

Noncoding Genome-Wide Association Studies Variant for Obesity: Inroads Into Mechanism

An Overview From the AHA's Council on Functional Genomics and Translational Biology

Connie Wu, PhD; Pankaj Arora, MD, FAHA

For the past decade in human genetics, genome-wide association studies (GWAS) have been the backbone for uncovering novel loci that could underlie a disease or phenotype of interest. However, the rapid growth in the field of human genetics has been plagued by a bottleneck in translating these GWAS findings into the development of potential therapeutics because of the challenges of identifying the “causal” variants and elucidating their underlying mechanisms of action.^{1,2} Recent advances in epigenomics (eg, the Roadmap Epigenomics Project^{3,4}) and genome editing (eg, the CRISPR/Cas9 technology⁵) have greatly facilitated our armory to dissect causality from GWAS data, equipping us with a range of resources that are necessary to understand the complex underpinnings of how specific variants can contribute to disease pathophysiology.

In the study by Claussnitzer et al⁶ published in the *New England Journal of Medicine*, the authors illustrate how researchers can use an integrative approach, taking advantage of novel epigenetic tools and genome editing techniques, as well as traditional cellular and mouse models, to identify a potential causal mechanism for the association of a noncoding variant in the fat mass and obesity-associated (*FTO*) locus with obesity (Figure).

From the Department of Medicine, Cardiovascular Research Center (C.W.) and Department of Anesthesia, Anesthesia Center for Critical Care Research, Critical Care and Pain Medicine (C.W.), Massachusetts General Hospital, Harvard Medical School, Boston, MA; Cardiology Division, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL (P.A.).

Correspondence to: Connie Wu, PhD, Cardiovascular Research Center, Anesthesia Center for Critical Care Research, Massachusetts General Hospital, 55 Fruit St, GRB-444, Boston, MA 02114. E-mail: connie.wu02@gmail.com and Pankaj Arora, MD, FAHA, Division of Cardiology, University of Alabama at Birmingham, 1808 7th Ave S, BDB 201, Birmingham, AL 35294. E-mail: parora@uabmc.edu

J Am Heart Assoc. 2016;5:e003060 doi: 10.1161/JAHA.115.003060.

© 2016 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley Blackwell. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

The initial reports demonstrating the association of the *FTO* locus with body mass index (BMI) and the risk of obesity were published in 2007 by Frayling et al⁷ and Dina et al.⁸ These studies revealed that the strongest genome-wide association signal is located in a 47-kb region of the first 2 introns of the *FTO* gene.^{7,8} The large number of genetic variants in linkage disequilibrium in this region (89 in total), as well as the lack of protein-altering variants, has made it challenging to pinpoint the “causal” variant and its underlying mechanism of action. Now, nearly a decade later, in a tour de force series of experiments, Claussnitzer et al⁶ identify a causal variant (rs1421085) in intron 1 of the *FTO* gene and present evidence that this variant modulates the binding of a transcriptional repressor, which plays a role in regulating the expression of 2 homeobox regulatory genes that are involved in the development and function of preadipocytes.

Using recently available chromatin state annotation data from the Roadmap Epigenomics Project,^{3,4} Claussnitzer et al⁶ found that intron 1 of the *FTO* gene contains a 12.8-kb-long enhancer that is most active in mesenchymal adipocyte progenitors. Subsequent enhancer luciferase reporter assays revealed a difference in the activity of this enhancer depending on the presence or absence of the *FTO* risk (ie, obesity-associated) haplotype, with the risk haplotype having 2.4 times higher enhancer activity than the nonrisk haplotype in human adipocytes. These results suggested that the *FTO* variants could be controlling enhancer activity. In an expression quantitative trait locus analysis, the authors observed that the mRNA levels of *IRX3* and *IRX5* were higher in adipose progenitor cells isolated from carriers of the *FTO* risk haplotype compared with carriers of the nonrisk haplotype. In addition, by using the chromosome conformation capture technique, the authors found that both *IRX3* and *IRX5* can directly interact with the *FTO* locus via long-range (1.2-Mb) chromatin interactions. These findings suggest that *IRX3* and *IRX5* could be putative target genes of the *FTO*-associated enhancer.

