

# A genetic variant alters the secondary structure of the lncRNA H19 and is associated with Dilated Cardiomyopathy

Leonie Martens<sup>1</sup>, Frank Rühle<sup>1</sup>, Anika Witten<sup>1</sup>, Benjamin Meder<sup>3,4,5</sup>, Hugo A. Katus<sup>3,4</sup>, Eloisa Arbustini<sup>6</sup>, Gerd Hasenfuß<sup>7,8</sup>, Moritz F. Sinner<sup>9,10</sup>, Stefan Kääh<sup>9,10</sup>, Sabine Pankuweit<sup>11</sup>, Christiane Angermann<sup>12</sup>, Erich Bornberg-Bauer<sup>13</sup>, Monika Stoll<sup>1,2\*</sup>

<sup>1</sup>Department of Genetic Epidemiology, Institute of Human Genetics, University of Münster, Münster, Germany. <sup>2</sup>Department of Biochemistry, Genetic Epidemiology and Statistical Genetics, CARIM School for Cardiovascular Diseases, Maastricht Center for Systems Biology (MaCSBio), Maastricht University, Maastricht, The Netherlands. <sup>3</sup>German Centre for Cardiovascular Research (DZHK), partner site Heidelberg/Mannheim, Heidelberg, Germany. <sup>4</sup>Department of Cardiology, Heidelberg University, Heidelberg, Germany. <sup>5</sup>Genome Technology Center Stanford, Department of Genetics, Stanford University, United States. <sup>6</sup>Centre for Inherited Cardiovascular Diseases, IRCCS Foundation, University Hospital Policlinico San Matteo, Pavia, Italy. <sup>7</sup>Department of Cardiology and Pneumology, University Medical Center Göttingen, Robert-Koch-Str. 40, 37075, Göttingen, Germany. <sup>8</sup>German Center for Cardiovascular Research (DZHK) Partner Site, Göttingen, Germany. <sup>9</sup>Department of Medicine I, University Hospital Munich, Ludwig-Maximilian's University, Munich, Germany. <sup>10</sup>German Centre for Cardiovascular Research (DZHK), partner site: Munich Heart Alliance. <sup>11</sup>University Hospital Giessen and Marburg, Department of Cardiology, Marburg, Germany. <sup>12</sup>Comprehensive Heart Failure Center, University Hospital and University of Würzburg, Würzburg, Germany. <sup>13</sup>Institute for Evolution and Biodiversity, University of Münster, Münster, Germany.

## Abstract

lncRNAs are at the core of many regulatory processes and have also been recognized to be involved in various complex diseases. They affect gene regulation through direct interactions with RNA, DNA or proteins. Accordingly, lncRNAs structure is likely to be essential for their regulatory function. Point mutations, which manifest as SNPs (single nucleotide polymorphisms) in genome screens, can substantially alter their function and, subsequently, the expression of their down-stream regulated genes. To test the effect of SNPs on structure, we investigated lncRNAs associated with dilated cardiomyopathy. Among 322 hu-

man candidate lncRNAs we demonstrate first the significant association of a SNP located in lncRNA H19 using data from 1084 diseased and 751 control patients. H19 is generally highly expressed in the heart, with a complex expression pattern during heart development. Next, we used MFE (minimum free energy) folding to demonstrate a significant refolding in the secondary structure of this 861 nt long lncRNA. Since MFE folding may overlook the importance of sub-optimal structures, we showed that this refolding also manifests in the overall Boltzmann structure ensemble. There, the composition of structures is tremendously affected in their thermodynamic probabilities through the genetic variant. Finally, we confirmed these results experimentally, using SHAPE-Seq, corroborating that SNPs

\*To whom correspondence should be addressed.  
Email: mstoll@uni-muenster.de

affecting such structures may explain hidden genetic variance not accounted for through genome wide association studies. Our results suggest that structural changes in lncRNAs, and lncRNA H19 in particular, affect regulatory processes and represent optimal targets for further in-depth studies probing their molecular interactions.

**Keywords:** RNA structure, lncRNA, cardiovascular disease, RiboSNitch, Boltzmann ensemble, minimum free energy, SHAPE-Seq, H19, rs217727, SNP

## 1 Introduction

lncRNAs are defined as RNA molecules which are  $> 200$  nucleotides (nt) long and, usually, not translated into proteins [1]. lncRNAs have garnered much attention over the last decade since comparative eukaryotic genomics and transcriptome analyses demonstrated that lncRNAs are abundant and often play important roles for gene regulation in general [2, 3, 4]. So far, lncRNAs have been implicated in several biological and developmental functions through regulation of gene transcription, post-transcriptional RNA processing and chromatin modification [5, 6, 7, 8, 9]. Naturally, several lncRNAs have also been linked to various pathological processes [10, 11].

