Biochemical and Biophysical Research Communications 379 (2009) 702–705

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

Recombination rates of human microRNA

Huizhi Zhao [a,b,1,2,3], Dong Wang [b,c,1,2], Bing Liu [a,3], Xingpeng Jiang [a,3], Jing Zhang [a,3], Ming Fan [a,3], Zhengjie Fan [a,3], Ying Chen [a,3], Sonya Wei Song [a,3], Wei Gao [b,c,2], Tianzi Jiang [a,*,3], Qinghua Cui [b,c,*,2]

[a] LIAMA Center for Computational Medicine and National Lab of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] Department of Medical Informatics, Peking University Health Science Center, 38 Xueyuan Rd., Beijing 100083, China
[c] MOE key Lab of Molecular Cardiology, Peking University, Beijing 100191, China

ARTICLE INFO

Article history:
Received 8 December 2008
Available online 6 January 2009

Keywords:
MicroRNAs
Recombination rate
Tissue specificity
Disease
MicroRNA expression level

ABSTRACT

The fact that microRNAs play a role in almost all biological processes is well established, as is the importance of recombination in generating genome variability. However, the association between microRNAs and recombination remains largely unknown. In order to investigate the recombination patterns of microRNAs, we performed a comprehensive analysis of the recombination rate of human microRNAs. We observed that microRNAs that are expressed in several tissues tend to have lower recombination rates than tissue-specific microRNAs. Additionally, microRNAs that are associated with a number of diseases are also likely to have lower recombination rates. Furthermore, microRNAs with higher expression levels are found to have fewer recombination events. These findings reveal patterns in recombination rates of microRNAs that could help in understanding the function, evolution, and disease-related roles of microRNAs.

MicroRNAs (miRNAs) are a class of ~22 nucleotide non-coding RNAs. It is estimated that 1–4% of the genes in the human genome code for miRNAs [1]. A single miRNA can regulate as many as 200 targets [1]. MiRNAs are important gene regulators at the post-transcriptional level in nearly all critical biological processes, such as cell growth, proliferation, differentiation, development, and apoptosis [2] and human signaling network [3]. Studies have revealed that miRNAs have also been associated with differences in gene expression between species [4], and are involved in various diseases [5]. Because of their importance, it is important to investigate miRNAs from many perspectives. For example, Liang et al. measured the expression profiles of 345 miRNAs from 40 normal human tissues and provided a global view of tissue distributions of the miRNAs [6]; Lu et al. reported that the SNP density of miRNAs can be associated with various diseases [5]; Zhang et al. found an increased SNP density in human-specific miRNAs and interpreted this as an indication of evolutionary acceleration in miRNA regions [7]. Borenstein and Ruppin and Shu et al. studied the evolution of the genetic robustness of miRNAs by analyzing the secondary structures of different species [8,9]; Davis et al. investigated the role of GC content on the prediction of the number of miRNA binding sites [10].

Recombination exchanges genetic information between homologous chromosomes during prophase I of meiosis [11]. This process ensures the proper segregation of chromosomes and plays an important role in shaping patterns of genetic diversity in populations [12]. In general, the recombination rate is substantially different between sexes and is often suppressed near centromeres and elevated near telomeres, but neither of these observations is true for all chromosomes [12]. Several studies have reported that the recombination rate varies greatly in different genomic regions and is associated with genome component features such as GC content or CpG frequency [11–15]. A few studies have also investigated the association between the recombination rate and certain external phenotypes of genes. For example, Nachman suggested that markers in regions with a low recombination rate might be non-randomly associated with disease susceptibility genes [12]. Boyle and Noor discussed the impact of variations in recombination rates across the genome on human genetic disease [16]. Kato et al. performed microarray experiments to identify the tissue specificity of human genes and investigated trends in recombination rates in tissue-specific genes [17]. However, the recombination rate of the genes that code for miRNAs can also be associated with genetic features remains unknown. Finding this information is important for better understanding miRNAs. Therefore, in this study, in order to understand the recombination of miRNA-encoding genes (miRNA genes), we performed a comprehensive analysis of recombination rates and human miRNAs.

* Corresponding authors.
E-mail addresses: jiangtz@nlpr.ia.ac.cn (T. Jiang), cuiqinghua@bjmu.edu.cn (Q. Cui).
[1] These authors contributed equally to this work.
[2] Fax: +86 10 82801001.
[3] Fax: +86 10 62551993.

0006-291X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2008.12.144

Methods and results

The recombination rate of miRNAs

We obtained the chromosome position data for the genes that encode for 701 human miRNAs from miRBase [18]. We downloaded datasets containing recombination rates (release 22, build 36) from the HapMap Project [19,20]. We defined the recombination rate of an miRNA as the recombination rate of the gene region in which the miRNA gene is located (Supplementary File).

