Regional variation in cardiovascular genes enables a tractable genome editing strategy

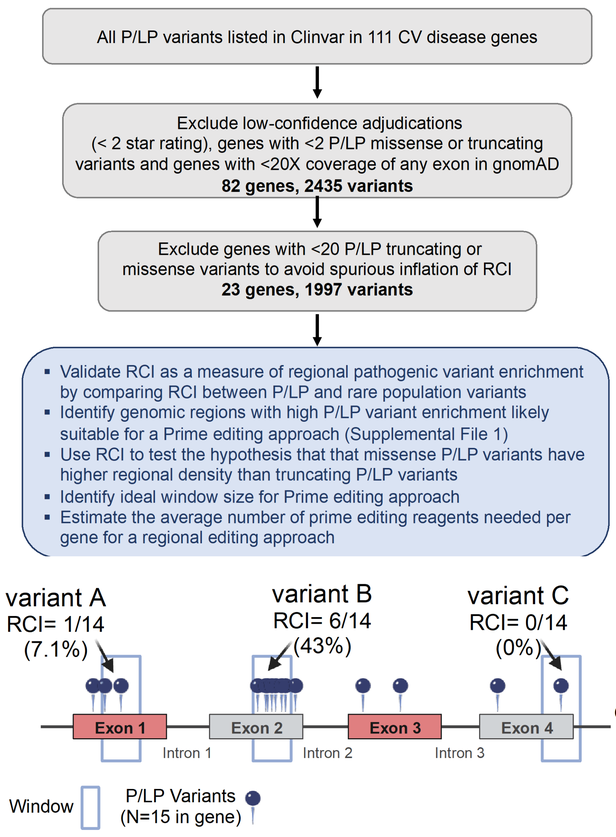
Vikki A. Krysov, BS[1,2*], Rachel H. Wilson, PhD[1*], Nicholas S. Ten, BS[1], Nathan Youlton, BS[1], Hannah N. De Jong, PhD[3]; Shirley Sutton, MS[1], Yong Huang, MD[1], Chloe M. Reuter, MS LCGC[1,6], Megan E. Grove, MS LCGC[4], Matthew T. Wheeler, MD PhD[1,6]; Euan A. Ashley, FRCP DPhil[1,5,6], and Victoria N. Parikh, MD[1,6]

[1] Division of Cardiovascular Medicine, Stanford University School of Medicine; [2] University of California, Davis School of Medicine; [3] Maze Therapeutics, Inc; [4] Color Health, Burlingame, CA; [5] Department of Genetics, Stanford University School of Medicine; [6] Stanford Center for Inherited Cardiovascular Disease, Stanford Medicine, Stanford, CA

Introduction

Recent rise in the number and therapeutic potential of genome engineering technologies has generated excitement for their application in cardiovascular therapeutics. One significant barrier to their implementation is costly and time-consuming reagent development for novel variants. We have previously shown that disease-associated variants cluster in functional protein domains where variants are found in the general population at lower frequency.[1,2] These findings may provide an opportunity to pre-emptively target multiple pathogenic variant clusters ("pathogenic hotspots") that address a majority of pathogenic and likely pathogenic (P/LP) variants for a given disease with a small number of pre-designed reagents. Prime editing, a search-and-replace editing technology, can overwrite genomic sequences with a reverse transcriptase guided by a Cas9 nickase and prime editing guide (peg)RNA.[3] We hypothesized that most cardiovascular disease-relevant genes in ClinVar would display regional variant clustering, and that multiple variants within a regional hotspot could be targeted with a limited number of prime editing reagents.

