Investigating Obesity at the Population, Pathway, and Gene Levels

Aashna Soni[1] and Biao Wang, PhD[1]

[1]Cardiovascular Research Institute, University of California, San Francisco

Motivation

- There is an obesity epidemic in the United States
  - More than 2 in 5 adults have obesity.[1]
  - Obesity rates rose from 30.5% in 1999-2002 to 41.9% in 2017-March 2020, while severe obesity rates almost doubled, from 4.7% to 9.2%.[1]
  - Obesity (for which BMI is the standard measure) is a significant risk factor for diseases such as type 2 diabetes, hypertension, and atherosclerosis.[2,3]
- Obesity causative factors
  - Obesity onset is influenced by genetic and environmental factors.[4]
  - Known causative genes for monogenic obesity: *LEP, LEPR, POMC, PCSK1, MC4R, SIM1, BDNF, NTRK2*.[4]
  - Pathways involved in pro-obesity mechanisms: tissue inflammation, insulin resistance, and adipogenesis.[4,5]
- The drug hydroxychloroquine, used to treat the autoimmune disease lupus (which causes inflammation), has been shown to improve insulin sensitivity in diabetic patients.[6]

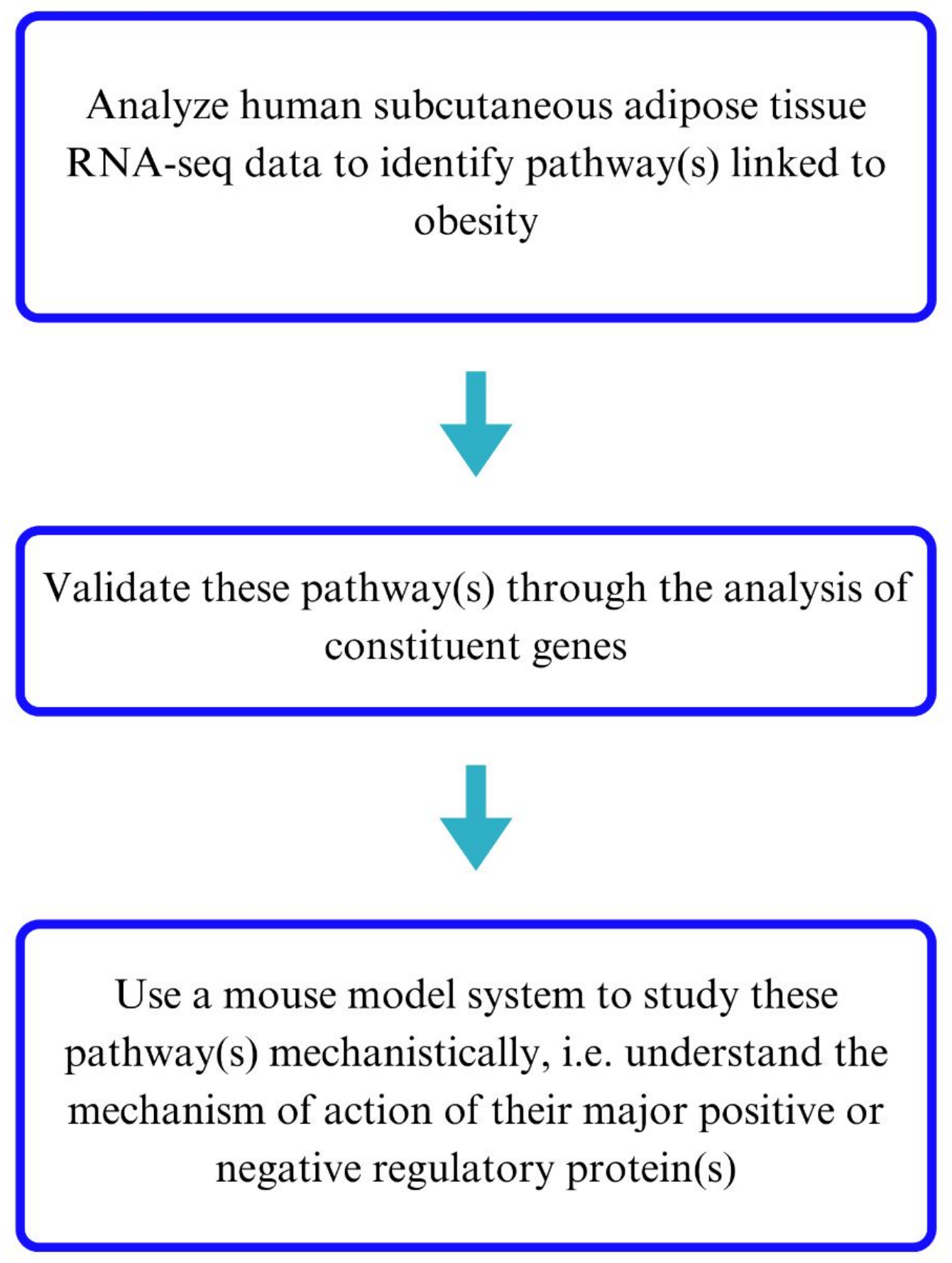
We are interested in further exploring links between inflammatory pathways, fat cell functionality, and metabolic syndromes (specifically obesity) to identify genes that could be targets of personalized treatments in the future.