The recombination rate of miRNA genes and their neighboring regions

The recombination rate varies greatly in different genomic regions [14]. However, functional regions, such as protein-coding genes, are reported to have lower recombination rates [11]. However, it is unknown how the recombination rates of miRNA genes vary across the genome. We hypothesize that, because miRNAs are important functional elements, the recombination rates of miRNA genes may be lower than those of randomly selected segments from neighboring regions. In order to discern the recombination patterns of miRNAs, we investigated the local recombination pattern around miRNA genes. We compared the recombination rates of miRNA genes with those of randomly selected segments from five upstream and five downstream segments. The result was that no significant differences were found between miRNA genes and their neighboring regions (Wilcoxon Test, data not shown). Two things could potential cause this. First, miRNA genes are very short, but the recombination segment provided by the HapMap database is much longer. Therefore, calculations of recombination rates of miRNA genes using long segments that contain these shorter miRNA genes may not be accurate. Second, recombination rate patterns vary greatly across the genome. Local recombination rate patterns differ for different miRNA genes and their neighboring regions, making it difficult to find a consistent result for all miRNA genes. In the future as more accurate descriptions of the characteristics of recombination rate patterns are revealed, we can further investigate the distributions of the recombination rates of miRNA genes.

Recombination rates and tissue specificity of miRNAs

Tissue specificity can provide important clues to the physiological and pathological functions of protein-coding genes according to Kato et al. [17], who used protein-coding gene analysis to suggest that it is possible that natural selection forms recombination rate tendencies according to the biological functions of specific tissues. However, it is unclear whether tissue specificity and recombination rates are associated for miRNAs genes. In order to investigate this, we obtained the tissue specificity index value for miRNAs [5], which is calculated based on miRNA profiles across 40 human tissues [6]. We first classified 351 miRNAs into two groups: tissue-specific miRNAs and non-tissue-specific miRNAs according to their specificity index values. MiRNAs with tissue specificity index values larger than or equal to 0.7 were regarded as tissue-specific miRNAs, of which there were 103; the rest were non-tissue-specific ones (248). We next investigated the difference between the recombination rates of the 103 tissue-specific miRNA genes and those of the 248 non-tissue-specific miRNA genes. We found that the recombination rates of the tissue-specific miRNA genes were higher than those of the non-tissue-specific ones ($P = 1.228 \times 10^{-6}$, Wilcoxon Test).

We then performed a correlation analysis between the tissue specificity index values and the recombination rates of the miRNA genes. These two factors were found to be positively correlated

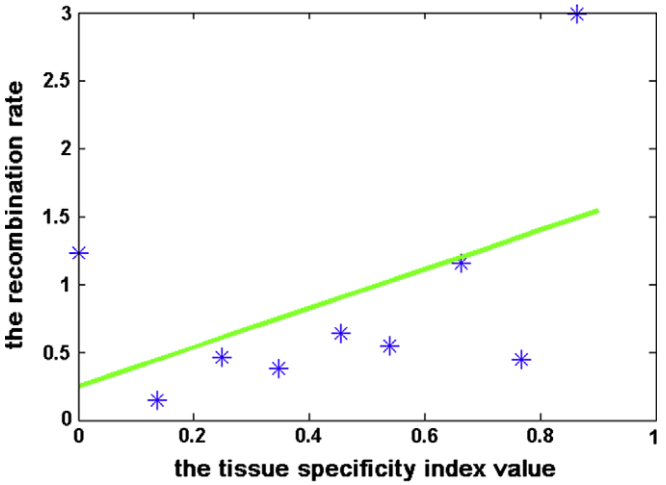

Fig. 1. The correlation between the recombination rate of miRNA genes and the tissue specificity of miRNAs. We further classified miRNAs into nine groups according to their tissue specificity index values. Each miRNA was assigned to one of the tissue specificity intervals. Each star represents a miRNA group whose *x* and *y* coordinates are the median tissue specificity index value and the median recombination rate of the group. The line is a linear least-squares fit of the two factors.

(Fig. 1, $R = 0.295$, $P = 1.742 \times 10^{-8}$, Spearman's Correlation). This correlation may imply that organisms are intolerant of changes in genes that code for miRNAs that are expressed in a number of tissues because of the possibility of disrupting some finely tuned function or functions. On the other hand, tissue-specific miRNAs may be under relatively fewer constraints and thus be able to exchange genetic information between homologous chromosomes more freely allowing them to be potentially able to adapt in changing environments. As a result, genes for miRNAs that are expressed in multiple tissues are likely to be less active during recombination and have lower degrees of recombination rate than those for tissue-specific miRNAs. From Fig. 1, we can see that the first point, the totally non-tissue-specific miRNA groups (tissue-specific index value = 0) and the last point, the most tissue-specific miRNA group (tissue-specific index value > 0.8) departed from the other points. These deviations from expectations may possibly be caused by the incompleteness of the tissue specificity data of the miRNAs. The miRNAs that are completely non-tissue specific may not express in tissues that we have not investigated, and therefore may be less tissue non-specific than we are aware of. The most tissue-specific miRNA groups may also express in some tissues that we did not identify. This may have caused the two ends of the line to vary from the other points.

The recombination rate and diseases associated with miRNAs

MiRNAs have been demonstrated to play critical roles in a number of diseases [5,21,22]. However, it remains unknown whether the recombination rate of miRNA genes can be associated with diseases that are known to be related to miRNAs. In order to answer this question, we performed an association analysis between the gene recombination rates and the numbers of associated diseases related to miRNAs. We downloaded miRNA and disease association data in which 210 miRNAs are reported to be associated with at least one disease from the HMDD database [5]. The remaining 491 miRNAs on the site have not been reported to be associated with any disease (called non-disease associated miRNAs here).

We grouped the miRNAs according to the numbers of associated diseases and measured the correlation between the gene recombination rates and the associated diseases of the groups. We used a