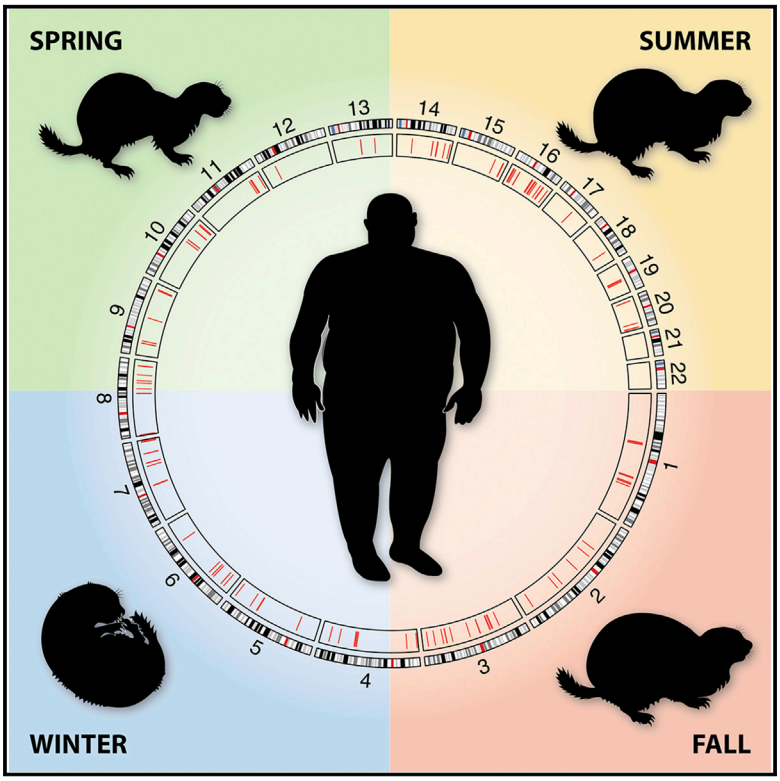
Article

Cell Reports

Parallel Accelerated Evolution in Distant Hibernators Reveals Candidate *Cis* Elements and Genetic Circuits Regulating Mammalian Obesity

Graphical Abstract



Authors

Elliott Ferris, Christopher Gregg

Correspondence

chris.gregg@neuro.utah.edu

In Brief

Obesity is a clinical problem but also an important adaptation in hibernators. By using comparative genomics approaches to analyze the genomes of hibernators from different clades and contrasting the results with human obesity risk loci, Ferris and Gregg found 364 conserved *cis* elements with putative roles in regulating obesity and hibernation.

Highlights

- Discovery of parallel accelerated regions (pARs) in distant hibernators
- Hibernator pARs reveal candidate *cis* elements regulating obesity
- Identification of obesity and hibernation genetic circuits

Ferris & Gregg, 2019, Cell Reports *29*, 2608–2620
November 26, 2019
https://doi.org/10.1016/j.celrep.2019.10.102

CellPress

OPEN ACCESS
CellPress

Cell Reports
Article

Parallel Accelerated Evolution in Distant Hibernators Reveals Candidate *Cis* Elements and Genetic Circuits Regulating Mammalian Obesity

Elliott Ferris[1] and Christopher Gregg[1,2,3,4,*]
[1]Department of Neurobiology and Anatomy, University of Utah, Salt Lake City, UT 84132-3401, USA
[2]Department of Human Genetics, University of Utah, Salt Lake City, UT 84132-3401, USA
[3]New York Stem Cell Foundation Robertson Neuroscience-Investigator
[4]Lead Contact
*Correspondence: chris.gregg@neuro.utah.edu
https://doi.org/10.1016/j.celrep.2019.10.102

SUMMARY

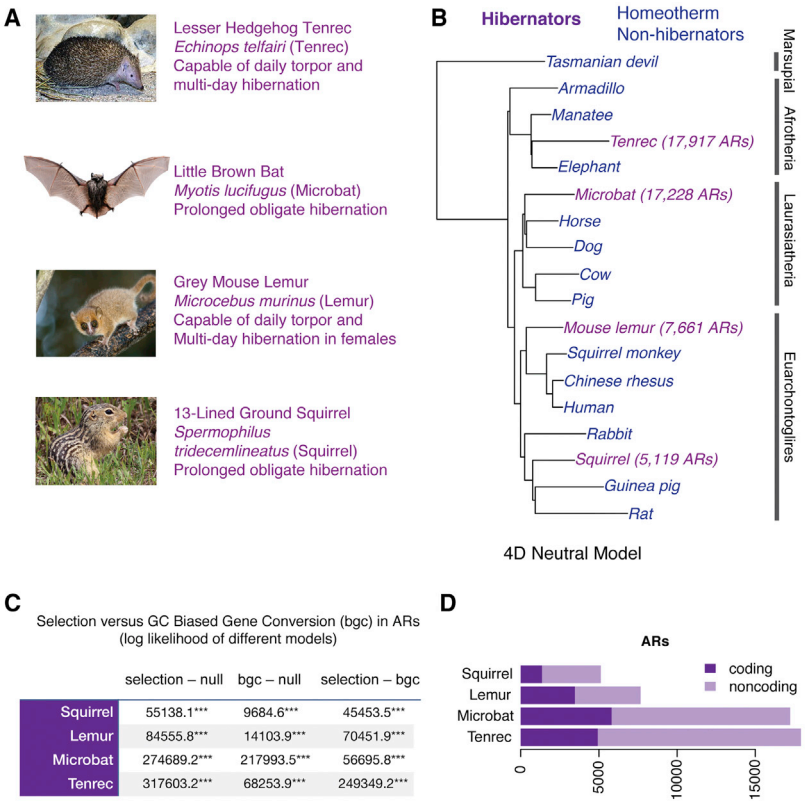
Obesity is a clinical problem and an important adaptation in many species. Hibernating mammals, for example, become obese, insulin resistant, and hyperinsulinemic to store fat. Here, we combine comparative phylogenomics with large-scale human genome data to uncover candidate *cis* elements regulating mammalian obesity. Our study examines genetic elements conserved across non-hibernating mammals to identify genome-wide patterns of accelerated evolution in hibernators from different clades. The results reveal the existence of parallel accelerated regions (pARs) in distant hibernators. Hibernator pARs are disproportionately located near human obesity susceptibility genes compared to random conserved regions, hibernator ARs that are not parallel, and non-hibernator pARs. We found 364 candidate obesity-regulating *cis* elements and genetic circuits in different cell types. The *Fat Mass and Obesity* (*FTO*) locus, the strongest genetic risk factor for human obesity, is an enriched site for hibernator pARs. Our results uncover noncoding *cis* elements with putative roles in obesity and hibernation.

INTRODUCTION

Obesity is a worldwide epidemic and influences risks for a range of different human diseases, including type II diabetes, age-related dementia, cardiovascular disease, and cancer (Hruby and Hu, 2015; Mazon et al., 2017; Procaccini et al., 2016). Thus, there is major interest in defining the mechanisms involved. The factors influencing obesity risks are complex and include genetic risk factors. Previous large-scale human genome-wide association studies (GWASs) uncovered ~250 genomic loci and 123 genes significantly associated with human BMI variance and obesity susceptibility (Ghosh and Bouchard, 2017; Locke et al., 2015; Turcot et al., 2018). The strongest obesity susceptibility locus is the *Fat Mass and Obesity* (*FTO*) gene locus (Dina et al., 2007; Frayling et al., 2007; Loos and Yeo, 2014; Yang et al., 2017). Variants of the *FTO* locus promoting obesity are noncoding and affect long-range regulatory interactions between *cis*-regulatory elements and genes in the *FTO* region, including *IRX3* and *IRX5* (Claussnitzer et al., 2015a; Landgraf et al., 2016; Smemo et al., 2014; Yang et al., 2017). Investigating noncoding *cis*-regulatory mechanisms can help us understand the basis of obesity. We uncover candidate *cis*-regulatory elements and genetic circuits for controlling mammalian obesogenic phenotypes.

We recently demonstrated that comparative, genome-wide analyses of accelerated evolution in species with highly distinctive and clinically relevant phenotypes can reveal candidate regulatory elements controlling those phenotypes (Ferris et al., 2018). Our approach uncovered candidate *cis*-regulatory elements for various phenotypes, such as cancer and mutation resistance from accelerated regions (ARs) in the elephant (Ferris et al., 2018). ARs are best known from studies of human ARs and are evolutionarily conserved elements with significantly increased nucleotide substitution rates due to the effects of positive selection, relaxed purifying selection, or GC-biased gene conversion in a particular lineage (Hubisz and Pollard, 2014; Kostka et al., 2012; Pollard et al., 2010). Elements exhibiting accelerated evolution in a particular lineage are proposed to have roles in shaping the unique traits of that lineage (Bird et al., 2007; Booker et al., 2016; Boyd et al., 2015; Capra et al., 2013; Eckalbar et al., 2016; Hubisz et al., 2011; Kim and Pritchard, 2007; Lindblad-Toh et al., 2011; Pollard et al., 2006a, 2006b, 2010; Prabhakar et al., 2006). Indeed, individual regulatory elements can have important phenotypic effects. For example, changes to regulatory elements contributed to the loss of limbs in snakes and penile spines in humans (Kvon et al., 2016; McLean et al., 2011). Here, we attempt to uncover candidate *cis*-regulatory elements and genetic circuits in the mammalian genome that regulate obesity by comparing genome-wide patterns of accelerated evolution in hibernating mammals.

Hibernator ARs could reveal important *cis*-regulatory elements controlling obesity-related phenotypes because hibernators evolved distinctive behavioral and physiological adaptations, including seasonal regulation of feeding behavior, metabolic activity, and adiposity (Dark, 2005; Martin, 2008). In the autumn season, obligate hibernators increase their body mass by 30%–50% (Dark, 2005; Hohn and Marshall, 1966; Kunz et al., 1998; Martin, 2008). Hibernators become obese, insulin resistant, and hyperinsulinemic, exhibiting some



	selection – null	bgc – null	selection – bgc
Squirrel	55138.1***	9684.6***	45453.5***
Lemur	84555.8***	14103.9***	70451.9***
Microbat	274689.2***	217993.5***	56695.8***
Tenrec	317603.2***	68253.9***	249349.2***

Figure 1. Genome-wide Identification of ARs in Distantly Related Hibernating and Non-hibernating Mammals

(A) Images and descriptions of the distantly related mammalian hibernators in the UCSC mm10-rooted multiple species alignment we tested for accelerated evolution. All four species evolved seasonal obesity for hibernation.

(B) Phylogenetic tree of species in the hibernator AR study. Branch lengths are based on the 4D neutral model. Hibernators are shown in purple, and the non-hibernator background species used for conserved region identification is shown in blue. The clades are shown to the right. The numbers of ARs found in each hibernator lineage are shown in brackets (FDR 10%). See Figure S1 for the comparative analysis of non-hibernating homeothermic species.

(C) Hibernator ARs are best explained by adaptive selective effects rather than GC-biased gene conversion. The table shows the log likelihood results for statistical models testing for significant effects of selection (selection – null), biased gene conversion (bgc – null), and the difference between the selection and bgc models (selection – bgc) across all ARs identified for each hibernator. The selection and bgc models both yield statistically significant results compared to the null hypothesis, but the log likelihood of the selection model is much greater in all cases, indicating that this model explains most nucleotide substitutions. ***$p < 1 \times 10^{-100}$.

(D) Number of ARs residing in coding versus non-coding background conserved regions for each hibernating species.

phenotypic similarities to human metabolic syndrome (Dark, 2005; Martin, 2008). However, unlike human metabolic syndrome, obesity in hibernators is not associated with chronic hypertension or inflammation and insulin resistance is reversible (Dark, 2005; Martin, 2008). Elucidating the mechanisms involved could improve understanding and potentially treatment of human metabolic diseases, including obesity and type II diabetes (Bouma et al., 2010; Carey et al., 2003; Cooper et al., 2014; Geiser, 2004; Jani et al., 2013; Martin, 2008).

Hibernation has arisen independently in different species across at least 7 orders of mammals, ranging from monotremes to primates (Geiser, 1998). The repeated appearance of hibernation across different lineages suggests the genes involved are present in both hibernators and non-hibernators and the phenotypic differences may be explained by regulatory changes (Srere et al., 1992; Villanueva-Cañas et al., 2014). Indeed, epigenetic changes are associated with the different seasonal phases of hibernation (Alvarado et al., 2015; Morin and Storey, 2009). However, *cis*-regulatory elements that control hibernation phenotypes are not well understood. Hibernation could have arisen in different lineages through parallel evolution, in which the same *cis*-regulatory elements changed independently to promote a similar phenotype. Alternatively, different lineages could have converged on a similar hibernation phenotype through different regulatory elements and/or parallel evolutionary effects at gene or pathway levels. Here, we test the hypothesis that the development of hibernation involved parallel changes to mammalian *cis*-regulatory elements conserved in non-hibernators and linked to obesity control genes. Our findings provide foundations to functionally dissect the noncoding regulatory mechanisms controlling obesity and hibernation.

RESULTS

Hibernators Display Accelerated Evolution due to Selection in Genomic Elements Conserved across Non-hibernating Homeothermic Mammals

In order to investigate potential adaptive changes in hibernators in discrete *cis*-regulatory elements of clinical significance, we tested whether hibernators exhibit accelerated evolution due to selective effects in elements conserved across non-hibernating mammals. Additionally, to understand the unique properties of these elements, we compared ARs in hibernating versus non-hibernating, homeothermic mammals. We first investigated the obligate hibernators available in the mm10 60-way University of California at Santa Cruz (UCSC) multiple genome alignment file and found data for hibernators from different clades that evolved hibernation independently (Figure 1A; Table S1). We found genome alignments for the 13-lined ground squirrel (squirrel; Ruf and Geiser, 2015) and little brown bat (microbat; Geiser and Stawski, 2011), extreme, obligate hibernators from the *laurasiatheria* and *euarchontoglires* clades, respectively. Both species are hyperphagic in the fall and

become obese and insulin resistant prior to entering long periods of torpor in the winter. In addition, we identified the grey mouse lemur (mouse lemur), a primate lineage exhibiting weight gain and hibernation. This phenomenon is predominantly observed in females, though both sexes are capable of prolonged bouts of torpor (Schmid, 1999, 2000; Schmid and Ganzhorn, 2009). Finally, the lesser hedgehog tenrec, which can exhibit weight gain followed by multi-day hibernation, was included (Lovegrove et al., 2014). The microbat and squirrel evolved hibernation to survive winter months (Carey et al., 2003), while mouse lemurs and tenrecs evolved hibernation to survive cool, dry seasons in Madagascar (Schmid, 2000; Schmid and Ganzhorn, 2009). These four lineages are separated by 85–102 Ma (Hedges et al., 2015) and evolved hibernation independently (Ruf and Geiser, 2015).

To find ARs in these hibernating species, we identified 537,189 50-bp genomic regions of the mammalian genome (1) conserved across non-hibernating, homeothermic mammals ranging from humans to marsupials and (2) present in the four hibernators (Figure 1A; Table S1). We tested these regions for significant accelerated evolution in the hibernators individually using methods previously described (Ferris et al., 2018; Hubisz and Pollard, 2014; Pollard et al., 2010). We found 17,228 microbat, 5,119 squirrel, 7,661 mouse lemur, and 17,917 tenrec significant ARs (false discovery rate [FDR] 10%; Figure 1B; see numbers in brackets). We then tested whether these ARs are best explained by adaptive effects (positive selection or relaxed purifying selection) or GC-biased gene conversion, as previously described (Ferris et al., 2018; Kostka et al., 2012). Nucleotide substitution patterns in the ARs from each hibernator are best explained by adaptive effects due to selection; the adaptive model has a greater log-likelihood (log likelihood adaptive model − log likelihood GC-biased conversion model; Figure 1C; likelihood ratio test; see STAR Methods). Therefore, in genomic elements conserved across non-hibernating, homeothermic mammals, the hibernators exhibit significant accelerated evolution due to selection effects, indicating elements important for their adaptive traits.

