Journal of Molecular and Cellular Cardiology xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Journal of Molecular and Cellular Cardiology

journal homepage: www.elsevier.com/locate/yjmcc

Review article

Bioinformatics of cardiovascular miRNA biology

Meik Kunz [a,b], Ke Xiao [b,c], Chunguang Liang [a], Janika Viereck [b], Christina Pachel [d], Stefan Frantz [d], Thomas Thum [b,e,f], Thomas Dandekar [a,g]

[a] Functional Genomics and Systems Biology Group, Department of Bioinformatics, Biocenter, Würzburg, Germany
[b] Institute for Molecular and Translational Therapeutic Strategies (IMTTS), Hannover Medical School, Hannover, Germany
[c] Plant Breeding Institute, Christian-Albrechts-University of Kiel, Olshausenstr. 40, 24098 Kiel, Germany
[d] Department of Internal Medicine I, University Hospital Würzburg, Germany and Comprehensive Heart Failure Center, University of Würzburg, Germany
[e] Excellence Cluster REBIRTH, Hannover Medical School, Hannover, Germany
[f] National Heart and Lung Institute, Imperial College London, London, UK
[g] EMBL Heidelberg, BioComputing Unit, Meyerhofstraße 1, 69117 Heidelberg, Germany

ARTICLE INFO

Article history:
Received 3 September 2014
Received in revised form 5 November 2014
Accepted 29 November 2014
Available online xxxx

Keywords:
MiRNAs
Multiple sequence alignment
RNA secondary structure
Interaction
Database
Omics
Heart

ABSTRACT

MicroRNAs (miRNAs) are small ~22 nucleotide non-coding RNAs and are highly conserved among species. Moreover, miRNAs regulate gene expression of a large number of genes associated with important biological functions and signaling pathways. Recently, several miRNAs have been found to be associated with cardiovascular diseases. Thus, investigating the complex regulatory effect of miRNAs may lead to a better understanding of their functional role in the heart. To achieve this, bioinformatics approaches have to be coupled with validation and screening experiments to understand the complex interactions of miRNAs with the genome. This will boost the subsequent development of diagnostic markers and our understanding of the physiological and therapeutic role of miRNAs in cardiac remodeling. In this review, we focus on and explain different bioinformatics strategies and algorithms for the identification and analysis of miRNAs and their regulatory elements to better understand cardiac miRNA biology. Starting with the biogenesis of miRNAs, we present approaches such as LocARNA and miRBase for combining sequence and structure analysis including phylogenetic comparisons as well as detailed analysis of RNA folding patterns, functional target prediction, signaling pathway as well as functional analysis. We also show how far bioinformatics helps to tackle the unprecedented level of complexity and systemic effects by miRNA, underlining the strong therapeutic potential of miRNA and miRNA target structures in cardiovascular disease. In addition, we discuss drawbacks and limitations of bioinformatics algorithms and the necessity of experimental approaches for miRNA target identification. This article is part of a Special Issue entitled 'Non-coding RNAs'.

Contents

1. Introduction

MicroRNAs (miRNAs) are highly conserved among different species. They are small ~22 nucleotide non-coding RNAs [9,18,30,46,79,81]. They have been found to regulate gene expression of a large number of human genes by binding to the 3′-untranslated region (3′-UTR) of messenger RNAs (mRNAs) and also influence protein synthesis through interacting with the protein translation machinery. MiRNAs

http://dx.doi.org/10.1016/j.yjmcc.2014.11.027
0022-2828/

are associated with many biological processes and diseases, including aging, cardiac function, metabolism and cancer [8,9,15,18,31,46,65,79]. Moreover, a single miRNA can target different mRNAs and a single mRNA can also be regulated by different miRNAs [14,31,57,65], pointing to a complex regulatory network. MiRNAs influence different signaling pathways and are useful as diagnostic markers as well as potential new therapeutic targets for cardiovascular diseases [15]. Cardiovascular diseases combine together to be the leading cause of death [15]. Several miRNAs have been known to be involved in cardiovascular diseases and also play a potential therapeutic role in cardiac remodeling (139 cardiac-related miRNAs and their role in cardiovascular diseases extensively reviewed in [22,52]). Specific miRNAs are not only deregulated in various cardiovascular cell types of diseased hearts [77], but also directly involved in pathologic reactions of the heart. For example, miRNA-1 is associated with myocardial infarction and miRNA-21 and miRNA-212/132 with cardiac fibrosis and hypertrophy respectively (extensively reviewed in [15,24,78,80]). Importantly, some miRNAs are transcribed as part of a cistron and regulated by cardiac transcription factors (TFs), *e.g.* miR-1/miR-133 by myogenic transcription factor (MyoD) and serum response factor (SRF) or miR-143/145 by cardiac NK-2 transcription factor (Nkx2–5) and SRF [65,76]. On the other hand, miRNAs can directly regulate cardiac associated TF and signaling pathways, *e.g.* miR-212/132, the anti-hypertrophic TF forkhead box O3 and the CN–NFAT signaling pathway [80]. To understand the complex effects of cardiovascular miRNAs their genomic localization including promoter analysis as well as interaction partners all have to be taken into account. It is thus of high interest to understand the complex role and function of cardiovascular miRNAs for a better understanding of their regulatory effects as a basis for future therapeutic approaches. For this purpose, different bioinformatics methods and search programs are useful. Owing to their small length as well as their specific cardiovascular expression profiles (cell type and development dependent), experimental methods alone cannot fulfill the detection and analysis of these miRNAs, *e.g.* regarding cardiac miRNAs with low expression levels or detecting sequence–structure-conservation [3,31,49,63]. For this, the combined use of experimental and computational approaches has revolutionized the identification and analysis of miRNAs, and in particular, their selective function in the heart.

