Taylor & Francis
Taylor & Francis Group

DATE NOTE

8 OPEN ACCESS | Check for updates

Tissue-specific gene expression in obese hyperglycemic mice

Mahmoud Ahmed[a], Omar Elashkar[a], Jong Youl Lee[b], Eun Ae Jeong[b], Kyung Eun Kim[b], Gu Seob Roh[b] and Deok Ryong Kim [a]

[a]Department of Biochemistry and Convergence Medical Science, Institute of Health Sciences, College of Medicine, Gyeongsang National University, Jinju, Korea; [b]Department of Anatomy and Convergence Medical Science, Bio Anti-aging Medical Research Center, Institute of Health Sciences, College of Medicine, Gyeongsang National University, Jinju, Korea

ABSTRACT

Ob/ob mice are *leptin*-deficient animals with uninhibited food intake and susceptibility to gain weight and develop type 2 diabetes. The mice have been used to study obesity and diabetes. We generated a dataset of different tissue gene expressions from wild-type and ob/ob mice with a normal diet (ND) or high-fat diet (HFD). The gene expression was profiled at a genome-scale using RNA-seq. We deposited the raw data to the short read archive and the processed data to the gene expression omnibus. In this manuscript, we describe generating the dataset and technical validation of the gene expression profiles. We assessed the quality of the reads, alignment, and the quantification of gene expression. We found that the tissue of origin explained the most variance between samples. Noncoding features differed in their contribution to the mice profiles. Gene expression profiles diverged between the experimental groups. To sum, this dataset can be used to study tissue-specific gene expression in weight gain susceptible mice and the response to HFD.

Introduction

The ob/ob mouse (*ob/ob* hereafter) is a mutant (original) strain in which the gene coding for leptin (*Lep*) is deficient (Ingalls et al. 1950). As a result, the mouse has uninhibited food intake and becomes obese as it ages (Lindström 2007). Dysregulated metabolism results in hyperglycemia and hyperlipidemia in this mouse. Moreover, in response to insulin resistance, the mouse adapts by increasing the number of pancreatic beta cells and insulin production. *Ob/ob* mouse is a frequently used model of type 2 diabetes and obesity.

Gene expression datasets of ob/ob mice were generated mostly using microarrays or rarely RNA-seq (Yu-Han et al. 2019; Wilhelmi et al. 2021). These studies focus on a single tissue (pancreas, Lupo et al. 2020; Wilhelmi et al. 2021; glomerulus, Sugahara et al. 2020; or myocardium, Eduardo Rame et al. 2011). One published dataset profiled the gene expression of adipose tissue in ob/ob mice using microarrays at 5 and 16 weeks (Prieur et al. 2011). Another study measured the abundance of non coding RNAs in the mice liver by sequencing (Yu-Han et al. 2019). Simultaneous profiling of gene expression from different tissue in the ob/ob is missing. Similarly, no dataset systematically evaluated the effect of HFD on the mice compared to the WT.

We generated a gene expression dataset in different wild-type (WT) and *ob/ob* mice tissues with a normal diet (ND) or high-fat diet (HFD). Changes at the gene expression level of lipid metabolism could be assessed using the epididymal fat, liver, and muscle tissues while hormonal changes in the brain (hypothalamus and hippocampus). We also expected to assess the changes in expression in tissues with comorbidities in the model (heart and aorta). Researchers could use this dataset to study the difference between normal mice and those susceptible to weight gain. In addition, the dataset can be used to study the response to HFD. In each case, the response can be dissected at the tissue level. Here, we present a description of the mice, data generation, raw and processed data. Besides, we provide a technical validation of the dataset.

CONTACT Deok Ryong Kim ✉ drkim@gnu.ac.kr

Materials and methods

Animals and diet

Three-week-old male C57BL/6 mice and *ob/ob* mice were purchased from Central Laboratory Animal Inc. (Seoul, South Korea). Animal experiments were performed in accordance with the National Institutes of Health guidelines on the use of laboratory animals. The Animal Care Committee for Animal Research at Gyeongsang National University approved the study protocol (GNU-160530-M0025). The mice were individually housed under a 12-h light/dark cycle. C57BL/6 wild-type (WT) or *ob/ob* mice were fed ad libitum with normal standard diet chow (ND, $n = 3$) or a high-fat diet (HFD, $n = 3$, 60% kcal from fat; Research Diet, Inc., New Brunswick, NJ, USA) for 20 weeks. The body fat contents of animals were quantified by EchoMRI (Whole Body Composition Analyzer, Houston, TX, USA). The overall experimental scheme is depicted in Figure 1(a).