We sought to compare hibernator ARs with regions accelerated in non-hibernating mammals. We identified seven homeothermic mammals in the 60-way alignment from the same clades as the hibernators that do not exhibit daily torpor and have genome sequence depth and quality scores similar to the hibernators (Figure S1A). We defined a background of conserved genomic elements in non-hibernating mammalian species to identify significant non-hibernator ARs, as performed for hibernators (Figures 1A and S1A). In a set of control analyses, we determined that the number of ARs found in each species is not significantly related to genome assembly quality (N50 score; Figure S1C). The total number of ARs is more strongly linked to the distance to the most closely related background species (Figure S1D). The majority of the ARs found in the eleven species profiled are noncoding (66.7%), revealing candidate *cis*-regulatory elements that changed to shape the unique phenotypes in each lineage (Figures 1D and S1B).

Discovery of Parallel Accelerated Regions in Distantly Related Hibernators

The development of altered obesity, metabolism, and hibernation phenotypes in the squirrel, microbat, tenrec, and lemur could have involved independent changes in parallel to conserved *cis*-regulatory elements with important roles in controlling these phenotypes—which we refer to as parallel accelerated regions (pARs). To test this, we asked whether two or more species possess significantly more shared ARs then expected by chance and computed the number of shared ARs for all 55 possible hibernator and non-hibernator species pairings (Figure 2A). For each species pairing, we tested whether the number of shared ARs is greater than expected by chance using a hypergeometric test and corrected for multiple testing errors using the Benjamini-Hochberg method. All hibernator:hibernator (hib:hib) species pairings are statistically significant, and the two extreme hibernators in our study, the microbat and squirrel, show the most significant enrichment effect (Figure 2B) despite the fact that the squirrel and lemur are more closely related (Figure 1B), suggesting the shared ARs reveal biology for shared traits rather than evolutionary relatedness. In contrast, we found only 52% of the non-hibernator species pairings and 46% hibernator:non-hibernator pairings have significantly shared ARs (Figure 2A). Our results reveal significant AR overlap in distant hibernating mammals.

We further investigated the nature of these parallel effects at the level of two-way and three-way species overlaps. All of the three-way hibernator intersections yield more pARs than expected by chance compared to only 5 of the 35 (14%) non-hibernator three-way intersections (Figure S2). We tested whether hibernators have significantly more two-way and/or three-way pARs compared to non-hibernators after controlling for the evolutionary distance between the species. We used generalized linear modeling to test whether the number of pARs for the two-way (Figure 2C) or three-way (Figure 2D) species intersections is significantly explained by a main effect of the species class (hibernators versus non-hibernators) by comparing to a nested model with the main effects for the number of pARs expected for each species combination and the mean evolutionary distance between the species (4D neutral model). The number of pARs is significantly related to hibernation status for two-way ($p = 0.04$; likelihood ratio test; quasi-poisson distribution) and three-way ($p = 0.03$) pARs (Figures 2C and 2D), indicating a significantly higher degree of AR overlap between hibernators as compared to non-hibernators. An example triplet hibernator pAR is shown in Figure 2E. We summarized the number of hibernator pARs uncovered from each two-way and three-way species intersection in a flower plot (Figure 2B). We did not observe four-way species hibernator pARs (Figure 2B) but uncover four-way parallel changes at the level of regulatory protein binding site motifs below (and see Discussion). In conclusion, pARs in distant hibernators (Table S2) suggest candidate *cis*-regulatory elements that shape hibernation-related phenotypes, such as changes to obesity regulation.

Hibernator pARs Are Disproportionately Located Near Known Obesity Susceptibility Loci and Reveal Candidate Obesity Regulating *Cis* Elements

Hibernators evolved pronounced changes to the regulation of obesity and metabolic activity compared to non-hibernators (Martin, 2008). We hypothesized that hibernator pARs reveal

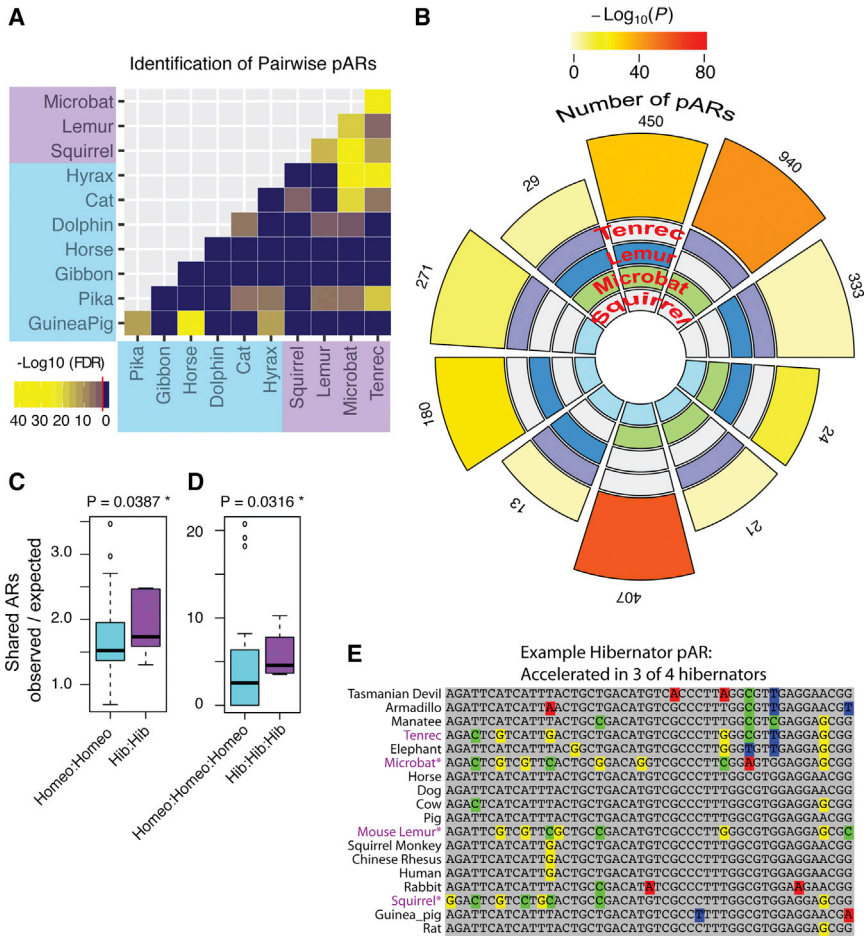


Figure 2. Discovery of Significant pARs in Distantly Related Hibernating Mammals

(A) The heatmap shows the results of pairwise testing for significant numbers of shared ARs between different hibernating and/or non-hibernating species. The heatmap data are based on the log10-scaled Benjamin-Hochberg adjusted p value for the hypergeometric tests, and the red line in the legend indicates the threshold for significance (FDR 5%). All pairwise comparisons between hibernators are significant.

(B) The flower plot shows the number of pARs identified and significance of the overlap for each of the different hibernator species intersections. The statistical significance of each intersection is shown by the color of the outer segment of the plot according to the legend. All cases are statistically significant (log 10 p value shown; SuperExactTest package in R). The number of pARs discovered from each species intersection is indicated by the outer numbers and the size of the outer blocks. The species intersected in each comparison is shown by colored blocks on the inner region (tenrec, purple; lemur, dark blue; microbat, green; squirrel, light blue). The data show that all two-way and three-way intersections between hibernators yield significantly more pARs than expected by chance.

(C and D) The boxplot compares the number of two-way (C) and three-way (D) species pARs for the different hibernator (purple) versus non-hibernator (blue) species intersections (see Figure S2). The plot shows the number of observed pARs normalized to the number expected for each intersection. Hibernators have significantly more two-way and three-way pARs than non-hibernators after controlling for evolutionary distance. p values were computed using generalized linear modeling on a quasipoisson distribution and likelihood ratio test (see STAR Methods).

(E) An example of a significant triplet hibernator pAR. The 50-bp sequence is well conserved across non-hibernating mammals from humans to marsupials but significantly accelerated in the hibernating squirrel, mouse lemur, and microbat (purple text).

cis-regulatory elements in the genome for controlling obesogenic phenotypes. To test this hypothesis, we first defined hibernator pARs as all elements exhibiting accelerated evolution in 2 or more of the four hibernators. These criteria revealed 2,407 hibernator pARs in the genome, of which 30.7% are coding and 69.3% are noncoding (Figure 3A). For comparison studies below, we also identified hibernator ARs that are not parallel (Figure 3B) and non-hibernator pARs, which constitute 2,283 elements identified in 2 or more of the seven non-hibernators (Figure 3A). Similar to hibernator pARs, we found that 33.1% and 66.9% of non-hibernator pARs are coding and noncoding, respectively (Figure 3A). Our approach therefore yielded similar total numbers of hibernator and non-hibernator pARs. From the perspective of hibernation, our methodology is conservative because non-hibernator ARs have a greater chance of being categorized as parallel in two or more of seven species compared to hibernator pARs, which are found in two or more of four species. We assigned hibernator and non-hibernator pARs to genes based on proximity using the GREAT algorithm (McLean et al., 2010; Table S2). The results suggest 1,919 genes that may be linked to hibernator pARs.

To determine whether hibernator pARs reveal candidate *cis*-regulatory elements for controlling obesity, we tested whether they are disproportionately located near genes known to influence human obesity susceptibility. Large-scale GWASs of human adult (Ghosh and Bouchard, 2017; Locke et al., 2015; Turcot et al., 2018) and childhood BMI (Felix et al., 2016) and body fat percentage (Lu et al., 2016) previously identified obesity risk genes. An analysis of these genome-wide significant SNPs using the data-driven expression-prioritized integration for complex traits (DEPICT) tool identified 123 obesity GWA genes (Ghosh and Bouchard, 2017; Turcot et al., 2018; Table S3). However, the *cis* elements controlling obesity are not well defined. We found that hibernator pARs are significantly enriched in conserved regions proximal to obesity GWA genes compared to randomly sampled conserved regions of similar size and GC content (p = 0.0002; Figures S3A and S3B). We then tested whether hibernator pARs are statistically enriched near obesity GWA genes compared to non-parallel hibernator ARs. The results show that hibernator AR associations with obesity genes depend significantly on whether or not the ARs are parallel ($p = 4.1 \times 10^{-4}$; chi-square test of independence; Figure 3C).

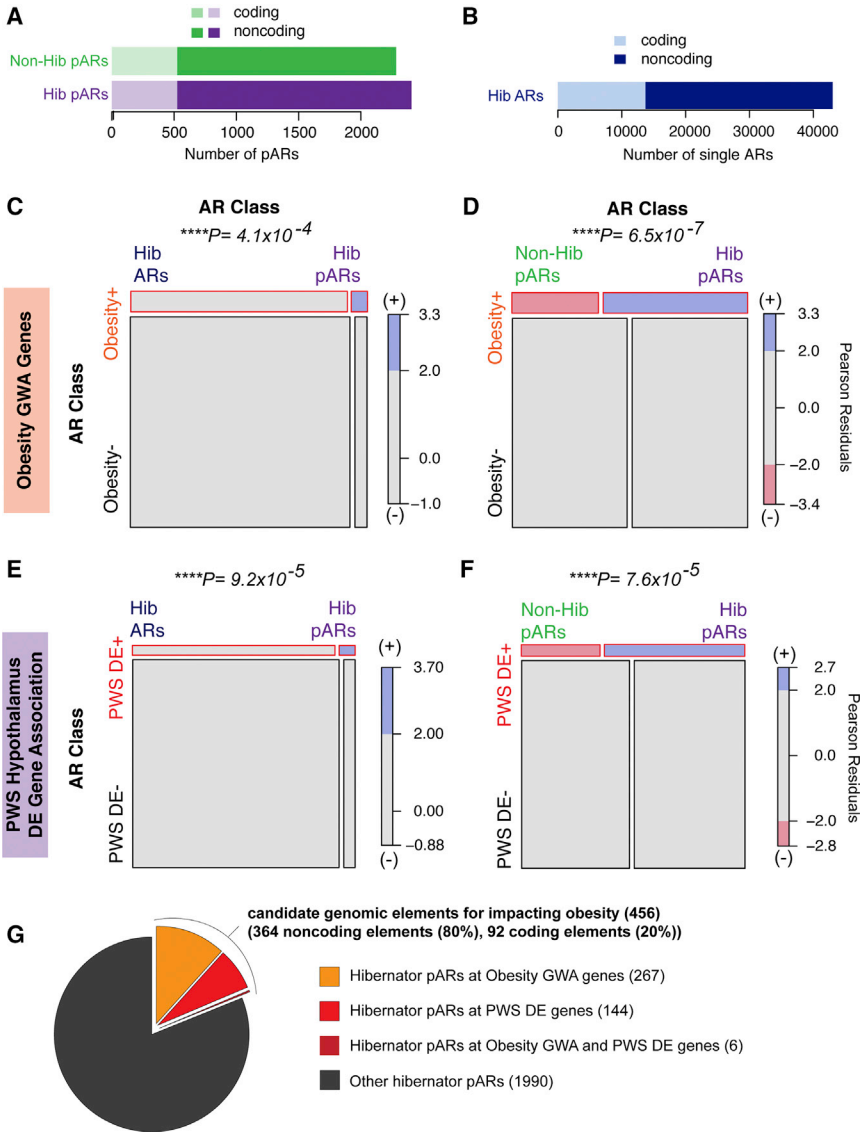


Figure 3. Hibernator pARs Predominantly Impact Noncoding Elements in the Mammalian Genome and Are Disproportionately Enriched Near Obesity Susceptibility Genes

(A and B) The stacked barplot shows the number of coding and noncoding pARs identified from 2 or more of the 4 hibernators (A, purple bar; hib ARs) and the number found in 2 or more of the seven non-hibernator homeotherms (A, green bar; non-hib ARs). The number of ARs that are not parallel and found in individual hibernators is also shown (B). The proportion of pARs or ARs in coding (light) versus noncoding (dark) elements is shown by shading in each plot (see legend and main text).

(C and D) The mosaic plots show hibernator pARs are significantly enriched near obesity risk genes identified from human GWASs. Each box is proportional to the contingency count data table of ARs assigned to obesity GWA genes (obesity+) and ARs assigned to other genes (obesity−), as well as the subsets of the ARs that are individual hibernator ARs versus pARs (C) or non-hibernator pARs versus hibernator pARs (D). The chi-square test is significant (p values shown above plot), showing that hibernator AR associations with obesity+ genes depend significantly on whether the ARs are parallel or not (C) and whether the ARs are parallel in hibernators versus non-hibernators (D). The mosaic plot also shows the chi-square Pearson residuals, indicating the positive (blue) and negative (red) associations in the data, and reveal that hibernator pARs (Hib pARs) are positively associated with obesity+ genes compared to non-parallel ARs (C) and non-hibernator pARs (D).

(E and F) Mosaic plots testing for associations between different classes of ARs and genes differentially expressed in the hypothalamus of PWS patients (PWS DE+) versus genes that are not (PWS DE−). The data show that AR associations with PWS DE genes significantly depend on AR class (E and F; chi-square p values above plots). Hibernator pARs are positively associated with PWS DE+ genes (blue), although ARs that are not parallel in hibernators (E) or parallel in non-hibernator (F; non-hib pARs) show a negative association with PWS DE+ genes (red).

(G) The pie chart shows the proportion of hibernator pARs enriched near obesity GWA genes and/or PWS DE genes. The absolute numbers of pARs are shown in brackets.

A mosaic plot of the data and the Pearson residuals reveals the hibernator pARs are positively associated with obesity GWA genes (blue, Figure 3C). We next tested whether the enrichment is characteristic of hibernator pARs compared to non-hibernator pARs (Figure 3D). Indeed, hibernator pARs are disproportionately near obesity GWA genes (Figure 3D; $p = 6.5 \times 10^{-7}$). We conclude that hibernator pARs enrich for candidate *cis*-regulatory elements for controlling obesity and identified 364 candidate elements in the mammalian genome (Table S3).

