



Published in final edited form as:

Cell Metab. 2016 October 11; 24(4): 627–639. doi:10.1016/j.cmet.2016.08.019.

Integrative Transcriptome Analyses of Metabolic Responses in Mice Define Pivotal LncRNA Metabolic Regulators

Ling Yang¹, Ping Li¹, Wenjing Yang², Xiangbo Ruan¹, Kurtis Kiewewetter¹, Jun Zhu², and Haiming Cao¹

¹Center for Molecular Medicine, National Heart, Lung and Blood Institute, NIH, Bethesda, MD 20892, USA

²Systems Biology Center, National Heart, Lung and Blood Institute, NIH, Bethesda, MD 20892, USA

SUMMARY

To systemically identify long non-coding RNAs (lncRNAs) regulating energy metabolism, we performed transcriptome analyses to simultaneously profile mRNAs and lncRNAs in key metabolic organs in mice under pathophysiologically representative metabolic conditions. Of 4759 regulated lncRNAs, function-orientated filters yield 359 tissue-specifically regulated and metabolically sensitive lncRNAs which are predicted by lncRNA-mRNA correlation analyses to function in diverse aspects of energy metabolism. Specific regulations of liver metabolically sensitive lncRNAs (lncLMS) by nutrients, metabolic hormones and key transcription factors were further defined in primary hepatocytes. Combining genome-wide screens, bioinformatics function predictions and cell-based analyses, we developed an integrative roadmap to identify lncRNA metabolic regulators. An lncLMS was experimentally confirmed in mice to suppress lipogenesis by forming a negative feedback loop in the SREBP1c pathway. Taken together, this study supports that a class of lncRNAs function as important metabolic regulators, and establishes a framework for systemically investigating the role of lncRNAs in physiological homeostasis.

Graphical Abstract

To whom correspondence should be addressed: Haiming Cao, Ph.D., Center for Molecular Medicine, National Heart, Lung and Blood Institute, NIH, Bethesda, MD 20892, USA, Phone: 301-402-9032, Fax: 301-480-0360, haiming.cao@nih.gov.

ACCESSION NUMBERS

The accession number for the microarray data reported in this paper is NCBI GEO: GSE85439.

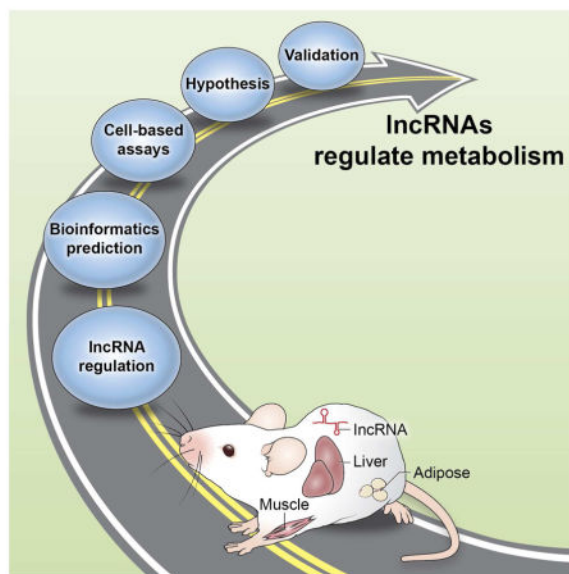
SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and nine tables can be found with this article online at

AUTHOR CONTRIBUTIONS

L.Y., P.L., X.B.R., and K.K. performed the experiments and analyzed the data. L.Y., W.Y., and J.Z. performed the bioinformatics analyses. L.Y., P.L., and H.C. wrote the manuscript. H.C. conceived and supervised the study.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



INTRODUCTION

Maintenance of metabolic homeostasis is one of the most fundamental processes required for life, and energy metabolism is intrinsically connected to nearly all essential biological and physiological activities in cells and in whole animals (Lempradl et al., 2015). From a systemic perspective, key metabolic organs in mammals engage in constant dialogues using endocrine factors to regulate overall energy balance. For example, adipose tissues release a range of adipokines to modulate the activities of the central nervous system and peripheral metabolic organs, particularly liver and muscle (Cao, 2014; Rosen and Spiegelman, 2006). Moreover, the states of systemic energy metabolism also have broad impacts on immune responses, reproduction and the pathophysiology of many disease conditions (Gregor and Hotamisligil, 2011; Metwally et al., 2007). Although great strides have been made in understanding the makeup and regulation of individual metabolic pathways in the past two decades, there are still considerable gaps in our knowledge of the complex regulatory networks governing metabolic physiology, and it remains very challenging to design effective therapies against major metabolic disorders such as obesity, diabetes and cardiovascular disease that have reached epidemic proportion globally (Eckel et al., 2005; Swinburn et al., 2011; Zimmet et al., 2014). Encouragingly, the ongoing efforts of decoding the remaining human genome have been continuously unraveling novel regulatory circuitries for fundamental biological processes, which represent a vital source of insight for better understanding the pathophysiology of metabolic disease.

Contrary to the longstanding assumption that mammalian genomes mainly encode protein-coding genes, nearly two thirds of all transcripts are noncoding RNAs primarily derived from genomic regions previously regarded as gene deserts (Harrow et al., 2012). Among the thousands of noncoding RNAs identified so far, the largest and probably also the least understood group is long noncoding RNAs, which are transcripts of 200nt or longer that lack coding potential. LncRNAs have been identified in all model organisms (Marques and

Ponting, 2014) and their number has been growing continuously (Cabili et al., 2011; Guttman et al., 2009; Iyer et al., 2015; Luo et al., 2013). It has been demonstrated that lncRNAs impact several aspects of cellular function, from chromatin modification and transcription regulation, to RNA stability, to translational control (Batista and Chang, 2013; Geisler and Collier, 2013). Evolution studies suggest that over a thousand lncRNAs could have conserved functions in mammals (Hezroni et al., 2015). Moreover, there are a significant number of human lncRNAs that overlap disease-associated single nucleotide polymorphisms (SNPs), further supporting that they could have important functions in tissue physiology and disease processes (Iyer et al., 2015). Understanding how lncRNAs engage in metabolic pathways could quickly fill in some of the gaps in our understanding of the complexity of metabolic pathophysiology and may yield valuable therapeutic targets for metabolic disorders. However, before these intriguing potentials can be effectively explored and realized, there is a need for a practical approach to efficiently identify functional lncRNAs that regulate energy metabolism in vivo. A comparative genomics approach based on information of homologous genes or protein motifs has been very successful in uncovering functions of novel protein coding genes but so far has proved to be ineffective in identifying lncRNA functions (Ulitsky and Bartel, 2013). This is because most lncRNAs have not been functionally studied, and more critically, lncRNAs are much less conserved than mRNAs even between closely related model organisms such as rat and mouse (Kutter et al., 2012). In fact, most lncRNAs are mammal-specific, thus it is difficult to infer their functional roles based on their evolutionary histories (Necsulea et al., 2014). The lack of a practical approach to predict potential functions for lncRNAs makes it particularly challenging to identify and characterize lncRNAs that regulate energy metabolism in vivo. Since energy metabolism is centrally important to most biological systems, all organisms, especially mammals, use delicate mechanisms to maintain metabolic homeostasis, and any perturbation to key metabolic pathways in an organ is often strongly compensated by rewiring of metabolic fluxes in the same organ or systemically. Thus a stringent “in vivo function” test is often needed to identify truly critical nodes in metabolic regulation. However, considering the significant efforts required to generate an lncRNA animal model, information of inferred functions for lncRNAs is often needed to choose the right target to start with. Moreover, such information can also aid in designing targeted and sensitive assays to rapidly identify specific metabolic phenotypes in a given lncRNA animal model that are often embedded in complex networks of organ communications or dampened by compensatory mechanisms. Thus, an effective approach to predict metabolic functions for lncRNAs or associate them with metabolic pathways could greatly accelerate the pace of identifying important lncRNA metabolic regulators.

To systemically evaluate the significance of lncRNAs in metabolic homeostasis, we have established an integrative roadmap to identify functional lncRNAs in metabolic regulation by combining genome-wide screens of lncRNAs transcriptomes in key metabolic tissues under pathophysiologically important metabolic conditions, extensive lncRNA-mRNA correlation analyses for lncRNA function prediction, and cellular assays connecting lncRNAs to specific metabolic pathways. Using this pipeline, we established an annotated catalog of tissue-specific lncRNA metabolic regulators and also identified a liver-specifically regulated lncRNA that robustly regulates lipid metabolism in vivo. Overall, our data support

that lncRNAs are a bona fide component of nutrient sensing and metabolic responses, and our work provides a framework for systemically identifying and mechanistically characterizing functional lncRNAs that regulate metabolic homeostasis in vivo.

RESULTS

lncRNA transcriptomes constitute robust signatures of metabolic responses in major metabolic organs

We have recently demonstrated that correlation analysis with co-regulated mRNAs can associate an lncRNA with specific metabolic pathways, which can be subsequently studied to define an lncRNA's metabolic function (Li et al., 2015). However, our previous approach relied on re-annotation of a microarray platform that was designed years ago and only includes information for a small number of currently known lncRNAs (Liao et al., 2011). To determine if a similar approach can be utilized to infer functions for currently documented lncRNAs on a genome-wide scale and to systemically identify functional lncRNA metabolic regulators, we set out to perform a range of new screens to simultaneously profile all mRNAs and lncRNAs in metabolic organs in mice.