2. Biogenesis, structure and miRNA biology

MiRNAs are located either in intronic regions of coding-genes, in non-coding genes or in intragenic regions of the genome. They are transcribed by RNA-Polymerase II (RNA-Pol II) as a primary-miRNA transcript (pri-miRNA) [31,50,53,55,65]. The pre-miRNA contains a characteristic hairpin structure, which is recognized by the RNase III enzyme Drosha. By binding of the RNase III enzyme Drosha and its cofactor DiGeorge Syndrome Critical Region 8 (DGCR8), a dsRNA binding protein, the ~70 nucleotide long hairpin precursor-miRNA (pre-miRNA) is processed and further transported into the cytoplasm by the nucleocytoplasmic shuttle protein Exportin 5 [15,31,65]. In the cytoplasm, another RNase III enzyme, Dicer, cleaves and unwinds the pre-miRNA to form the ~22 bp double stranded miRNA [31,65]. Finally, this miRNA duplex forms a single-stranded guiding RNA (mature miRNA), which associates with the RNA-induced silencing complex (RISC) to regulate its mRNA targets, whereas the second single-stranded passenger RNA strand is mostly degraded [15,31,32,56,65]. In general, miRNAs regulate the gene expression by binding to the 3′-UTR of mRNAs, whereas few studies also reported that they also bind to the coding region or 5′-UTR [31,54,65]. As a result of a complete or incomplete complementary binding, a single miRNA can target multiple mRNAs and a single mRNA may be regulated by multiple miRNAs [14,57,65,31]. This clarifies their complex regulatory effect and high targeting potential (about 30% to 60%) of mammalian genes [30,65], at the same time pointing out their potentially important therapeutic role. However, the experimental identification of miRNA targets is very complex and elaborate, indicating the necessity of computational prediction tools. As a result, several computational tools were developed, mainly using miRNA length, sequence and structural information (*e.g.* hairpin structure and minimal folding free energy; [31,58]), which are very efficient in the identification and analysis of miRNAs. Algorithms such as RNAfold and Mfold quickly and accurately predict the putative secondary structure of an miRNA based on the principle of minimum free energy and are used in different computational tools [31,37,90]. MiRNA detection tools can be divided into comparative and non-comparative methods [10,31,36]. Comparative algorithms use the sequence conservation for the miRNA prediction and help to identify miRNAs among species, whereas non-comparative algorithms only use the intrinsic miRNA structure without any sequence conservation and are therefore able to identify evolutionarily distant species or species-specific miRNAs (reviewed in [31]) (Fig. 1).

3. Novel miRNA discovery through NGS platforms and experimental identification of miRNA targets

3.1. Novel miRNA discovery

Microarrays have been widely and extensively used as an efficient method for miRNA expression profiling on a genome-wide level. However, the discovery of novel miRNAs is still an inherent weakness of this hybridization-based technology. The short length of miRNA and the high similarity between miRNA family members make specific probe design for microarrays challenging.

The development of next-generation sequencing technologies and the drop of costs in recent years open an efficient route for the rapid discovery of novel or low-expressed miRNAs. The miRDeep algorithm [29] was first introduced in 2008 and is currently widely used to detect and quantify miRNA from small RNA sequencing. This tool has been further developed as an integrated program named miRDeep* [4] which is freely available with a user-friendly interface. Sequence reads archived in FastQ and alignment profiles in BAM/SAM format can be used directly for further analysis such as miRNA detection and expression profiling. A java-based sRNA analysis tool is the UEA sRNA workbench [74]. It provides biologists with an easy solution, with a nice graphical user interface, for handling their RNA-seq data, starting from quality filters for the reads to target predictions. A major advantage of high throughput RNA sequencing in cardiovascular disease is the unambiguous and sensitive detection ability for novel miRNAs. On the other hand, deep sequencing is a comparatively new approach and no standard data analysis strategy has been suggested. Furthermore, substantial computational support is necessary for a more precise prediction and expression quantification.

3.2. Experimental miRNA target identification

The experimental identification of miRNA targets can be done by expression profiling (*e.g.* microarray analysis after miRNA overexpression or knockdown, proteomics) or biochemical isolation of the miRISC complex using immunoprecipitation (different experimental methods extensively reviewed in [75,87]). Generally, miRNAs regulating the mRNA level and miRNA and mRNA expression are often negatively correlated [28,31,33]. Therefore, miRNA profiling microarray experiments to identify deregulated mRNAs after ectopic miRNA expression or antagonism are useful for experimental identification of putative miRNA targets and can be further combined with bioinformatics [75]. However, results from Matkovich et al. show for cardiovascular research that compared to mRNAs, cardiac miRNAs are more sensitive to the acute functional status of end-stage heart failure [61], exemplifying that changes in mRNA level are not always a reliable method for miRNA target prediction [87]. Therefore, different additional techniques for miRNA target identification in cardiac tissues are available, such as RISC-IP and proteomics. The RISC-IP is a new biochemical method for

Please cite this article as: Kunz M, et al, Bioinformatics of cardiovascular miRNA biology, J Mol Cell Cardiol (2014), http://dx.doi.org/10.1016/j.yjmcc.2014.11.027