Next, the authors examined the effects of knocking down or overexpressing *IRX3* and *IRX5* in preadipocytes isolated

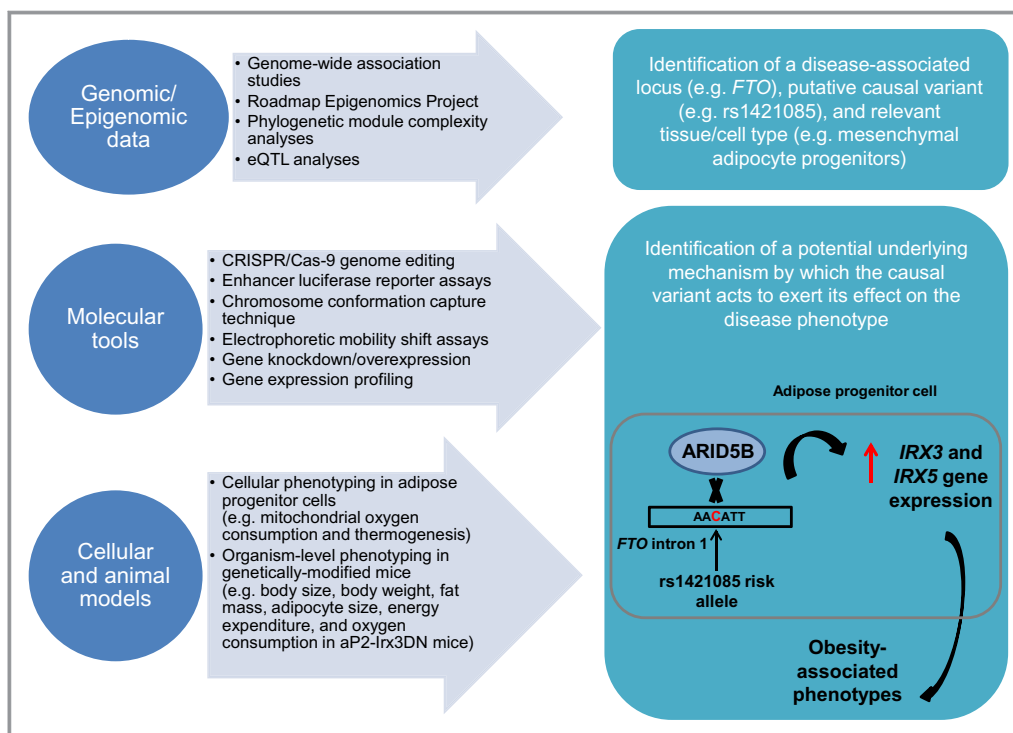


Figure. Integrative approach for deciphering the mechanistic basis of a disease-associated noncoding genetic variant. A combination of genomic/epigenomic data, molecular tools, and cellular and animal models can be used to enable the identification of a causal variant in a disease-associated locus and its potential mechanism of action in a relevant tissue/cell type. Specific examples of the types of data, tools, and models used in Clausnitzer et al⁶ are listed in the figure. Using this integrative approach, Clausnitzer et al⁶ established a putative causal variant (rs1421085) in the *FTO* obesity-associated locus and its potential underlying mechanism of action in adipose progenitor cells. Specifically, the rs1421085 risk allele disrupts the binding of ARID5B to a repressor motif in intron 1 of the *FTO* gene, resulting in increased enhancer activity and increased expression levels of target genes *IRX3* and *IRX5*, which subsequently leads to obesity-associated phenotypes.

from both carriers of the *FTO* risk haplotype and carriers of the nonrisk haplotype. They found that silent interfering (si) RNA-mediated knockdown of either *IRX3* or *IRX5* in preadipocytes from the risk haplotype carriers restored oxygen consumption and thermogenesis response to levels seen in the nonrisk haplotype carriers. In contrast, doxycycline-mediated overexpression of either *IRX3* or *IRX5* in preadipocytes from nonrisk haplotype carriers reduced mitochondrial function and thermogenesis response to levels seen in risk haplotype carriers. The authors then investigated the effects of the repression of *Irx3* in a mouse model, which expresses a dominant-negative form of *Irx3* specifically in adipose tissue (aP2-Irx3DN mice). These mice were found to exhibit a smaller body size and reduced body weight, fat mass, and adipocyte size compared with control mice. Moreover, the aP2-Irx3DN mice had increased energy expenditure and oxygen consumption and were resistant to weight gain when fed a high-fat diet. All of these features are in line with an antiobesity phenotype and provide support for the human observations. Taken together, these results suggest that *IRX3*

and *IRX5* may regulate obesity-associated phenotypes in preadipocytes via an *FTO*-mediated mechanism.