To further test our hypothesis and potentially uncover a more comprehensive set of *cis*-regulatory elements for controlling obesogenic phenotypes, we used published transcriptome data of genes differentially expressed in the hypothalamus of Prader-Willi syndrome (PWS) patients compared to age-matched controls (Bochukova et al., 2018). PWS is a genetic obesity syndrome characterized in part by hyperphagia and childhood morbid obesity (Bochukova et al., 2018; Cassidy and Driscoll, 2009). Genes differentially expressed in the hypothalamus of PWS patients (PWS DE genes) are expected to have roles in regulating feeding behavior, the neuroendocrine system, and body weight (Bochukova et al., 2018). We tested whether hibernator pARs are disproportionately enriched near PWS DE genes. Hibernator pARs are significantly enriched in non-hibernator conserved regions proximal to PWS DE genes compared to randomly sampled conserved regions ($p = 0.013$; Figure S3C). Moreover, hibernator pARs are

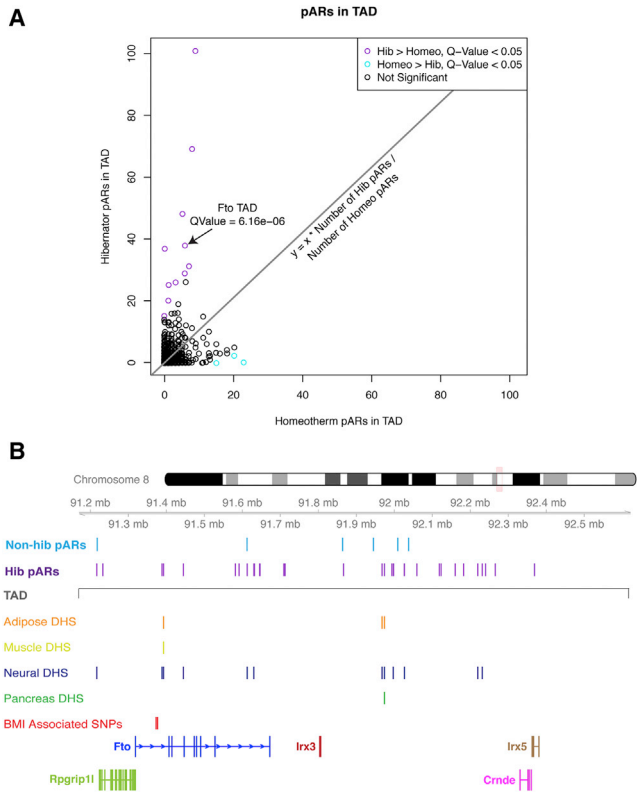


Figure 4. Identification of TADs in the Genome with Disproportionate Numbers of Hibernator pARs and Discovery that the *FTO* TAD Is among the Most Enriched Sites

(A) The scatterplot shows the relative numbers of hibernator versus non-hibernator pARs in different TADs of the mouse genome. TADs significantly enriched for hibernator pARs are shown in purple (chi-square test; q-value < 0.05), those enriched for non-hibernator pARs are shown in cyan, and other TADs are shown in black (not significant). For each TAD, the numbers of hibernator pARs (y axis) versus non-hibernator pARs (x axis) are shown.

(B) The plot shows the locations of 38 hibernator (hib pARs) and 6 non-hibernator (non-hib pARs) pARs in the *Fto* TAD. The boundaries of the TAD are indicated (gray line). The tracks below indicates the DNase I hypersensitivity (DHS) of human homologs of the hibernator pARs in different tissues. The data reveal tissue-dependent activity patterns and many *cis* elements active in neural tissue. Multiple hibernator pARs lie in the first intron of *Fto*/*FTO* near known obesity risk variant sites in the first intron (BMI-associated SNPs: rs1421085 and rs1558902; red track), which can influence the expression of *Irx3*/*IRX3* through long-range regulatory contacts.

significantly positively associated with PWS DE genes compared to ARs that are not parallel in hibernators (Figure 3E; $p = 9.2 \times 10^{-5}$) and compared to non-hibernator pARs (Figure 3F; $p = 7.6 \times 10^{-5}$). Given that changes to feeding, weight gain, and metabolic activity are distinctive attributes of hibernators, our findings support the conclusions that hibernator pARs reveal mechanisms for hibernation and candidate *cis*-regulatory elements for controlling obesogenic phenotypes. By aggregating the hibernator pARs located at BMI-linked obesity loci with those near PWS DE genes, we identified 456 genomic elements, including 364 candidate *cis*-regulatory elements for controlling obesogenic phenotypes (Figure 3G). These elements are near to 114 obesity susceptibility genes, including 65 obesity GWA genes, 51 PWS DE genes, and 1 obesity/PWS DE gene (Table S3).

Hibernator pARs Are Disproportionately Enriched in Specific Topologically Associated Domains (TADs) in the Mammalian Genome, and the *FTO-IRX3-IRX5* Gene Locus Is a Primary Enrichment Site

The genome is organized into spatially restricted, TADs that delineate blocks of genes with coordinated regulation and *cis*-regulatory elements that functionally interact (Acemel et al., 2017). Regulatory interactions between different TADs are generally inhibited. We previously found that ARs from different species cluster at different sites in the mammalian genome, which can indicate important candidate loci for shaping traits unique to a species (Ferris et al., 2018). Here, we asked whether hibernator pARs are enriched in specific TADs in the genome. Using the coordinates for 1,817 known mouse TADs (Dixon et al., 2012), we counted the pARs in each TAD (Table S4). For each TAD, we performed a chi-square test on the counts of hibernator versus non-hibernator pARs to determine whether a given TAD harbors significantly more hibernator or non-hibernator pARs. False positives due to multiple testing of different TADs were controlled using the q-value approach (Storey, 2002). We found 11 TADs significantly enriched for hibernator pARs (Figure 4A; $q < 0.05$), revealing TADs in the mammalian genome with predicted roles in the development of hibernation.

Among the top 5 TADs enriched for hibernator pARs, we identified the TAD harboring the *FTO* locus (Figure 4A; Table S4), the strongest genetic risk factor for human obesity (Claussnitzer et al., 2015a; Meyre et al., 2009; Yang et al., 2017). The TAD encompassing *FTO* extends from *RPGRIP1L* to *IRX5* and harbors 38 hibernator pARs but only 6 non-hibernator pARs (Figure 4B), a 15-fold, significant enrichment effect ($p = 6.1 \times 10^{-6}$; chi-square test). Given these previous findings and the changes to obesogenic phenotypes in hibernators, the hibernator pARs in the *FTO* TAD reveal conserved noncoding elements that may influence mammalian obesity. By comparing the locations of the human homologs of these hibernator pARs to available DNase sequencing (DNase-seq) data for different human tissues (chromatin immunoprecipitation [ChIP]-Atlas), we found many (39.5%) located in noncoding elements active in neural tissue (Figure 4B, blue track); pAR homologs are also active in adipose, muscle, and pancreas (Figure 4B). Two hibernator pARs are located in the first *FTO* intron, which harbors known noncoding obesity risk variants, including rs1421085 and rs1558902 (Figure 4B, red track; Yang et al., 2017). Overall, the hibernator pARs in the *FTO-IRX3-IRX5* TAD predominantly uncovered candidate regulatory elements with putative roles in the brain. The other TADs and conserved elements we identified as having putative roles in shaping hibernation-related phenotypes are presented as a resource in Table S4.

Hibernator pARs Change Conserved Regulatory Protein-Binding Site Motifs and Reveal Candidate Genetic Circuits Controlling Obesity

We do not know the functional effects of the changes within hibernator pARs. The changes could eliminate binding site motifs conserved across non-hibernators or add new binding sites not

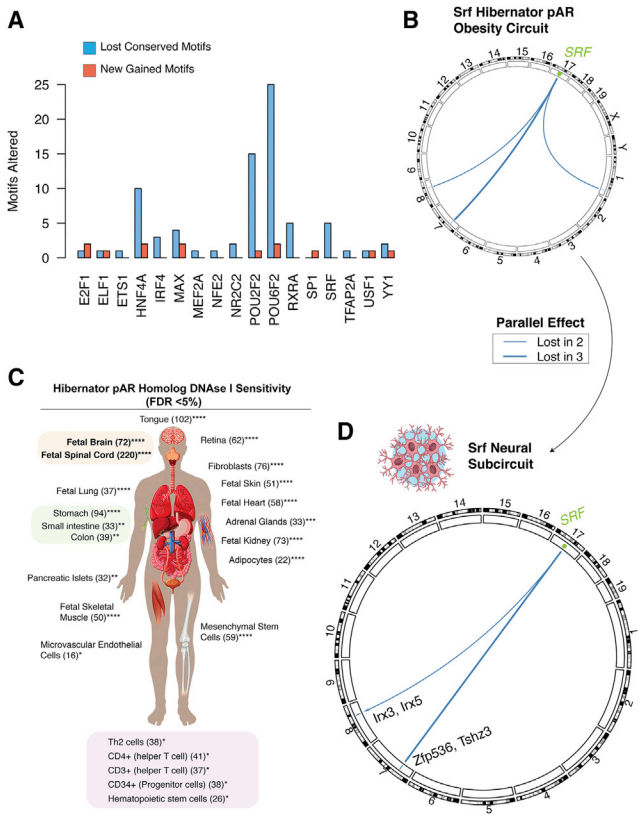


Figure 5. Hibernator pARs Eliminate and Add Specific Regulatory Protein-Binding Site Motifs, Indicating Candidate Genetic Circuits in Different Tissues for Controlling Obesity

(A) This barplot summarizes the number of binding site motifs lost and gained in 2 or more hibernators within pARs assigned to obesity GWA genes and PWS DE genes. Numbers of conserved binding site motifs lost are indicated by blue bars, and new motifs gained are indicated in red. The data show that motif loss is more frequent than new motif gains and that the regulatory proteins with the largest number of motif changes in obesity-linked hibernator pARs are POU2F2 and POU6F2.

(B) Lost and gained motif sites in obesity-associated hibernator pARs reveal candidate genetic circuits for shaping mammalian obesity and hibernation phenotypes. A circos plot shows the genetic circuit linking the regulatory protein, Srf, to the lost Srf binding site motifs in hibernator pARs near known obesity risk genes. Connections in the circuit are weighted according to the number of hibernating species showing the parallel motif loss (blue lines; see legend for number of hibernators with lost motif site).

(C) The schematic shows human tissue and cell types with DNase I hypersensitive sites that overlap hibernator pARs elements more than expected by chance (FDR < 5%; *in silico* ChIP analysis; ChIP-Atlas). The number of DNase I hypersensitive sites that overlap hibernator pARs is shown in brackets for each tissue/cell type. *FDR < 0.05; **FDR < 0.01; ***FDR < 0.001; ****FDR < 1×10^{-4}.

(D) The refined Srf obesity genetic subcircuit for neural tissues was determined by integrating information from the analyses in (B) and (C). The subcircuit now includes only obesity-linked hibernator pARs with DNase I hypersensitivity in human neural tissue (also see Figure S5).

found in the non-hibernator sequence. Such changes may alter expression of neighboring genes or remodel genetic circuits involved in the development of hibernation-related phenotypes. To begin to test this hypothesis, we collected available binding motif data and threshold data for 55 different known regulatory protein binding site motifs with threshold binding data available (Jolma et al., 2013). The threshold scores allow one to evaluate whether nucleotide changes are likely to disrupt a conserved motif or introduce a new motif (Dabrowski et al., 2015). We defined the motifs present in the conserved non-hibernator background sequence for each of the 2,407 hibernator pARs, revealing 1,048 motif sites for 29 different regulatory proteins (Table S5). These results identify upstream regulatory proteins putatively acting on the pAR elements in non-hibernators.

At each motif site, we calculated the binding score for the hibernator sequence. Motifs with scores below the binding threshold score were counted as lost. For example, a binding site motif for the transcription factor POUF2 is conserved in non-hibernators, ranging from guinea pig to elephant, but lost in all four hibernators (Figure S4A). We also identified motifs gained in 2 or more hibernators. For example, we uncovered an HNF4A binding site present in squirrel and microbat sequence but absent in non-hibernators (Figure S4B). We uncovered all parallel lost and gained motifs in two or more hibernators within hibernator pARs. Conserved motifs were lost in parallel ~3× more frequently than new motifs were gained in parallel (Figure S4C).

Regulatory protein-binding sites in the genome create genetic circuits connecting regulatory proteins to the downstream networks of genes they regulate. These genetic circuits control the development of different phenotypes and shape physiological and behavioral responses to environmental signals. The motif changes caused by hibernator pARs for specific regulatory proteins open the possibility of defining candidate genetic circuits for controlling mammalian obesity and hibernation. To define these circuits, we focused on the 267 hibernator pARs enriched near obesity GWA genes (Figure 3G), thereby defining a network of genes and potential *cis*-regulatory elements. For each *cis* element in this network, we identified the parallel lost and gained binding site motifs in the hibernators (Table S5). We found 19 regulatory proteins with motifs lost in 2 or more hibernators and 11 with motifs gained in 2 or more hibernators (Figure 5A). The largest number of motif changes occurred for the transcription factors Pou2f2 and Pou6f2 (Figure 5A). Interestingly, Pou6f2 is known to have important roles in regulating gene expression changes in response to increased BMI (Glastonbury et al., 2016). We next assembled genetic circuits for each regulatory protein. For example, the transcription factor Srf has four parallel lost motif sites in hibernator pARs near obesity susceptibility genes (Figure 5A). Candidate Srf-mediated genetic circuits for obesity connect Srf to obesity susceptibility genes on four different mouse chromosomes (Figure 5B). We found 110 candidate genetic circuits for obesity and hibernation with this approach (Table S5).

Different cell types and tissues have different metabolic rates and roles in shaping obesogenic phenotypes, and the activity of regulatory elements is often cell type dependent. We thus sought to refine our genetic circuits for specific tissues. To achieve this, we first compared the human genome coordinates of the 2,407 hibernator pARs to available DNase-seq data for different human tissue and cell types using *in silico* ChIP in the ChIP-Atlas. The results uncovered 22 cell classes with significant hibernator pAR homolog enrichments in DNase I hypersensitive sites

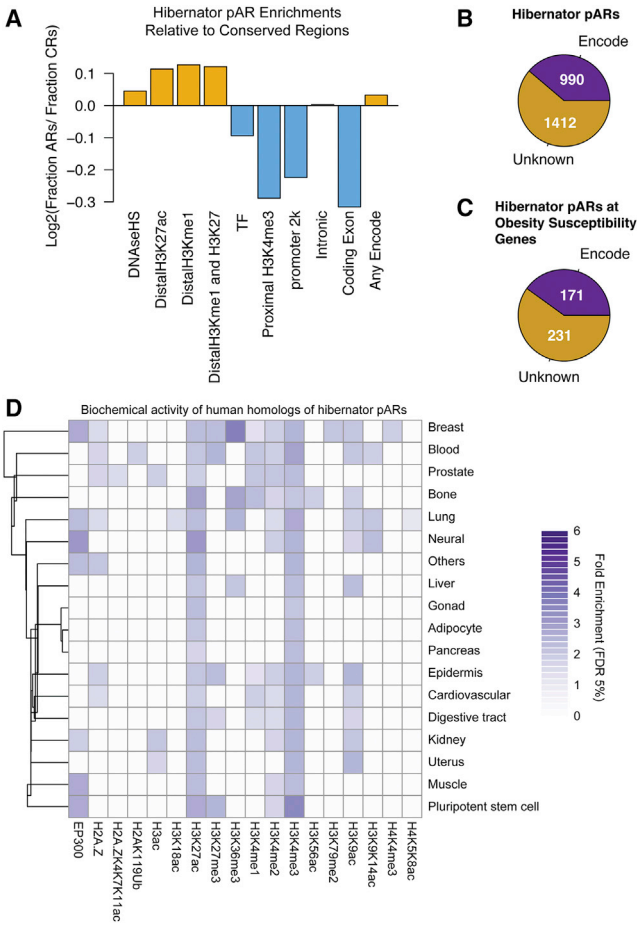


Figure 6. Hibernator pARs Impact Known and Putative Novel Functional *cis*-Regulatory Elements in the Mammalian Genome

(A) The barplot shows the proportion of hibernator pARs in different classes of human genome elements compared to background non-hibernator conserved regions (CRs). Hibernator pARs are disproportionally enriched (yellow bars) in distal regulatory elements and under-represented (blue bars) in proximal regulatory elements and coding regions compared to CRs. Hibernator pARs were converted from mm10 to hg19 coordinates. 2,364 of the 2,370 hibernator pARs lifted over. These elements were compared to human ENCODE DNase-seq and ChIP-seq datasets for biochemical markers of enhancers (distal H3K27ac+; distal H3K27ac+/H3K4me1+), transcription factor (TF) binding sites, promoters (proximal H3K4me3+; promoter 2k), transcriptionally active H3K9ac+ elements, intronic regions, and coding regions.

