Clinical Epigenetics

CORRESPONDENCE Open Access

Combining human tissue and iPSC-derived cardiomyocyte eQTL datasets to understand noncoding genetic variants: boosting the cardiogenetics toolbox

Saif Dababneh[1,2], Kyoung-Han Kim[3,4] and Glen F. Tibbits[2,5,6*]

Introduction

Advancements in next-generation sequencing and genome-wide association studies (GWAS) have revealed hundreds of loci associated with various cardiovascular diseases, highlighting the important role genetic variants play in disease pathogenesis and identifying potential therapeutics. Notably, most GWAS variants are located in noncoding genomic regions, which do not directly affect protein function. Instead, these variants are often found in genomic regions containing regulatory elements, such as promoters, enhancers, and silencers. Consequently, they regulate gene expression levels and the cell-type specificity of transcripts via modulation of transcription factor binding and chromatin accessibility [1]. Unlike variants in coding regions, where the pathogenic effect of the variant could be predicted by changes in amino acid sequence, understanding the impact of noncoding variants requires comprehensive transcriptomic and epigenomic investigations, rendering the process more challenging and costly. Additionally, the pathogenicity of noncoding variants is more difficult to interpret clinically due to our limited understanding and the scarcity of noncoding variant risk prediction tools.

One common methodology used to assess the impact of a GWAS-identified noncoding variant on gene expression is expression quantitative trait locus (eQTL), in which the expression of a gene of interest is stratified by the number of loci carrying the variant of interest [1]. Databases providing the scientific community with eQTLs are critical for researchers to dissect the impact of noncoding variants. By analyzing changes in variant-mediated gene expression, researchers can identify target genes likely affected and perform biological validation experiments to elucidate the transcriptional and epigenetic mechanisms involved using various disease models.

The genotype-tissue expression (GTEx) biobank, established in 2013, has become an immense resource for transcriptomic and eQTL analyses in cardiogenetics. It encompasses gene expression (via RNA-sequencing) data on 429 and 432 left atrial (LA) appendage and left ventricle (LV) samples from males and females of various ages and ancestries, 372 and 386 of which were genotyped, respectively (Fig. 1A) [2]. However, given that many eQTLs are cell-type specific, wide-scale interpretation of noncoding variants has remained challenging. In parallel, there has been a steep rise in the use of human

*Correspondence:
Glen F. Tibbits
tibbits@sfu.ca
[1] Department of Cellular and Physiological Sciences, University of British Columbia, Vancouver, BC V5Z 4H4, Canada
[2] Cellular and Regenerative Medicine Centre, BC Children's Hospital Research Institute, 938 West 28th Avenue, Vancouver, BC V5Z 4H4, Canada
[3] Department of Cellular and Molecular Medicine, Faculty of Medicine, University of Ottawa, Ottawa, ON K1H 8M5, Canada
[4] University of Ottawa Heart Institute, Ottawa, ON K1Y 4W7, Canada
[5] Departments of Molecular Biology and Biochemistry / Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC V5A 1S6, Canada
[6] School of Biomedical Engineering, University of British Columbia, Vancouver, BC V6T 1Z3, Canada