In the context of comparative human genomics, genome-wide association studies (GWAS) have provided a wealth of genetic associations for common complex diseases. Surprisingly, around one third of all SNPs in the most recent version of the GWAS catalog [12] are located in annotated lncRNAs. Accordingly, lncRNAs are emerging as prime candidates for explaining the missing heritability of complex traits. Several prominent lncRNAs have been described in the

context of the cardiovascular system and its diseases [13, 14, 15, 16, 17] which account for the highest mortality in the developed world, exceeding by far the annual deaths from cancer [18, 19]. However, in most cases, the functional and structural description of implicated lncRNAs is rudimentary at best and suitable functional assays are commonly lacking. Since lncRNAs do not code for a protein, their structure is assumed to be essential for their function. Certain RNAs can undergo strong structural changes upon single nucleotide substitutions [20, 21, 22, 23] and it is likely that a group of SNPs affect lncRNA function accordingly. Indeed, a recent study investigating the transcriptome of a family trio (father, mother, child)[22] reported that 15% of the transcribed SNPs altered RNA secondary structure. Furthermore, structure-altering SNPs were linked to altered gene expression and a set of disease phenotypes. These SNPs are designated as RiboSNitches [23]. So far, there are only few well characterized functional examples of structure altering SNPs in regulatory RNAs. One such example was described for a RiboSNitch which, in the wake of a SNP, can no longer attach to a regulatory protein (IREBP, Iron Response Element Binding Protein). As a consequence, the FTL gene, which encodes a subunit of the Ferritin complex, can no longer store excess iron which leads to hyperferritinemia cataract syndrome [24, 25], causing clouded eye lenses at early ages.

Unfortunately, predicting and determining RNA structures remains notoriously difficult and almost impossible for long molecules such as lncRNAs. However, a recent surge in experimental techniques coupled with next generation sequencing (NGS) has significantly improved predictions. These techniques comprise enzymatic and chemical approaches, depending on the reagent used to modify certain nucleotides before probing the struc-

ture. One prominent chemical approach is called SHAPE (selective 2-hydroxyl acylation by primer extension)-Seq [26], which binds to unpaired nucleotides and interrupts the reverse transcription process. Subsequent sequencing and analysis of the differences in read length compared to a control is used for computational structure prediction, significantly increasing the reliability [27].

Here, we set out to investigate the influence of dilated cardiomyopathy (DCM) associated RiboSNitches on lncRNA structure and function. First, we identified 14 disease associated SNPs located in several lncRNA transcripts. Next, we investigated them for their structure through MFE structure prediction algorithms. To further support these results, we also compared the change in the suboptimal structure ensemble. Here, the lncRNAs Carmn, H19 and MLIP-AS1 demonstrated that the according SNP had a massive effect on the relative probabilities of likely sub-optimal structures. Finally, we used SHAPE-Seq analyses to experimentally confirm the *in silico* predictions of these potential RiboSNitches.

Indeed, we found the SNP rs217727, located in the lncRNA H19, to be disease associated and inducing a remarkable shift in the structural ensemble. This may impair lncRNA function, and thus contribute to the pathogenesis of complex diseases such as heart failure.

## 2 Materials and Methods

### 2.1 Association Analysis

The samples for sequencing and genotyping were part of the German National Genome Research Network (NGFN) call "The genetics of heart failure - from populations to disease mechanisms". First, 96 DCM samples were screened for genetic variants in lncRNAs. Customized probes for the Illumina Nextera rapid

capture enrichment protocol were used for the lncRNA exonic regions. Reads were mapped to the hg19 build using BWA mem [28], followed by variant calling using GATK 3.7 [29]. Association analysis was then performed using Plink 1.7 [30] logistic regression assuming an additive model with sex as covariate and matching controls from the 1000 Genomes project[31]. Promising variants were subsequently confirmed by genotyping of a larger cohort comprising 1084 DCM cases and 751 disease-free controls. The study was conducted in accordance with the principles of the Declaration of Helsinki. All participants of the study gave written informed consent and the study was approved by the local ethic committees at the participating study centers.

### 2.2 Secondary Structure Predictions

RNA secondary structure prediction algorithms are based on thermodynamics and most often the minimum free energy (MFE) structure is determined. The MFE calculation is aiming to determine the predicted structure with the lowest free energy, since it is assumed that the lower the value, the more stable and likely the structure. For this, most algorithms are using the dynamic programming approach based on Zuker et al. [32]. Here, MFE structure predictions were performed using the ViennaRNA package RNAfold [33]. Since the MFE structure is not necessarily the only naturally occurring structure, we also considered suboptimal structures. For this, structures from the Boltzmann distribution were sampled with the ViennaRNA package RNAsubopt [33]. The resulting structures were then transformed into a binary vector, followed by t-SNE for visualization (according to [25, 34]).