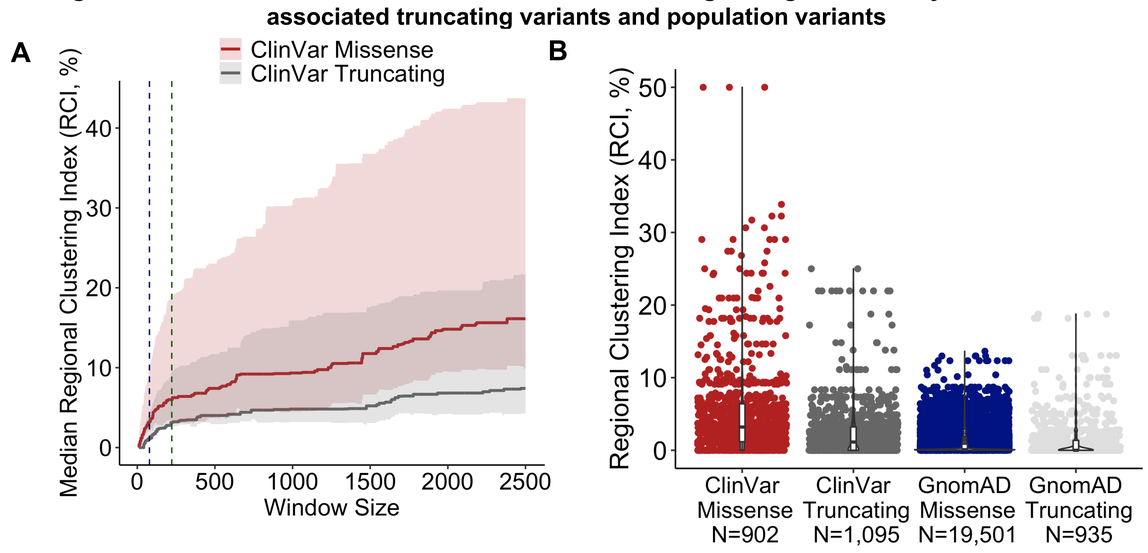
Methods



Results

We hypothesized that the regional clustering is more pronounced in P/LP missense variants than truncating variants, which are often associated with loss of function regardless of their position in a transcript. We further sought to understand how editing window size might impact the mean percent of variants per gene (and therefore number of editing reagents required). Overall, missense P/LP had a higher mean RCI as compared to truncating variants regardless of window size (**Figure 1A**). Further, the percent of simultaneously targetable variants (RCI) initially rises exponentially with window size, followed by diminishing returns after a window size of ~254 bp, at which a median of 6% of missense P/LP variants in each gene is targeted (**Figure 1A**).

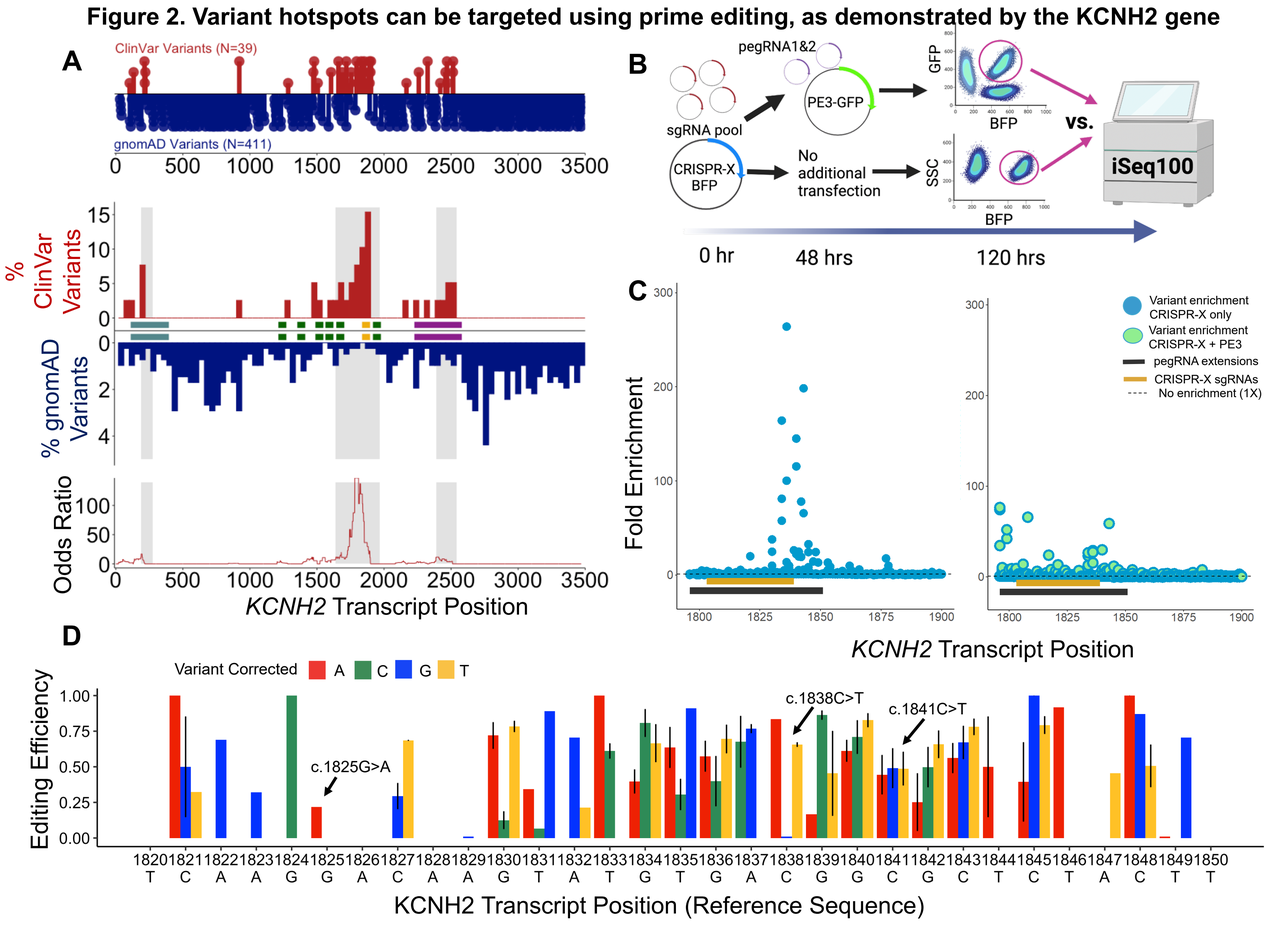
Figure 1. Disease-associated missense variants exhibit higher regional density than disease-associated truncating variants and population variants