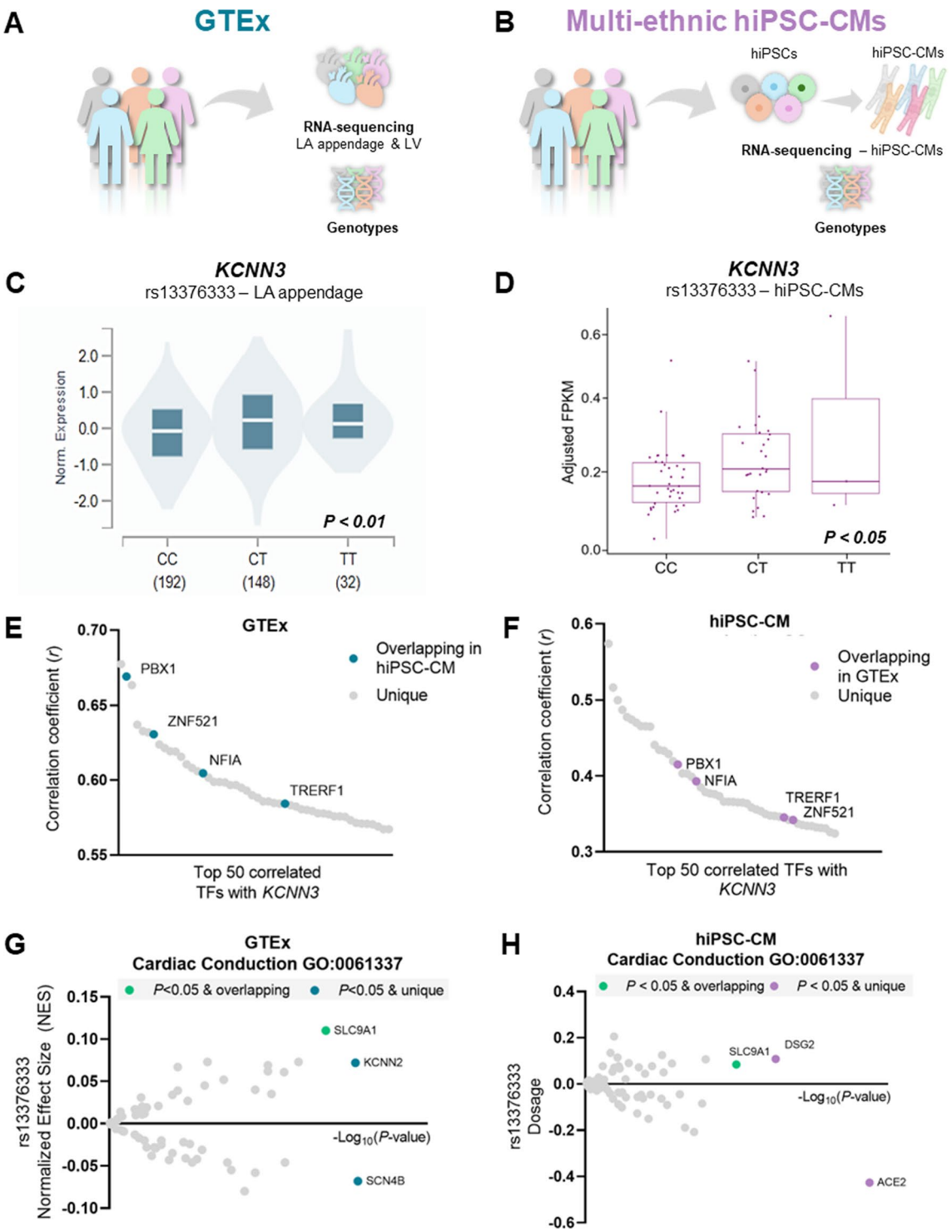

Fig. 1 Investigating rs13376333 using GTEx and multi-ethnic hiPSC-CM databases. **A** Summary of genotype-tissue expression (GTEx) database and **B** multi-ethnic hiPSC-CM database. The impact of rs13376333 on *KCNN3* expression in **C** GTEx left atrial appendage samples and **D** hiPSC-CMs. The top 50 correlated transcription factors (TFs) with *KCNN3* expression in **E** GTEx left atrial appendage samples and **F** hiPSC-CMs, showing unique TFs and TFs overlapping in both databases. The impact of rs13376333 on genes in the gene ontology (GO) term Cardiac Conduction in **G** GTEx left atrial appendage samples and **H** hiPSC-CMs, showing unique and overlapping genes that are significantly affected ($P < 0.05$) by rs13376333

induced pluripotent stem-cell derived cardiomyocytes (hiPSC-CMs) to model cardiac disorders and delineate the pathogenicity and mechanisms of noncoding variants. The use of hiPSC-CMs facilitates this research in a personalized and scalable manner which captures patients' genomic diversity, as done for heart failure, congenital heart disease and arrhythmia [3–5]. Thus, generating a genotype-expression biobank with hiPSC-CMs from diverse patients is critical to advance the study of noncoding variants in the era of personalized medicine.

Recently, Lv *et al.* [6] have successfully constructed a multi-ethnic hiPSC-CM eQTL database, which includes 71 hiPSC-CMs derived from male and female patients of diverse ethnicities and ages (Fig. 1B). As noted by the authors, eQTLs derived from cardiac tissue are not cell-type specific, with only ~20% cardiomyocyte composition in cardiac tissue. This resource represents an important advance in understanding the mechanisms of cardiomyocyte-specific noncoding variants, which is crucial for downstream biological validation in hiPSC-CMs. In this perspective article, we highlight the potential of integrating tissue and hiPSC-CM eQTL datasets to deepen our understanding of noncoding variants.