### 2.3 In vitro SHAPE-Seq

Genetic variants that were significantly associated with our phenotype and were predicted to influence the structural properties of a lncRNA were subjected to validation through SHAPE-Seq structural probing. First, the RNA transcripts were generated in vitro from a ThermoFisher Gene Synthesis Plasmid with the T7 RNA polymerase. In vitro SHAPE-Seq was performed according to the protocol by Watters et al. [26] using the 1M7 reagent. The samples were sequenced 80 cycles in a paired-end mode with the NextSeq500 system and v2 chemistry. The resulting reads were analyzed with the accompanying Spats software for generating a reactivity profile. The resulting reactivities were then incorporated as an additional constraint during secondary structure prediction. RNAstructure [35] uses the approach first suggested by Deigan et al. [36] which transforms the reactivity information into pseudo-free energy using the equation  $\Delta G = m \ln[r+1] + b$ , where  $m$  and  $b$  are heuristically determined comparing probing data to crystallization structures, and  $r$  are the reactivity values. The SHAPE-probing information guided sampling of the structural ensemble was performed through the RNAstructure [35] pipeline consisting of Rsample, stochastic, and RsampleCluster. This pipeline is based on the algorithm by Ding and Lawrence [34, 37] for sampling and centroid determination and diana [38] followed by Calinski-Harabasz index [39] for finding the optimal  $k$  for clustering. In order to visualize the structures, the t-SNE [40] algorithm was utilized for dimensional reduction through the R package Rtsne [41].

## 3 Results

We found 14 associated SNPs ( $p \leq 0.05$ , coverage  $\geq 30$ ) that overlapped with lncRNA transcripts (limited to 1000 nt length). We

prioritized transcripts that were expressed in human myocardium samples (SRR1957191, SRR3151752, SRR3151758) [42, 43] and demonstrated a larger change in the structural ensemble. Genotyping of selected RiboSNitches (rs13158382 (Carmn), rs217727 (H19) and rs12527071 (MLIP-AS1)) in our confirmation case-control cohort, confirmed the association between DCM and rs217727, located in lncRNA H19.

### 3.1 Structural prediction

The MFE structure prediction of the H19 reference and rs217727 RiboSNitch sequences showed a local change in the structure located around the altered nucleotide (Figure 1).

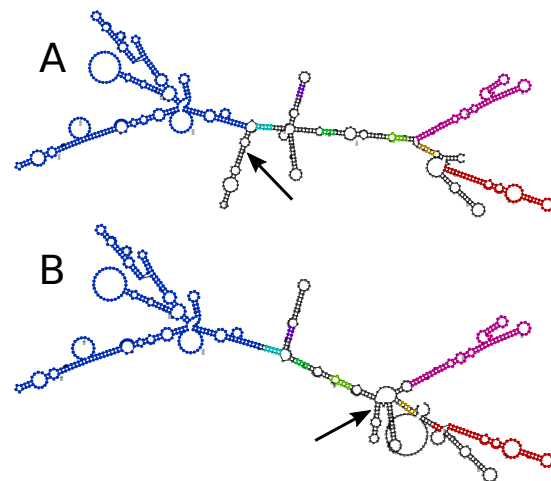


Figure 1: Minimum Free Energy (MFE) structure of the lncRNA H19. Secondary structure prediction was done using RNAfold. **A** is the reference structure and **B** is the structure changed through the potential RiboSNitch rs217727. Arrows indicate the changed nucleotides and identical colors highlight unaffected substructures. Reference structure free energy: -306.50 kcal/mol; RiboSNitch structure free energy: -308.50 kcal/mol.

Because most RNAs can assume multiple or

different structures from the MFE structure itself, these predictions are not completely accurate particularly for long RNA structures. Therefore, we considered the ensemble of suboptimal RNA structures. The entire ensemble of structures, sampled from the Boltzmann distribution, shows a clear shift, which is another indicator of a potential RiboSNitch (Figure 2).

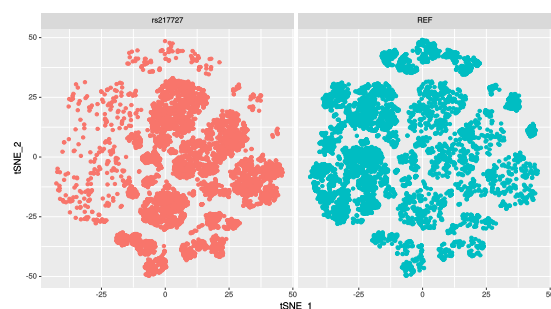


Figure 2: tSNE of suboptimal structures for H19 sampled from the Boltzmann ensemble ( $n=10000$ ). **Left** are the structures for the RNA with the genetic variant rs217727 and **Right** of the reference RNA. The shift shows that due to the genetic variant, other structures become more likely, which can lead to a change of function.

