UC Irvine
UC Irvine Previously Published Works

Title

GWAS SNPs: decoders of disease-associated traits

Permalink

Journal

ISSN

Authors

Stamos, Michael J
Brady, Matthew T

Publication Date

DOI

GWAS SNPs: decoders of disease-associated traits

Shilpa Buch, Guoku Hu, Susmita Sil

Department of Pharmacology and Experimental Neuroscience, University of Nebraska Medical Center, Omaha, NE, USA

Correspondence to: Shilpa Buch. Department of Pharmacology and Experimental Neuroscience, University of Nebraska Medical Center, Omaha, NE, 68198-5880, USA. Email: sbuch@unmc.edu.

Comment on: Madelaine R, Notwell JH, Skariah G, *et al.* A screen for deeply conserved non-coding GWAS SNPs uncovers a MIR-9-2 functional mutation associated to retinal vasculature defects in human. Nucleic Acids Res 2018;46:3517-31.

The quest for identifying the genetic basis of common illnesses is deep rooted in the search for disease prevention and treatment. The advancement of genome-wide sequencing technology has linked thousands of genetic regions with hundreds of diseases and trait. This method searches the genome for small variations, called single nucleotide polymorphisms or single nucleotide polymorphisms (SNPs) that occur more frequently in people with a particular disease than in people without the disease. In fact, widespread genome-wide association studies (GWAS) have identified many human diseases associated SNPs, specifically in the protein coding regions. SNPs in the coding regions result in the loss or a change in the protein sequence and structure, which, in turn, affects its function, ultimately causing disease. A classic example is sickle-cell anemia which is caused by a SNP, rs334 (GAG codon changing to GTG), in a gene coding for the beta chain of hemoglobin protein. rs334(A) encodes the normal Hb A form of hemoglobin, while rs334(T) encodes the sickling form of hemoglobin, Hb S that disrupts hemoglobin formation resulting in sickle cell anemia (1). It is well recognized that SNPs in the protein coding regions do not directly cause diseases, rather are associated with disease susceptibility via alterations in protein structure and function (2). In addition to the presence of SNPs in the coding regions, recent GWAS studies have revealed that more than 90% of disease-associated SNPs are actually located in the non-coding regions including microRNAs, long non-coding RNAs as well as promoter regions of the genes (3). The role of SNPs in the non-exonic regions, however, remains largely elusive.

Recent report by Madelaine *et al.* describes the development of a novel methodology for screening regulatory GWAS non-coding SNPs that are conserved between the human and the zebrafish, and the associated syntenic genes (4). Using this approach, the authors identified 45 risk-associated SNPs located in the deeply conserved non-exonic elements (CNEs) across multiple species. To determine the function of these CNE/SNP pairs, the authors generated a chimeric, transgenic zebrafish expressing eGFP under the control of human CNE and demonstrated the transcriptional enhancer function for five of these conserved non-coding elements including CNE1/rs17421627; CNE8/rs1568679; CNE10/rs12431307; CNE15/rs11190870; and CNE18/rs16932455. Intriguingly, a single base pair change in three of the five CNEs, (CNE1/rs17421627, CNE8/rs1568679 and CNE18/rs16932455) resulted in significant reduction of transcriptional enhancer activity, thereby underpinning the pivotal role of these SNPs in regulating the expression of the adjacent genes.

Furthermore, the authors also performed a series of elegant experiments and demonstrated the function of the CNE1/rs17421627 pair in retinal vasculature defects. Firstly, the authors depleted the CNE1 enhancer in the zebrafish genome using the CRISPR/Cas-9 technology which resulted in disrupted blood vessel network and vascular defects in the retina of the homozygous mutant fish without any apparent morphological change in the eye. The function of this CNE/SNP pair was further validated in a kdrl (vegfr2) reporter line. These results strongly underpinned the function of CNE1/s17421627 in blood vessel development in the retina. Secondly, the authors

sought to examine the underlying mechanism(s) involved in this process. They found that one of the genes adjacent to CNE1, that for miR-9-2 in humans (miR-9-5 in zebrafish) had a similar expression pattern in the brain and retina as that for CNE1:egfp. Additionally, in the homozygous ΔCNE1 mutant zebrafish, miR-9-5 precursor was significantly decreased compared with the control siblings, indicating thereby that CNE1/rs17421627 mutation-induced downregulation of miR-9-2 could contribute to the associated retinal vasculature defects. Thirdly, in the miR-9 depleted kdrl:mCherry fish, the authors found a dramatic reduction of the hyaloid vasculature in the retina without any morphological change in the brains or eyes. Taken together this study clearly underscores the role of CNE1 as an enhancer of miR-9-2, and that rs17421627 risk nucleotide abolishes CNE1 enhancer activity. Importantly, deletion of CNE1 (ΔCNE1) in the zebrafish genome induced downregulation of miR-9, which in turn, led to retinal vasculature defects.

In summary, in this study, the authors have designed a computational tool for predicting non-coding GWAS SNPs that are conserved during evolution. Using the approach 45 non-coding CNE/SNP pairs that were deeply conserved across the vertebrate evolution and that were associated with human disease traits were identified. Interestingly, 5 out of these 45 CNE/SNP pairs act as transcriptional enhancers for the respective adjacent genes. This study paves the way for future examination into the role of other non-coding SNPs with human diseases. In addition to serving as enhancer, non-coding SNP containing regions can also be transcribed into non-coding RNAs, including miRNAs and lncRNAs. For example, Miao *et al.* have identified almost 7,260,238 SNPs in ~141,353 human lncRNA transcripts and almost 3,921,448 SNPs in 117,405 mouse lncRNA transcripts (5). It is thus possible that in addition to the presence of SNPs in the non-exonic regions, that there could also be presence of SNPs in the non-coding RNA transcripts that, in turn, could also contribute to SNP-associated diseases. Given that 90% of SNPs are located in the non-coding regions and that, most of them are not yet explored, efforts aimed at deciphering their functional properties and the underlying mechanism(s) warrants investigation. These studies could have immense clinical ramifications for developing better prevention and treatment strategies for various genetic diseases with a focus on personalized medicine.

Footnote

Provenance and Peer Review: This article was commissioned and reviewed by the Section Editor Dr. Rongrong Gao (Department of Cardiology, the First Affiliated Hospital of Nanjing Medical University, Nanjing, China).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/ncri.2018.07.03). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Ingram VM. Gene mutations in human haemoglobin: the chemical difference between normal and sickle cell haemoglobin. Nature 1957;180:326-8.
2. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285-91.
3. Chen L, Jin P, Qin ZS, et al. DIVAN: accurate identification of non-coding disease-specific risk variants using multi-omics profiles. Genome Biol 2016;17:252.
4. Madelaine R, Notwell JH, Skariah G, et al. A screen for deeply conserved non-coding GWAS SNPs uncovers a MIR-9-2 functional mutation associated to retinal

vasculature defects in human. Nucleic Acids Res 2018;46:3517-31.

5. Miao YR, Liu W, Zhang Q, et al. lncRNASNP2: an updated database of functional SNPs and mutations in human and mouse lncRNAs. Nucleic Acids Res 2018;46:D276-80.