Biological Background

ADAR1: RNA-editing enzyme that converts adenosine to inosine in cellular double-stranded RNA (dsRNA) to mark it as "self."[7]

- Acts like a master switch, prohibiting activation of the type 1 interferon response; proper functioning is required to prevent autoinflammation.[7]
- While ADAR1 has been well-studied in the context of viral infections and cancer, little is known about its role in adipose tissue.

Approach: High Level



Analyze human subcutaneous adipose tissue RNA-seq data to identify pathway(s) linked to obesity

Validate these pathway(s) through the analysis of constituent genes

Use a mouse model system to study these pathway(s) mechanistically, i.e. understand the mechanism of action of their major positive or negative regulatory protein(s)

Human Study

Dataset: Cohort of METabolic Syndrome In Men (METSIM) population study (GEO accession GSE135134): transcripts-per-million (TPM) RNA-seq data for abdominal subcutaneous fat tissue samples from 434 male subjects of varying ages and BMIs.[8]

Methods:

1. Data processing: genes for which greater than 80% of the patient samples had an expression value of 0.0 were removed, TPM values were log-transformed (log2(TPM + 0.001)), and patients were categorized into obese and healthy weight based on BMI values.
2. Gene set enrichment analysis (GSEA)[9,10] was conducted to identify pathways enriched in obese and healthy weight phenotypes.
3. Validation of select pathway(s):
   a. Log-transformed TPM expression values of genes within the pathway were plotted against BMI values of patients, linear regression was performed, and p-values were obtained.
   b. Average log2 fold-changes in expression between obese and healthy weight phenotypes of genes known to be significant in the pathway(s) were calculated.

Results:

While most of the highly-enriched pathways found through GSEA are known to be correlated with obesity, interestingly, the interferon alpha response (part of type 1 interferon signaling) ranked 21st in enrichment in the obese phenotype. This pathway was the focus of further investigation.

The expression values of many genes within this pathway strongly correlate with BMI. Specific genes (p-value < 0.05): *IRF7, IFI30, ISG15, ISG20, IFITM2, GBP2,* and *IFITM3*. See Figure 1 for example plot.

Genes known to be regulated by ADAR1 showed significant log2 fold-change in expression between healthy weight and obese phenotypes. See Table 1.