### 3.2 Structural probing

Although computational prediction of secondary structures has been widely used, the accuracy is often insufficient [44] and decreases for longer sequences. Therefore, experimentally derived data is needed to guide prediction algorithms. After successfully establishing and following the SHAPE-Seq protocol for H19 with the wildtype and the minor allele of rs217727, we received between 9 and 19 million paired-end reads. This was sufficient for the 861 nucleotides long sequence of lncRNA H19 to reach an average coverage of 5186.54 (min-

imum coverage of 106.5). These resulting reactivities were then incorporated during structure prediction. When assessing the accuracy of the predicted MFE structure compared to the SHAPE-Seq aided prediction through the RNAstructure [35] scorer function, we found very low accordance (sensitivity 45.95%; positive predictive value (PPV) 41.98%). As a next step, we investigated the structural ensemble, while taking the SHAPE-seq data into account. In Figure 3, suboptimal structures show clear differences between the reference structures and the RiboSNitch structures. Cluster analysis identified two and three clusters, respectively. Each cluster was assigned a centroid structure, which is the most representative structure.

## 4 DISCUSSION

Here, we investigated the influence of genetic variants on the abundant group of functional lncRNAs. In particular, we were interested in SNPs affecting secondary structure of lncRNAs, potentially contributing to disease. DCM associated SNPs were investigated with respect to their potential as RiboSNitches. Since computationally predicted structures alone are not sufficiently reliable, we combined sophisticated computational and experimental methods. We found that the variant rs217727, located in the lncRNA H19 was not only significantly associated with the DCM disease phenotype, but also highly affected RNA secondary structures. In a previous study, H19 together with the genetic variant rs217727 was associated with an increased risk of coronary artery disease in a Chinese population [45]. The lncRNA H19 is evolutionary conserved across several species, highly expressed in the adult heart and located in an imprinted gene cluster, which also includes genes like IGF2. Whereas H19 is solely



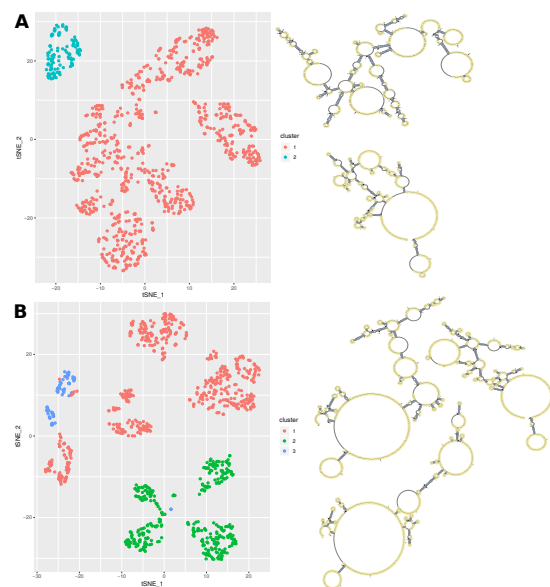


Figure 3: Visualization of the structural ensemble taking the SHAPE-seq data into account. For this 1000 structures were sampled from the Boltzmann distribution, that were colored by their optimal cluster affiliation. t-SNE was used for dimensionality reduction and plotted in 2D. **A** Cluster analysis of the reference structures reveals two distinct clusters with different centroid structures. **B** Cluster analysis of the RiboSNitch structures reveals three distinct clusters with different centroid structures.

expressed from the maternal allele, the growth factor IGF2 is expressed paternally. Despite their reciprocal expression, both genes are still able to regulate each other through the methylation machinery. For example, a mutation in the CTCF target site upstream of H19 has been shown to lead to loss of IGF2 imprinting [46] and such a loss of imprinting in murine embryos resulted in major cardiac abnormalities [47]. Furthermore, H19 has been found to be dysregulated in cardiac hypertrophy and heart failure [48, 49]. In particular, it was shown