Combining databases to explore the atrial fibrillation-associated variant rs13376333

Atrial fibrillation (AF) is largely co-morbid with heart failure, worsens the prognosis, and in some cases, can be the root cause of heart failure [7]. To demonstrate the utility of these two publicly available resources in understanding potential mechanisms of noncoding variants, we focus on the common AF-associated variant rs13376333. This variant is located within the intronic region between exons 1 and 2 of the *KCNN3* gene, which is the most significantly associated with AF from the 1q21 locus on chromosome 1 (Odds ratio = 1.56) [8]. Given the proximity of rs13376333 to the *KCNN3* promoter, we first tested the association of rs13376333 with *KCNN3* expression, showing a significantly positive cis-eQTL association in both LA appendage samples (Fig. 1C) and hiPSC-CMs (Fig. 1D). To identify potential transcriptional regulators of *KCNN3*, we then measured the correlation of *KCNN3* expression with the expression of all known human transcription factors (TFs) in each dataset, revealing 4 shared hits (*PBX1*, *ZNF521*, *NFIA*, *TRERF1*) among the top 50 correlated TFs (Fig. 1E, F). Although these data may not be informative in isolation, they can provide candidate transcriptional regulators in combination with motif enrichment, epigenomics, and biological assays. Furthermore, given that AF is a rhythm disorder, we assessed the impact of rs13376333 (eQTL) on genes in the gene ontology term 'Cardiac Conduction' in LA appendage samples (Fig. 1G) and hiPSC-CMs (Fig. 1H), and identified a common hit, *SLC9A1*, encoding Na^+/H^+ exchanger 1 (NHE1). NHE1 is upregulated in heart failure and AF and is a potential target of SGLT2 inhibitors [9], suggesting a potential implication of *SLC9A1* in AF risk in rs13376333 carriers. Given the emerging role of SGLT2 inhibitors in the treatment of all-cause heart failure [10], our analysis pinpointed a candidate mechanism that can be further investigated to understand its contribution to AF in rs13376333-carrying hiPSC-CMs and heart failure patients. Moreover, this analysis suggested that the expression of genes beyond the 1q21 locus (e.g., *KCNN2* on Chr 5, *SCN4B* on Chr 11, *DSG2* on Chr 18) may be altered with rs13376333, implicating both cis- and trans-eQTLs associated with this variant, although it remains to be clarified whether the trans-eQTLs are mediated through the cis-altered genes. Collectively, these analyses underscore the value of eQTL databases in hypothesis generation and candidate target identification, but biological validation remains necessary to fully elucidate disease mechanisms.

Future directions

We have demonstrated that integrating human cardiac tissue and hiPSC-CM eQTL datasets is feasible and can be of great value. To further enhance the utility of these resources, we recommend enabling the analysis of large gene lists simultaneously in the hiPSC-CM eQTL database. Currently, up to 400 genes can be tested at once in the GTEx dataset; expanding this capability would allow for more comprehensive investigations of cis- and trans-acting variants and their genome-wide effects. With rapid advancements in single-cell genomics and QTL methods, such as histone acetylation QTL (haQTL), methylation QTL (mQTL), and splicing QTL (sQTL), the investigation of noncoding variants is expected to become increasingly feasible, reliable, and specific to both cell-type and cell-state. Future research should leverage multi-ethnic and diverse hiPSC-CMs to combine single-cell omics with advanced QTL studies in combination with biological validation. By better understanding the role of noncoding variants, patients carrying such variants can be better risk stratified and novel pharmacological targets can be developed by the unraveling of novel disease mechanisms.

Acknowledgements

We would like to thank the contributors of the datasets utilized in this manuscript for making these valuable datasets publicly available and user-friendly.

Author contributions

S.D. conceived the study design, performed all analyses, and wrote the first draft of the manuscript. K.-H.K. and G.F.T. provided input and revised the manuscript. All authors approve the final version of the manuscript.