Figure 1: Scatterplot for *IFI30*

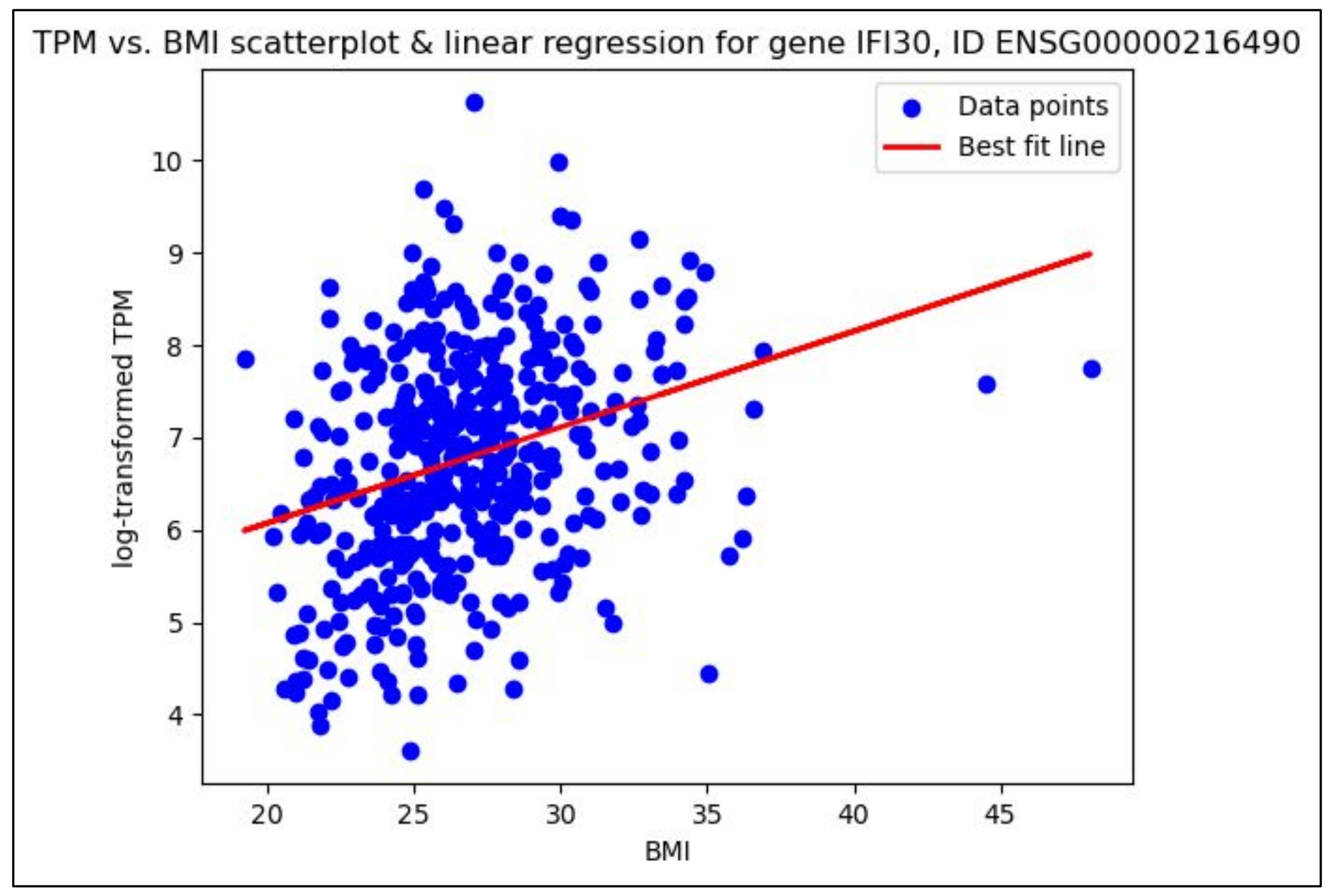


Table 1: Expression fold change for ADAR1-regulated genes

Gene name	Log2 fold change (obese / healthy weight)
Gvin1	2.66
NMES1	2.16
Sass6	2.04
IFI30	1.91

Mouse Study

Sought to study interferon alpha response using a mouse model. ADAR1 wildtype (WT) and knockout (KO) mice were used as model systems to identify ADAR1's RNA-editing sites to understand its regulatory mechanism of action.

Dataset: raw RNA-sequencing data (FASTQ format): 3 fat tissue samples of mice under WT ADAR1 and 3 under KO ADAR1 conditions.

Methods:

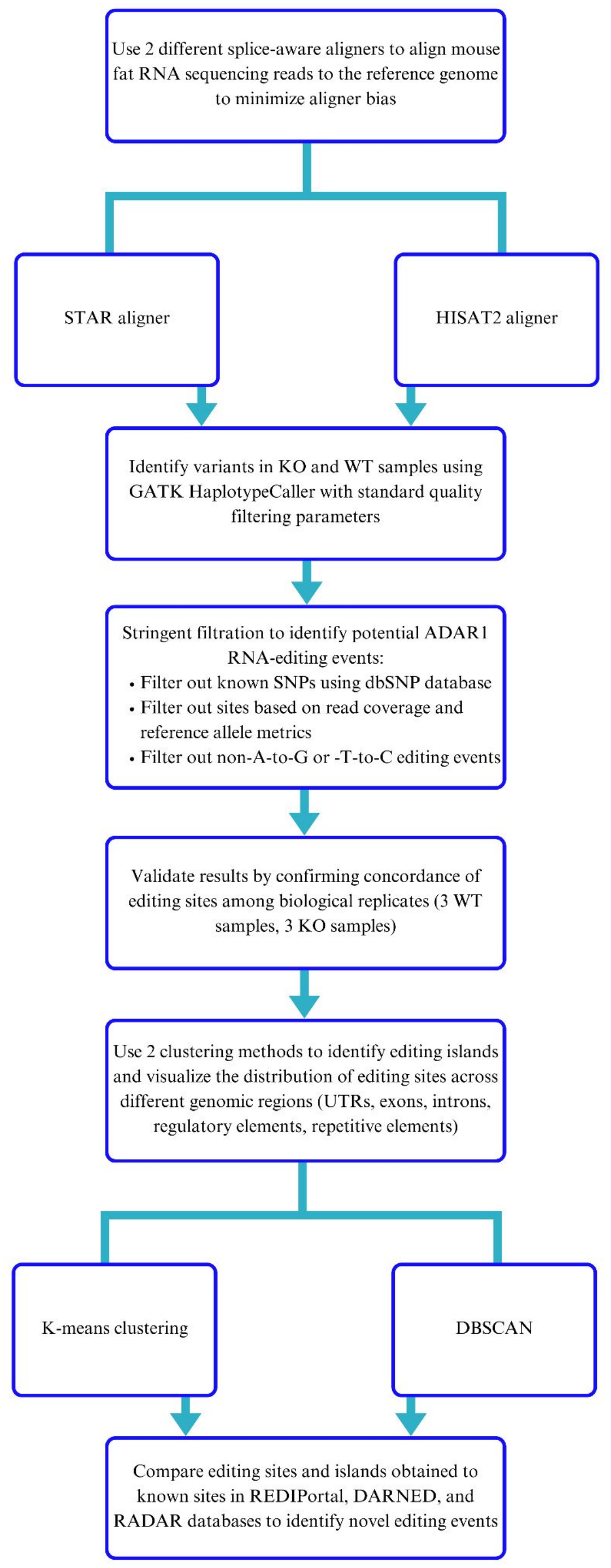
1. Raw sequencing reads were processed based on their quality metrics.
2. Reads were aligned to the reference genome using STAR.
3. Multiple open-source packages designed to identify RNA-editing sites were tested on control datasets and the mouse data, and limitations were noted:
   a. RNAEditor[11,12]: aligner it uses (BWA) is not splice-aware and is thus suboptimal for analyzing RNA sequencing data.
   b. SPRINT[13]: produced small overlap in RNA-editing sites common among biological replicates, showing that the results were unreliable.
   c. REDItools Denovo pipeline[14]: limited in that it works optimally only when a small portion of the genome is probed. So, coding regions were probed for RNA-editing sites. While the number of editing sites was consistent among biological replicates, there was no statistically significant change in the number of coding region editing sites between KO and WT conditions.

As ADAR1's editing sites are mostly located in noncoding regions (99% of the genome), there is a need for a tool that can accurately identify sites in these regions and produce results that are reproducible among biological replicates.

Future Plans

Implement, test, and revise a newly-designed statistical framework to identify ADAR1's RNA-editing sites in coding and noncoding regions.

Proposed Framework



Use 2 different splice-aware aligners to align mouse fat RNA sequencing reads to the reference genome to minimize aligner bias

STAR aligner

HISAT2 aligner

Identify variants in KO and WT samples using GATK HaplotypeCaller with standard quality filtering parameters

Stringent filtration to identify potential ADAR1 RNA-editing events:

- Filter out known SNPs using dbSNP database
- Filter out sites based on read coverage and reference allele metrics
- Filter out non-A-to-G or -T-to-C editing events

Validate results by confirming concordance of editing sites among biological replicates (3 WT samples, 3 KO samples)

Use 2 clustering methods to identify editing islands and visualize the distribution of editing sites across different genomic regions (UTRs, exons, introns, regulatory elements, repetitive elements)

K-means clustering

DBSCAN

Compare editing sites and islands obtained to known sites in REDIPortal, DARNED, and RADAR databases to identify novel editing events