that knock-down of H19 induced cardiomyocyte hypertrophy, whereas over-expression reversed the effect [49]. H19 can also function as a microRNA-sponge in important regulatory mechanisms involved in cardiac fibrosis. It is competing through binding of miR-455 with connective tissue growth factor (CTGF) and knockdown of H19 confirmed the antifibrotic influences of miR-455 and CTGF amongst others [50]. Thus, in case of a conformational change of the RNA due to a genetic variant, the ability of acting as a microRNA-sponge would be impaired and the expression of the microRNA would consequently increase. This is in accordance with a previous study investigating the miRNOME by deep sequencing of the human heart, where miR-445 was upregulated in DCM patients compared to healthy controls [51]. Another direct target of H19 in the development of fibrosis was identified by Tao et al. [52]. H19 negatively regulates DUSP5 through epigenetic mechanisms and decreased DUSP5 levels lead to an increase in proliferation of cardiac fibroblasts via the MEK/ERK signaling pathway. Additionally, H19 plays an important role in regulating apoptosis. Experiments in rat models for adriamycin-induced DCM showed that H19 is upregulated, whereas knockdown decreased cardiomyocyte apoptosis, resulting in improved left ventricular structure and function [53]. In contrast, Li et al. [54] reported that overexpression of H19 in diabetic cardiomyopathy improved left ventricular function by reducing oxidative stress, inflammation, and apoptosis. Fibrosis and apoptosis are relevant mechanisms in cardiac remodeling and are important processes with an impact on ventricular function in cardiac pathologic conditions [55]. In fact, genetic variants located in H19 and its host microRNA miR-675 have been shown to play a significant role in cardiac pathophysiology in previous studies, increasing the risk for developing hypertrophic cardiomyopathy [56] and coronary artery dis-

ease [45]. Since H19 is non-coding, the only way its variants may influence downstream regulation is through structural changes resulting in altered lncRNA function. For example, in rat models of diabetic cardiomyopathy, RNA-binding protein immunoprecipitation experiments were able to demonstrate the ability of H19 to bind directly to EZH2, which is part of the Polycomb Repressive Complex 2 (PRC2), affecting the epigenetic regulation of DIRAS3, a growth suppressor [57]. The murine and human H19 was shown to be able to act as an anti-hypertrophic lncRNA through interaction with PRC2, demonstrating its potential for gene therapy in cardiac hypertrophy [58]. In conclusion, we identified the secondary structure of H19, a lncRNA of importance for cardiovascular function and disease. Further we demonstrated, that a RiboSNitch, associated with the disease phenotype of DCM, significantly alters this structure. It is likely that through a structural change the essential region becomes inaccessible for binding interactions, potentially contributing to a pathological outcome. Further validation of the genetic variant in vivo demonstrating the altered lncRNA interactions in animal models would be of great interest.

## 5 Data Availability

All relevant data are accessible through BioProject ID PRJNA614999.

## 6 Acknowledgements

We would like to acknowledge the excellent technical assistance by the staff of the Core Facility Genomics of the Medical Faculty at the University Münster. This work was supported by the fund Innovative Medical Research of the University of Münster Medical School to F.R. (RÜ121510) and the Deutsche Forschungsge-

meinschaft (DFG, German Research Foundation) - 281125614/GRK 2220 to M.S.. B.M. is supported by an Excellence Fellowship of the Else Kröner Fresenius Foundation.

## Disclosure statement

The authors report no conflict of interest.

## References

- [1] Thomas Derrien, Rory Johnson, Giovanni Busotti, Andrea Tanzer, Sarah Djebali, Hagen Tilgner, Gregory Guernec, David Martin, Angelika Merkel, David G Knowles, et al. The gencode v7 catalog of human long noncoding rnas: analysis of their gene structure, evolution, and expression. *Genome research*, 22(9):1775–1789, 2012.
- [2] Matthew K Iyer, Yashar S Niknafs, Rohit Malik, Udit Singhal, Anirban Sahu, Yasuyuki Hosono, Terrence R Barrette, John R Prensner, Joseph R Evans, Shuang Zhao, et al. The landscape of long noncoding rnas in the human transcriptome. *Nature genetics*, 47(3):199, 2015.
- [3] Igor Ulitsky and David P Bartel. lincrnas: genomics, evolution, and mechanisms. *Cell*, 154(1):26–46, 2013.
- [4] Mitchell Guttman, Ido Amit, Manuel Garber, Courtney French, Michael F Lin, David Feldser, Maite Huarte, Or Zuk, Bryce W Carey, John P Cassidy, et al. Chromatin signature reveals over a thousand highly conserved large non-coding rnas in mammals. *Nature*, 458(7235):223–227, 2009.
- [5] Tim R Mercer, Marcel E Dinger, and John S Mattick. Long non-coding rnas: insights into functions. *Nature reviews genetics*, 10(3):155, 2009.
- [6] MD Ballantyne, RA McDonald, and AH Baker. lncrna/microrna interactions in the vasculature. *Clinical Pharmacology & Therapeutics*, 99(5):494–501, 2016.
- [7] Miao-Chih Tsai, Ohad Manor, Yue Wan, Nima Mosammaparast, Jordon K Wang, Fei Lan, Yang Shi, Eran Segal, and Howard Y Chang. Long non-coding rna as modular scaffold of histone modification complexes. *Science*, 329(5992):689–693, 2010.
- [8] Graeme D Penny, Graham F Kay, Steven A Sheardown, Sohaila Rastan, and Neil Brockdorff. Requirement for xist in x chromosome inactivation. *Nature*, 379(6561):131, 1996.