Measurement of metabolic parameters

After fasting overnight, Mice were euthanized using intraperitoneal Zoletil (50 mg/ml; Virbac Korea) and Rompun (2.3 mg/ml; Bayer Korea). A mouse of 60 grams received 30 ul Zoletil and 30 ul Rompun and showed no movement at 2 min and no response to stimuli at 8 min. Blood samples were collected transcardially through the left ventricle with a 1-mL syringe. Serum glucose, total cholesterol, alanine aminotransferase (ALT), and aspartate aminotransferase (AST) levels were determined using the Green Cross Reference Laboratory kits (Yongin-si, South Korea).

Library preparation and sequencing

After blood collection, seven different tissues of from each mouse: hippocampus, hypothalamus, heart, aorta, liver, skeletal muscle (gastrocnemius muscle), and epididymal fat pads, were quickly removed and stored at −80°C. Total RNA from each tissue was isolated using the TRIzol reagent (Invitrogen, CA, USA). Complementary DNA was then synthesized using a reverse-transcription kit (Thermo Scientific, Waltham, MA, USA), according to the manufacturer's instructions. The RNA-seq analyzes were performed by C&K Genomics (C&K Genomics Inc., Seoul, South Korea). The sequencing libraries were constructed using Illumina's TruSeq RNA Prep kit (Illumina Inc., San Diego, CA, USA), and DNA sequencing data were generated using the NextSeq 500 platform (Illumina Inc.).

Quality control assessment, alignment, and counting

Quality assessment and control were applied to the raw reads using Trimmomatic (Bolger et al. 2014). Adaptors, over-represented sequences, and low-quality reads were removed. Raw reads were aligned to the UCSC mouse genome mm10 using HISAT2 (Kim et al. 2015). Aligned reads were counted in the same genome known features using FeatureCount (Liao et al. 2014). Feature counts to estimate the gene expression. Variance stabilization transformation (VST) was applied using vst (Love et al. 2014). Known batch effects were removed using removeBatchEffect for purposes of visualization, and unknown variables were estimated SVA and added to the metadata table (Ritchie et al. 2015; Leek et al. 2019). Transformed counts, phenotype data, and features information were packaged in a SummarizedExperiment object. Figure 1(b) shows the steps of the pre-processing and processing pipeline.

Methods for technical validation

The technical validation of the data was performed in two steps-first, the quality assessment and control of the raw and intermediate data (discussed in Methods). Second, the validation of the experimental design and expected biology in the gene expression and the external measurements (presented in detail in this manuscript). The transformed counts were used to partition the variance in principal component analysis and correlation analysis. Variance partitioning applies a random-effects model to the data to estimate the amount of variance in each gene explained by the model variables (Hoffman and Schadt 2016). The top 1000 genes whose variance is explained by a certain variable were used to calculate over-representation in gene ontology terms (Ashburner et al. 2000, may; Yu et al. 2012). Principal component analysis represents the variance between the samples in two dimensions. Pearson's correlation coefficient was calculated between pairs of variables to decide how well the gene expression in one sample/group resembles others.
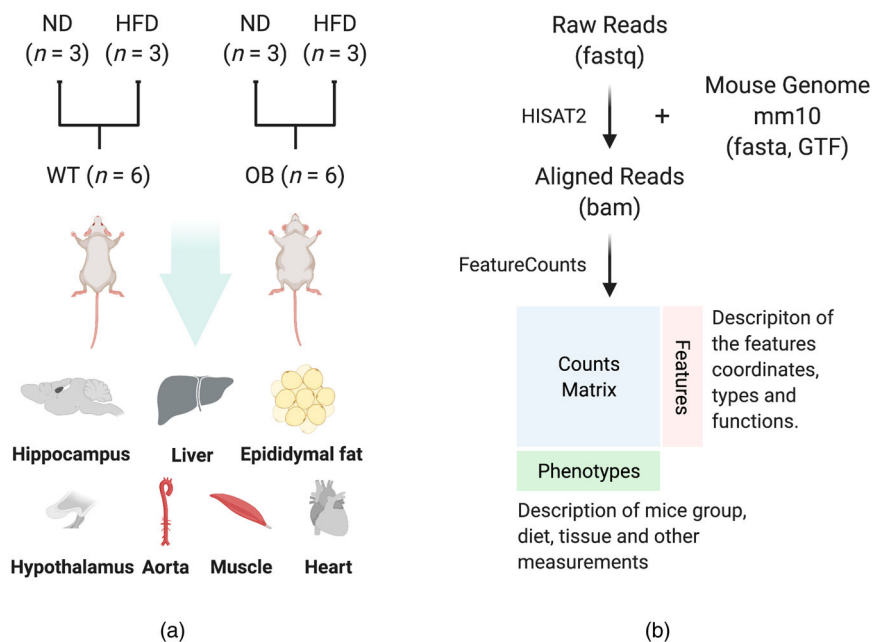
Figure 1. Experimental design and data processing pipeline. (a) Wild-type (WT) mice ($n = 6$) were divided into two equal groups and fed either a normal diet (ND) or a high-fat diet (HFD). An equal number of ob/ob mice ($n = 6$) fed either ND or HFD. After 20 weeks, mice were sacrificed for tissue collection. Six types of tissue were collected from each mouse, including the hippocampus, hypothalamus, heart, liver, aorta, muscle, and epididymis. (b) Raw reads were aligned to the mouse genome (mm10) using HISAT2. Reads were counted in known features using FeatureCounts. Counts were packaged along with phenotype and feature information in a Bioconductor experimental data package (ObMiTi).

