Filling in the Gaps

Deciphering the Function of Noncoding DNA

Andreas S. Barth, MD, PhD

1. *May D, Blow MJ, Kaplan T, McCulley DJ, Jensen BC, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Afzal V, Simpson PC, Rubin EM, Black BL, Bristow J, Pennacchio LA, Visel A. Large-scale discovery of enhancers from human heart tissue.* Nat Genet. *2011;44:89–93.*

Study Hypothesis

Less than 2% of human DNA codes for proteins, whereas the vast majority of DNA in the human genome consists of sequences with unknown function. A substantial portion of the noncoding sequences harbors regulatory motifs important for gene expression. Recent genome-wide association studies (GWAS) suggest that the majority of single nucleotide polymorphisms (SNPs) associated with human traits and diseases are found in noncoding genomic regions, potentially involving regulatory elements. The critical role of distal regulatory elements, that is, enhancers, is highlighted in GWAS of expression quantitative trait loci (eQTLs) showing that 70–80% of regulatory variants are found at large distances from protein-coding genes. In this report, May et al[1] generated a genome-wide catalogue of enhancers from human myocardium. As it is well established that sequence variations in a noncoding regulatory sequence can lead to abnormal gene expression, mapping of the genomic coordinates of enhancers in a given tissue is the first step toward elucidating the function of these noncoding DNA sequences and gaining a better understanding of the mechanisms by which sequence variations in regulatory sequences affect human biology.

How Was the Hypothesis Tested?

The transcriptional coactivator family p300-CBP is ubiquitously expressed and interacts with numerous transcription factors in promoters and enhancers. The authors used chromatin immunoprecipitation with p300 from fetal and adult human myocardial tissues combined with massively parallel sequencing (ChIP-Seq) to predict the genomic location of heart-specific enhancers with a genomic distance of at least 2.5 kb from the nearest transcription start site. For validation, data from human hearts were compared with ChIP-Seq data from mouse hearts. Additionally, in vivo activity of 65 putative p300-based predictions of human heart enhancers was tested in transgenic mouse assays.

Principal Findings

The authors identified more than 6200 candidate enhancer sequences in adult and fetal human hearts. Interestingly, comparison of putative enhancers from human fetal tissue and age-matched mouse myocardium revealed considerable species-specific differences, as only about one-fifth of fetal human heart enhancer candidates coincided with significant peaks at the orthologous site in the mouse genome. Despite the poor sequence conservation between human and mouse enhancers, the top candidates in each species shared a similar collection of transcription factor binding sites, providing a possible explanation for the activity of even poorly conserved human heart enhancers in transgenic mouse assays, which validated 43 of 65 candidate enhancers tested (66%), a significantly larger proportion than what can be expected by chance alone. Additional support that these genomic regions function as enhancers in human myocardial tissue came from statistical enrichment analysis of functional gene annotations that demonstrated that the predicted enhancers were enriched near genes with known cardiac functions. Importantly, more than twice as many putative enhancers (≈5000) were found in fetal heart tissue compared with adult myocardium (≈2200) and only half of the enhancers identified in adult tissue were concomitantly found in fetal hearts, suggesting considerable developmental-specific regulation.

Testing the hypothesis that genomic variation within the candidate enhancers carries functional significance, the authors identified several SNPs for gene expression in primary human cells. Because functional variants—particularly eQTLs that map far away from their parent genes (trans eQTLs)—commonly operate in a tissue-dependent manner, the authors examined 81 human fetal candidate enhancers located within 50 kb of 30 selected cardiac genes that are associated with various cardiovascular diseases. They found more than 1500 SNPs in these genomic regions; however, in-depth follow-up studies are needed to elucidate the role of these SNPs for the in vivo activity of the cardiac enhancers and for their potential role in disease processes.

From the Department of Medicine, Johns Hopkins Bayview Medical Center and Department of Medicine, Division of Cardiology, Johns Hopkins University, Baltimore, MD.

Correspondence to Andreas S. Barth, MD, PhD, Department of Medicine, Johns Hopkins Bayview Medical Center and Department of Medicine, Division of Cardiology, Johns Hopkins University, 720 Rutland Ave, Ross Bldg 844, Baltimore, MD 21205. E-mail abarth3@jhmi.edu

(*Circ Cardiovasc Genet*. 2012;5:151-152.)

***Circ Cardiovasc Genet* is available at http://circgenetics.ahajournals.org** **DOI: 10.1161/CIRCGENETICS.111.962670**

Implications

This study has several important implications: First, the genome-wide annotation of the relatively uncharted noncoding genome is expected to aid in the interpretation of GWAS data and facilitate the functional exploration of how noncoding enhancer variants contribute to cardiac disease. Second, this study highlights the dynamic regulatory architecture of the genome and supports a remarkable heterogeneity of enhancers, characterized by profound species-, tissue-, and developmental-specific activity. As an indication of significant species-specific differences in enhancer sequences, neither direct comparison to mouse ChIP-Seq data nor computational methods based on comparative genomics and sequence conservation between mouse and human DNA were able to accurately predict more than 20–30% of human enhancers identified in the present study. This highlights that mouse enhancers may be of limited value to localize human enhancers and underscores the need for sequencing-based experiments directly from human myocardium. In addition, the differences between enhancers in fetal versus adult myocardium also suggest important implications for disease states, as heart failure is commonly characterized by a reversal to a fetal gene expression pattern. Thus, one can only speculate that substantial differences in enhancer activity also exist between the failing and nonfailing myocardium. Deciphering the role of enhancers in nonfailing and failing hearts will be the next step to understand the complexity of gene expression networks in the failing heart.

Acknowledgments

Dr Barth is a member of the Early Career Committee of the American Heart Association Functional Genomics and Translational Biology Council.

Disclosures

None.

Circulation
Cardiovascular Genetics

Filling in the Gaps: Deciphering the Function of Noncoding DNA

Andreas S. Barth

Circ Cardiovasc Genet. 2012;5:151-152
doi: 10.1161/CIRCGENETICS.111.962670
Circulation: Cardiovascular Genetics is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Print ISSN: 1942-325X. Online ISSN: 1942-3268