- [9] Radha Raman Pandey, Tanmoy Mondal, Faizaan Mohammad, Stefan Enroth, Lisa Redrup, Jan Komorowski, Takashi Nagano, Debora Mancini-DiNardo, and Chandrasekhar Kanduri. Kcnq1ot1 antisense noncoding rna mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Molecular cell*, 32(2):232–246, 2008.
- [10] Geng Chen, Ziyun Wang, Dongqing Wang, Chengxiang Qiu, Mingxi Liu, Xing Chen, Qipeng Zhang, Guiying Yan, and Qinghua Cui. Lncrnadisease: a database for long-non-coding rna-associated diseases. *Nucleic acids research*, 41(D1):D983–D986, 2012.
- [11] Rajnish A Gupta, Nilay Shah, Kevin C Wang, Jee-won Kim, Hugo M Horlings, David J Wong, Miao-Chih Tsai, Tiffany Hung, Pedram Argani, John L Rinn, et al. Long non-coding rna hotair reprograms chromatin state to promote cancer metastasis. *Nature*, 464(7291):1071–1076, 2010.
- [12] Annalisa Buniello, Jacqueline A L MacArthur, Maria Cerezo, Laura W Harris, James Hayhurst, Cinzia Malangone, Aoife McMahon, Joannella Morales, Edward Mountjoy, Elliot Sollis, et al. The nhgri-ebi gwas catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic acids research*, 47(D1):D1005–D1012, 2019.
- [13] Steffie Hermans-Beijnsberger, Marc van Bilsen, and Blanche Schroen. Long non-coding rnas in the failing heart and vasculature. *Non-coding RNA research*, 3(3):118–130, 2018.
- [14] Shizuka Uchida and Stefanie Dimmeler. Long non-coding rnas in cardiovascular diseases. *Circulation research*, 116(4):737–750, 2015.
- [15] Yvan Devaux, Jennifer Zangrando, Blanche Schroen, Esther E Creemers, Thierry Pedrazzini, Ching-Pin Chang, Gerald W Dorn II, Thomas Thum, Stephane Heymans, et al. Long noncoding rnas in cardiac development and ageing. *Nature Reviews Cardiology*, 12(7):415, 2015.
- [16] Kun Wang, Fang Liu, Lu-Yu Zhou, Bo Long, Shu-Min Yuan, Yin Wang, Cui-Yun Liu, Teng Sun, Xiao-Jie Zhang, and Pei-Feng Li. The long noncoding rna chrh regulates cardiac hypertrophy by targeting mir-489. *Circulation research*, 114(9):1377–1388, 2014.
- [17] JE Kontaraki, ME Marketou, GE Kochiadakis, Spyros Maragkoudakis, John Konstantinou, PE Vardas, and Fragiskos I Parthenakis. The long non-coding rnas mhrt, fendrr and carmen, their expression levels in peripheral blood mononuclear cells in patients with essential hypertension and their relation to heart hypertrophy. *Clinical and Experimental Pharmacology and Physiology*, 45(11):1213–1217, 2018.
- [18] Nick Townsend, Lauren Wilson, Prachi Bhatnagar, Kremlin Wickramasinghe, Mike Rayner, and Melanie Nichols. Cardiovascular disease in europe: epidemiological update 2016. *European heart journal*, 37(42):3232–3245, 2016.
- [19] Emelia J Benjamin, Paul Muntner, and Márcio Sommer Bittencourt. Heart disease and stroke statistics-2019 update: a report from the american heart association. *Circulation*, 139(10):e56–e528, 2019.
- [20] Xiaofei Yang, Minglei Yang, Hongjing Deng, and Yiliang Ding. New era of studying rna secondary structure and its influence on gene regulation in plants. *Frontiers in plant science*, 9:671, 2018.
- [21] Janine K Flores and Sandro F Ataíde. Structural changes of rna in complex with proteins in the srp. *Frontiers in molecular biosciences*, 5:7, 2018.
- [22] Yue Wan, Kun Qu, Qiangfeng Cliff Zhang, Ryan A Flynn, Ohad Manor, Zhengqing Ouyang, Jiajing Zhang, Robert C Spitale, Michael P Snyder, Eran Segal, et al. Landscape and variation of rna secondary structure across the human transcriptome. *Nature*, 505(7485):706–709, 2014.
- [23] Matthew Halvorsen, Joshua S Martin, Sam Broadaway, and Alain Laederach. Disease-associated mutations that alter the rna structural ensemble. *PLoS genetics*, 6(8):e1001074, 2010.
- [24] Amanda C Solem, Matthew Halvorsen, Silvia BV Ramos, and Alain Laederach. The potential of the ribosnitch in personalized medicine. *Wiley Interdisciplinary Reviews: RNA*, 6(5):517–532, 2015.
- [25] Joshua S Martin, Matthew Halvorsen, Lauren Davis-Neulander, Justin Ritz, Chetna Gopinath, Arthur Beauregard, and Alain Laederach. Structural effects of linkage disequilibrium on the transcriptome. *Rna*, 18(1):77–87, 2012.
- [26] Kyle E Watters and Julius B Lucks. Mapping rna structure in vitro with shape chemistry and next-generation sequencing (shape-seq). In *RNA Structure Determination*, pages 135–162. Springer, 2016.
- [27] David Loughrey, Kyle E Watters, Alexander H Settle, and Julius B Lucks. Shape-seq 2.0: systematic optimization and extension of high-throughput chemical probing of rna secondary structure with next generation sequencing. *Nucleic acids research*, 42(21):e165–e165, 2014.