(B and C) Pie charts show the number of human homologs of hibernator pARs (B) and the subset of hibernator pARs at obesity susceptibility genes (C) that lie within ENCODE annotated functional elements in the human genome (Encode, purple). Many hibernator pARs impact functionally uncharacterized (unknown, brown) conserved genomic elements.

(D) The heatmap shows human homologs of hibernator pARs are significantly enriched in different classes of biochemically active putative functional elements identified in different tissues by ChIP-seq in the ChIP-Atlas database. The color indicates fold enrichment (see legend). Colored blocks in the heatmap indicate a statistically significant enrichment (FDR < 5%). White blocks indicate the insignificant enrichments or unavailable datasets.

compared to randomly sampled pARs (FDR < 5%; Figure 5C). These data show that neural tissue has the largest number of DNase-I-sensitive hibernator pAR homologs, suggesting many of these elements have functions in the brain (Figure 5C). Further data reveal subsets of hibernators pAR homologs with evidence of activity in other cell types, including adipose, pancreatic islets, different immune cells, skeletal muscle, and others (Figure 5C). We then refined candidate obesity circuits for specific tissues. For example, the genetic circuit for Srf was refined based on the DNase I sensitivity results to reveal a putative subcircuit for neural tissue (Figures 5B and 5D). The subcircuit indicates a Srf regulatory connection to a neural active hibernator pAR homolog near *Irx3* and *Irx5* in the *Fto* TAD on mouse chromosome 8 and to another near *Tshz3* on chromosome 7 (Figure 5D). Both *Irx3* and *Tshz3* are obesity susceptibility genes (Locke et al., 2015; Smemo et al., 2014). We also dissected obesity and hibernation subcircuits for Pou6f2 (Glastonbury et al., 2016) in neural, adipose, and skeletal muscle (Figure S5). We applied this approach across our data, identifying candidate obesity and hibernation genetic circuits for different cell types (Table S5).

To further uncover upstream regulatory mechanisms acting on obesity and hibernation genetic circuits, we determined whether specific regulatory protein-binding site motifs are significantly enriched in hibernator pARs near obesity susceptibility genes by applying the PWMenrich R package and Clover method (Frith et al., 2004) to the DNA sequences for each hibernator. We tested for motif enrichments relative to human DNA sequence derived from conserved regions near obesity susceptibility genes. Our analysis found significantly enriched regulatory protein-binding site motifs in the hibernator pARs for each hibernating species (FDR 5%; Table S5). A Venn diagram analysis comparing the top enriched protein motifs for the four hibernators revealed that 85% of the top enriched regulatory proteins (FDR 10%) are found in 2 or more hibernators, including 13 found in all four hibernators (Table S5). We therefore found top candidate upstream regulatory proteins controlling genetic circuits for obesity and hibernation.

Finally, we identified ENCODE annotated elements in the human genome impacted by hibernator pARs and compared relative enrichments in different element classes to the background non-hibernator conserved regions (Figure 6A). Human homologs of hibernator pARs are located in known enhancers (H3K27ac+, HeK4me1+/H3K27ac+), promoters (proximal H3K4me3+, promoter 2k), transcription factor binding sites, and other functional elements (Figure 6A). Hibernator pARs are disproportionately located within distal noncoding elements compared to the background conserved regions and under-represented in coding exons and promoter regions (Figure 6A). Many of the conserved elements impacted by hibernator pARs (60%; Figure 6B) and hibernator pARs at obesity susceptibility genes (57%; Figure 6C) remain unclassified. Indeed, only seven hibernator pARs are located in VISTA annotated enhancers: two are associated with the *IRX3* gene in the *FTO* obesity risk locus (vista IDs: 1211 and 1651). Six hibernator pARs overlap hypothalamus expression quantitative trait loci (eQTLs) identified by the Genotype-Tissue Expression Consortium. Finally, we ascertained the biochemical activity patterns of human homologs of hibernator pARs across different human tissues (Figure 6D). The data show statistically significant enrichments for specific histone modifications in a tissue-dependent manner (Figure 6D). Thus,

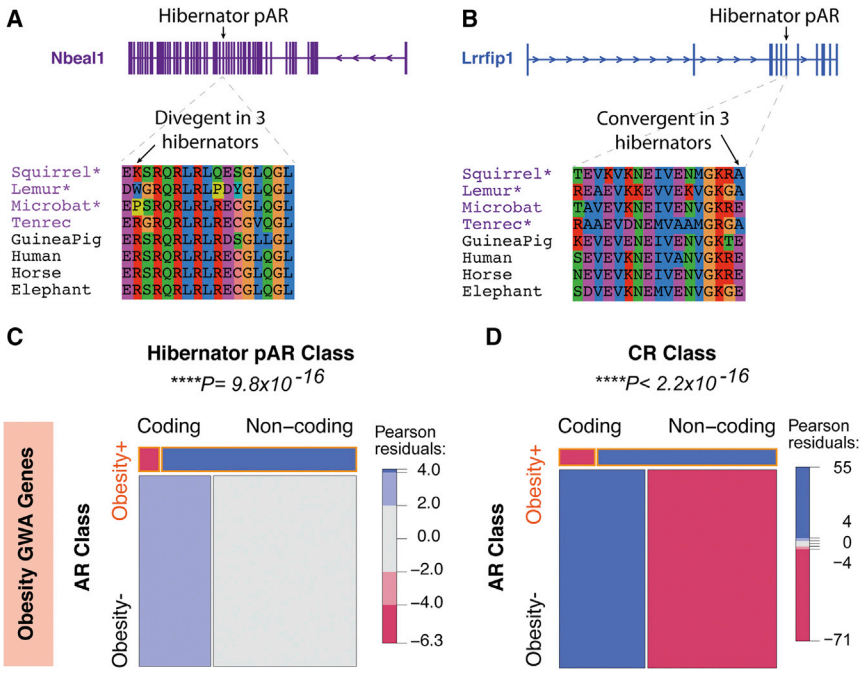


Figure 7. Parallel Coding Changes in Hibernators Are Relatively Rare in Obesity GWA Genes Compared to Changes to Noncoding *cis* Elements

(A and B) Examples of parallel coding changes to amino acid sequences in hibernator pARs are shown. The two classes of parallel changes detected are divergent changes (A; Nbeal1), which involve changing the same non-hibernator conserved amino acid to different residues, and convergent changes (B; Lrrfip1), which involve changes to the same residue. Hibernators are highlighted by purple text, and non-hibernators from different clades are in black text. The exon harboring the hibernator pAR is indicated above the gene models. Hibernators with the parallel coding change are indicated by asterisks.

(C) The mosaic plot shows noncoding hibernator pARs are significantly enriched near human obesity GWA genes compared to coding pARs. Each box is proportional to the contingency count table of obesity GWA genes (obesity+) and other genes (obesity−), and subsets of obesity+ and obesity− genes linked to coding versus noncoding hibernator pARs. The chi-square test is significant (p values shown above plot), indicating that hibernator pAR associations with obesity+ genes depend significantly on whether the pARs are noncoding or coding. The mosaic plot also shows the chi-square Pearson residuals, indicating the positive (blue) and negative (red) associations in the data, and reveal that noncoding hibernator pARs are positively associated with obesity+ genes relative to coding pARs.

(D) The mosaic plot shows noncoding non-hibernator CRs are significantly enriched near human obesity GWA genes compared to CRs in coding sequence.

human homologs of hibernator pARs are enriched for putative functional elements in different tissues, but many are in conserved elements not yet functionally characterized.

Parallel Coding Changes in Obesity GWA Genes Are Relatively Rare, Implicating Noncoding Mechanisms in the Evolution of Hibernation-Linked Obesity Phenotypes

It is unclear whether noncoding or coding changes are likely to play the major roles in the development of altered obesity and metabolic phenotypes in hibernators. Above, we found that 30.7% of hibernator pARs are located in coding regions, suggesting that parallel amino acid changes might occur for some proteins in hibernators at sites conserved in non-hibernators. We found 184 genes with hibernator pARs predicted to cause amino acid changes in 2 or more hibernators at non-hibernator conserved sites (Table S6). Because pARs indicate nucleotide sequence changes in a 50-bp conserved region, they do not necessarily indicate that the same amino acid residue is changed in hibernators in a coding region. An analysis of amino acid changes in hibernator pARs revealed divergent amino acid changes, in which the same site changed to different amino acids in different hibernators (Figure 7A), as well as convergent changes, in which the same site changed to the same amino acid in multiple hibernators (Figure 7B). Divergent and convergent changes in 2 or more hibernators occurred for 31% (230) and 25% (184) of coding hibernator pARs, respectively. The remaining 44% of coding pARs either resulted in changes to different codons in the 50-bp region or caused parallel synonymous changes.

Having identified parallel coding changes in the hibernators compared to non-hibernators, we tested whether obesity GWA genes are impacted. Remarkably, only a single obesity GWA gene, Lrp1b (low-density lipoprotein receptor-related protein 1B), has either convergent or divergent parallel amino acid changes in hibernators (Table S6). This result suggests that parallel changes are rare for hibernators in the coding regions of obesity GWA genes. Indeed, a chi-square analysis comparing the proportion of coding versus noncoding hibernator pARs linked to obesity GWA genes revealed a significant enrichment for noncoding pARs and a relative under-representation of coding pARs ($p = 9.8 \times 10^{-16}$; Figure 7C). These results suggest especially important roles for *cis* elements near obesity GWA genes in shaping hibernation-related phenotypes. Moreover, we found that non-hibernator conserved regions are also significantly enriched near obesity GWA genes (Figure 7D), which indicates evolutionary constraint on many *cis* elements in these regions of the mammalian genome. These effects were not observed for PWS DE genes, which instead showed relative enrichments for pARs and conserved regions in coding regions (Figure S6). Overall, our data support a major role for *cis*-regulatory mechanisms in controlling hibernation and obesity phenotypes, particularly at obesity risk genes found through human GWASs.

DISCUSSION

Hibernation has arisen independently in several mammalian lineages, though the mechanisms involved are poorly understood.

Most studies have focused on understanding adult hibernator physiology and protein coding changes. We found putative *cis*-regulatory elements conserved across non-hibernating homeotherms and accelerated in hibernators due to selection. We uncovered pARs shared across distantly related hibernators, revealing candidate *cis*-regulatory elements for controlling mammalian hibernation phenotypes. Hibernator pARs are disproportionately located near known human obesity susceptibility genes compared to random conserved regions in the genome, non-parallel hibernator ARs or homeotherm pARs. Our analyses also reveal TADs in the mammalian genome with an excess of hibernator pARs, which implicates functional units of genes and regulatory elements in the development of hibernation, such as the TAD encompassing the *FTO* obesity susceptibility locus. Hibernator pARs include changes in regulatory protein-binding site motifs conserved in the genomes of non-hibernators. This allowed us to construct candidate genetic circuits in different tissues for controlling obesity and hibernation. This shows the strength of combining comparative phylogenomics with human genetics data to find candidate *cis*-regulatory mechanisms for specific clinical phenotypes. These findings lay foundations for targeted functional studies of individual *cis* elements and genetic circuits to elucidate noncoding mechanisms controlling obesity, hypometabolic states, and feeding behaviors.

Parallel Hibernator ARs Reveal Candidate *Cis*-Regulatory Mechanisms Controlling Obesity

Obesity is an epidemic impacting ~30% of the world's population, and the etiology is complex. Obesity can arise due to altered impulsivity, satiety, meal frequency, food choice, reward seeking, activity, and anxiety responses (Hruby and Hu, 2015; Rowland et al., 2008). Obesity can also be influenced by physiological traits, including metabolic rate, inflammation, BMI, and predispositions to adiposity (Hruby and Hu, 2015; Singla et al., 2010). Strikingly, obesity risks are strongly influenced by developmental factors rather than adult lifestyle choices (Reynolds et al., 2015). We expect many elements impacted by hibernator pARs to function during development to shape obesity and hibernation-related phenotypes (Booker et al., 2016; Eckalbar et al., 2016). We expect this because many human (Capra et al., 2013) and bat ARs are developmental enhancers (Booker et al., 2016; Eckalbar et al., 2016). Moreover, we found human homologs of hibernator pARs are frequently active in fetal tissues, suggesting roles in development. Indeed, the development of new morphological, physiological, and behavioral phenotypes frequently involves changes to developmental *cis*-regulatory architecture rather than protein-coding changes (Carroll, 2008; Wray, 2007). The field is working to understand the genetic circuits that create different phenotypes (Booker et al., 2016; Erwin and Davidson, 2009). Studies of hibernation have predominantly focused on adults (Carey et al., 2003; Geiser, 2004; Jastroch et al., 2016). Fewer have investigated the developmental origins of hibernation. Our study reveals relevant and developmentally active *cis*-regulatory elements and genetic circuits, supporting further work in this area.

The mouse will be an effective genetic model organism for future functional studies and targeted genome/epigenome editing of the *cis* elements we uncovered. Mice exhibit short bouts of daily torpor that can be induced by food restriction or high foraging costs (Heldmaier et al., 2004; Schubert et al., 2010); its phenotype is intermediate to hibernation and homeothermy, and perturbing the conserved elements and circuits we uncovered could result in informative changes to obesogenic phenotypes and torpor. Top candidate elements for future studies are hibernator pARs in the *FTO-IRX3-IRX5* locus. This locus is the strongest genetic risk factor for human obesity (Yang et al., 2017). Genes in this region have functions in regulating impulsivity, food intake, reward, and dopamine signaling (Castellini et al., 2017; Chuang et al., 2015; Hess et al., 2013; Karra et al., 2013; Sevgi et al., 2015), adipocyte development (Claussnitzer et al., 2015a; Landgraf et al., 2016; Smemo et al., 2014; Yang et al., 2017), and hypothalamic functions (Smemo et al., 2014; Yang et al., 2017). However, the regulatory architecture of the *FTO* TAD is complex, and important *cis* elements have yet to be discovered (Claussnitzer et al., 2015a; Smemo et al., 2014). Dissecting the *cis*-regulatory logic of the *FTO* TAD is not only important for understanding obesity but also for age-related brain atrophy (Ho et al., 2010). Thus, learning how hibernator versus non-hibernator variants of pARs shape different phenotypes through changes to the expression of genes in the *FTO* TAD may illuminate how this locus shapes disease risks. Overall, although large GWASs have uncovered important genomic loci for controlling obesity (Speliotes et al., 2010; Locke et al., 2015), the resolution of these studies is limited, and it remains challenging to identify discrete *cis*-regulatory elements controlling obesity. Our study shows the power of combining comparative genomics with large-scale human genetics to help solve this problem.