We first tested a microarray platform, which could detect 25376 mRNA and 18431 non-redundant lncRNA transcripts. From transcriptomes in the livers of mice subject to a fasting and refeeding regimen (Figure S1A), 1328 mRNAs and 663 lncRNAs were identified to be regulated by 24h fasting *vs* ad libitum, as well as 2085 mRNAs and 1723 lncRNAs by 4h refeeding following 24h fasting *vs* fasting alone (Figure S1A). Among these fasting-regulated transcripts, 285 mRNAs and 237 lncRNAs were completely reversed by refeeding (Figure S1A), suggesting that their expression levels are stringently governed by nutrient availability and the metabolic states of the animals. The dynamic regulation of several randomly selected lncRNAs by fasting and refeeding was confirmed by quantitative real-time PCR, attesting to the quality of this method (Figure S1B). This microarray platform has been thoroughly evaluated for detecting mRNAs (Hughes et al., 2001) but to further confirm its accuracy in detecting lncRNA expression at a genome-wide level, we also performed RNA-seq for the same ad libitum and fasting liver samples as described above. For those differentially expressed lncRNAs (fasting *vs* ad libitum) identified by microarray, 93.4% could be detected by RNA-seq and the Pearson correlation between microarray and RNA-seq of expressional changes for these lncRNAs is 0.71 (Figure S1C and Table S1). The correlation for lncRNAs with higher RPKM is also higher, reaching 0.75 for lncRNAs with RPKM>0.1 and 0.8 for lncRNAs with RPKM >0.5 (Figure S1C). Taken together, these results support that this microarray profiling platform can effectively quantify both lncRNA and mRNA transcriptomes associated with metabolic responses on a genome-wide scale.

Since all coding and non-coding transcripts were quantified in parallel, our expression profile also allows the assessment and comparison of the global impact of nutrient status on the lncRNA and mRNA transcriptomes in the liver. We first performed hierarchical clustering analyses on all transcripts that were differentially expressed. As expected, mRNA expression profiles readily separate all samples into three distinct groups based on their feeding status, and samples cluster tightly within each treatment group (Figure 1A top left). Interestingly, a nearly identical pattern of sample clustering was observed for regulated

lncRNAs (Figure 1A top right), indicating that expression profiles of lncRNAs could serve as a metabolic signature in a manner similar to those of protein coding mRNAs. Consistently, principal component analyses (PCA) on all regulated transcripts, either for mRNAs or lncRNAs, readily separate all samples into distinct groups (Figure 1A bottom). These patterns suggest that regulated lncRNA and mRNA transcriptomes might function coordinately in related physiological processes and their intrinsic functional connections could be defined by performing correlation analyses of samples under multiple metabolic conditions where the lncRNA-mRNA networks are sufficiently dynamically regulated.

To systematically identify regulated lncRNAs in major metabolic organs and to establish an inclusive dataset to predict lncRNA functions in energy metabolism based on their correlation with mRNAs, we simultaneously profiled all mRNAs and lncRNAs in liver, adipose and muscle tissue in mice under the following five pathophysiologically representative metabolic conditions (Figure 1B): 24h fasting and 4h refeeding following 24h fasting, representing the extreme ends of caloric cycling, and affecting the circulating amounts of a wide array of nutrients, metabolites and metabolic hormones; short term (48 hours) high fat diet (HFD) feeding, a robust model of acute metabolic response to fatty meals (Lin et al., 2005); long-term (12 weeks) high-fat diet feeding, a widely used model mimicking dietary obesity in humans; and ob/ob mouse, a genetic model of overt obesity and steatosis (Cohen et al., 2002). In total, we performed 106 microarray analyses of liver, adipose, and muscle tissue from mice under these representative metabolic conditions to comprehensively profile lncRNA and mRNA transcriptomes (Figure S2A). First, to globally analyze metabolism-associated transcripts for all three tissues, we filtered all probes and retained 15337 mRNAs and 7975 lncRNAs that have significant expression levels in at least one biological condition regardless of tissue type. These lncRNA transcripts retained for our later analysis contain varying numbers of exons, with the majority of them containing only one exon (Figure S2B). Based on their genomic locations relative to nearby coding genes, these lncRNA transcripts are classified as bidirectional, intergenic, antisense intronic, sense intronic, antisense, and sense overlapping lncRNAs (Guttman et al., 2009; Rinn and Chang, 2012), with over 54% of them as intergenic lncRNAs (Figure S2C). Both mRNAs and lncRNAs group the 106 samples tightly by tissue type through principle component analysis (PCA) (Figure 1C), supporting that lncRNA transcriptomes, similar to those of mRNA, could effectively define the metabolic state of a key metabolic organ. Next, probes in each tissue with significant expression were filtered, and 12911, 12237, 11877 mRNAs and 6926, 5576, 5812 lncRNAs were selected for further analyses for liver, adipose and muscle tissue respectively. Gene Ontology (GO) term analyses of regulated mRNAs in each tissue under these metabolic conditions reveal that these regulated mRNA transcriptomes clearly reflect the condition- and tissue-specific metabolic responses. For example, mRNAs in the livers of ob/ob and fasted mice share very similar GO terms such as oxidation reduction and lipid metabolism (Figure S2D). Regulated mRNAs in the adipose tissues of mice fed with a HFD for 48 hours (48h) are represented by genes involved in lipid metabolism and the cell cycle, and those in the 12 weeks (12w) HFD and ob/ob groups are often involved in immune responses (Figure S2E). Interestingly, in skeletal muscle, mRNAs regulated by a 48h HFD are represented by genes involved in fat cell differentiation and immune responses, while those responsive to a 12w HFD are often genes related to RNA metabolism, which might

explain why skeletal muscles in the 12w HFD group have much more regulated mRNAs and lncRNAs as compared to other four metabolic conditions (Figures S2F and 1D). Globally, in all five conditions and three tissues, a total of 7390 mRNAs and 4759 lncRNAs were found to be regulated. Tissue-wise, 3872 (30.0%), 3197 (26.1%), 2324 (19.6%) mRNAs and 2734 (35.1%), 1831 (28.9%), 1578 (24.2%) lncRNAs were regulated by at least one of five metabolic conditions in liver, adipose, and muscle tissue respectively, suggesting that there are a large number of lncRNAs in these metabolic organs that are responsive to changes in nutrient environment and metabolic states.

Taken together, our results support that the regulation of lncRNA expression is a bona fide component of nutrient sensing and metabolic responses in major metabolic organs and that the significance of lncRNAs to metabolic homeostasis could potentially be systemic and go beyond the scope that has currently been recognized.

Integrative transcriptome analyses identify tissue-specifically regulated and metabolically sensitive lncRNAs as potential metabolic regulators

Since our ultimate goal for this comprehensive profiling and analysis is to identify functional lncRNAs in vivo, we reason that lncRNAs that are regulated by multiple metabolic conditions in our screens would be more likely to be functional in animals. Moreover, those more dynamically regulated lncRNAs would be ideal for functional prediction by lncRNA-mRNA expression correlations across different metabolic conditions. Therefore, we filtered and selected all lncRNAs that are regulated by at least three out of the five metabolic conditions in each tissue from the total 4759 regulated lncRNAs, and obtained 359 highly regulated non-redundant lncRNAs (Figure 2A and Table S2). Interestingly, we found that these lncRNAs exhibited very strong tissue-specific regulation (Figure 2A), suggesting that they might primarily play tissue-specific roles, and lncRNA-mRNA correlation analyses within each tissue would be required to pinpoint their functions. GO term analyses of those similarly selected mRNAs that are regulated by at least three conditions indicate that they are functionally implicated in diverse metabolic processes (Figures S3A–S3C, and Table S2). Furthermore, those lncRNAs regulated in at least three conditions cluster metabolic conditions of each tissue into a pattern similar to that of mRNAs (Figure 2B), suggesting that these lncRNAs could faithfully mirror the features of different metabolic conditions reflected by the mRNA signatures. We consider that this selected pool of lncRNAs are most likely to be functional as lncRNA metabolic regulators in vivo, and have termed them as liver, adipose or muscle metabolically sensitive lncRNAs (lncLMS, lncAMS or lncMMS) respectively (Table S2).