Table 1. Samples and experimental groups.

Group	Diet	Tissue (N)	Samples (N)
ob/ob	HFD	7	21
	ND	7	21
WT	HFD	7	21
	ND	7	21

Data Records

This dataset includes 84 samples from twelve mice in four groups of three and seven tissue each. Raw data were deposited in the short read archive (SRA) under the project number (PRJNA701378) and the processed data in the gene expression omnibus (GEO) under the series number (GSE167264). Table 1 summarizes the data records.

Dataset validation

Quality assessment of raw, aligned, and assigned reads

To assess the quality of the sequencing data, we examined the reads at different steps of the pipeline. Raw reads were of sufficient quality, with most reads (95%) above 30/40 quality score (Figure 2(a)). The overall alignment of the reads was always above 95%, with only a few samples below that (Figure 2(b)). Concordance between pairs of reads was similarly high ($>65\%$) for most samples. The majority of reads ($>70\%$) were assigned to known features (Figure 2(c)). Causes for the assignment included ambiguity, multi-mapping, or unmapped reads, which were few.

Tissue explains the variance between samples

After removing the batch effects, the genetic background of the individual mice did not contribute greatly to the variance in gene expression as indicated by principal component analysis (Figure 3(a)). WT mice fed ND were clustered together and were significantly distinct from other mice. In contrast, the mice on HFD showed the greatest variance. Ob/ob mice on either diet were similar to each other. A large percentage of the variance between the samples in a given tissue (Epididymal fat pads) was explained by two experimental variables mice group and diet.