- [28] Heng Li. Aligning sequence reads, clone sequences and assembly contigs with bwa-mem. *arXiv preprint arXiv:1303.3997*, 2013.
- [29] Aaron McKenna, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, David Altshuler, Stacey Gabriel, Mark Daly, et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation dna sequencing data. *Genome research*, 20(9):1297–1303, 2010.
- [30] Shaun Purcell, Benjamin Neale, Kathe Todd-Brown, Lori Thomas, Manuel AR Ferreira, David Bender, Julian Maller, Pamela Sklar, Paul IW De Bakker, Mark J Daly, et al. Plink: a tool set for whole-genome association and population-based linkage analyses. *The American journal of human genetics*, 81(3):559–575, 2007.
- [31] 1000 Genomes Project Consortium et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*, 491(7422):56, 2012.
- [32] Michael Zuker and Patrick Stiegler. Optimal computer folding of large rna sequences using thermodynamics and auxiliary information. *Nucleic acids research*, 9(1):133–148, 1981.
- [33] Ronny Lorenz, Stephan H Bernhart, Christian Höner Zu Siederdissen, Hakim Tafer, Christoph Flamm, Peter F Stadler, and Ivo L Hofacker. Viennarna package 2.0. *Algorithms for molecular biology*, 6(1):26, 2011.
- [34] YE Ding, Chi Yu Chan, and Charles E Lawrence. Rna secondary structure prediction by centroids in a boltzmann weighted ensemble. *Rna*, 11(8):1157–1166, 2005.
- [35] Jessica S Reuter and David H Mathews. Rnastructure: software for rna secondary structure prediction and analysis. *BMC bioinformatics*, 11(1):129, 2010.
- [36] Katherine E Deigan, Tian W Li, David H Mathews, and Kevin M Weeks. Accurate shape-directed rna structure determination. *Proceedings of the National Academy of Sciences*, 106(1):97–102, 2009.
- [37] Ye Ding and Charles E Lawrence. A statistical sampling algorithm for rna secondary structure prediction. *Nucleic acids research*, 31(24):7280–7301, 2003.
- [38] Martin Maechler, Peter Rousseeuw, Anja Struyf, Mia Hubert, Kurt Hornik, et al. Cluster: cluster analysis basics and extensions. *R package version*, 1(2):56, 2012.
- [39] Tadeusz Caliński and Jerzy Harabasz. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3(1):1–27, 1974.
- [40] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605, 2008.
- [41] Jesse H Krijthe. Rtsne: T-distributed stochastic neighbor embedding using barnes-hut implementation. *R package version 0.13*, URL <https://github.com/jkrijthe/Rtsne>, 2015.
- [42] Michael O Duff, Sara Olson, Xintao Wei, Sandra C Garrett, Ahmad Osman, Mohan Bolisetty, Alex Plocik, Susan E Celniker, and Brenton R Graveley. Genome-wide identification of zero nucleotide recursive splicing in drosophila. *Nature*, 521(7552):376–379, 2015.
- [43] Qiupeng Zheng, Chunyang Bao, Weijie Guo, Shuyi Li, Jie Chen, Bing Chen, Yanting Luo, Dongbin Lyu, Yan Li, Guohai Shi, et al. Circular rna profiling reveals an abundant circchipk3 that regulates cell growth by sponging multiple mirnas. *Nature communications*, 7(1):1–13, 2016.
- [44] Leonie Martens, Frank Rühle, and Monika Stoll. Lncrna secondary structure in the cardiovascular system. *Non-coding RNA research*, 2(3-4):137–142, 2017.
- [45] Wei Gao, Meng Zhu, Hao Wang, Shan Zhao, Di Zhao, Yang Yang, Ze-Mu Wang, Fang Wang, Zhi-Jian Yang, Xiang Lu, et al. Association of polymorphisms in long non-coding rna h19 with coronary artery disease risk in a chinese population. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 772:15–22, 2015.
- [46] Vinod Pant, Sreenivasulu Kurukuti, Elena Pugacheva, Shaharum Shamsuddin, Piero Mariano, Rainer Renkawitz, Elena Klenova, Victor Lobanov, and Rolf Ohlsson. Mutation of a single ctf target site within the h19 imprinting control region leads to loss of igf2 imprinting and complex patterns of de novo methylation upon maternal inheritance. *Molecular and cellular biology*, 24(8):3497–3504, 2004.
- [47] MM Lau, CE Stewart, Ziyang Liu, Harshida Bhatt, Peter Rotwein, and Colin L Stewart. Loss of the imprinted igf2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes & Development*, 8(24):2953–2963, 1994.
- [48] Simona Greco, Germana Zaccagnini, Alessandra Perfetti, Paola Fuschi, Rea Valaperta, Christine