Since we have quantified the expression levels of all lncRNAs and mRNAs in parallel in several metabolic tissues under representative pathophysiological conditions, we could predict lncRNA functions based on their co-expression patterns with protein coding transcripts, which could be subsequently used to formulate hypotheses for experimentally testing their role in metabolic regulation in vivo. Therefore we evaluated the co-expression of these metabolically sensitive lncRNAs with all detectable mRNAs in liver, adipose and muscle tissue respectively by correlation analysis (Figures S3D–S3I, Tables S3). To provide an example of examining the specific roles of these metabolically sensitive lncRNAs, we

have selected eight metabolically sensitive lncRNAs from each tissue for further analyses. As shown in S4, multiple nutrient-sensing pathways, metabolic processes and reactions are often among the top inferred functions for these lncRNAs based on GO terms of correlated mRNAs. Specifically, one representative metabolically sensitive lncRNA from liver, adipose and muscle tissue was shown to be associated with steroid and triglyceride metabolism, cell cycle and division, or fat differentiation and fatty acid synthesis, respectively (Figures 2C–2E). Intriguingly, our work also captured a number of lncRNAs that have recently been shown to be involved in adipogenesis and energy metabolism, and our correlation analyses could provide further insights into their functions in cell differentiation and metabolism. For example, one lncLMS, ENSMUST00000138256, and one lncAMS, NR_030715, have both been recently shown to be related to the thermogenic gene program in brown adipose tissue (Zhao et al., 2014), and both predicted to associate with lipid transport, localization and storage in our analysis (Figure 2F and data not shown). Additionally, one lncAMS, uc009csb.1 (AK079912), and two lncLMSs, AK045415 and AK040954, have been shown to be robustly increased during white adipose differentiation in a separated report (Sun et al., 2013a). Uc009csb.1 was predicted to be associated with fat cell differentiation in adipose tissue (Figure 2G), and AK045415 and AK040954 associated with sterol and fatty acid metabolic processes in liver respectively in our analyses (Figure 2H and data not shown). These results indicate that the co-expression network we built can efficiently predict the potential metabolic functions for tissue-specifically regulated and metabolically sensitive lncRNAs.

Specific regulation of lncRNAs by nutrients, hormones and transcription factors *in vitro*

Although our expression profiles and lncRNA-mRNA correlation analyses have provided crucial information regarding each metabolically sensitive lncRNA's potential function, these predictions are often broad and could implicate multiple biological processes or metabolic pathways. Considering the challenging nature of characterizing a potential lncRNA metabolic regulator *in vivo*, we sought to acquire further experimental evidence in cells to connect these lncRNAs to specific metabolic or signaling pathways that could aid hypothesis generation and experimental design for functional tests in animals. To this end, we used mouse primary hepatocytes as a model to experimentally analyze the regulation of a set of lncLMSs by specific nutrients, hormonal factors and master transcription factors of key metabolic signaling pathways. These representative lncLMSs are strongly regulated by the fasting-refeeding regimen. AK085787, uc009kuu.1, AK016911, and AK020503 were induced by fasting and reduced by refeeding whereas Gm16551, uc008txr.1, AK033690, and AK009289 showed the opposite pattern (Figure 3A). As part of our selection criteria for metabolically sensitive lncRNAs, these lncRNAs often exhibited altered expressions in mice fed with a HFD for 48h or 12w or mice with genetic obesity (Figure 3B and 3C). From our lncRNA-mRNA correlation analyses, these lncRNAs are correlated with coding genes involved in multiple metabolism GO terms including glycolysis, gluconeogenesis, lipid biosynthesis, oxidation reduction and additional metabolic processes (Table S4).

To determine the specific effects of each nutrient and metabolic hormone on lncRNAs in hepatocytes, key nutrients and metabolic hormones were applied individually. We used glucagon and cAMP to mimic fasting signals in primary hepatocytes. As expected, the levels

of gluconeogenesis genes glucose-6-phosphatase, catalytic (G6P) and phosphoenolpyruvate carboxykinase 1 (PCK1) were significantly stimulated by glucagon or cAMP (Figures 3D and 3E). Interestingly, two fasting-induced lncRNAs, AK085787 and uc009kuu.1, and one fasting suppressed lncRNA, uc008txr.1, were either increased or reduced by glucagon, respectively (Figure 3D). The same pattern was observed in cells treated with cAMP (Figure 3E). We used fetal bovine serum, high glucose and insulin to mimic nutrient-sensitive responses, and although we observed that refeeding-induced lncRNAs tend to be induced by insulin, the regulatory patterns of these lncRNAs are not always a direct reflection of their regulation during the fasting-refeeding treatment, suggesting that a combination of nutrients and metabolic hormones might need to be used to more closely mimic a physiological change (Figures 3F–3H). Nonetheless, our results indicate that metabolically sensitive lncRNAs identified *in vivo* are often subject to the regulation of specific nutrients and hormones, and such information could help further connect them to specific metabolic processes.

Cellular nutrient sensing involves a myriad of signaling cascades that ultimately convene on a number of key transcription factors to activate or suppress their target metabolic genes (Lempradl et al., 2015). The regulation of lncRNAs by nutrients and hormones suggests that these lncRNAs might be controlled by known master transcription factors of energy metabolism. To test this idea, we expressed a number of representative metabolic transcription factors in primary hepatocytes, including HNF4 α , FOXO1, ChREBP and SREBP1c. HNF4 α is a master transcription factor of liver genes, and is known to recruit PGC1 α to regulate the gluconeogenesis pathway during fasting (Gonzalez, 2008). FOXO1 is a crucial fasting-induced transcription factor that activates the gluconeogenesis pathway and suppresses lipogenic genes (Gross et al., 2008), whereas ChREBP is a glucose-sensing transcription factor that regulates gene expressions in glucose metabolism and the lipogenic pathway during refeeding (Uyeda and Repa, 2006). SREBP1c is a master regulator of lipogenic genes and it also suppresses gluconeogenesis genes during refeeding (Horton et al., 2002). Mouse HNF4 α , or constitutively active forms of mouse FOXO1, ChREBP or SREBP1c were independently expressed in mouse primary hepatocytes, and their known targets of coding genes were examined to verify their transcriptional activity (Figures 3I–3L). Intriguingly, six out of eight lncRNAs were regulated by these transcription factors (Figures 3I–3L). For example, AK085787 and Gm16551 were strongly regulated by HNF4 α and SREBP1c respectively. This specific regulation suggests that these lncRNAs might be involved in the metabolic processes controlled by these transcription factors. Our results here suggest that physiological regulation of lncRNA expressions *in vivo* can be further connected to specific nutrient- or hormone-sensing responses and metabolic signaling pathways in cells, which could aid the hypothesis generation and experimental design for further functional tests *in vivo*.

An lncRNA metabolic regulator identified by our integrative roadmap regulates lipid metabolism *in vivo*

We have now built a roadmap to identify functional lncRNA metabolic regulators (Figure 5) by combining genome-wide screens of lncRNAs in major metabolic organs under multiple pathophysiological conditions (Figure 1), functional predictions for metabolically sensitive

lncRNAs (Figure 2), and further cell-based assays to test their regulation by individual metabolic cues and transcription factors (Figure 3). To assess if this integrative approach is sufficiently robust to identify lncRNAs that function as metabolic regulators *in vivo*, we inferred and tested the specific metabolic function for one lncRNA, Gm16551 (ENSMUST00000160017), in animals. Gm16551 was initially characterized as a capped, spliced, and polyadenylated intergenic noncoding transcript in the Ensembl database (Figure S4A). We found Gm16551 to be exclusively expressed in the liver among several tissues we examined (Figure 4A). The histone marks surrounding the Gm16551 gene also indicate that it is actively transcribed in the liver and is silenced in the heart (Figure 4B). Gm16551 expression was reduced by fasting and recovered upon refeeding in mouse liver (Figure 3A), and it was also down-regulated in the livers of dietary and genetically obese mice (Figures 3B and 3C). We also found that Gm16551 was strongly up-regulated by SREBP1c nearly 50 folds in isolated primary hepatocytes (Figure 3L). In our lncRNA-mRNA correlation analyses, Gm16551 in the liver negatively correlates with the lipid biosynthesis process including a *de novo* lipogenic gene ATP citrate lyase (ACLY), an SREBP1c target gene (Figure 4C and Table S3). We thus hypothesize that SREBP1c might induce Gm16551 expression which functions as a negative feedback regulation of the SREBP1c-induced *de novo* lipogenesis pathway in mouse liver. To test this hypothesis, we first over-expressed SREBP1c in mouse livers by adenoviruses, and consistent with our findings in cells, Gm16551 was up-regulated by SREBP1c in mice alongside known SREBP1c target genes such as fatty acid synthase (FAS) (Figures 4D and S5B). To specifically examine Gm16551's role in regulating the lipogenic pathway *in vivo*, we knocked it down in mouse livers by administering two independent Gm16551 shRNA adenoviruses (Figures 4E and S5C). In support of our hypothesis, three key enzymes in the *de novo* lipogenesis pathway, ACLY, FAS, and stearoyl-Coenzyme A desaturase 1 (SCD1), were found to be increased at both mRNA and protein levels in the livers of Gm16551 knockdown mice compared with those of control mice (Figures 4F, 4G and S5C). Consistently, plasma triglyceride (TG) levels were also elevated in Gm16551 knockdown mice although their hepatic TG levels were not altered (Figures 4H, S5D and S5E). To understand how Gm16551 KD mediates these effects, we went on to study the mechanism of action of Gm16551. We noticed that the closest neighboring gene of Gm16551 is *Onecut1*, which is also known as *Hnf6a* whose function could potentially explain the observations associated with Gm16551 KD. To test this possibility, we defined the full length of the Gm16551 transcript in mouse liver using rapid amplification of complementary DNA ends (RACE), and found that it has no overlap with the transcripts from *Onecut1* (Figure S4A).