At a window size commensurate with maximally reported prime editing extension length (78 bp)[3], mean RCI in missense variants remained higher than for truncating variants, and rare variants observed in the same genes in gnomAD showed reduced RCI compared to ClinVar P/LP variants (**Figure 1B**). In the average cardiovascular disease gene, it is possible that 20 or fewer pegRNAs can address all known pathogenic missense variants.

We next tested the feasibility of prime editing for multiple variants in a single hotspot in *KCNH2*, a cause of long and short QT syndrome with a high mean missense and truncating variant RCI. Positions c.1800-1850 showed significant P/LP enrichment and overlap the H5 intramembrane pore-forming domain (**Figure 2A**). We first introduced multiple variants in this hotspot in HEK293 cells using CRISPR-X.[4] This was followed by prime editing with generation PE3 including 2 pegRNAs[3] and FACS (**Figure 2B**). CRISPR-X alone showed up to ~300-fold increase per variant compared to untreated cells (**Figure 2C).** In CRISPR-X+PE3, a much smaller fold increase was observed (**Figure 2C**). The mean prime editing efficiency across CRISPR-X-enriched variants within this hotspot was 57±27%, including 3 P/LP variants (**Figure 2D**).

Figure 2. Variant hotspots can be targeted using prime editing, as demonstrated by the KCNH2 gene



Discussion

- We show that regional clustering of P/LP variants in CV genes represents an opportunity to develop off-the-shelf therapeutics that re-write the genome for patients with inherited cardiovascular disease.
- The clustering of P/LP missense rather than truncating variants in these hotspots suggests that prime editing technology is particularly valuable for dominant negative disease.

References

1. Parikh VN, Caleshu C, Reuter C, Lazzeroni LC, Ingles J, Garcia J, McCaleb K, Adesiyun T, Sedaghat-Hamedani F, Kumar S, Graw S, Gigli M, Stolfo D, Dal Ferro M, Ing AY, Nussbaum R, Funke B, Wheeler MT, Hershberger RE, Cook S, Steinmetz LM, Lakdawala NK, Taylor MRG, Mestroni L, Merlo M, Sinagra G, Semsarian C, Meder B, Judge DP, Ashley E. Regional Variation in RBM20 Causes a Highly Penetrant Arrhythmogenic Cardiomyopathy. *Circ Heart Fail*. 2019;12:e005371.

2. Dries AM, Kirillova A, Reuter CM, Garcia J, Zouk H, Hawley M, Murray B, Tichnell C, Pilichou K, Protonotarios A, Medeiros-Domingo A, Kelly MA, Baras A, Ingles J, Semsarian C, Bauce B, Celeghin R, Basso C, Jongbloed JDH, Nussbaum RL, Funke B, Cerrone M, Mestroni L, Taylor MRG, Sinagra G, Merlo M, Saguner AM, Elliott PM, Syrris P, van Tintelen JP, James CA, Haggerty CM, Parikh VN. The genetic architecture of Plakophilin 2 cardiomyopathy. *Genet Med*. 2021;23:1961–1968.

3. Anzalone AV, Randolph PB, Davis JR, Sousa AA, Koblan LW, Levy JM, Chen PJ, Wilson C, Newby GA, Raguram A, Liu DR. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 2019;576:149–157.

4. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, Gauthier LD, Brand H, Solomonson M, Watts NA, Rhodes D, Singer-Berk M, England EM, Seaby EG, Kosmicki JA, Walters RK, Tashman K, Farjoun Y, Banks E, Poterba T, Wang A, Seed C, Whiffin N, Chong JX, Samocha KE, Pierce-Hoffman E, Zappala Z, O'Donnell-Luria AH, Minikel EV, Weisburd B, Lek M, Ware JS, Vittal C, Armean IM, Bergelson L, Cibulskis K, Connolly KM, Covarrubias M, Donnelly S, Ferriera S, Gabriel S, Gentry J, Gupta N, Jeandet T, Kaplan D, Llanwarne C, Munshi R, Novod S, Petrillo N, Roazen D, Ruano-Rubio V, Saltzman A, Schleicher M, Soto J, Tibbetts K, Tolonen C, Wade G, Talkowski ME, Genome Aggregation Database Consortium, Neale BM, Daly MJ, MacArthur DG. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581:434–443.