Candidate Noncoding Mechanisms Underlying the Evolution of Hibernation

The origins of mammalian hibernation are debated. For most lineages, it is unclear whether obligate hibernation, daily torpor, or homeothermy (non-hibernating/stable body temperature) are ancestral or derived traits (Geiser, 2008; Grigg et al., 2004; Lovegrove et al., 2014; Nowack and Dausmann, 2015). Most mammals exhibit some capacity for a hypometabolic state even without major changes to body temperature, which helps manage the high energetic costs of endothermy (Heldmaier et al., 2004). The capacity for hypometabolism may be a basic property of mammalian physiology, and hibernation may be an extreme extension of this trait (Geiser, 1998; Heldmaier et al., 2004). Our data show that genomic elements conserved in non-hibernating homeotherms underwent parallel accelerated evolution in distant hibernators, which suggests hibernation is a derived rather than ancestral trait in these lineages. Our findings suggest that these elements are involved in hibernation adaptations. They (1) are accelerated in hibernators primarily due to selection, (2) are disproportionately parallel in species that independently evolved hibernation, and (3) are enriched near obesity susceptibility loci. Targeted functional studies could determine how these elements affect gene expression, when and where they are active, and the nature of phenotypes they control. We prospectively linked pARs to genes based on

proximity; however, long-range regulatory contacts are frequent and more detailed chromatin conformation studies could determine how pARs regulate gene expression. One could swap endogenous hibernator DNA sequence for the conserved non-hibernator sequence (or vice versa) and test for gene expression differences. New CRISPR-based methods that permit multiplex epigenome editing will facilitate the functional dissection of the genetic circuits impacted by hibernator pARs (Carleton et al., 2017). The *cis* elements we uncovered near obesity susceptibility loci could shape obesity-related phenotypes. Obesity, hypometabolic, and torpor states are linked to foraging costs and food intake (Heldmaier et al., 2004; Schubert et al., 2010). Thus, some pARs could regulate behavior patterns and metabolic rates. Parallel evolutionary effects may occur at the level of pathways, genes, or regulatory elements. We expect that changes to regulatory elements only partially explain hibernation, and changes at pathway and gene levels also play important roles.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- METHOD DETAILS
  - Identification of ARs in the Target Species
  - Assigning pARs to TADs
  - Regulatory Binding Site Motif and Genetic Circuit Analysis
  - Analysis of Parallel Coding Changes
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Genome Quality Control Studies
  - Analysis of GC-biased Gene Conversion Versus Adaptive Nucleotide Substitutions in Hibernator ARs
  - Statistical Identification of pARs and Calculation of Observed versus Expected Shared ARs
  - Statistical Analysis of AR Enrichments at Obesity Susceptibility Genes
  - Identification of ARs in Human Putative Functional Genomic Elements
- DATA AND CODE AVAILABILITY
  - Software and Databases
  - Datasets

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.10.102.

ACKNOWLEDGMENTS

We thank Drs. Jan Christian, Joshua Schiffman, Lisa Abegglen, Hilary Coon, Jared Rutter, and Joseph Yost and members of the Gregg lab for reviewing earlier versions of the manuscript. We thank Dr. Melissa J. Hubisz for help with RPHAST. This work was supported by National Institutes of Health (NIH) grants R01MH109577 and R01AG064013 (C.G.).

AUTHOR CONTRIBUTIONS

Conceptualization, E.F. and C.G.; Methodology, E.F. and C.G.; Investigation, E.F. and C.G.; Software, E.F.; Formal Analysis, E.F. and C.G.; Data Visualization, E.F. and C.G.; Writing, C.G. and E.F.; Supervision, C.G.; Funding, C.G.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 19, 2019
Revised: June 21, 2019
Accepted: October 24, 2019
Published: November 26, 2019

REFERENCES

Acemel, R.D., Maeso, I., and Gómez-Skarmeta, J.L. (2017). Topologically associated domains: a successful scaffold for the evolution of gene regulation in animals. Wiley Interdiscip. Rev. Dev. Biol. *6*, e265.

Alvarado, S., Mak, T., Liu, S., Storey, K.B., and Szyf, M. (2015). Dynamic changes in global and gene-specific DNA methylation during hibernation in adult thirteen-lined ground squirrels, Ictidomys tridecemlineatus. J. Exp. Biol. *218*, 1787–1795.

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

Bochukova, E.G., Lawler, K., Croizier, S., Keogh, J.M., Patel, N., Strohbehn, G., Lo, K.K., Humphrey, J., Hokken-Koelega, A., Damen, L., et al. (2018). A transcriptomic signature of the hypothalamic response to fasting and BDNF deficiency in Prader-Willi syndrome. Cell Rep. *22*, 3401–3408.

Booker, B.M., Friedrich, T., Mason, M.K., VanderMeer, J.E., Zhao, J., Eckalbar, W.L., Logan, M., Illing, N., Pollard, K.S., and Ahituv, N. (2016). Bat accelerated regions identify a bat forelimb specific enhancer in the HoxD locus. PLoS Genet. *12*, e1005738.

Bouma, H.R., Carey, H.V., and Kroese, F.G.M. (2010). Hibernation: the immune system at rest? J. Leukoc. Biol. *88*, 619–624.

Boyd, J.L., Skove, S.L., Rouanet, J.P., Pilaz, L.-J., Bepler, T., Gordân, R., Wray, G.A., and Silver, D.L. (2015). Human-chimpanzee differences in a FZD8 enhancer alter cell-cycle dynamics in the developing neocortex. Curr. Biol. *25*, 772–779.

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

Carey, H.V., Andrews, M.T., and Martin, S.L. (2003). Mammalian hibernation: cellular and molecular responses to depressed metabolism and low temperature. Physiol. Rev. *83*, 1153–1181.

Carleton, J.B., Berrett, K.C., and Gertz, J. (2017). Multiplex enhancer interference reveals collaborative control of gene regulation by estrogen receptor α-bound enhancers. Cell Syst. *5*, 333–344.e5.

Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell *134*, 25–36.

Cassidy, S.B., and Driscoll, D.J. (2009). Prader-Willi syndrome. Eur. J. Hum. Genet. *17*, 3–13.

Castellini, G., Franzago, M., Bagnoli, S., Lelli, L., Balsamo, M., Mancini, M., Nacmias, B., Ricca, V., Sorbi, S., Antonucci, I., et al. (2017). Fat mass and obesity-associated gene (FTO) is associated to eating disorders susceptibility and moderates the expression of psychopathological traits. PLoS ONE *12*, e0173560.

Chuang, Y.-F., Tanaka, T., Beason-Held, L.L., An, Y., Terracciano, A., Sutin, A.R., Kraut, M., Singleton, A.B., Resnick, S.M., and Thambisetty, M. (2015). FTO genotype and aging: pleiotropic longitudinal effects on adiposity, brain function, impulsivity and diet. Mol. Psychiatry *20*, 133–139.

Claussnitzer, M., Dankel, S.N., Kim, K.-H., Quon, G., Meuleman, W., Haugen, C., Glunk, V., Sousa, I.S., Beaudry, J.L., Puviindran, V., et al. (2015a). FTO obesity variant circuitry and adipocyte browning in humans. N. Engl. J. Med. *373*, 895–907.

Cooper, S., Howard, D., Koch, T., Lin, X., Johnston, N., Dobbs, K., Zuberbuehler, M., Cooper, A., Lamont, C., Mortimer, S., et al. (2014). Platelet levels and activity are decreased in hibernating 13-lined ground squirrels which may prevent venous thromboembolism (879.17). FASEB J. *28*, 829.17.

Dabrowski, M., Dojer, N., Krystkowiak, I., Kaminska, B., and Wilczyński, B. (2015). Optimally choosing PWM motif databases and sequence scanning approaches based on ChIP-seq data. BMC Bioinformatics *16*, 140.

Dark, J. (2005). Annual lipid cycles in hibernators: integration of physiology and behavior. Annu. Rev. Nutr. *25*, 469–497.

Dina, C., Meyre, D., Gallina, S., Durand, E., Körner, A., Jacobson, P., Carlsson, L.M.S., Kiess, W., Vatin, V., Lecoeur, C., et al. (2007). Variation in FTO contributes to childhood obesity and severe adult obesity. Nat. Genet. *39*, 724–726.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Eckalbar, W.L., Schlebusch, S.A., Mason, M.K., Gill, Z., Parker, A.V., Booker, B.M., Nishizaki, S., Muswamba-Nday, C., Terhune, E., Nevonen, K.A., et al. (2016). Transcriptomic and epigenomic characterization of the developing bat wing. Nat. Genet. *48*, 528–536.

Erwin, D.H., and Davidson, E.H. (2009). The evolution of hierarchical gene regulatory networks. Nat. Rev. Genet. *10*, 141–148.

Felix, J.F., Bradfield, J.P., Monnereau, C., van der Valk, R.J., Stergiakouli, E., Chesi, A., Gaillard, R., Feenstra, B., Thiering, E., Kreiner-Møller, E., et al.; Bone Mineral Density in Childhood Study (BMDCS); Early Genetics and Lifecourse Epidemiology (EAGLE) consortium; Early Growth Genetics (EGG) Consortium; Bone Mineral Density in Childhood Study BMDCS (2016). Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. Hum. Mol. Genet. *25*, 389–403.

Ferris, E., Abegglen, L.M., Schiffman, J.D., and Gregg, C. (2018). Accelerated evolution in distinctive species reveals candidate elements for clinically relevant traits, including mutation and cancer resistance. Cell Rep. *22*, 2742–2755.

Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M., Perry, J.R.B., Elliott, K.S., Lango, H., Rayner, N.W., et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science *316*, 889–894.

Frith, M.C., Fu, Y., Yu, L., Chen, J.-F., Hansen, U., and Weng, Z. (2004). Detection of functional DNA motifs via statistical over-representation. Nucleic Acids Res. *32*, 1372–1381.

Geiser, F. (1998). Evolution of daily torpor and hibernation in birds and mammals: importance of body size. Clin. Exp. Pharmacol. Physiol. *25*, 736–739.

Geiser, F. (2004). Metabolic rate and body temperature reduction during hibernation and daily torpor. Annu. Rev. Physiol. *66*, 239–274.

Geiser, F. (2008). Ontogeny and phylogeny of endothermy and torpor in mammals and birds. Comp. Biochem. Physiol. A Mol. Integr. Physiol. *150*, 176–180.

Geiser, F., and Stawski, C. (2011). Hibernation and torpor in tropical and subtropical bats in relation to energetics, extinctions, and the evolution of endothermy. Integr. Comp. Biol. *51*, 337–348.

Ghosh, S., and Bouchard, C. (2017). Convergence between biological, behavioural and genetic determinants of obesity. Nat. Rev. Genet. *18*, 731–748.

Glastonbury, C.A., Viñuela, A., Buil, A., Halldorsson, G.H., Thorleifsson, G., Helgason, H., Thorsteinsdottir, U., Stefansson, K., Dermitzakis, E.T., Spector, T.D., and Small, K.S. (2016). Adiposity-dependent regulatory effects on multitissue transcriptomes. Am. J. Hum. Genet. *99*, 567–579.

Grigg, G.C., Beard, L.A., and Augee, M.L. (2004). The evolution of endothermy and its diversity in mammals and birds. Physiol. Biochem. Zool. *77*, 982–997.

Hedges, S.B., Marin, J., Suleski, M., Paymer, M., and Kumar, S. (2015). Tree of life reveals clock-like speciation and diversification. Mol. Biol. Evol. *32*, 835–845.

Heldmaier, G., Ortmann, S., and Elvert, R. (2004). Natural hypometabolism during hibernation and daily torpor in mammals. Respir. Physiol. Neurobiol. *141*, 317–329.

Hess, M.E., Hess, S., Meyer, K.D., Verhagen, L.A.W., Koch, L., Brönneke, H.S., Dietrich, M.O., Jordan, S.D., Saletore, Y., Elemento, O., et al. (2013). The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nat. Neurosci. *16*, 1042–1048.

Ho, A.J., Stein, J.L., Hua, X., Lee, S., Hibar, D.P., Leow, A.D., Dinov, I.D., Toga, A.W., Saykin, A.J., Shen, L., et al.; Alzheimer's Disease Neuroimaging Initiative (2010). A commonly carried allele of the obesity-related FTO gene is associated with reduced brain volume in the healthy elderly. Proc. Natl. Acad. Sci. USA *107*, 8404–8409.

Hohn, B.M., and Marshall, E.H. (1966). Annual and seasonal weight changes in a thirteen-lined ground squirrel population, Itasca State Park, Minnesota. J. Minn. Acad. Sci. *33*, 102–106.

Hruby, A., and Hu, F.B. (2015). The epidemiology of obesity: a big picture. Pharmacoeconomics *33*, 673–689.

Hubisz, M.J., and Pollard, K.S. (2014). Exploring the genesis and functions of Human Accelerated Regions sheds light on their role in human evolution. Curr. Opin. Genet. Dev. *29*, 15–21.

Hubisz, M.J., Pollard, K.S., and Siepel, A. (2011). PHAST and RPHAST: phylogenetic analysis with space/time models. Brief. Bioinform. *12*, 41–51.

Jani, A., Martin, S.L., Jain, S., Keys, D., and Edelstein, C.L. (2013). Renal adaptation during hibernation. Am. J. Physiol. Renal Physiol. *305*, F1521–F1532.

Jastroch, M., Giroud, S., Barrett, P., Geiser, F., Heldmaier, G., and Herwig, A. (2016). Seasonal control of mammalian energy balance: recent advances in the understanding of daily torpor and hibernation. J. Neuroendocrinol. *28*. https://doi.org/10.1111/jne.12437.

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. Cell *152*, 327–339.

Karra, E., O'Daly, O.G., Choudhury, A.I., Yousseif, A., Millership, S., Neary, M.T., Scott, W.R., Chandarana, K., Manning, S., Hess, M.E., et al. (2013). A link between FTO, ghrelin, and impaired brain food-cue responsivity. J. Clin. Invest. *123*, 3539–3551.

Kim, S.Y., and Pritchard, J.K. (2007). Adaptive evolution of conserved noncoding elements in mammals. PLoS Genet. *3*, 1572–1586.

Kostka, D., Hubisz, M.J., Siepel, A., and Pollard, K.S. (2012). The role of GC-biased gene conversion in shaping the fastest evolving regions of the human genome. Mol. Biol. Evol. *29*, 1047–1057.

Kunz, T.H., Wrazen, J.A., and Burnett, C.D. (1998). Changes in body mass and fat reserves in pre-hibernating little brown bats (Myotis lucifugus). Écoscience *5*, 8–17.

Kvon, E.Z., Kamneva, O.K., Melo, U.S., Barozzi, I., Osterwalder, M., Mannion, B.J., Tissières, V., Pickle, C.S., Plajzer-Frick, I., Lee, E.A., et al. (2016). Progressive loss of function in a limb enhancer during snake evolution. Cell *167*, 633–642.e11.

Landgraf, K., Scholz, M., Kovacs, P., Kiess, W., and Körner, A. (2016). FTO obesity risk variants are linked to adipocyte IRX3 expression and BMI of children - relevance of FTO variants to defend body weight in lean children? PLoS ONE *11*, e0161739.

Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M.F., Parker, B.J., Washietl, S., Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., et al.; Broad Institute Sequencing Platform and Whole Genome Assembly Team; Baylor College of Medicine Human Genome Sequencing Center Sequencing Team; Genome Institute at Washington University (2011). A high-resolution map of human evolutionary constraint using 29 mammals. Nature *478*, 476–482.

Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al.; LifeLines Cohort Study; ADIPOGen Consortium; AGEN-BMI Working Group; CARDIOGRAMplusC4D Consortium; CKDGen Consortium; GLGC; ICBP; MAGIC Investigators; MuTHER Consortium; MIGen Consortium; PAGE Consortium; ReproGen Consortium; GENIE Consortium; International Endogene Consortium (2015).