Furthermore, we examined expression levels of *Onecut1* upon Gm16551 KD and found there was no change (Figures 4E and S5C). Therefore, Gm16551 does not appear to *cis* regulate *Onecut1*. Since Gm16551 is robustly induced by SREBP1c, Gm16551 might be directly involved in the regulation of SREBP1c activity. To test this possibility, we first simultaneously delivered adenoviruses to express Gm16551 shRNAs and a dominant negative SREBP1c into mice, and found that the up-regulation of lipogenic genes by Gm16551 KD is abolished by the presence of the dominant negative SREBP1c (Figures 4I and S5F). This result indicates that enhanced SREBP1c activity is required for Gm16551 KD to mediate its effects on lipogenesis. Next, we overexpressed both an active form of

SREBP1c and Gm16551 and found that lipogenic gene expression as well as elevated circulation triglycerides induced by the active SREBP1c were suppressed by increased expression of Gm16551 (Figures 4J and 4K), supporting that Gm16551 is able to suppresses SREBP1c activities in vivo. Taken together, our findings support a model where SREBP1c induces Gm16551 to activate an lncRNA-mediated negative feedback loop to regulate SREBP1c activity in mouse liver. These results support that our lncRNA function prediction strategy combining correlation analysis of in vivo expression profiles and information of nutrient and hormonal regulation in cells could specifically identify connections between a metabolically sensitive lncRNA and a major metabolic process, which could be quickly verified by performing hypothesis-based metabolic assays in vivo.

The significant effects of Gm16551 on de novo lipogenesis also prompt us to search for its human homolog, which would allow for defining its impact on human lipid metabolism. We used the mouse Gm16551 sequence to blast against the Ensembl human transcripts database and identified one human lncRNA transcript that shows high sequence similarity with mouse Gm16551 (Figure S4I). We designed RT-PCR primers to quantify this human sequence, and found that it shows liver-specific expression in humans, as was seen in mice. Expression of the Gm16551 human homolog is only detected in human adult and fetal livers (Figure S4J). Interestingly, the Gm16551 human homolog has recently been shown to be significantly increased in the livers of patients with non-alcoholic fatty liver diseases (Figure S4K, GEO accession No. GSE72756) (Sun et al., 2015). Taken together, Gm16551 has conserved expression, tissue distribution and metabolic disease-associated regulation in both human and mouse, which indicate that the human homolog of Gm16551 might have a similar function to what we have characterized in mice.

Finally, to leverage the full potential of the functional predictions for these lncRNAs regulated in multiple pathophysiological mouse models for studying important human lncRNAs in metabolic homeostasis and disease, we also provide a catalog of human homologs for mouse metabolically sensitive lncRNAs (Table S5). This database could serve as a foundation for in-depth analyses of mouse lncRNAs and their human homologs in vivo, which represents a vital step in translating the exciting new developments in lncRNA biology into therapeutic advances in metabolic disorders.

DISCUSSION

We have utilized genome-wide screens to establish a comprehensive catalog of lncRNAs implicated in metabolic regulation and a battery of bioinformatics and experimental analyses to develop a roadmap for identifying and characterizing functional lncRNAs in metabolic homeostasis (Figure 5). Our results support that lncRNAs are a bona fide component of nutrient sensing and metabolic responses, and a class of metabolically sensitive lncRNAs could constitute important metabolic regulators in vivo. Several features of lncRNAs involved in energy metabolism were reaffirmed from this study. First, lncRNA expression patterns change systemically and coordinately in response to altered metabolic state, in a way similar to those of mRNAs. Second, expression levels of lncRNAs are regulated by key nutrient factors and metabolic hormones, and are under the control of master transcription factors of energy homeostasis. Third, there are core groups of metabolism-associated

lncRNAs regulated in major metabolic organs by metabolic conditions, and they often exhibit strikingly altered expression in mice with metabolic disorders, supporting their potential pathophysiological significance. Finally, although the conservation between mouse and human lncRNAs is much lower than that of mRNAs, we were able to identify a short list of lncRNAs that are implicated in metabolic regulation in both human and mouse that could serve as a starting point for characterizing the role of lncRNAs in the pathogenesis of human metabolic disorders.

Currently, it is still very challenging to identify and functionally characterize lncRNAs regulating metabolic homeostasis in animals, and the very small number of lncRNAs shown to regulate energy metabolism in vivo precludes an objective assessment of the role of lncRNAs in metabolic physiology. To overcome this major hurdle, we developed a functional lncRNA detection pipeline by first comprehensively profiling lncRNAs in important metabolic organs under representative metabolic conditions, and then selecting a group of dynamically regulated lncRNAs that demarcate tissue-specific metabolic responses for functional prediction through correlation analyses. This integrative approach allows us to effectively reduce 4759 lncRNAs regulated by representative pathophysiological conditions to a concise and annotated catalog of 359 putative tissue-specific lncRNA metabolic regulators. It is worth noting that lncRNA functions predicted from co-expressed mRNAs are often broad or associated with multiple metabolic pathways. Thus, complementary to these analyses, we used cell-based assays to further define individual nutrients, metabolic hormones and key transcription factors that regulate an lncRNA's expression, which could directly connect a putative lncRNA metabolic regulator to specific metabolic signaling pathways. Finally, all information was combined to develop a hypothesis to guide in vivo metabolic assay design to define an lncRNA's function in systemic energy metabolism (Figure 5). Using this approach, we have identified a liver metabolically sensitive lncRNA, Gm16551, which was strongly regulated by the master transcription factor of lipogenesis, SREBP1c, in both isolated hepatocytes and in mouse liver. Subsequently, we generated liver-specific Gm16551 knockdown mice and identified increased expression of key lipogenic genes and elevated circulating levels of triglycerides in these mice. These results validate the robustness of our pipeline, and provide a proof-of-principle that specific functions of lncRNA metabolic regulators can be identified by combining regulation information and correlation analyses to infer functions with cell-based analyses to identify their links with specific signaling and metabolic pathways. Furthermore, demonstration of an in vivo function for Gm16551 is also a significant step in connecting lncRNAs to the maintenance of metabolic homeostasis, since so far there are only a few lncRNAs that have been shown to be able to significantly regulate an important metabolic pathway in vivo (Lan et al., 2015; Li et al., 2015; Sallam et al., 2016; Zhao and Lin, 2015). We fully expect that there are many more lncRNAs critical to metabolic homeostasis awaiting to be discovered and characterized, and we hope that our functional lncRNA detection pipeline could significantly speed up the process of identifying and mechanistically characterizing these lncRNA metabolic regulators which could provide fresh insights into the complex network of metabolic regulation and pathogenesis of metabolic disorders.

It is currently unknown how many lncRNAs are expressed in each metabolic tissue in mice and how they are regulated under physiological or disease conditions. As part of our efforts

to identify functional lncRNAs in energy metabolism, we have established a comprehensive and annotated catalog of lncRNAs in key metabolic organs in mice under representative pathophysiological conditions. Of note, several lncRNAs that were reported recently to be involved in adipogenesis and brown fat function (Sun et al., 2013a; Zhao et al., 2014) are successfully identified as metabolically sensitive lncRNAs by our roadmap and our correlation analyses could provide additional insights into their potential functions in energy metabolism. Furthermore, our work also set the stage to investigate the significance of human lncRNAs in human metabolic disease. Accumulating evidence supports that human lncRNAs harbor a large number of disease-associated genetic variants (Iyer et al., 2015), but causally connecting these variants to specific disease conditions often requires detailed mechanistic analyses in an animal model. Since the mouse is one of most widely used pre-clinical animal models, it is important to understand if sequence-conserved lncRNAs can carry out similar functions in mouse and human and if genetic mouse models of lncRNAs can provide insights into disease relevance for their human counterparts. To facilitate this line of investigation, we have performed a careful database search to build a list of human homologs for all mouse metabolically sensitive lncRNAs (Table S5). Regulation information of these lncRNAs in human and mouse metabolic disease could provide initial clues on their relevance to pathological conditions in humans, and further study of genetic mouse models for selected lncRNAs could define their roles in the pathogenesis of human metabolic disorders.

In conclusion, our work has systemically characterized lncRNA expression in metabolic homeostasis and disease and established a roadmap to identify and functionally characterize metabolism-associated lncRNAs in vivo and in vitro (Figure 5). We have also generated a catalog of potential lncRNA metabolic regulators with predicted functions. In junction with the list of potential lncRNA metabolic regulators that are conserved in human and mouse, this work could offer useful information for delineating the metabolic functions of lncRNAs in mouse and human and for further evaluating their therapeutic potentials in human disease.

EXPERIMENTAL PROCEDURES

Animal experiments

All animal experiments were performed in accordance and with approval from the NHLBI Animal Care and Use Committee. Male C57BL/6 mice from the Jackson Laboratory were used for all experiments. For in vivo lncRNA functional test, mice were purchased at 8 weeks of age and housed 3–5 mice per cage with free access to water and normal chow diet (24% kcal from protein, 14% kcal from fat, and 62% kcal from carbohydrate, NIH-31, Harlan Teklad), and animals were acclimatized to the housing conditions for at least 10–14 days before experiments. Groups of co-housed mice were randomly assigned to experimental groups with age and weight matched between groups. Models with different metabolic conditions were established for profiling the transcriptomes in liver, white adipose and muscle tissue. For the fasting and refeeding regimen, mice of around 20 weeks of age were either allowed free access to food or were subjected to a 24-hour fast before being euthanized for tissue harvest (sacrificed at 10am). A third group was fasted for 24 hours and then allowed to feed ad libitum for another 4 hours before tissue harvest (sacrificed at 2pm).

