2308 (2013).

74. J. G. Doench, N. Fusi, M. Sullender, M. Hegde, E. W. Vaimberg, K. F. Donovan, I. Smith, Z. Tothova, C. Wilen, R. Orchard, H. W. Virgin, J. Listgarten, D. E. Root, Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol*. **34**, 184–191 (2016).

75. T. D. Schmittgen, K. J. Livak, Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc*. **3**, 1101–1108 (2008).

76. M. Haeussler, A. S. Zweig, C. Tyner, M. L. Speir, K. R. Rosenbloom, B. J. Raney, C. M. Lee, B. T. Lee, A. S. Hinrichs, J. N. Gonzalez, D. Gibson, M. Diekhans, H. Clawson, J. Casper, G. P. Barber, D. Haussler, R. M. Kuhn, W. J. Kent, The UCSC Genome Browser database: 2019 update. *Nucleic Acids Res.* **47**, D853–D858 (2019).

77. P. H. L. Krijger, G. Geeven, V. Bianchi, C. R. E. Hilvering, W. de Laat, 4C-seq from beginning to end: A detailed protocol for sample preparation and data analysis. *Methods*. **170**, 17–32 (2020).

78. H. Wang, N. Su, L.-C. Wang, X. Wu, S. Bui, K.-J. Chang, A. Nielsen, H.-T. Vo, Y. Luo, X.-J. Ma, in *In Situ Hybridization Methods*, G. Hauptmann, Ed. (Springer New York, New York, NY, 2015; https://doi.org/10.1007/978-1-4939-2303-8_21), pp. 405–414.

79. M. Martinez-Losa, T. E. Tracy, K. Ma, L. Verret, A. Clemente-Perez, A. S. Khan, I. Cobos, K. Ho, L. Gan, L. Mucke, M. Alvarez-Dolado, J. J. Palop, Nav1.1-Overexpressing Interneuron Transplants Restore Brain Rhythms and Cognition in a Mouse Model of Alzheimer's Disease. *Neuron*. **98**, 75-89.e5 (2018).

80. S. F. Owen, J. D. Berke, A. C. Kreitzer, Fast-Spiking Interneurons Supply Feedforward Control of Bursting, Calcium, and Plasticity for Efficient Learning. *Cell*. **172**, 683-695.e15 (2018).

81. S.-I. Hagiwara, E. Kaushal, S. Paruthiyil, P. J. Pasricha, B. Hasdemir, A. Bhargava, Gastric corticotropin-releasing factor influences mast cell infiltration in a rat model of functional dyspepsia. *PLoS ONE*. **13**, e0203704 (2018).

82. C. Tai, C.-W. Chang, G.-Q. Yu, I. Lopez, X. Yu, X. Wang, W. Guo, L. Mucke, Tau Reduction Prevents Key Features of Autism in Mouse Models. *Neuron*. **106**, 421-437.e11 (2020).

83. K. Scearce-Levie, E. D. Roberson, H. Gerstein, J. A. Cholfin, V. S. Mandiyan, N. M. Shah, J. L. R. Rubenstein, L. Mucke, Abnormal social behaviors in mice lacking Fgf17. *Genes Brain Behav.* **7**, 344–354 (2008).

84. G. Krabbe, S. S. Minami, J. I. Etchegaray, P. Taneja, B. Djukic, D. Davalos, D. Le, I. Lo, L. Zhan, M. C. Reichert, F. Sayed, M. Merlini, M. E. Ward, D. C. Perry, S. E. Lee, A. Sias, C. N.

Parkhurst, W.-B. Gan, K. Akassoglou, B. L. Miller, R. V. Farese, L. Gan, Microglial NFκB-TNFα hyperactivation induces obsessive-compulsive behavior in mouse models of progranulin-deficient frontotemporal dementia. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 5029–5034 (2017).

85. H. Belinson, J. Nakatani, B. A. Babineau, R. Y. Birnbaum, J. Ellegood, M. Bershteyn, R. J. McEvilly, J. M. Long, K. Willert, O. D. Klein, N. Ahituv, J. P. Lerch, M. G. Rosenfeld, A. Wynshaw-Boris, Prenatal β-catenin/Brn2/Tbr2 transcriptional cascade regulates adult social and stereotypic behaviors. *Mol. Psychiatry*. **21**, 1417–1433 (2016).

86. C. C. Cheung, W. C. Krause, R. H. Edwards, C. F. Yang, N. M. Shah, T. S. Hnasko, H. A. Ingraham, Sex-dependent changes in metabolism and behavior, as well as reduced anxiety after eliminating ventromedial hypothalamus excitatory output. *Mol Metab*. **4**, 857–866 (2015).

87. L. Wang, J. Simms, C. J. Peters, M. Tynan-La Fontaine, K. Li, T. M. Gill, Y. N. Jan, L. Y. Jan, TMEM16B Calcium-Activated Chloride Channels Regulate Action Potential Firing in Lateral Septum and Aggression in Male Mice. *J. Neurosci.* **39**, 7102–7117 (2019).

88. W. T. Swaney, B. N. Dubose, J. P. Curley, F. A. Champagne, Sexual experience affects reproductive behavior and preoptic androgen receptors in male mice. *Horm Behav*. **61**, 472–478 (2012).

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:
Norman Andrew
2478BD29827042A...
Author Signature

8/31/2020
Date