Funding
This work was supported by the Heart and Stroke Foundation of Canada (HSFC) to G.F.T. S.D. was supported by the MD/PhD studentship (University of British Columbia) and the Canada Graduate Scholarship—Doctoral Award (Canadian Institutes of Health Research, 193439). K.-H.K. was supported by the National New Investigator Award from the HSFC and the Early Researcher Award from the Government of Ontario, Canada.

Availability of data and materials
Both datasets used in our analysis are publicly available: GTEx eQTL calculator (https://gtexportal.org/home/testyourown) and Multi-ethnic hiPSC-CMs eQTL (https://guerratylab.org/ipsc/). The web user interface was used to calculate eQTL associations. Pearson correlation analyses were performed using Microsoft Excel. Each gene from the gene ontology term 'Cardiac Conduction' was manually entered into each eQTL calculator and subsequently compared across the datasets.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 14 July 2024 Accepted: 5 August 2024
Published online: 14 August 2024

References

1. Zhang F, Lupski JR. Non-coding genetic variants in human disease. Hum Mol Genet. 2015;24:102–10. https://doi.org/10.1093/hmg/ddv259.
2. GTEx Consortium. The genotype-tissue expression (GTEx) project. Nat Genet. 2013;45:580–5. https://doi.org/10.1038/ng.2653.
3. Xiao F, Zhang X, Morton SU, Kim SW, Fan Y, Gorham JM, et al. Functional dissection of human cardiac enhancers and noncoding de novo variants in congenital heart disease. Nat Genet. 2024;56:420–30. https://doi.org/10.1038/s41588-024-01669-y.
4. van Eif VWW, Protze SI, Bosada FM, Yuan X, Sinha T, van Duijvenboden K, Ernault AC, Mohan RA, Wakker V, de Gier-de VC, Hooijkaas IB, Wilson MD, Verkerk AO, Bakkers J, Boukens BJ, Black BL, Scott IC, Christoffels VM. Genome-wide analysis identifies an essential human TBX3 pacemaker enhancer. Circ Res. 2020;127(12):1522–35. https://doi.org/10.1161/CIRCRESAHA.120.317054.
5. Gacita AM, Fullenkamp DE, Ohiri J, Pottinger T, Puckelwartz MJ, Nobrega MA, McNally EM. Genetic variation in enhancers modifies cardiomyopathy gene expression and progression. Circulation. 2021;143(13):1302–16. https://doi.org/10.1161/CIRCULATIONAHA.120.050432.
6. Lv W, Babu A, Morley MP, Musunuru K, Guerraty MA. Resource of gene expression data from a multiethnic population cohort of induced pluripotent stem cell-derived cardiomyocytes. Circ Genom Precis Med. 2024;17:e004218. https://doi.org/10.1161/CIRCGEN.123.004218.
7. Santhanakrishnan R, Wang N, Larson MG, Magnani JW, McManus DD, Lubitz SA, Ellinor PT, Cheng S, Vasan RS, Lee DS, Wang TJ, Levy D, Benjamin EJ, Ho JE. Atrial fibrillation begets heart failure and vice versa: temporal associations and differences in preserved versus reduced ejection fraction. Circulation. 2016;133(5):484–92. https://doi.org/10.1161/CIRCULATIONAHA.115.018614.
8. Ellinor PT, Lunetta KL, Glazer NL, Pfeufer A, Alonso A, Chung MK, et al. Common variants in KCNN3 are associated with lone atrial fibrillation. Nat Genet. 2010;42:240–4. https://doi.org/10.1038/ng.537.
9. Trum M, Riechel J, Lebek S, Pabel S, Sossalla ST, Hirt S, Arzt M, Maier LS, Wagner S. Empagliflozin inhibits Na^+/H^+ exchanger activity in human atrial cardiomyocytes. ESC Heart Fail. 2020;7:4429–37. https://doi.org/10.1002/ehf2.13024.
10. Vaduganathan M, Docherty KF, Claggett BL, Jhund PS, de Boer RA, Hernandez AF, Inzucchi SE, Kosiborod MN, Lam CSP, Martinez F, Shah SJ, Desai AS, McMurray JJV, Solomon SD. SGLT-2 inhibitors in patients with heart failure: a comprehensive meta-analysis of five randomised controlled trials. Lancet. 2022;400(10354):757–67. https://doi.org/10.1016/S0140-6736(22)01429-5.

Publisher's Note