For high fat diet feeding (20% kcal from protein, 60% kcal from fat, and 20% kcal from carbohydrate, D12492, Research Diets Inc.), 6 week old mice were fed for a total of 12 weeks, a second group of mice of 18 weeks old were fed for 48h. Mice fed with a normal chow diet were used as controls (sacrificed at 2pm). For the genetic obesity model, 10 week old Ob/Ob mice and their lean control were used. All mice were subjected to a 4–5h food withdrawal before tissue harvest unless specified otherwise (sacrificed at 2pm).

Microarray

Total RNA was isolated from tissue samples using Trizol reagent (Invitrogen). For transcriptome profiling, an Agilent array platform was used (8 × 60K, ArrayStar), and sample preparation and microarray hybridization were performed based on the manufacturer's standard protocols with minor modifications (ArrayStar, Rockville, MD). Briefly, mRNA was purified from 1 µg total RNA after removal of rRNA (mRNA-ONLY™ Eukaryotic mRNA Isolation Kit, Epicentre). Each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcript without 3' bias utilizing a random priming method. The labeled cRNAs were hybridized onto the Mouse LncRNA Array V2.0 (8 × 60K, ArrayStar). After slides were washed, the arrays were scanned by the Agilent Scanner G2505B. Agilent Feature Extraction software (version 10.7.3.1) was used to analyze acquired array images. Quantile normalization and subsequent data processing were performed using the GeneSpring GX v11.5.1 software package (Agilent Technologies). Since our understanding of the lncRNAs has evolved very quickly in most recent years, we used the NONCODE 2016 mouse database (<http://www.noncode.org/>) to re-annotate all of the lncRNA probes on the microarray platform. By cross-referencing with the latest NONCODE database, we found our microarray platform could detect total 18431 non-redundant lncRNAs, which were used for downstream analysis.

Bioinformatics Analysis of lncRNA and mRNA transcriptomes and individual lncRNAs

Targets with at least 4 out of 12 samples having flags in Present or Marginal were chosen for further analyses for the initial 12 fasting and refeeding liver samples. For the entire sample set, normalization was performed across the 106 samples. We used a cutoff of normalized array values (log2-transformed values) ≥ 7.64 to filter out the very low- or non-expressing targets for subsequent analyses of all array datasets. We performed Student's t test to identify differentially expressed transcripts with $p < 0.05$ and fold change ≥ 2 . The normalized intensities of the differentially expressed genes calculated by one-way ANOVA in the initial 12 liver samples were used for principal component analysis (PCA) and hierarchical clustering by R packages.

The normalized intensities of all expressed genes in 106 samples were used for PCA using R packages. The hierarchical clustering of mRNAs and lncRNAs regulated at least three times was prepared with Multiexperiment Viewer. 2-circle or 3-circle venn diagrams were drawn by BioVenn, and 5-circle venn diagrams were created by adding numbers of specially and commonly regulated lncRNAs and mRNAs in each condition to a pre-drawn template. Gene Ontology (GO) analyses were performed by DAVID Bioinformatics Resources 6.7 (<https://david.ncifcrf.gov/>) ($p < 0.05$ unless otherwise specified). The correlation coefficients of mRNAs and lncRNAs were made by R packages based on Pearson Correlation method. The

mRNAs that show correlation coefficients with each metabolically sensitive lncRNA >0.7 (positive) or <-0.7 (negative) were used to perform GO analysis separately.

The protein-coding potential of representative liver metabolically sensitive lncRNAs was evaluated by two widely used algorithms, Coding-Non-Coding Index (CNCI) (Sun et al., 2013b) and Coding Potential Assessment Tool (CPAT) (Wang et al., 2013). For CNCI, the transcripts with scores more than 0 are classified as coding while less than 0 are non-coding. For CPAT, mouse coding probability (CP) cutoff is 0.44, so transcripts with scores more than 0.44 are classified as “coding” and less than 0.44 as “noncoding”.

For the human homolog search, mouse lncRNA sequences were aligned to human transcript sequences using NCBI blast + (blastn). Those lncRNAs that can align with Human lncRNAs with $e\text{-value} < 1e-5$ and (alignment > 50 or alignment/human_RNA_length > 0.15) were identified and the human lncRNAs with the longest aligned sequence were listed.

Adenovirus production and in vivo adenovirus administration

ShRNAs for Gm16551 were designed to act against the mouse sequence using the online i-Score shRNA designer. Two pair of shRNAs (shRNA1: GATCTTTCCTCAGACTCAA; shRNA2: AGAACTCAATAGTAAATTA) were shown to be able to efficiently knock down Gm16551 in cells. The hairpin template oligonucleotides were synthesized by Integrated DNA Technologies and were subsequently cloned into the adenoviral vector of the pAd/Block-it system (Invitrogen) to produce adenoviruses according to the manufacturer's protocols. YFP (Addgene plasmid #15302), constitutively active mouse FOXO1 (Addgene plasmid #17547), constitutively active mouse SREBP1c (Addgene plasmid #8883), and dominant negative Rat SREBP1c (Addgene plasmid #8885) were subcloned into the adenoviral vector pAd/CMV/V5-DEST™ (Invitrogen) from plasmids purchased from Addgene. A constitutively active mouse ChREBP lacking the N-terminal low glucose inhibitory domain was PCR amplified from mouse liver cDNA (Poupeau and Postic, 2011). The full length of the Gm16551 transcript was cloned into pAd/CMV/V5-DEST™ based on the RACE results. (Forward Primer: TCAGAGCTTTCTCGGGCCCA, Reverse Primer: TCAATATCGAAATTAGATTTATTGTC). Adenoviruses were amplified in HEK293A cells and purified by CsCl gradient centrifugation. Purified viruses were desalted with PD10 columns (GE Healthcare Life Sciences) and were tittered with an Adeno-X Rapid Titer Kit (Clontech). Mouse HNF4α adenovirus was a gift from Dr. Yanqiao Zhang (NEOMED, Ohio). Adenoviruses were delivered into mice intravenously at $1-2 \times 10^9$ pfu/mouse. After seven to twelve days, animal experiments were performed, and tissue and plasma samples were harvested for analyses.

Statistical analysis

Values represent mean \pm SEM. Statistical significance of differences was determined by Student's t test or One-way ANOVA with Bonferroni's post-hoc comparison where appropriate. P values less than 0.05 were considered to be significant.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Division of Intramural Research of the National Heart Lung and Blood Institute, NIH (HL006103 and HL006159). We thank Yanqin Yang in NHLBI DNA sequencing core facility and Gangqing Hu in the laboratory of Keji Zhao for advises on bioinformatics and statistical analyses and thank Yoshiyuki Wakabayashi and Poching Liu in the NHLBI DNA sequencing core facility for their help with RNA-seq analysis.

References

- Batista PJ, Chang HY. Long Noncoding RNAs: Cellular Address Codes in Development and Disease. *Cell*. 2013; 152:1298–1307. [PubMed: 23498938]
- Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011; 25:1915–1927. [PubMed: 21890647]
- Cao H. Adipocytokines in obesity and metabolic disease. *J Endocrinol*. 2014; 220:T47–59. [PubMed: 24403378]
- Cohen P, Miyazaki M, Socci ND, Hagge-Greenberg A, Liedtke W, Soukas AA, Sharma R, Hudgins LC, Ntambi JM, Friedman JM. Role for stearoyl-CoA desaturase- 1 in leptin-mediated weight loss. *Science*. 2002; 297:240–243. [PubMed: 12114623]
- Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet*. 2005; 365:1415–1428. [PubMed: 15836891]
- Geisler S, Collier J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol*. 2013; 14:699–712. [PubMed: 24105322]
- Gonzalez FJ. Regulation of hepatocyte nuclear factor 4 alpha-mediated transcription. *Drug Metab Pharmacokinet*. 2008; 23:2–7. [PubMed: 18305369]
- Gregor MF, Hotamisligil GS. Inflammatory mechanisms in obesity. *Annu Rev Immunol*. 2011; 29:415–445. [PubMed: 21219177]
- Gross DN, van den Heuvel AP, Birnbaum MJ. The role of FoxO in the regulation of metabolism. *Oncogene*. 2008; 27:2320–2336. [PubMed: 18391974]
- Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009; 458:223–227. [PubMed: 19182780]
- Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S, et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*. 2012; 22:1760–1774. [PubMed: 22955987]
- Hezroni H, Koppstein D, Schwartz MG, Avrutin A, Bartel DP, Ulitsky I. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep*. 2015; 11:1110–1122. [PubMed: 25959816]
- Horton JD, Goldstein JL, Brown MS. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J Clin Invest*. 2002; 109:1125–1131. [PubMed: 11994399]
- Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR, et al. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol*. 2001; 19:342–347. [PubMed: 11283592]
- Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet*. 2015; 47:199–208. [PubMed: 25599403]
- Kutter C, Watt S, Stefflova K, Wilson MD, Goncalves A, Ponting CP, Odom DT, Marques AC. Rapid turnover of long noncoding RNAs and the evolution of gene expression. *PLoS Genet*. 2012; 8:e1002841. [PubMed: 22844254]
- Lan X, Yan J, Ren J, Zhong B, Li J, Li Y, Liu L, Yi J, Sun Q, Yang X, et al. A Novel Long Non-coding RNA Lnc-HC Binds hnRNP A2B1 to Regulate Expressions of Cyp7a1 and Abca1 in Hepatocytic Cholesterol Metabolism. *Hepatology*. 2015
- Lempradl A, Pospisilik JA, Penninger JM. Exploring the emerging complexity in transcriptional regulation of energy homeostasis. *Nat Rev Genet*. 2015; 16:665–681. [PubMed: 26460345]