To attempt to identify the causal variant in the *FTO* locus, the authors used the phylogenetic module complexity analysis (PMCA),⁹ which is a tool that was developed in their laboratory to analyze “the evolutionary conservation of cis-regulatory modules across related species.”⁶ Based on this analysis, the authors honed in on the rs1421085 (T/C) variant, which was reported to have the highest PMCA score. This variant was also found to lie in a conserved binding motif for the AT-rich interaction domain (ARID) family of transcriptional regulators. By conducting electrophoretic mobility shift assays, the authors showed that the risk allele (C) of rs1421085 disrupts the binding of the ARID5B transcriptional repressor, which is highly expressed in adipose tissue and adipocytes, resulting in increased enhancer activity at the *FTO* locus. To study the potential role of ARID5B in regulating the expression of *IRX3* and *IRX5*, the authors examined the effects of knocking down or overexpressing *ARID5B* in preadipocytes isolated from carriers of the *FTO* risk haplotype and carriers of

the nonrisk haplotype. They found that siRNA-mediated knockdown of *ARID5B* in preadipocytes from nonrisk haplotype carriers not only increased the *IRX3* and *IRX5* mRNA levels to levels seen in risk haplotype carriers but also reduced basal oxygen consumption and lipolysis and altered expression patterns from mitochondrial to lipid markers. Conversely, doxycycline-mediated overexpression of *ARID5B* in preadipocytes from nonrisk haplotype carriers resulted in a reduction in the mRNA levels of *IRX3* and *IRX5*. Taken together, these findings support their model in which the rs1421085 risk allele disrupts the ARID5B binding motif, preventing ARID5B from binding to the DNA and leading to increased enhancer activity and subsequent increases in the expression levels of the target genes *IRX3* and *IRX5*.

In an elegant series of genome editing experiments, the authors used CRISPR/Cas9 technology to examine the effects of altering the predicted causal variant rs1421085, either changing the rs1421085 T allele to the C allele in preadipocytes from nonrisk haplotype carriers (which would be expected to disrupt the ARID5B motif) or changing the rs1421085 C allele to the T allele in preadipocytes from risk haplotype carriers (which would be expected to rescue the disrupted ARID5B motif). They demonstrated that genome editing of the T allele to the C allele increased the mRNA levels of *IRX3* and *IRX5* in preadipocytes from the nonrisk carriers and, interestingly, reediting the C allele back to the T allele restored low *IRX3* and *IRX5* mRNA levels in the preadipocytes. On the other hand, genome editing of the C allele to the T allele decreased the mRNA levels of *IRX3* and *IRX5* in preadipocytes from the risk carriers, and this decrease was not observed when *ARID5B* was knocked down. These findings suggest that the function of the rs1421085 variant is mediated by ARID5B, which regulates the expression of *IRX3* and *IRX5* in preadipocytes. Moreover, the authors found that the C-to-T genome editing in preadipocytes from the risk carriers increased the mRNA levels of a number of thermogenesis regulators and mitochondrial markers, decreased the mRNA levels of lipid storage and lipolytic markers, and increased basal metabolic rate, oxygen consumption, and thermogenesis. Taken together, these results provide evidence that rs1421085 could be the long-sought-after causal variant underlying the association between the *FTO* locus and obesity.