Genetic studies of body mass index yield new insights for obesity biology. Nature *518*, 197–206.

Loos, R.J.F., and Yeo, G.S.H. (2014). The bigger picture of FTO: the first GWAS-identified obesity gene. Nat. Rev. Endocrinol. *10*, 51–61.

Lovegrove, B.G., Lobban, K.D., and Levesque, D.L. (2014). Mammal survival at the Cretaceous-Palaeogene boundary: metabolic homeostasis in prolonged tropical hibernation in tenrecs. Proc. Biol. Sci. *281*, 20141304.

Lu, Y., Day, F.R., Gustafsson, S., Buchkovich, M.L., Na, J., Bataille, V., Cousminer, D.L., Dastani, Z., Drong, A.W., Esko, T., et al. (2016). New loci for body fat percentage reveal link between adiposity and cardiometabolic disease risk. Nat. Commun. *7*, 10495.

Martin, S.L. (2008). Mammalian hibernation: a naturally reversible model for insulin resistance in man? Diab. Vasc. Dis. Res. *5*, 76–81.

Mazon, J.N., de Mello, A.H., Ferreira, G.K., and Rezin, G.T. (2017). The impact of obesity on neurodegenerative diseases. Life Sci. *182*, 22–28.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

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

Meyer, D., Zeileis, A., and Hornik, K. (2006). The Strucplot Framework: Visualizing Multi-Way Contingency Tables with vcd. Journal of Statistical Software *17*, 1–48.

Meyre, D., Delplanque, J., Chèvre, J.C., Lecoeur, C., Lobbens, S., Gallina, S., Durand, E., Vatin, V., Degraeve, F., Proença, C., et al. (2009). Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nat. Genet. *41*, 157–159.

Morin, P., Jr., and Storey, K.B. (2009). Mammalian hibernation: differential gene expression and novel application of epigenetic controls. Int. J. Dev. Biol. *53*, 433–442.

Nowack, J., and Dausmann, K.H. (2015). Can heterothermy facilitate the colonization of new habitats? Mammal Rev. *45*, 117–127.

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

Pollard, K.S., Salama, S.R., Lambert, N., Lambot, M.-A., Coppens, S., Pedersen, J.S., Katzman, S., King, B., Onodera, C., Siepel, A., et al. (2006b). An RNA gene expressed during cortical development evolved rapidly in humans. Nature *443*, 167–172.

Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., and Meno, C. (2018). ChIP-Atlas: A data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. *19*, e46255. https://doi.org/10.15252/embr.201846255.

Pollard, K.S., Hubisz, M.J., Rosenbloom, K.R., and Siepel, A. (2010). Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. *20*, 110–121.

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

Procaccini, C., Santopaolo, M., Faicchia, D., Colamatteo, A., Formisano, L., de Candia, P., Galgani, M., De Rosa, V., and Matarese, G. (2016). Role of metabolism in neurodegenerative disorders. Metabolism *65*, 1376–1390.

Reynolds, C.M., Gray, C., Li, M., Segovia, S.A., and Vickers, M.H. (2015). Early life nutrition and energy balance disorders in offspring in later life. Nutrients *7*, 8090–8111.

Rowland, N.E., Vaughan, C.H., Mathes, C.M., and Mitra, A. (2008). Feeding behavior, obesity, and neuroeconomics. Physiol. Behav. *93*, 97–109.

Ruf, T., and Geiser, F. (2015). Daily torpor and hibernation in birds and mammals. Biol. Rev. Camb. Philos. Soc. *90*, 891–926.

Schmid, J. (1999). Sex-specific differences in activity patterns and fattening in the gray mouse lemur (*Microcebus murinus*) in Madagascar. J. Mammal. *80*, 749–757.

Schmid, J. (2000). Torpor in the tropics: the case of the gray mouse lemur (*Microcebus murinus*). Basic Appl. Ecol. *1*, 133–139.

Schmid, J., and Ganzhorn, J.U. (2009). Optional strategies for reduced metabolism in gray mouse lemurs. Naturwissenschaften *96*, 737–741.

Schubert, K.A., Boerema, A.S., Vaanholt, L.M., de Boer, S.F., Strijkstra, A.M., and Daan, S. (2010). Daily torpor in mice: high foraging costs trigger energy-saving hypothermia. Biol. Lett. *6*, 132–135.

Sevgi, M., Rigoux, L., Kühn, A.B., Mauer, J., Schilbach, L., Hess, M.E., Gruendler, T.O.J., Ullsperger, M., Stephan, K.E., Brüning, J.C., and Tittgemeyer, M. (2015). An obesity-predisposing variant of the FTO gene regulates D2R-dependent reward learning. J. Neurosci. *35*, 12584–12592.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. *15*, 1034–1050.

Singla, P., Bardoloi, A., and Parkash, A.A. (2010). Metabolic effects of obesity: a review. World J. Diabetes *1*, 76–88.

Smemo, S., Tena, J.J., Kim, K.-H., Gamazon, E.R., Sakabe, N.J., Gómez-Marín, C., Aneas, I., Credidio, F.L., Sobreira, D.R., Wasserman, N.F., et al. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature *507*, 371–375.

Speliotes, E.K., Willer, C.J., Berndt, S.I., Monda, K.L., Thorleifsson, G., Jackson, A.U., Lango Allen, H., Lindgren, C.M., Luan, J., Mägi, R., et al.; MAGIC; Procardis Consortium (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat. Genet. *42*, 937–948.

Srere, H.K., Wang, L.C., and Martin, S.L. (1992). Central role for differential gene expression in mammalian hibernation. Proc. Natl. Acad. Sci. USA *89*, 7119–7123.

Storey, J.D. (2002). A direct approach to false discovery rates. J. R. Stat. Soc. B *64*, 479–498.

Turcot, V., Lu, Y., Highland, H.M., Schurmann, C., Justice, A.E., Fine, R.S., Bradfield, J.P., Esko, T., Giri, A., Graff, M., et al.; CHD Exome+ Consortium; EPIC-CVD Consortium; ExomeBP Consortium; Global Lipids Genetic Consortium; GoT2D Genes Consortium; EPIC InterAct Consortium; INTERVAL Study; ReproGen Consortium; T2D-Genes Consortium; MAGIC Investigators; Understanding Society Scientific Group (2018). Protein-altering variants associated with body mass index implicate pathways that control energy intake and expenditure in obesity. Nat. Genet. *50*, 26–41.

Villanueva-Cañas, J.L., Faherty, S.L., Yoder, A.D., and Albà, M.M. (2014). Comparative genomics of mammalian hibernators using gene networks. Integr. Comp. Biol. *54*, 452–462.

Wray, G.A. (2007). The evolutionary significance of cis-regulatory mutations. Nat. Rev. Genet. *8*, 206–216.

Yang, Q., Xiao, T., Guo, J., and Su, Z. (2017). Complex relationship between obesity and the fat mass and obesity locus. Int. J. Biol. Sci. *13*, 615–629.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
PHAST and RPHAST	Hubisz et al., 2011	http://compgen.cshl.edu/phast/; http://compgen.cshl.edu/rphast/
BioMart	Bioconductor	http://bioconductor.org/packages/release/bioc/html/biomaRt.html
VCD	Meyer et al., 2006	https://cran.r-project.org/web/packages/vcd/index.html
Biostrings	Bioconductor	http://bioconductor.org/packages/release/bioc/html/Biostrings.html
LOLA	Bioconductor	https://bioconductor.org/packages/release/bioc/html/LOLA.html
GenomicRanges	Bioconductor	https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html
MotifDB	Bioconductor	http://bioconductor.org/packages/release/bioc/html/MotifDb.html
PWMEnrich	Bioconductor	http://bioconductor.org/packages/release/bioc/html/PWMEnrich.html
Seaview	Seaview	http://doua.prabi.fr/software/seaview
GREAT	McLean et al., 2010	http://bejerano.stanford.edu/great/public/html/
Galaxy	Galaxy	https://usegalaxy.org
Circlize	CRAN	https://cran.r-project.org/web/packages/circlize/index.html
ChIP-Atlas	Oki et al., 2018	http://chip-atlas.org
SuperExactTest	CRAN	https://cran.r-project.org/web/packages/SuperExactTest/index.html
Custom python and R scripts	*Gregg Lab, University of Utah*	Available upon request
Other		
Multiple alignments of 59 vertebrate genomes with Mouse	UCSC	http://hgdownload.cse.ucsc.edu/goldenpath/mm10/multiz60way/maf/
Multiple alignments of 59 vertebrate genomes with Mouse Neutral Model	UCSC	http://hgdownload.cse.ucsc.edu/goldenpath/mm10/phastCons60way/mm10.60way.phastCons.mod
BMI risk loci identified by GWAS	Ghosh and Bouchard, 2017; Turcot et al., 2018	N/A
PWS DE genes in the hypothalamus	Bochukova et al., 2018	N/A
ENCODE hg19 bedfiles for DNase HS (ENCFF751YPI), distal H3K27ac (ENCFF786PWS), H3K9ac (ENCFF690JTO), proximal H3K4me3 (ENCFF140OFS), distal and proximal TF binding sites (ENCFF787QYS), proximal H3K4me1 (ENCFF076KTT), USCS hg19 2K promoter table, intron table and coding exon table	UCSC Table Browser	https://genome.ucsc.edu/cgi-bin/hgTables

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Christopher Gregg (chris.gregg@neuro.utah.edu). This study did not generate new unique reagents.

METHOD DETAILS

Identification of ARs in the Target Species

Our study is based on the species available in the mm10 vertebrate 60-way UCSC multiple genome alignment file (MAF). From the background mammalian species, we identified 50 bp conserved regions using PhastCons (Hubisz et al., 2011; Siepel et al., 2005) with the following parameters: expected.length = 45, target.coverage = 0.3, rho = 0.31). From these conserved regions, we next defined regions accelerated in each of the target species using RPHAST (Hubisz et al., 2011; Pollard et al., 2010). Statistically significant ARs were defined with a false discovery rate threshold of 10% to define a relatively comprehensive set of elements for analyzing parallel ARs. The hibernation versus homeothermic status of the different species in our study was confirmed by exhaustive literature searches and only high-confidence species with published evidence were included (Table S1).

Assigning pARs to TADs

We used the R package GenomicRanges to assign pARs to TADs. We compared the numbers of hibernator and non-hibernator pARs within a given TAD to the numbers outside of the TAD with a Chi-square test.

Regulatory Binding Site Motif and Genetic Circuit Analysis

Position weight matrices (PWMs) detailing motifs bound by studied transcriptions factors (TFs) (Jolma et al., 2013) were accessed through the R package PWMEnrich. We used binding thresholds and methods described by Dabrowski et al. to write custom python scripts to find motifs altered in our 4 hibernating species relative to guinea pig, human, dog, and elephant (Dabrowski et al., 2015). These predicted gained and lost interactions were used to build predicted hibernation and obesity circuits (Table S5). Circos plots were created with the R package circlize. Multi-alignments were visualized with Seaview.

Analysis of Parallel Coding Changes

We defined coding pARs as those overlapping coding regions by >50% using bedtools. We downloaded the corresponding amino acid and nucleotide sequences from the UCSC table browser. We focused our study on amino acids and nucleotides conserved across 4 non-hibernator species from each clade: guinea pig, human, horse, and elephant. We wrote custom R scripts employing the R packages 'Biostrings' and 'limma' to identify amino acid and nucleotide changes at these conserved positions in the 4 hibernating species. Some of the amino acid changes are 'divergent'; different hibernators have different amino acid substitutions. Other changes are 'convergent'; multiple hibernators have the same amino acid substitution at a non-hibernator conserved position. In Table S6, we label amino acid changes found in two or more hibernators as 'Hib Consensus Change AA'. We used the packages biomaRt and ensembldb to assign amino acid positions to mouse transcripts and genes and to human proteins and amino acid positions.

QUANTIFICATION AND STATISTICAL ANALYSIS

Genome Quality Control Studies

To determine whether our AR discovery is impacted by variations in genome quality, we plotted N50, a measure of genome quality, for the various genomes against the fraction of conserved regions that are accelerated in each species. With an adjusted R-square value of 0.09, we concluded that variation in genome quality is not a significant driver of AR discovery. By plotting the fraction of conserved regions against the evolutionary distance (4D neutral model) between an animal and its closest relative among the background species, we found a significant relationship (adjust R-squared = 0.35) between this evolutionary distance and the fraction of conserved regions that are accelerated.

Analysis of GC-biased Gene Conversion Versus Adaptive Nucleotide Substitutions in Hibernator ARs

We used previously published methods (Kostka et al., 2012) to test for biased gene conversion versus adaptive patterns of nucleotide substitutions at ARs. We analyzed the nucleotide substitution patterns in ARs from each hibernator relative to the background species. Using RPhast and a likelihood ratio test, we compared the log likelihood of two models examining: (1) lineage-specific GC-biased substitution rate compared to the background (gBGC model) versus (2) lineage-specific increased nucleotide substitutions without GC bias (adaptive model) compared to the background. We tested if the difference between log-likelihoods of the adaptive model – the GC-biased substitution model is positive, which indicates that adaptive selection is a better fitting model. Further, we test whether a model of adaptive selection + GC-biased conversion explains significantly more of the variance in nucleotide substitutions compared to GC-biased conversion alone. Our test statistic is two times the difference of the log likelihood of the two nested models with 1 degree of freedom and is expected to asymptotically follow a chi-square distribution. We analyzed the ARs in aggregate rather than individually, because we expect that this approach results in more accurate modeling of substitution patterns across the ARs than an analysis of the individual 50bp ARs.

Statistical Identification of pARs and Calculation of Observed versus Expected Shared ARs

Expected overlaps between two AR sets were calculated as follows: Expected Overlap = ((AR1/CR)*(AR2/CR)) * CR where AR1 is the number of ARs defined for species 1 and AR2 is the number ARs defined for species2. CR is the number of conserved regions common to both animals. In the case of comparisons between species in the hibernation study and the homeotherm study; only the 280,154 CRs common to both studies and the contained ARs are considered. Similarly, expected overlaps for 3 AR sets were calculated as follows: Expected Overlap = ((AR1/CR)*(AR2/CR)(AR3/CR)) * CR. The observed overlap is the number ARs shared between species. We used the R package SuperExactTest to calculate Expected Overlap and p values for the various overlaps and corrected for multiple testing errors using the Benajmini-Hochberg method. To access the overall significance of hibernation as a variable impacting interspecies parallel accelerated evolution, p values were computed using generalized linear modeling on a quasipoisson distribution and likelihood ratio test. The full model where the response variable is the observed number of ARs shared between species and the explanatory variables are expected overlap, the neutral model distance between species, and the species pairing class (hib:hib, home:homeo) was compared to a nested model of the expected overlap and the neutral model distance between species, thereby testing the main effect of the species pairing.

Statistical Analysis of AR Enrichments at Obesity Susceptibility Genes

ARs were assigned to genes based on mm10 annotations using GREAT (McLean et al., 2010). We used the basal plus 1 MB extension default settings. Mouse homologs of the human obesity genes were determined using the R package biomaRt. Contingency tables for ARs and gene sets were analyzed and visualized with the R packages vcd.

Identification of ARs in Human Putative Functional Genomic Elements

To test whether ARs are located in putative functional elements in the human and mouse genomes and characterize the biochemical activity of these elements, we used publicly available DNase-Seq and ChIP-Seq datasets in the ChIP-Atlas (http://chip-atlas.org). Bedfiles of ARs were analyzed for human or mouse biochemically active genomic element enrichment effects in different tissues using *in silico ChIP* (https://chip-atlas.org/enrichment_analysis) using the default significance threshold of 100. Enrichments were computed relative to 100x random permutations of the input data, according to the *in silico ChIP* algorithm documentation (https://github.com/inutano/chip-atlas/wiki#virtual_chip_doc). A significant enrichment effect was thresholded based on $q < 0.05$, providing an estimated 5% false discovery rate for enrichment effects. To identify ARs associated with specific classes of biochemically active elements, the bedfiles for particular experiments detailed in the text were downloaded from the ChIP-Atlas for the peak call files and are based on a significance threshold of $q < 1x10^{-5}$ for a significant peak. Significant enrichment effects were frequently observed for multiple independent experiments annotated for the same cell class and biochemical mark in the ChIP-Atlas. In Figure 5C, we show the data for the dataset with the maximum enrichment effect compared to random permutations of the data.