- Li P, Ruan X, Yang L, Kiesewetter K, Zhao Y, Luo H, Chen Y, Gucek M, Zhu J, Cao H. A liver-enriched long non-coding RNA, lncLSTR, regulates systemic lipid metabolism in mice. *Cell Metab.* 2015; 21:455–467. [PubMed: 25738460]
- Liao Q, Liu C, Yuan X, Kang S, Miao R, Xiao H, Zhao G, Luo H, Bu D, Zhao H, et al. Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. *Nucleic Acids Res.* 2011; 39:3864–3878. [PubMed: 21247874]
- Lin J, Yang R, Tarr PT, Wu PH, Handschin C, Li S, Yang W, Pei L, Uldry M, Tontonoz P, et al. Hyperlipidemic effects of dietary saturated fats mediated through PGC-1 β coactivation of SREBP. *Cell.* 2005; 120:261–273. [PubMed: 15680331]
- Luo H, Sun S, Li P, Bu D, Cao H, Zhao Y. Comprehensive characterization of 10,571 mouse large intergenic noncoding RNAs from whole transcriptome sequencing. *PLoS One.* 2013; 8:e70835. [PubMed: 23951020]
- Marques AC, Ponting CP. Intergenic lncRNAs and the evolution of gene expression. *Curr Opin Genet Dev.* 2014; 27:48–53. [PubMed: 24852186]
- Metwally M, Li TC, Ledger WL. The impact of obesity on female reproductive function. *Obes Rev.* 2007; 8:515–523. [PubMed: 17868286]
- Necsulea A, Soumillon M, Warnefors M, Liechti A, Daish T, Zeller U, Baker JC, Grutzner F, Kaessmann H. The evolution of lncRNA repertoires and expression patterns in tetrapods. *Nature.* 2014; 505:635–640. [PubMed: 24463510]
- Poupeau A, Postic C. Cross-regulation of hepatic glucose metabolism via ChREBP and nuclear receptors. *Biochim Biophys Acta.* 2011; 1812:995–1006. [PubMed: 21453770]
- Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem.* 2012; 81:145–166. [PubMed: 22663078]
- Rosen ED, Spiegelman BM. Adipocytes as regulators of energy balance and glucose homeostasis. *Nature.* 2006; 444:847–853. [PubMed: 17167472]
- Sallam T, Jones MC, Gilliland T, Zhang L, Wu X, Eskin A, Sandhu J, Casero D, Vallim TQ, Hong C, et al. Feedback modulation of cholesterol metabolism by the lipid-responsive non-coding RNA LeXis. *Nature.* 2016; 534:124–128. [PubMed: 27251289]
- Sun C, Liu X, Yi Z, Xiao X, Yang M, Hu G, Liu H, Liao L, Huang F. Genome-wide analysis of long noncoding RNA expression profiles in patients with non-alcoholic fatty liver disease. *IUBMB Life.* 2015
- Sun L, Goff LA, Trapnell C, Alexander R, Lo KA, Hacisuleyman E, Sauvageau M, Tazon-Vega B, Kelley DR, Hendrickson DG, et al. Long noncoding RNAs regulate adipogenesis. *Proc Natl Acad Sci U S A.* 2013a; 110:3387–3392. [PubMed: 23401553]
- Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R, Zhao Y. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 2013b; 41:e166. [PubMed: 23892401]
- Swinburn BA, Sacks G, Hall KD, McPherson K, Finegood DT, Moodie ML, Gortmaker SL. The global obesity pandemic: shaped by global drivers and local environments. *Lancet.* 2011; 378:804–814. [PubMed: 21872749]
- Ulitisky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell.* 2013; 154:26–46. [PubMed: 23827673]
- Uyeda K, Repa JJ. Carbohydrate response element binding protein, ChREBP, a transcription factor coupling hepatic glucose utilization and lipid synthesis. *Cell Metab.* 2006; 4:107–110. [PubMed: 16890538]
- Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res.* 2013; 41:e74. [PubMed: 23335781]
- Zhao XY, Li S, Wang GX, Yu Q, Lin JD. A long noncoding RNA transcriptional regulatory circuit drives thermogenic adipocyte differentiation. *Mol Cell.* 2014; 55:372–382. [PubMed: 25002143]
- Zhao XY, Lin JD. Long Noncoding RNAs: A New Regulatory Code in Metabolic Control. *Trends Biochem Sci.* 2015; 40:586–596. [PubMed: 26410599]
- Zimmet PZ, Magliano DJ, Herman WH, Shaw JE. Diabetes: a 21st century challenge. *Lancet Diabetes Endocrinol.* 2014; 2:56–64. [PubMed: 24622669]

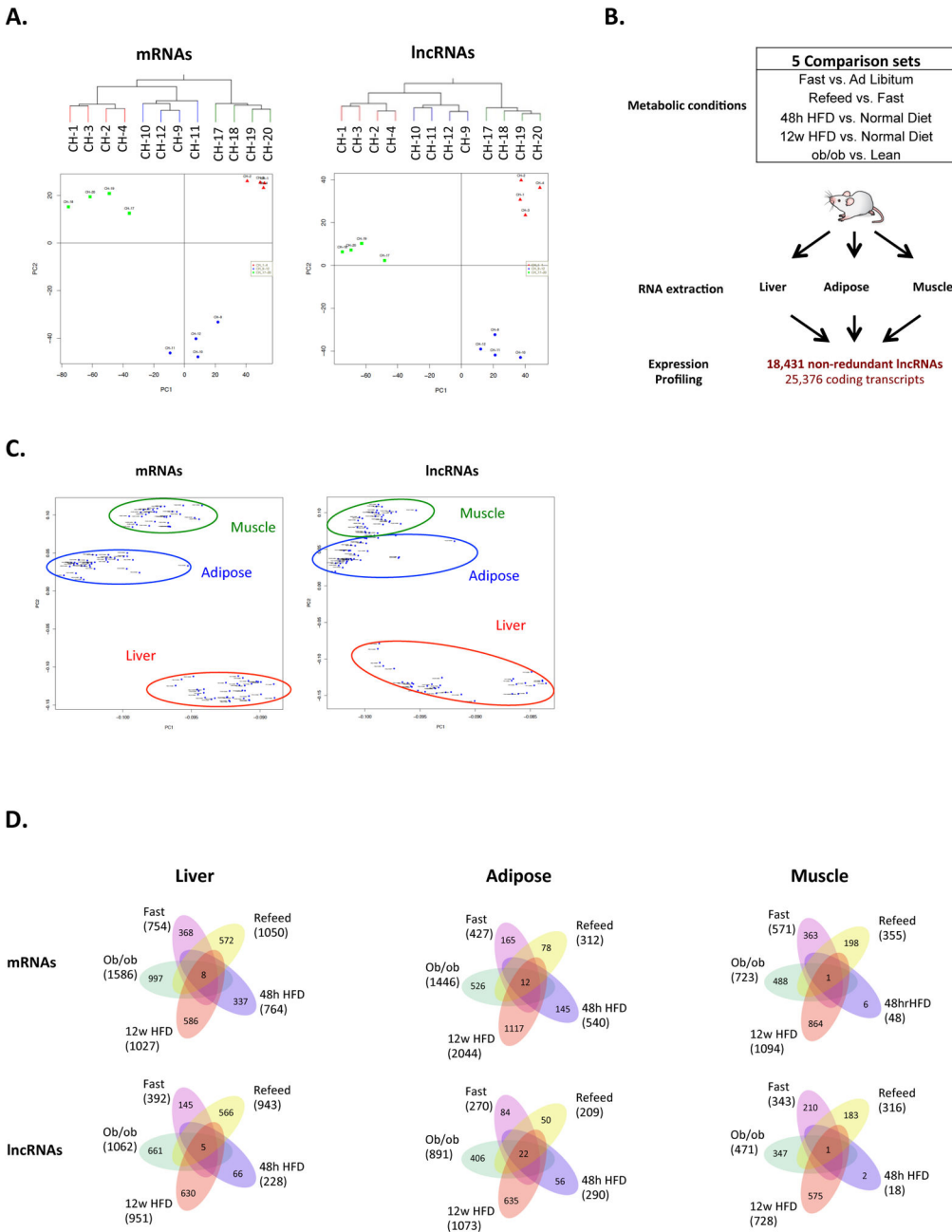


Figure 1. Dynamic regulation of lncRNAs and mRNAs in key metabolic organs under multiple metabolic conditions

(A) Hierarchical clustering (top) and principal components analyses (bottom) of differentially expressed mRNAs (left) and lncRNAs (right) in mouse livers during fasting and refeeding. The differentially expressed genes were defined by one-way ANOVA analysis to be significantly different among the three groups. (CH_1–4 are Ad libitum group shown as red, CH_9–12 are Fast group shown as blue, and CH_17–20 are Refeed group shown as green).