The study by Claussnitzer et al⁶ has uncovered a potential causal mechanism by which the rs1421085 variant in the *FTO* gene acts to exert its effect on obesity, advancing our understanding of this prominent obesity-associated locus. The fact that rs1421085 is an intronic variant highlights the importance of studying noncoding variants in our systematic efforts to follow GWAS signals to ascertain the “causal” variant underlying the disease of interest. In fact, it has been reported that the majority of the variants identified by GWAS are located in noncoding regions of the genome.¹⁰ Although

noncoding DNA sequences do not encode proteins, one should be cognizant that variants in noncoding DNA regions can still be functional by affecting transcriptional regulation, posttranscriptional regulation,¹¹ or epigenetic regulation as illustrated by Claussnitzer et al.⁶ The rs1421085 risk allele disrupts the binding of the transcriptional repressor, ARID5B, to an enhancer sequence in the DNA, leading to increased activity of the enhancer and subsequent upregulation of the expression levels of the downstream target genes *IRX3* and *IRX5*. Although the molecular mechanisms by which *IRX3* and *IRX5* affect mitochondrial oxygen consumption and thermogenesis remain to be elucidated, the discovery of the ARID5B-*FTO*-*IRX3*/*IRX5* regulatory axis provides an exciting opportunity to stimulate the development of innovative strategies that enhance the thermogenic properties of adipocytes as an approach to combat obesity. In addition to considering strategies to induce known regulators of thermogenesis, such as UCP1 and PGC1A, one might envisage discovering and developing drugs that would therapeutically perturb *IRX3*/*IRX5* or other targets in the ARID5B-*FTO*-*IRX3*/*IRX5* pathway to help promote the browning of adipocytes and increase thermogenesis and energy expenditure. Because it is likely that these targets play a role in other biological processes, a major challenge in therapeutic development will be to manipulate these targets in a specific way to avoid potential deleterious off-target effects in nonadipose tissue. In this regard, it may be informative to investigate whether there are other disease phenotypes or traits, besides obesity, that are associated with the rs1421085 genetic variant, as such knowledge could be used to gain a more comprehensive understanding of the potential untoward consequences, if any, of modulating the ARID5B-*FTO*-*IRX3*/*IRX5* pathway.

Last but not least, the study by Claussnitzer et al⁶ offers a roadmap for interpreting GWAS findings. By starting with a GWAS signal, researchers can consider using a combination of epigenomic, phylogenetic, and expression quantitative trait locus analyses to uncover the pathogenesis of the disease-associated locus in the relevant tissue/cell type. An array of cutting-edge molecular tools and state-of-the-art technologies could then be used to enable the identification of novel targets and pathways through which the variant acts to exert its effect on the disease phenotype. As demonstrated by Claussnitzer et al⁶ in their discovery of the ARID5B-*FTO*-*IRX3*/*IRX5* regulatory pathway, the advent of the chromosome conformation capture technique has facilitated our ability to identify “causal” genes that are not necessarily in the vicinity of the genetic variant in the DNA sequence but are, in fact, large genomic distances away. Moreover, Claussnitzer et al⁶ showed that CRISPR/Cas9 genome editing is a convenient and powerful tool to dissect causality of a genetic variant. The CRISPR/Cas9 technology allows researchers to introduce the risk allele of the variant in a

specific cell type of interest, evaluate the effect of that allele on cellular function/phenotype, and test whether correcting the allele back to the nonrisk allele would rescue the cellular function/phenotype. These single-nucleotide editing experiments can provide compelling evidence that the risk allele is sufficient to bring about the disease phenotype. Overall, the study by Claussnitzer et al⁶ is a great example of how researchers can capitalize on the genomic and technological advances of the past decade to gain mechanistic insights into observed genetic associations with human diseases.

Acknowledgments

Drs Wu and Arora are members of the Early Career Committee of the American Heart Association Council on Functional Genomics and Translational Biology.

Disclosures

None.