DATA AND CODE AVAILABILITY

A list of the software and datasets used and associated URLs or accession database IDs are provided here:

Software and Databases

1. PHAST and RPHAST: http://compgen.cshl.edu/phast/; http://compgen.cshl.edu/rphast/
2. BioMart: http://bioconductor.org/packages/release/bioc/html/biomaRt.html
3. VCD: https://cran.r-project.org/web/packages/vcd/index.html
4. Biostrings: http://bioconductor.org/packages/release/bioc/html/Biostrings.html
5. LOLA: https://bioconductor.org/packages/release/bioc/html/LOLA.html
6. GenomicRanges: https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html
7. MotifDB: http://bioconductor.org/packages/release/bioc/html/MotifDb.html
8. PWMEnrich: http://bioconductor.org/packages/release/bioc/html/PWMEnrich.html
9. Seaview: http://doua.prabi.fr/software/seaview
10. GREAT: http://bejerano.stanford.edu/great/public/html/
11. GALAXY: https://usegalaxy.org
12. Circlize: https://cran.r-project.org/web/packages/circlize/index.html
13. ChIP-Atlas: http://chip-atlas.org
14. SuperExactTest: https://cran.r-project.org/web/packages/SuperExactTest/index.html
15. Custom python and R scripts: GREGG Lab - available upon request

Datasets

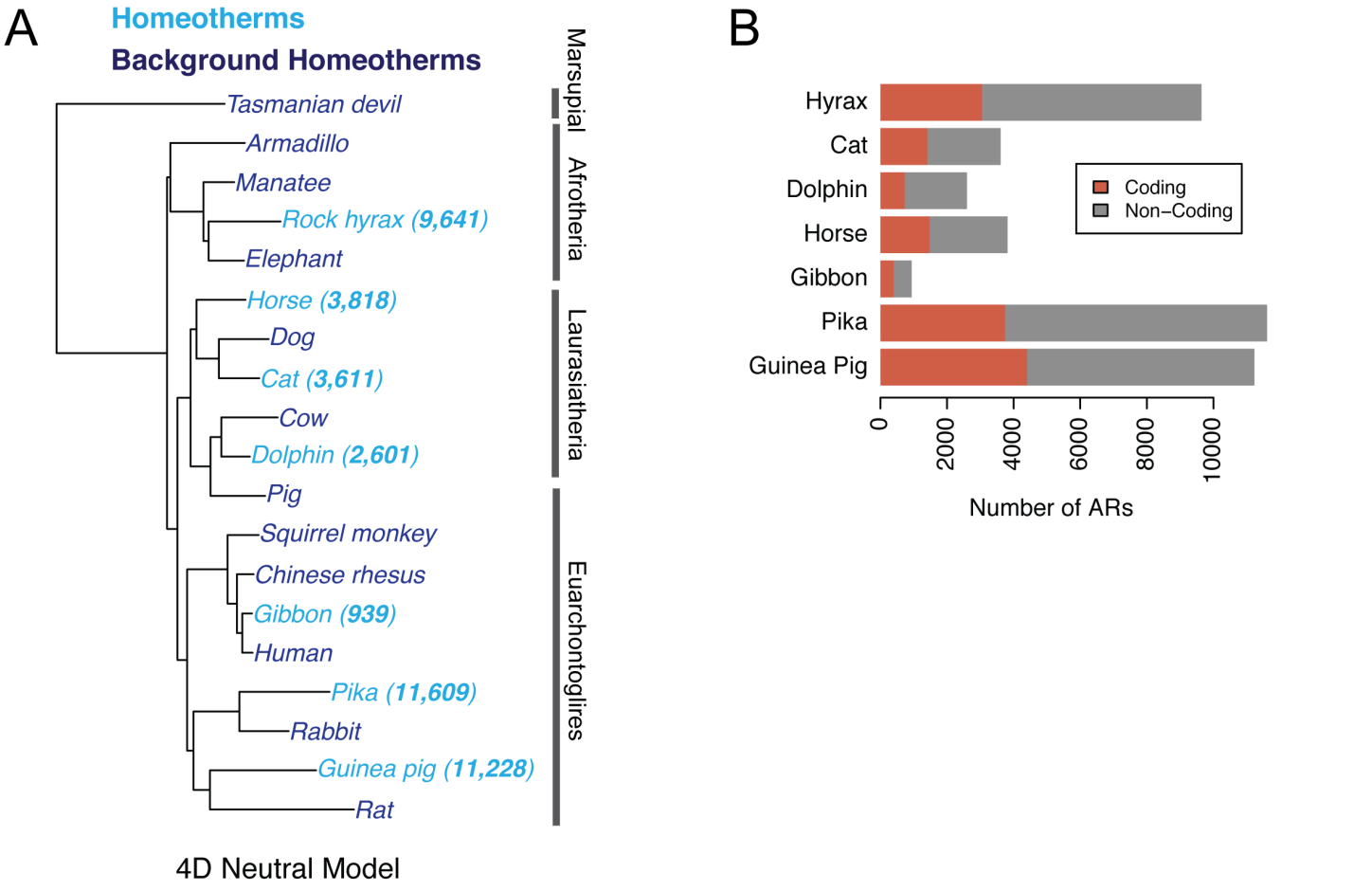
1. UCSC mm10 60-way multiple alignment file: http://hgdownload.cse.ucsc.edu/goldenPath/mm10/multiz60way/
2. 60 way mm10 Neutral Model: http://hgdownload.cse.ucsc.edu/goldenpath/mm10/phastCons60way/mm10.60way.phastCons.mod
3. BMI risk loci identified by GWAS (Ghosh and Bouchard, 2017; Locke et al., 2015; Turcot et al., 2018) – Figure 3
4. PWS DE genes in the hypothalamus (Bochukova et al., 2018) – Figure 3
5. ENCODE hg19 bedfiles for DNase HS (ENCFF751YPI), distal H3K27ac (ENCFF786PWS), H3K9ac (ENCFF690JTO), proximal H3K4me3 (ENCFF140OFS), distal and proximal TF binding sites (ENCFF787QYS), proximal H3K4me1 (ENCFF076KTT), USCS hg19 2K promoter table, intron table, and coding exon table – Figure 6

Cell Reports, Volume *29*

Supplemental Information

Parallel Accelerated Evolution in Distant Hibernators Reveals Candidate *Cis* Elements and Genetic Circuits Regulating Mammalian Obesity

Elliott Ferris and Christopher Gregg


A
Homeotherms
Background Homeotherms
Tasmanian devil
Armadillo
Manatee
Rock hyrax (9,641)
Elephant
Horse (3,818)
Dog
Cat (3,611)
Cow
Dolphin (2,601)
Pig
Squirrel monkey
Chinese rhesus
Gibbon (939)
Human
Pika (11,609)
Rabbit
Guinea pig (11,228)
Rat
Marsupial
Afrotheria
Laurasiatheria
Euarchontoglires
4D Neutral Model
B
Hyrax
Cat
Dolphin
Horse
Gibbon
Pika
Guinea Pig
Coding
Non−Coding
0
2000
4000
6000
8000
10000
Number of ARs
C
Genome Quality
and AR Number
Hibernators
Homeotherms
Fraction ARs/ Conserved Regions
0.005
0.015
0.025
50
500
5000
50000
N50 Chromosome Sizes (KB)
Adjusted R−Squared = 0.0895
P-Value = 0.1926
D
Nearest Neighbor Difference
and AR Number
Hibernators
Homeotherms
Fraction ARs/ Conserved Regions
0.005
0.015
0.025
0.1
0.2
0.3
0.4
Distance (4D Neutral Model)
Adjusted R−Squared = 0.353
P-value = 0.0316

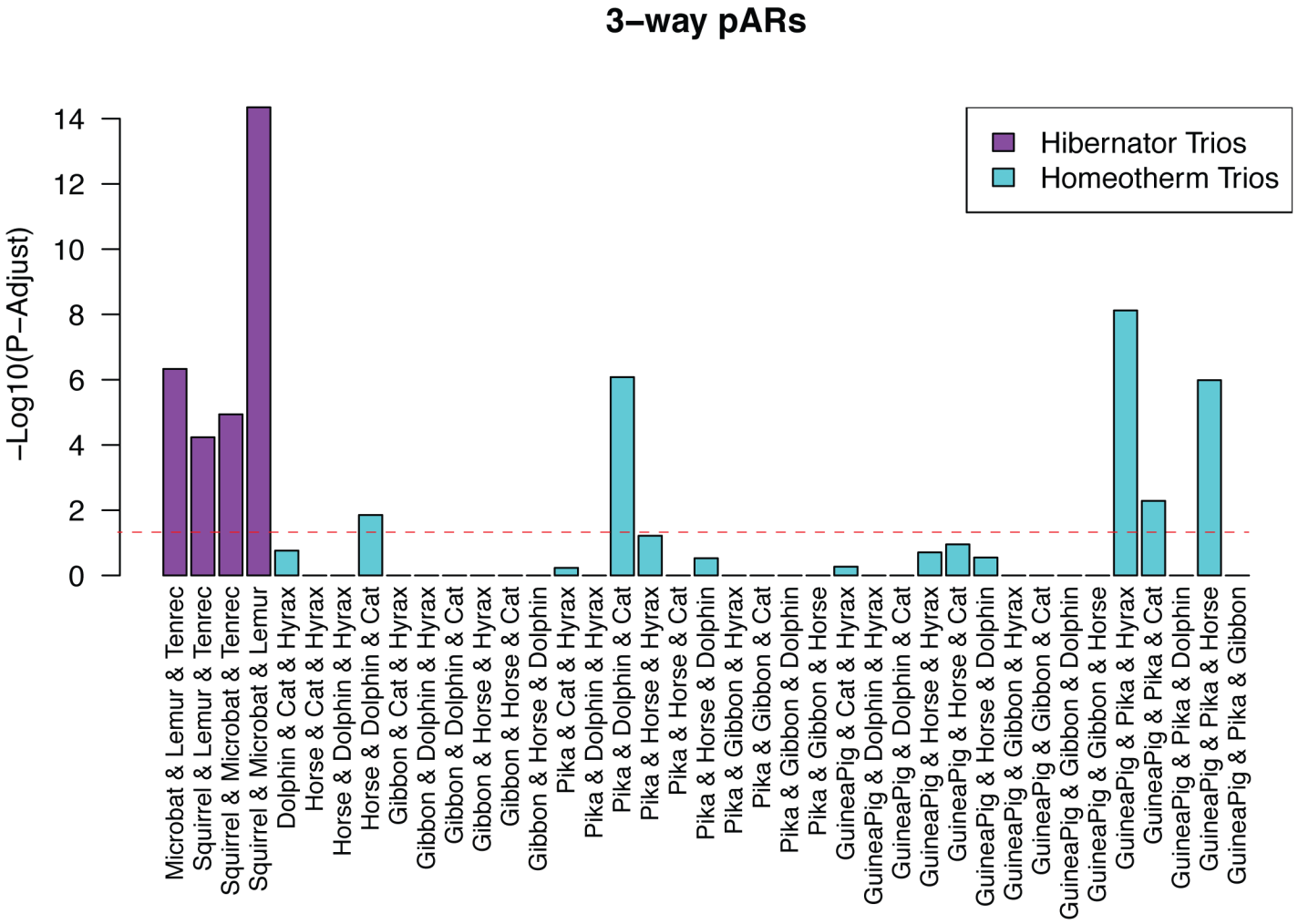
Supplemental Figure 1. Identification of ARs in distantly related non-hibernating homeotherms and the identification of factors driving AR numbers in different lineages. Related to Figure 1.

(A) Phylogenetic tree based on the 4D neutral model. Target non-hibernating homeothermic species are shown in cyan and the non-hibernator background species for conserved region identification are shown in dark blue. Numbers of ARs found in each homeotherm lineage are shown (FDR 10%). The clades are shown to the right. See Figure 1A for the hibernator species tree.

(B) Number of ARs residing in coding versus non-coding mammalian conserved regions for each homeotherm.

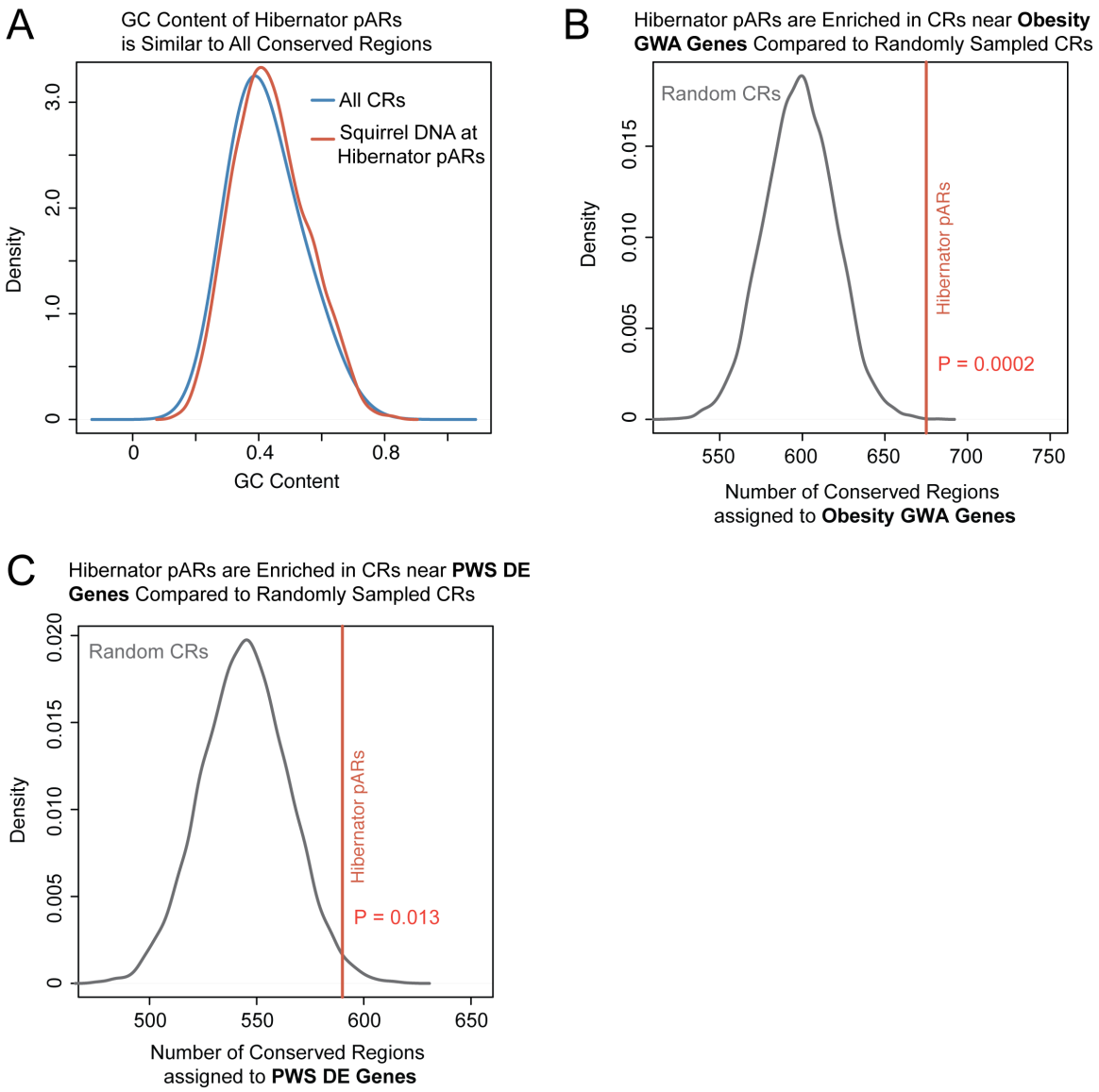
(C) Plot of the fraction of background conserved regions that are ARs in each species relative to the N50 metric of genome assembly quality. Linear regression analysis indicates that the number of ARs in each species in our study is not significantly related to N50 genome quality scores ($P>0.05$). Hibernator data are indicated by purple dots and homeotherm data are indicated by cyan dots.