References

[1] Centers for Disease Control and Prevention. (2024, May 14). Adult Obesity Facts. Centers for Disease Control and Prevention. https://www.cdc.gov/obesity/adult-obesity-facts/index.html
[2] Centers for Disease Control and Prevention. (2022, September 2). Adult BMI calculator. Centers for Disease Control and Prevention. https://www.cdc.gov/bmi/adult-calculator/index.html
[3] Centers for Disease Control and Prevention. (2024, May 15). Heart Disease Risk Factors. Centers for Disease Control and Prevention. https://www.cdc.gov/heart-disease/risk-factors/index.html
[4] Mahmoud, R., Kimonis, V., & Butler, M. G. (2022). Genetics of Obesity in Humans: A Clinical Review. International journal of molecular sciences, 23(19), 11005. https://doi.org/10.3390/ijms231911005
[5] Wen, X., Zhang, B., Wu, B., Xiao, H., Li, Z., Li, R., Xu, X., & Li, T. (2022). Signaling pathways in obesity: mechanisms and therapeutic interventions. Signal transduction and targeted therapy, 7(1), 298. https://doi.org/10.1038/s41392-022-01149-x
[6] Qiao, X., Zhou, Z.-C., Niu, R., Su, Y.-T., Sun, Y., Liu, H.-L., Teng, J.-L., Ye, J.-N., Shi, H., Yang, C.-D., & Cheng, X.-B. (2019). Hydroxychloroquine improves obesity-associated insulin resistance and hepatic steatosis by regulating lipid metabolism. Frontiers in Pharmacology, 10. https://doi.org/10.3389/fphar.2019.00855
[7] Song, B., Shiromoto, Y., Minakuchi, M., & Nishikura, K. (2022). The role of RNA editing enzyme ADAR1 in human disease. Wiley interdisciplinary reviews. RNA, 13(1), e1665. https://doi.org/10.1002/wrna.1665
[8] El-Sayed Moustafa, J. S., Jackson, A. U., Brotman, S. M., Guan, L., Villicaña, S., Roberts, A. L., Zito, A., Bonnycastle, L., Erdos, M. R., Narisu, N., Stringham, H. M., Welch, R., Yan, T., Lakka, T., Parker, S., Tuomilehto, J., Seow, J., Graham, C., Huettner, I., Acors, S., … Small, K. S. (2022). ACE2 expression in adipose tissue is associated with cardio-metabolic risk factors and cell type composition-implications f
COVID-19. International journal of obesity (2005), 46(8), 1478–1486. https://doi.org/10.1038/s41366-022-01136-w
[9] Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., Houstis, N., Daly, M. J., Patterson, N., Mesirov, J. P., Golub, T. R., Tamayo, P., Spiegelman, B., Lander, E. S., Hirschhorn, J. N., … Groop, L. C. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. N
Genetics, 34(3), 267–273. https://doi.org/10.1038/ng1180
[10] Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences, 102(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102
[11] John, D., Weirick, T., Dimmeler, S., & Uchida, S. (2017). RNAEditor: easy detection of RNA editing events and the introduction of editing islands. Briefings in bioinformatics, 18(6), 993–1001. https://doi.org/10.1093/bib/bbw087
[12] Stellos, K., Gatsiou, A., Stamatelopoulos, K., Perisic Matic, L., John, D., Lunella, F. F., Jaé, N., Rossbach, O., Amrhein, C., Sigala, F., Boon, R. A., Fürtig, B., Manavski, Y., You, X., Uchida, S., Keller, T., Boeckel, J. N., Franco-Cereceda, A., Maegdefessel, L., Chen, W., … Dimmeler, S. (2016). Adenosine-to-inosine RNA editing controls cathepsin S expression in atherosclerosis by enabling HuR-mediated post-transcriptional regulation. Nature medicine, 22(10), 1140–1150. https://doi.org/10.1038/nm.4172
[13] Zhang, F., Lu, Y., Yan, S., Xing, Q., & Tian, W. (2017). Sprint: An SNP-free toolkit for identifying RNA editing sites. Bioinformatics, 33(22), 3538–3548. https://doi.org/10.1093/bioinformatics/btx473
[14] Lo Giudice, C., Tangaro, M. A., Pesole, G., & Picardi, E. (2020). Investigating RNA editing in deep transcriptome datasets with REDItools and REDIportal. Nature Protocols, 15(3), 1098–1131. https://doi.org/10.1038/s41596-019-0279-7