(B) Experimental outline. Each metabolic condition contains four to five mice, and the expression profiles of mRNAs or lncRNAs from 37 liver samples, 37 adipose samples, and 32 muscle samples were analyzed.

(C) Principal components analyses of expressed genes from 106 samples: mRNAs (left), and lncRNAs (right). Samples circled in red are liver, in blue are adipose, and in green are muscle.

(D) Numerical distribution of condition-specifically and commonly regulated mRNAs (top row) and lncRNAs (bottom row) in liver (left), adipose (middle), and muscle tissue (right).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

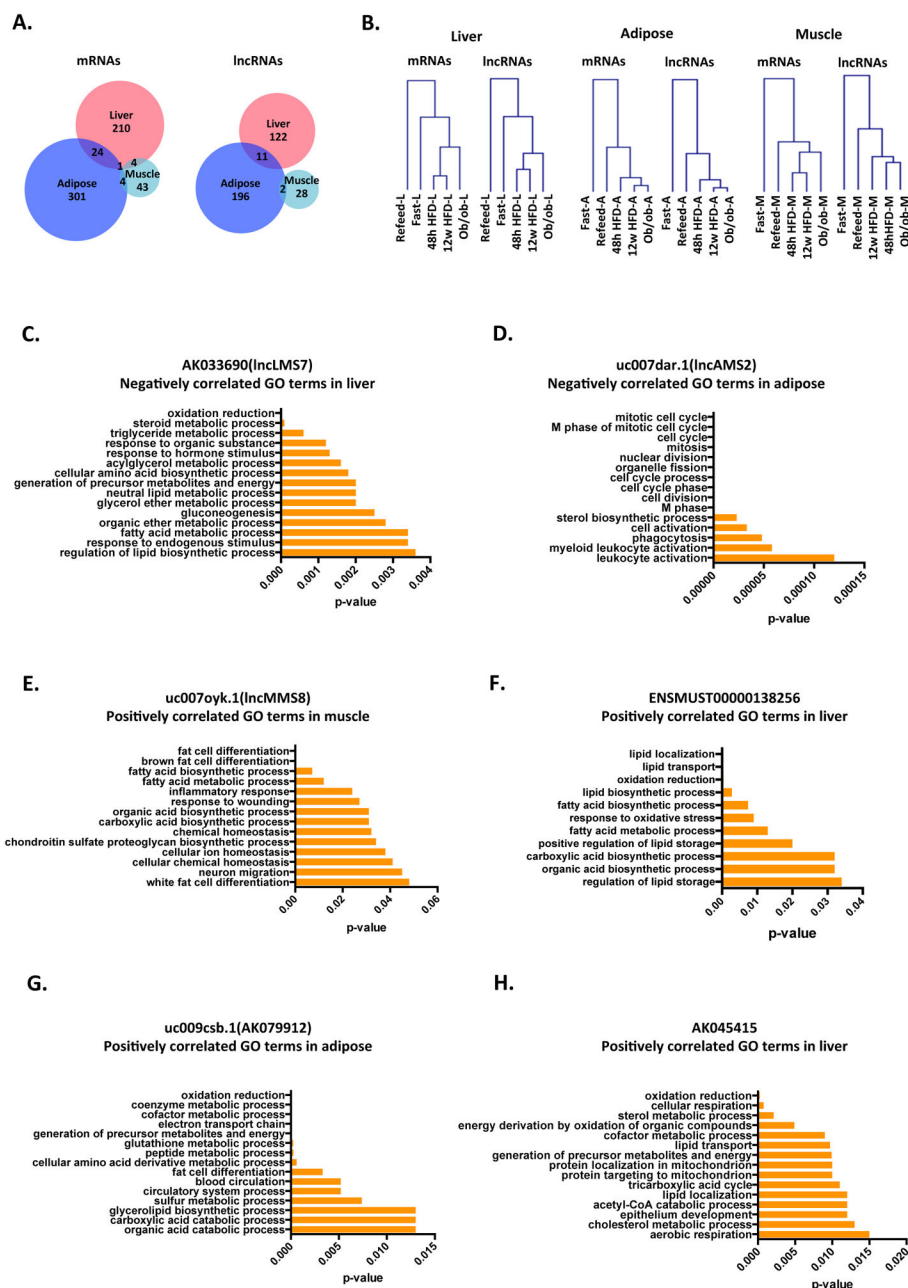


Figure 2. Functional prediction of metabolically sensitive lncRNAs by lncRNA-mRNA co-expression correlation under diverse metabolic conditions

(A) The numbers of condition-specific and commonly regulated mRNAs (left) and lncRNAs (right) by at least three metabolic conditions.

(B) Hierarchical clustering of mRNAs and lncRNAs regulated at least three times (Figure 2A) across all metabolic conditions in liver (left), adipose (middle), and muscle (right) tissue by Pearson's correlation.

(C–E) Top gene ontology (GO) biological process terms for mRNAs positively or negatively correlated with representative metabolically sensitive lncRNAs in liver (C), adipose (D), and muscle (E).

(F–H) Top GO terms of correlated mRNAs for recently reported lncRNAs that were shown to be involved in brown and white adipocyte differentiation.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

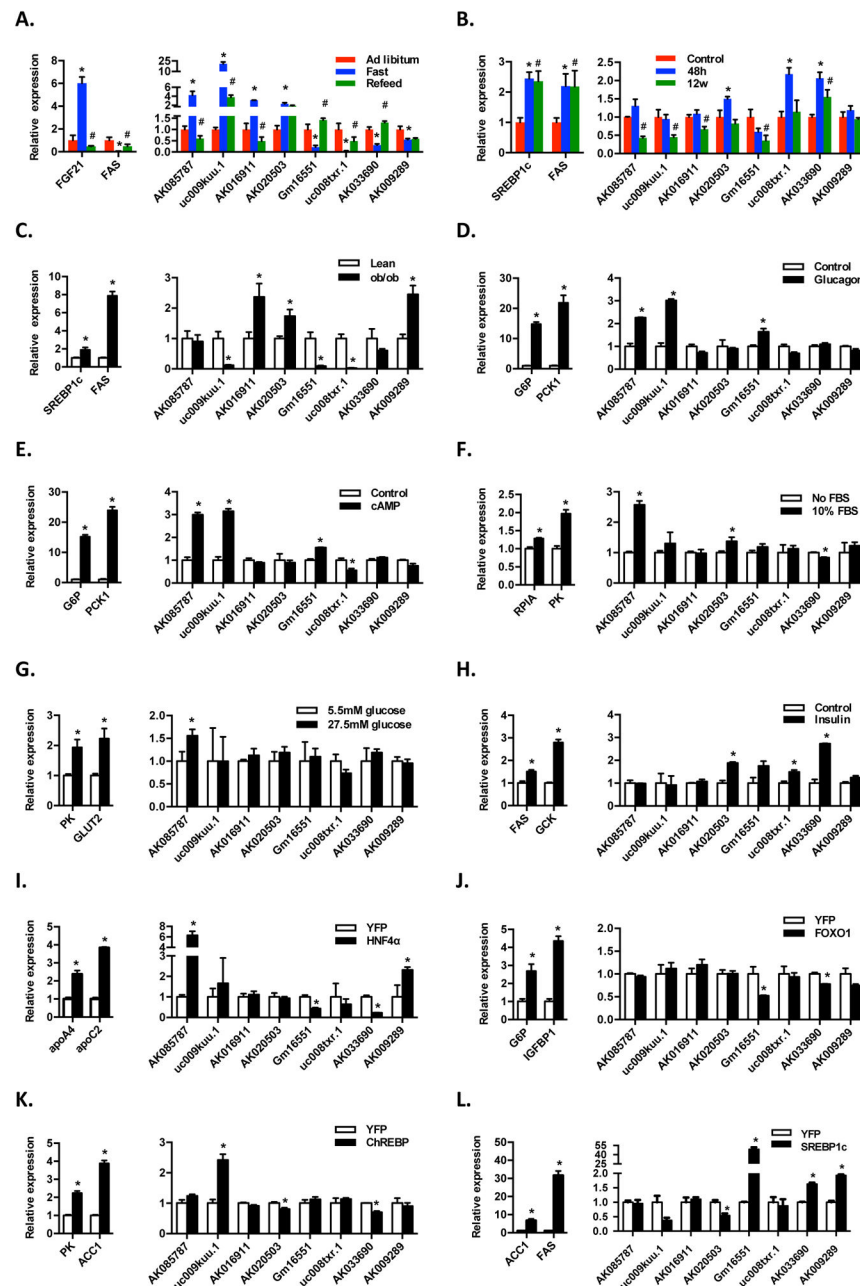


Figure 3. Liver metabolically sensitive lncRNAs are dynamically regulated by nutrients, hormones and metabolic transcription factors in mouse primary hepatocytes
(A–C) Quantitative real-time RCR analyses of liver metabolically sensitive lncRNAs (lncLMS) in mice fed ad libitum (Ad Libitum), subject to a 24-hour fast (Fast), or a 24-hour fast followed by a 4-hour refeeding (Refeed) (A), fed with normal diet (Control), high fat diet for 48 hours (48h) or 12 weeks (12w) (B), or in ob/ob and their control wild type (Lean) mice (C). Expression levels of protein-coding genes known to be regulated by these conditions are shown on the left. Error bars are SEM, n=4–5. *P<0.05 (Fast vs Ad libitum, 48h HFD vs Control, or ob/ob vs Lean), # p<0.05 (Refeed vs Fast or 12week HFD vs Control).