The greatest amount of variance in gene expression across tissues was explained by the tissue variable
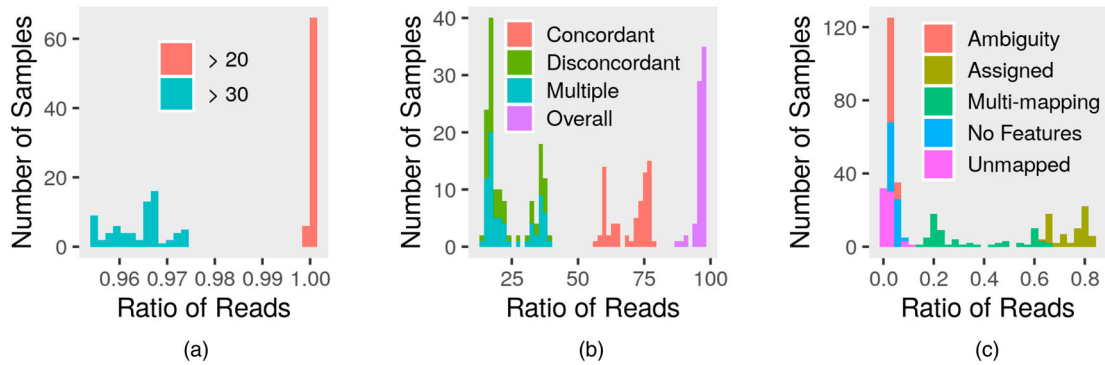
Figure 2. Quality assessment of the raw, aligned and counted reads. (a) A histogram of the ratio of reads with quality scores above 20 (red) or 30 (blue). (b) A histogram of the ratio of reads with a given overall alignment (purple), multiple alignments (blue), disconcordant (green), or concordant aligned pairs (red). (c) A histogram of the ratio of reads that were assigned to features (brown) or unassigned either due to ambiguity (red), no features (blue), multiple mapping (green), or unmapped (move).
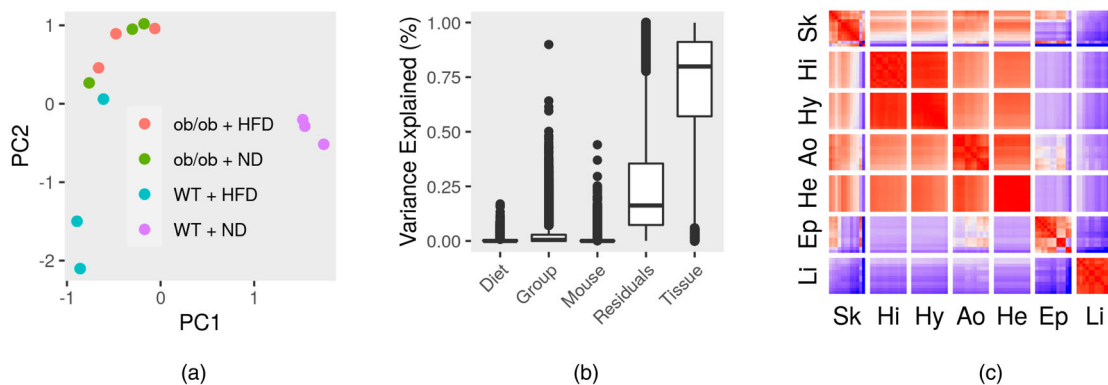

Figure 3. Explaining and partitioning variance between the experimental groups. (a) Principal component (PC) analysis of the read counts in coding genes in mice ($n = 12$) from Epididymal fat pad samples Experimental groups of different genotypes and diets are represented as colors. (b) The estimated percentage of variance in read counts of coding genes by variables in the random-effects fixed model. (c) Pearsons correlation coefficients between scaled gene expression in all samples ($n = 84$) are shown as color values (0, blue and 1, red).Skeletal muscle; Sk, Hippocampus; Hi, Hypothalamus; Hy, Aorta; Ao, Heart; He, Epididymal fat pad; Ep, Liver; Li.

(Figure 3(b)). Indeed, gene expression from the same tissues was strongly correlated across groups of mice and diets (Figure 3(c)). This was less obvious in adipose and skeletal muscle, where the greatest difference between the experimental groups is expected. These observations illustrate a critical aspect of this dataset, where any downstream analysis should take the tissue variable into account.

Genes whose variance is mostly explained by different variables were over-represented in different categories of gene ontology terms (Table 2). Tissue explained the variance of genes in terms related to cellular components and processes of particular cell types, for example, neurons (synapse, axon, dendrite), liver, and fat cells (fatty acid metabolic process and peroxisome). By contrast, the genes explained by group variables were over-represented in terms related to nucleic acid maintenance and transcription (DNA

Table 2. Over-representation of genes explained by variables in gene ontology terms.

Variable	Term	Ratio
Diet	immune system process	0.06
	innate immune response	0.05
	immune response	0.04
	inflammatory response	0.04
	chromosome	0.04
Group	cellular response to DNA damage stimulus	0.07
	mRNA processing	0.06
	RNA splicing	0.05
	nuclear speck	0.05
	DNA repair	0.05
Tissue	axon	0.09
	postsynaptic membrane	0.07
	postsynaptic density	0.07
	nervous system development	0.07
	chemical synaptic transmission	0.06