References

1. Paul DS, Soranzo N, Beck S. Functional interpretation of non-coding sequence variation: concepts and challenges. *Bioessays*. 2014;36:191–199.
2. Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. *Am J Hum Genet*. 2012;90:7–24.
3. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, Ziller MJ, Amin V, Whitaker JW, Schultz MD, Ward LD, Sarkar A, Quon G, Sandstrom RS, Eaton ML, Wu YC, Pfenning AR, Wang X, Claussnitzer M, Liu Y, Coarfa C, Harris RA, Shores N, Epstein CB, Gjoneska E, Leung D, Xie W, Hawkins RD, Lister R, Hong C, Gascard P, Mungall AJ, Moore R, Chuah E, Tam A, Canfield TK, Hansen RS, Kaul R, Sabo PJ, Bansal MS, Carles A, Dixon JR, Farh KH, Feizi S, Karlic R, Kim AR, Kulkarni A, Li D, Lowdon R, Elliott G, Mercer TR, Neph SJ, Onuchic V, Polak P, Rajagopal N, Ray P, Sallari RC, Siebenthal KT, Sinnott-Armstrong NA, Stevens M, Thurman RE, Wu J, Zhang B, Zhou X, Beaudet AE, Boyer LA, De Jager PL, Farnham PJ, Fisher SJ, Hausler D, Jones SJ, Li W, Marra MA, McManus MT, Sunyaev S, Thomson JA, Tlsty TD, Tsai LH, Wang W, Waterland RA, Zhang MQ, Chadwick LH,

- Bernstein BE, Costello JF, Ecker JR, Hirst M, Meissner A, Milosavljevic A, Ren B, Stamatoyannopoulos JA, Wang T, Kellis M. Integrative analysis of 111 reference human epigenomes. *Nature*. 2015;518:317–330.
4. Ernst J, Kellis M. Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nat Biotechnol*. 2015;33:364–376.
5. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol*. 2014;32:347–355.
6. Claussnitzer M, Dankel SN, Kim KH, Quon G, Meuleman W, Haugen C, Glunk V, Sousa IS, Beaudry JL, Puviindran V, Abdennur NA, Liu J, Svensson PA, Hsu YH, Drucker DJ, Mellgren G, Hui CC, Hauner H, Kellis M. FTO obesity variant circuitry and adipocyte browning in humans. *N Engl J Med*. 2015;373:895–907.
7. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, Perry JR, Elliott KS, Lango H, Rayner NW, Shields B, Harries LW, Barrett JC, Ellard S, Groves CJ, Knight B, Patch AM, Ness AR, Ebrahim S, Lawlor DA, Ring SM, Ben-Shlomo Y, Jarvelin MR, Sovio U, Bennett AJ, Melzer D, Ferrucci L, Loos RJ, Barroso I, Wareham NJ, Karpe F, Owen KR, Cardon LR, Walker M, Hitman GA, Palmer CN, Doney AS, Morris AD, Smith GD, Hattersley AT, McCarthy MI. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*. 2007;316:889–894.
8. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, Carlsson LM, Kiess W, Vatin V, Lecoeur C, Delplanque J, Vaillant E, Pattou F, Ruiz J, Weill J, Levy-Marchal C, Horber F, Potoczna N, Hercberg S, Le Stunff C, Bougneres P, Kovacs P, Marre M, Balkau B, Cauchi S, Chevre JC, Froguel P. Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet*. 2007;39:724–726.
9. Claussnitzer M, Dankel SN, Klocke B, Gallert H, Glunk V, Berulava T, Lee H, Oskolkov N, Fadista J, Ehlers K, Wahl S, Hoffmann C, Qian K, Ronn T, Riess H, Muller-Nurasyid M, Bretschneider N, Schroeder T, Skurk T, Horsthemke B, Spieler D, Klingenspor M, Seifert M, Kern MJ, Mejhert N, Dahlman I, Hansson O, Hauck SM, Bluher M, Arner P, Groop L, Illig T, Suhre K, Hsu YH, Mellgren G, Hauner H, Laumen H. Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms. *Cell*. 2014;156:343–358.
10. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE, Kang HM, Marth GT, McVean GA. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012;491:56–65.
11. Arora P, Wu C, Khan AM, Bloch DB, Davis-Dusenbery BN, Ghorbani A, Spagnoli E, Martinez A, Ryan A, Tainsh LT, Kim S, Rong J, Huan T, Freedman JE, Levy D, Miller KK, Hata A, Del Monte F, Vandenwijngaert S, Swinnen M, Janssens S, Holmes TM, Buys ES, Bloch KD, Newton-Cheh C, Wang TJ. Atrial natriuretic peptide is negatively regulated by microRNA-425. *J Clin Invest*. 2013;123:3378–3382.

Key Words: adipocyte • epigenetics • genome-wide association study • obesity