(D–H) Expression levels of lncLMSs in isolated mouse primary hepatocytes treated with 200nM glucagon (D), 100uM cAMP (E) for 6hrs or cultured in medium supplemented with or without 10% fetal bovine serum (FBS) (F), containing 5.5mM or 27.5mM glucose (G), or treated with 100nM insulin (H) for 16hrs. Expression levels of protein-coding genes known to be regulated by these conditions are shown on the left. Error bars are SEM, n=3. *P<0.05. (I–L) Expression levels of lncLMSs in isolated mouse hepatocytes 24hrs after being transduced with adenoviruses expressing mouse HNF4 α (I), constitutively active mouse FOXO1 (J), constitutively active mouse ChREBP (K), or constitutively active mouse SREBP1c (L). YFP adenovirus transduced hepatocytes are used as a control, and expression levels of protein-coding genes known to be regulated by these transcription factors are shown on the left. Error bars are SEM, n=3. *P<0.05.

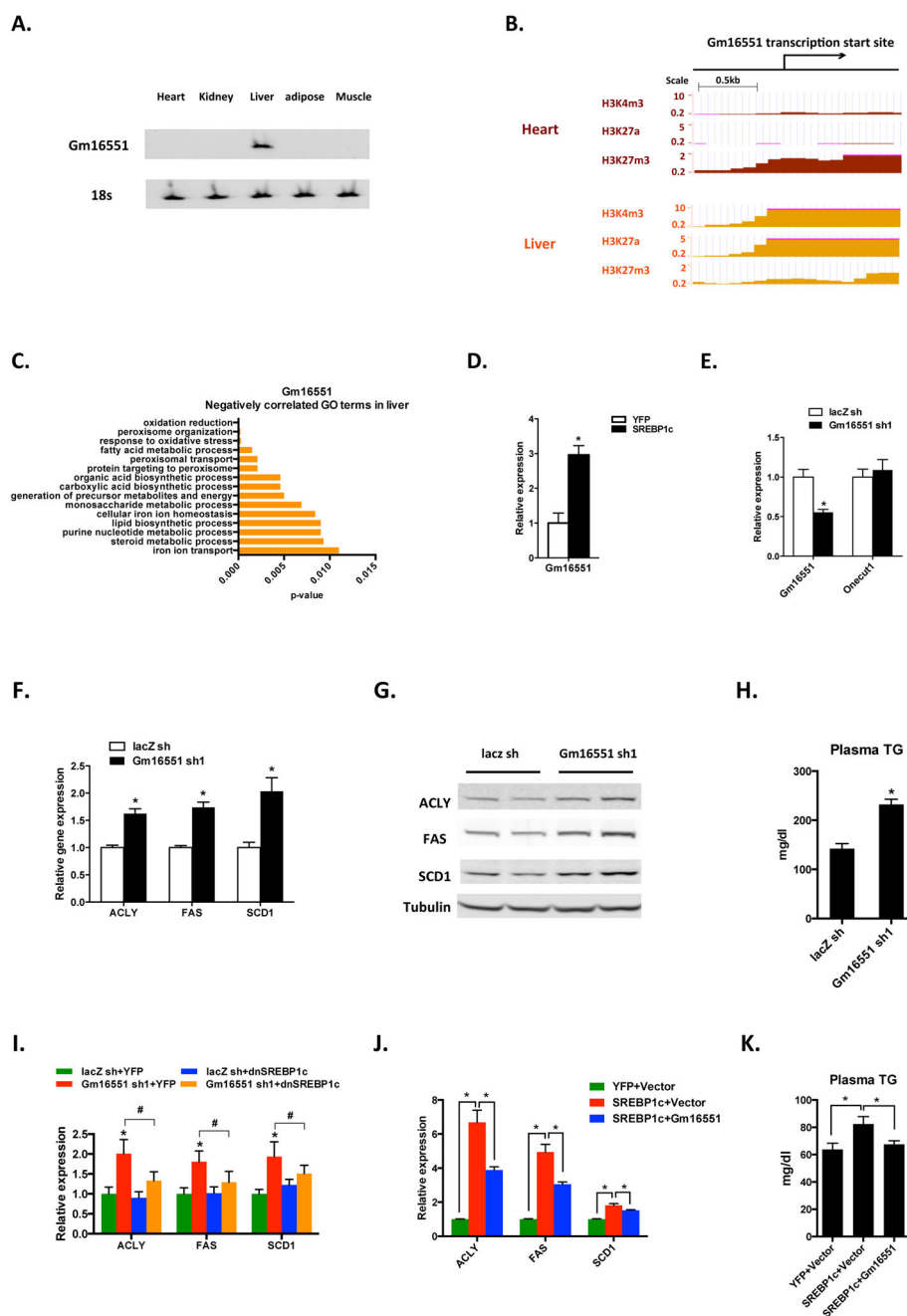


Figure 4. Identification of an lncRNA regulating lipogenesis in mouse liver by a functional lncRNA detection roadmap

(A) Semi-quantitative PCR analyses of Gm16551 expression levels in heart, kidney, liver, adipose, and muscle tissue from 8 week old male mice. 18s was used as a control.

(B) Histone modifications at the Gm16551 locus in mouse heart and liver tissue retrieved from ENCODE/LICR ChIP-seq data.

(C) Top gene ontology (GO) biological process terms for mRNAs negatively correlated with Gm16551.

- (D) Expression levels of Gm16551 in male mouse livers transduced with YFP or constitutively active mouse SREBP1c adenoviruses. Error bars are SEM, n=5. *P<0.05.
- (E) Expression levels of Gm16551 and its neighboring coding gene *Onecut1* in the livers of male mice that received shRNA adenoviruses for lacZ or Gm16551 respectively (sacrificed at 2pm, which is 4hr refeeding after 24hr fasting). Error bars are SEM, n=6. *P<0.05.
- (F) Expression levels of ATP citrate lyase (ACLY), fatty acid synthase (FAS) and stearoyl-Coenzyme A desaturase 1 (SCD1) in the livers of control (lacZ sh) and Gm16551 knockdown (Gm16551 sh) mice. Error bars are SEM, n=6. *P<0.05.
- (G) Immunoblot analyses of ACLY, FAS, and SCD1 proteins in the livers of control (lacZ sh) and Gm16551 knockdown (Gm16551 sh) mice.
- (H) Plasma triglyceride (TG) levels in control and Gm16551 knockdown mice. Error bars are SEM, n=6. *P<0.05.
- (I) Expression levels of ACLY, FAS, and SCD1 in the livers of mice receiving lacZ sh plus YFP, lacZ plus dominant negative SREBP1c, Gm16551 knockdown (Gm16551 sh1) plus YFP, and dominant negative SREBP1c plus Gm16551 knockdown adenoviruses (sacrificed at 2pm, which is 4hr refeeding after 24hr fasting). Error bars are SEM, n=5. *P<0.05, #P<0.05.
- (J) Expression levels of ACLY, FAS, and SCD1 in the livers of mice receiving YFP plus vector, active SREBP1c plus vector, and active SREBP1c plus Gm16551 adenoviruses. Error bars are SEM, n=7. *P<0.05.
- (K). Plasma triglyceride level in mice receiving YFP plus vector, active SREBP1c plus vector, and active SREBP1c plus Gm16551 adenoviruses. Error bars are SEM, n=7. *P<0.05.

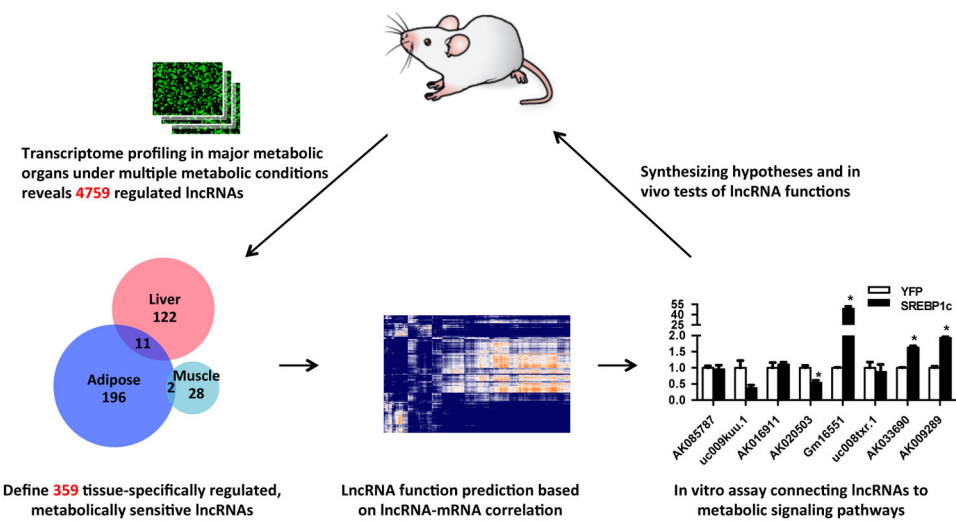


Figure 5.
An integrative roadmap defines functional tissue-specific lncRNA metabolic regulators.