repair, RNA splicing, and helicase activity). Finally, immune-related terms had genes whose variance was explained by the diet variable over-represented.
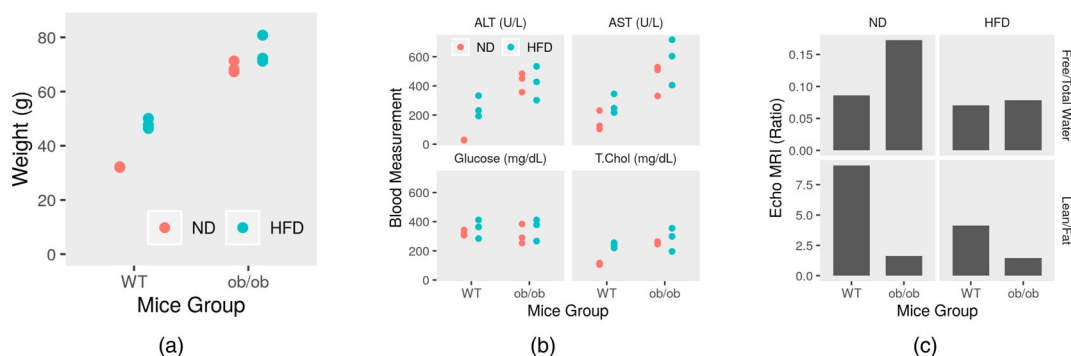
Figure 4. Biological measurement, blood work, and echo MRI measurements. (a) The final weight (g) of the mice ($n = 12$) at the end of the study period. (b) Mice measurements of liver functions AST (U/L) and ALT (U/L), blood glucose (mg/dL) and total cholesterol (mg/DL). (c) Mice body composition is shown as the ratio of lean/fat and free/total water as measured by echo MRI.

Gene expression reflects a divergent genetic response to diets

We measured several metabolic parameters of the mice in this study including body weight, body composition and blood levels of glucose, cholesterol, ALT and AST. As expected, the *ob/ob* mice and the HFD groups were heavier (Figure 4(a)). The liver function, blood glucose, and total cholesterol were worse in the HFD groups in both types of mice, although the difference compared to ND were more apparent in the case of the WT mice (Figure 4(b)). A similar pattern emerged in the composition of the body as measured by echo-MRI; HFD groups faired worse in terms of the percentage of body fat and free water compared to mice on ND (Figure 4(c)). This observation supports the notion that *ob/ob* and HFD groups accumulate more fat and weigh more.

The arcuate nucleus in the hypothalamus is sensetive to leptin and ghrelin (Nogueiras et al. 2008). The two hormones signal through specific receptors to regulate satiety and food intake through two other substances, NYP and AGRP (Figure 5(a)). *ob/ob* mice are deficient in Leptin, but it is unclear whether leptin is involved in the pathology of obesity in normal mice (Mayer et al. 1953). Leptin receptor (Lpr) seems to be completely shut down in *ob/ob* mice (Figure 5(b)). Only WT mice on ND seem to have relatively higher expression of the receptor compared to HFD. The analgous ghrelin receptor and effector were also regulated between the experimental groups (Figure 5(c)). Due to the lack of the inhibitory signal or the active stimulatory signals, NYP and AGRP was more expressed in ND, especially in WT mice (Figure 5(d)). Finally, we should consider the fact that *ob/ob* mice exhibit

changes in key signaling pathways at different ages, often in the form of adaptation or tolerance to deficient leptin or excess insulin (Danieisson et al. 1968).

Insulin signaling and glucose utilization are also dysregulated in *ob/ob* mice (Wang et al. 2019). Normally, insulin signals through its receptor and substrate to induce the formation of transport vesicles (Figure 5(c)). These vesicles transport glucose from the membrane transporters (Vargas and Carrillo Sepulveda 2019). As expected, we observed dysregulation in *ob/ob* mice similar to that, if not more extreme, of the response to HFD in WT mice. Insulin receptor substrate (Irs1), glucose transporter 4 (Slc2a4), Prkaa1, and Pik3r1 were down-regulated ob/ob adipose tissue and WT HFD fed mice compared to WT mice on ND (Figure 5(d)). A full list of differentially expressed genes in each tissue between genotypes and diets is available as an additional file (See data availability statment). In summary, the dataset shows a gene expression pattern reflective of insulin resistance and low glucose utilization in the adipose tissue.