(D) Plot of the fraction of background conserved regions that are ARs in each species relative to the evolutionary distance of the target species to its nearest neighboring species in the background phylogenetic tree (distance base on the 4D neutral model). A linear regression analysis revealed a significant relationship ($P=0.0316$), indicating that this is an important explanatory variable contributing to species differences in the number of observed ARs.



Supplemental Figure 2. Three-way species pARs are significant in all hibernators compared to only a minority of non-hibernators. Related to Figure 2.

The barplot shows the results of a statistical analysis of three-way AR overlaps in hibernators (purple bars) and non-hibernating homeotherms (cyan bars) using the SuperExactTest package in R for testing intersections between genomics datasets. The BH adjusted p-value for each three-way species pairing is shown (red line shows FDR = 5%). The data show that all hibernator pairings yielded significant triplet pARs, while only 5 out 35 non-hibernator pairings reached significance.

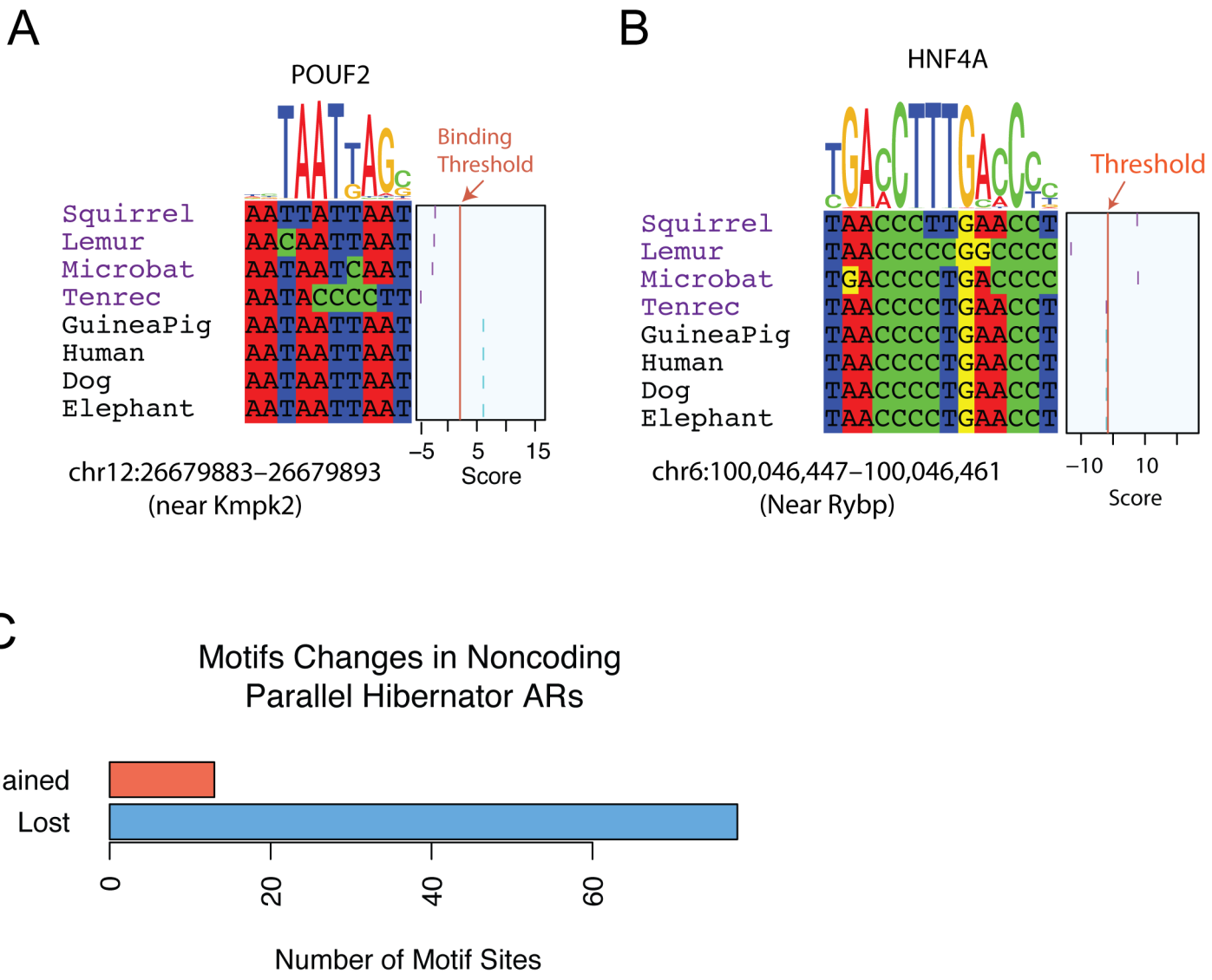


Supplemental Figure 3. Hibernator pARs are significantly enriched near obesity GWA genes and PWS DE genes compared to random conserved regions in the genome. Related to Figure 3.

(**A**) The plot shows the distribution of GC content calculated for each 50bp hibernator pAR compared to 50bp background non-hibernator conserved regions (CRs) (squirrel genome sequence data shown). The GC content is similar for the two classes of elements.

(**B** and **C**) The plots show the results of a bootstrap test to determine whether hibernator pARs are significantly enriched in CRs located proximal to obesity GWA genes (B) or PWS DE genes (C). CRs were resampled with replacement and the number proximal to the obesity GWA or PWS DE genes are shown for each sampling event (10,000 iterations). The random CR results are shown by the grey

line and compared to the number observed for hibernator pARs (red line). The results show significant enrichments for hibernator pARs near obesity GWA genes (B) and PWS DE genes (C) relative to random CRs in the genome. P-value shown in orange text.

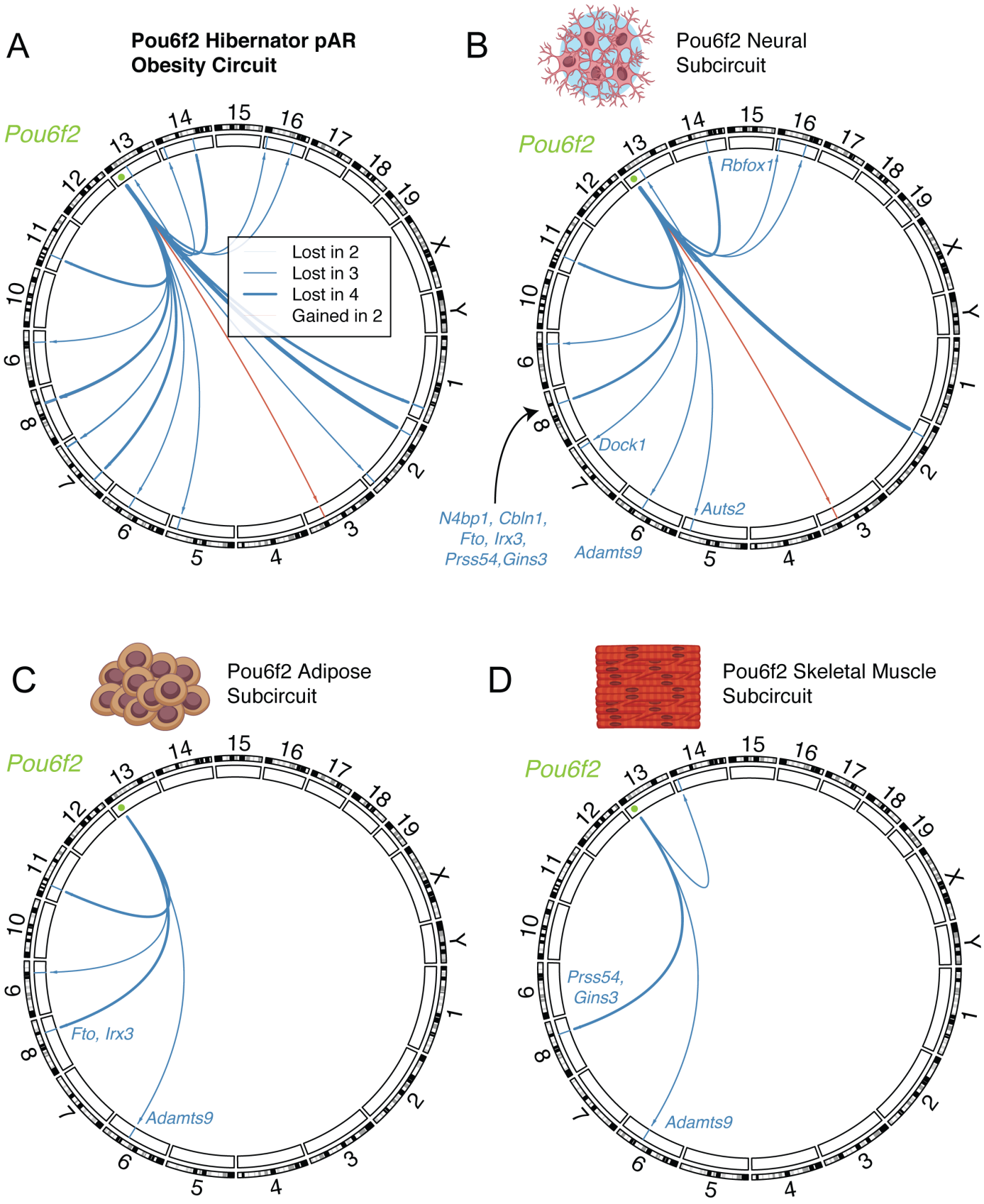


Supplemental Figure 4. Hibernator pARs change regulatory architecture and binding site motifs conserved across non-hibernating homeotherms. Related to Figure 5.

(A and B) The multiple alignments show examples of parallel lost (A) and gained (B) regulatory protein binding site motifs in hibernator pARs. The motifs for the regulatory proteins POUF2 (A) and HNF4A (B) are shown. Hibernators are highlighted in purple and non-hibernators from different clades are in black text. The computed binding score based on the observed DNA sequence for each species is plotted on the left. The threshold score for protein binding is shown by the red line. The data show that the POUF2 motif score is below the threshold in all four hibernators, but above the threshold in all four non-hibernators (A). The data therefore indicate parallel loss of the conserved

non-hibernator site in all four hibernators. For the shown HNF4A motif, the data show the binding score for the squirrel and microbat are above the threshold, but the scores from the other species are below the threshold for binding (B). Therefore, the data indicate a parallel gain of this motif in the two extreme hibernators.

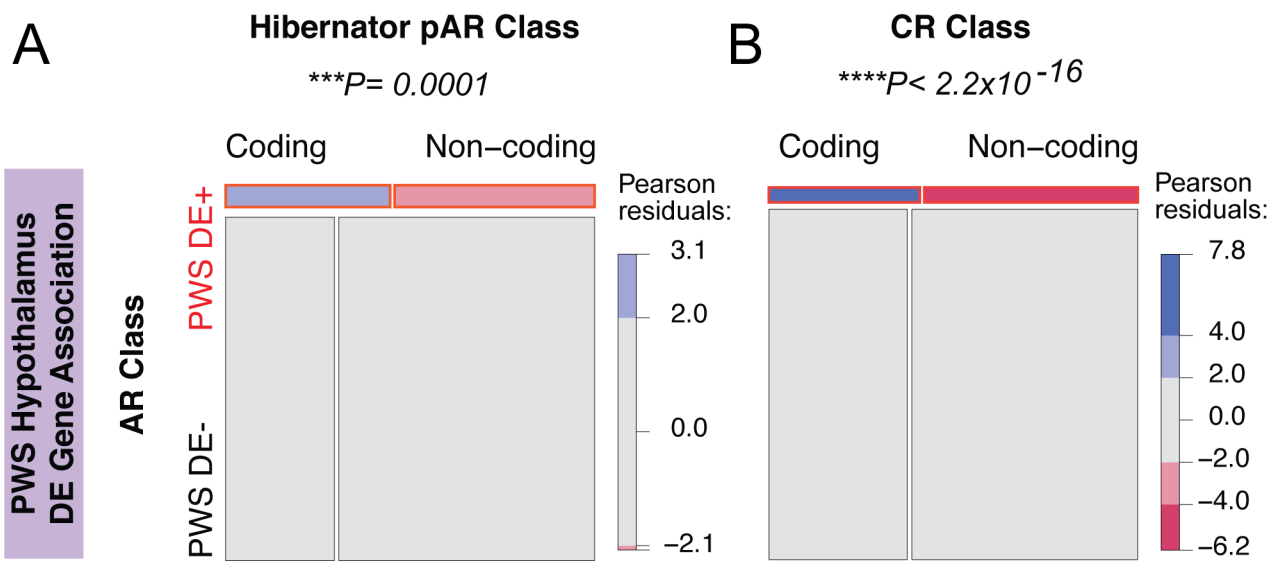
(**C**) The bar chart shows the total numbers of parallel lost conserved motifs (blue bar) and gained new motifs in the hibernators within the hibernator pARs. The data reveal that pARs most frequently remove motif sites conserved across non-hibernating homeotherms, indicating a putative loss of function effect.


A
Pou6f2 Hibernator pAR Obesity Circuit
Pou6f2
Lost in 2
Lost in 3
Lost in 4
Gained in 2
B
Pou6f2 Neural Subcircuit
Pou6f2
Rbfox1
Dock1
Auts2
Adamts9
N4bp1, Cbln1, Fto, Irx3, Prss54,Gins3
C
Pou6f2 Adipose Subcircuit
Pou6f2
Fto, Irx3
Adamts9
D
Pou6f2 Skeletal Muscle Subcircuit
Pou6f2
Prss54, Gins3
Adamts9

Supplemental Figure 5. Tissue-dependent DNAse I sensitivity of human homologs of hibernator pARs refine candidate genetic circuits for obesity and hibernation according to cell type. Related to Figure 5.

(A) The circos plot shows the genetic circuit connecting the regulatory protein Pou6f2 to the hibernator pARs near known obesity risk genes that have parallel lost or gained Pou6f2 motif sites. The connections are weighted according the number of hibernating species with the parallel motif loss (blue) or gain (red) (see legend).

(**B-D**) The circos plots show how the Pou6f2 circuit is refined into subcircuits for specific tissues by only including hibernator pARs for which the human homologs have evidence for DNAse I hypersensitivity in human neural tissue (B), adipose tissue (C) or skeletal muscle tissue (D). These results reveal candidate genetic circuits in different tissues for mammalian obesity and hibernation-related phenotypes and set foundations for targeted functional studies.



Supplementary Figure 6. Hibernator pARs and non-hibernator conserved regions associated with PWS DE genes are enriched in coding relative to non-coding regions. Related to Figure 7.

(**A**) Mosaic plot testing for associations between coding versus non-coding hibernator pARs and genes differentially expressed in the hypothalamus of PWS patients (PWS DE+) versus genes that are not (PWS DE-). The data show that pAR associations with PWS DE genes are significantly dependent on whether or not the pARs lie in coding versus noncoding sequence (Chi-Square p-values above plots). Coding hibernator pARs are positively associated with PWS DE+ genes (blue), while pARs that are non-coding show a negative association with PWS DE+ genes (red).

(B) The mosaic plot shows non-hibernator conserved regions (CRs) in coding sequence are significantly and disproportionately enriched near PWS DE genes compared to CRs in noncoding sequence.