- Voellenkle, Serenella Castelvechio, Carlo Gaetano, Nicoletta Finato, Antonio Paolo Beltrami, et al. Long noncoding rna dysregulation in ischemic heart failure. *Journal of translational medicine*, 14(1):183, 2016.
- [49] Lantao Liu, Xiangbo An, Zhenhua Li, Yao Song, Linling Li, Song Zuo, Nian Liu, Guan Yang, Haijing Wang, Xuan Cheng, et al. The h19 long non-coding rna is a novel negative regulator of cardiomyocyte hypertrophy. *Cardiovascular research*, 111(1):56–65, 2016.
- [50] Zhi-Wen Huang, Li-Hong Tian, Bin Yang, and Run-Min Guo. Long noncoding rna h19 acts as a competing endogenous rna to mediate ctgf expression by sponging mir-455 in cardiac fibrosis. *DNA and cell biology*, 36(9):759–766, 2017.
- [51] Stefanos Leptidis, Hamid el Azzouzi, Sjoukje I Lok, Roel de Weger, Serv Olieslagers, Natasja Kisters, Gustavo J Silva, Stephane Heymans, Edwin Cuppen, Eugene Berezikov, et al. A deep sequencing approach to uncover the mirnome in the human heart. *PloS one*, 8(2), 2013.
- [52] Hui Tao, Wei Cao, Jing-Jing Yang, Kai-Hu Shi, Xiao Zhou, Li-Ping Liu, and Jun Li. Long noncoding rna h19 controls dusp5/erk1/2 axis in cardiac fibroblast proliferation and fibrosis. *Cardiovascular Pathology*, 25(5):381–389, 2016.
- [53] Yanlin Zhang, Mengyao Zhang, Weiting Xu, Jianchang Chen, and Xiang Zhou. The long non-coding rna h19 promotes cardiomyocyte apoptosis in dilated cardiomyopathy. *Oncotarget*, 8(17):28588, 2017.
- [54] Xiangquan Li, Hao Wang, Biao Yao, Weiting Xu, Jianchang Chen, and Xiang Zhou. lncrna h19/mir-675 axis regulates cardiomyocyte apoptosis by targeting vdac1 in diabetic cardiomyopathy. *Scientific reports*, 6:36340, 2016.
- [55] A Piek, RA De Boer, and HHW Silljé. The fibrosis-cell death axis in heart failure. *Heart failure reviews*, 21(2):199–211, 2016.
- [56] Juan Gómez, Rebeca Lorca, Julián R Reguero, María Martín, César Morís, Belén Alonso, Sara Iglesias, Beatriz Díaz-Molina, Pablo Avanzas, and Eliecer Coto. Genetic variation at the long non-coding rna h19 gene is associated with the risk of hypertrophic cardiomyopathy. *Epigenomics*, 10(7):865–873, 2018.
- [57] Chuanjun Zhuo, Ronghuan Jiang, Xiaodong Lin, and Mingjing Shao. Lncrna h19 inhibits autophagy by epigenetically silencing of diras3 in diabetic cardiomyopathy. *Oncotarget*, 8(1):1429, 2017.
- [58] Janika Viereck, Anne Bürke, Ariana Foinquinos, Shambhabi Chatterjee, Jan A Kleeberger, Ke Xiao, Heike Janssen-Peters, Sandor Batkai, Deepak Ramanujam, Theresia Kraft, et al. Targeting muscle-enriched long non-coding rna h19 reverses pathological cardiac hypertrophy. *European Heart Journal*, 2020.