Usage notes

The processed data are distributed as Bioconductor experimental data (ObMiTi). The package contains a SummarizedExperiment object. This object contains three tables. (1) Read counts of the features (row) in samples (columns). (2) Phenotype data describing the variables and the measurements of the samples. (3) Information about the features, including identifiers, biotypes, and genomic coordinates. The object was optimized for speedy and flexible downstream analysis using R/Bioconductor packages.
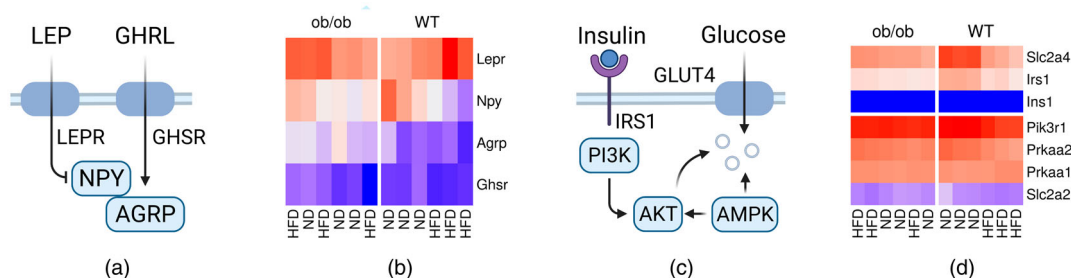
Figure 5. Leptin and Insulin signaling in hypothalamus and adipose tissue. (a) A diagram of leptin signaling in the hypothalamus. (b) Gene expression (low, blue and high, red) of coding genes in the signaling pathway of in the hypothalamus of WT and *ob/ob* mice with ND or HFD. (c) A diagram of insulin signaling in the adipose tissue. (d) Gene expression (low, blue and high, red) of coding genes in the signaling pathway of in the Epididymal fat pads of WT and *ob/ob* mice with ND or HFD. Leptin Receptor (LEPR), Neuropeptide Y (NPY), Ghrelin Secretagogue Receptor (GHSR), Agouti Related Neuropeptide (AGRP), Insulin (Ins1), Insulin Receptor Substrate 1 (IRS1), Phosphoinositide-3-Kinase Regulatory Subunit 1 (Pik3r1), Solute Carrier Family 2 Member 2 & 4 (Slc2a2 & 4), Solute Carrier Family 2 Member 2 (Slc2a2), Protein Kinase AMP-Activated Catalytic Subunit Alpha 1 & 2 (Prkaa1 & 2).

Limitations

Some researchers take issue with using the ob/ob mouse as a model for obesity. This is because the etiology of the disease in this mouse and human is different. Nonetheless, we believe these mice can help study the biology of different tissue changes once the condition is developed. It is also crucial for studying the role of hormone imbalance (leptin deficiency in particular) in predisposing the animal to weight grain or abnormal storage of fat.

The small sample size in this experiment makes it hard to estimate the occurring natural variance between the mice. However, the dataset can help find large effect sizes or as a starting point of investigation to guide the design of future experiments. In addition, this dataset lacks protein level measurement. Considering the nature of hormonal changes associated with the model, many of the relevant biology can only be studied at the level of protein. A large proportion of the variance between the samples was explained by the tissue variable. This is to be expected in gene expression data from different organs. Limiting the analysis to a single tissue can avoid the issue. However, across tissue comparisons require specialized methods to either adjust for the tissue of origin's effect or remove it entirely.

Acknowledgements

Data availability statement

The data described in this note can be accessed freely and openly. Raw data can be accessed on sequence read archive (SRA) under (PRJNA701378; https://www.ncbi.nlm.nih.gov/bioproject/?term = PRJNA701378) and the processed data on gene expression omnibus (GEO) under (GSE167264; https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc = GSE167264). The software environment used during the data processing is available as a Docker image (https://hub.docker.com/repository/docker/bcmslab/obesemice). The source code for the pre-processing and processing pipelines is available within the Bioconductor package (ObMiTi) (https://github.com/BCMSLab/ObMiTi). A full list of differentially expressed genes in response to HFD in ob/ob mice compared to WT is available on Figshare (https://doi.org/10.6084/m9.figshare.18666452.v1). The source code for reproducing the technical validation presented in this manuscript is available here (https://github.com/BCMSLab/obese_mice). All code is open-source (GPL-3).

Disclosure statement

Funding

ORCID

Deok Ryong Kim 🆔 http://orcid.org/0000-0002-3288-8257

References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. 2000, May. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 25(1):25–29. http://www.nature.com/doifinder/10.1038/75556.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 30(15):2114–2120.

Danieisson A, Hellman B, Taljedal IB. 1968. Glucose tolerance in the period preceding the appearance of the manifest obese-hyperglycemic syndrome in mice. Acta Physiol Scand. 72(1–2):81–84.

Eduardo Rame J, Barouch LA, Sack MN, Lynn EG, Abu-Asab M, Tsokos M, Kern SJ, Barb JJ, Munson PJ, Halushka MK, et al. 2011. Caloric restriction in leptin deficiency does not correct myocardial steatosis: failure to normalize PPARα/PGC1α and thermogenic glycerolipid/fatty acid cycling. Physiol Genomics. 43(12):726–738.

Hoffman GE, Schadt EE. 2016. variancePartition: interpreting drivers of variation in complex gene expression studies. BMC Bioinformatics. 17(1):483.

Ingalls AM, Dickie MM, Snell GD. 1950. Obese, a new mutation in the house mouse∗. J Hered. 41(12):317–318.

Kim D, Langmead B, Salzberg SL. 2015, Apr. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 12(4):357–360. http://www.nature.com/doifinder/10.1038/nmeth.3317.

Leek JT, Johnson WE, Parker HS, Fertig EJ, Jaffe AE, Storey JD, Zhang Y, Torres LC. 2019. sva: surrogate variable analysis. http://bioconductor.org/packages/sva/.

Liao Y, Smyth GK, Shi W. 2014, Apr. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics (Oxford, England). 30(7):923–930.

Lindström P. 2007, May. The physiology of obese-hyperglycemic mice [ob/ob mice]. TheScientificWorldJournal. 7:Article ID 804524.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550. http://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0550-8.

Lupo F, Piro G, Torroni L, Delfino P, Trovato R, Rusev B, Fiore A, Filippini D, De Sanctis F, Manfredi M, et al. 2020. Organoid-transplant model systems to study the effects of obesity on the pancreatic carcinogenesis in vivo. Front Cell Dev Biol. 8:308.

Mayer J, Russell RE, Bates MW, Dickie MM. 1953. Metabolic, nutritional and endocrine studies of the hereditary obesity-diabetes syndrome of mice and mechanism of its development. Metabolism. 2(1):9–21.

Nogueiras R, Tschöp MH, Zigman JM. 2008. CNS regulation of energy metabolism: ghrelin versus leptin NIH public access. Ann N Y Acad Sci. 1126:14–19.

Prieur X, Mok CY, Velagapudi VR, Núñez V, Fuentes L, Montaner D, Ishikawa K, Camacho A, Barbarroja N, O'Rahilly S, et al. 2011. Differential lipid partitioning between adipocytes and tissue macrophages modulates macrophage lipotoxicity and M2/M1 polarization in obese mice. Diabetes. 60(3):797–809.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43(7):e47.

Sugahara M, Tanaka S, Tanaka T, Saito H, Ishimoto Y, Wakashima T, Ueda M, Fukui K, Shimizu A, Inagi R, et al. 2020. Prolyl hydroxylase domain inhibitor protects against metabolic disorders and associated kidney disease in obese type 2 diabetic mice. J Am Soc Nephrol. 31(3):560–577.

Vargas E, Carrillo Sepulveda MA (2019). Physiology, glucose transporter type 4 (GLUT4). StatPearls.

Wang Y, Zhou H, Palyha O, Mu J. 2019. Restoration of insulin receptor improves diabetic phenotype in T2DM mice. JCI Insight. 4(15):e124945.

Wilhelmi I, Neumann A, Jähnert M, Ouni M, Schürmann A. 2021. Enriched alternative splicing in islets of diabetes-susceptible mice. Int J Mol Sci. 22(16):8597.

Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS: J Integr Biol. 16(5):284–287. http://online.liebertpub.com/doi/abs/10.1089/omi.2011.0118.

Yu-Han H, Kanke M, Kurtz CL, Cubitt R, Bunaciu RP, Miao J, Zhou L, Graham JL, Hussain MM, Havel P, et al. 2019. Acute suppression of insulin resistance-associated hepatic miR-29 in vivo improves glycemic control in adult mice. Physiol Genomics. 51(8):379–389.