Gene 534 (2014) 24–32

Contents lists available at ScienceDirect

Gene

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

Integrated evolutionary analysis of human miRNA gene clusters and families implicates evolutionary relationships

Li Guo, Yang Zhao, Hui Zhang, Sheng Yang, Feng Chen *

Department of Epidemiology and Biostatistics, Ministry of Education Key Lab for Modern Toxicology, School of Public Health, Nanjing Medical University, 211166, China

ARTICLE INFO

Article history:
Accepted 15 October 2013
Available online 24 October 2013

Keywords:
Integrated analysis
microRNA (miRNA)
miRNA gene cluster/family
Evolution

ABSTRACT

Many microRNAs (miRNAs) are clustered on chromosomes and co-transcribed as polycistronic transcripts. Here, an integrated evolutionary analysis of human miRNA gene clusters and families was performed. Generally, miRNA gene clusters include 2–8 members, but some larger clusters have been found to have more members (over 40 miRNAs). 62.22% of them have been shown to be involved in homologous miRNA genes, including multicopy pre-miRNAs and sense/antisense homologous miRNAs. Multicopy pre-miRNAs can enrich the distribution and relationship between miRNA clusters and families. An miRNA family may be located in one or more clusters, and a cluster may be involved in one or more families. Members of different families have been shown to be prone to appear in clusters, and vice versa. Reconstructed phylogenetic trees and networks may indicate potential evolutionary relationships, which also indicate duplication history in specific related gene clusters and families. Related miRNA families are always found to share common target mRNAs and biological pathways. Some clusters containing non-homologous miRNAs also tend to be clustered together as well as homologous miRNAs. In the present work, it is shown that homologous miRNAs are prone to appear in clusters based on functional and evolutionary pressures. The phenomenon of miRNA clusters containing homologous or genetic relationships is quite common. The integrative evolutionary analysis will provide more potential evolutionary and functional relationships between homologous and clustered miRNAs.

1. Introduction

MicroRNAs (miRNAs) are a class of small negative non-coding RNA (ncRNA) regulatory molecules. They play important biological roles via negatively regulating gene expression and translation processes (Bartel, 2004). The short ncRNAs tend to be extremely well conserved and can be found in genomic sequence datasets (Berezikov et al., 2005). They guide animal development through playing an evolutionary role (Grimson et al., 2008; Liu et al., 2008; Niwa and Slack, 2007; Sempere et al., 2006). miRNAs are widely studied especially because of their contributions to cancer development, and they may be novel biomarkers for diagnosis of cancer and other diseases (Cho, 2010; Wang et al., 2009).

miRNAs are not randomly distributed on chromosomes, and they are prone to cluster in a single polycistronic transcript (Lagos-Quintana et al., 2003; Lai et al., 2003; Lee et al., 2002; Mourelatos et al., 2002). They may be co-expressed and may play similar roles in the same biological processes, often through coordinate regulation of those processes (Bashirullah et al., 2003; Baskerville and Bartel, 2005; Seitz et al., 2004). However, they always have different levels of enrichment because of complex maturation and degradation processes (Guo and Lu, 2010; Viswanathan et al., 2009; Yu et al., 2006), even when they are transcribed at equal rates via co-transcription. Some clustered miRNAs share very similar sequences, and they are identified as members of the same miRNA gene family (Aravin et al., 2003). The phenomenon further complicates distribution of miRNA gene clusters and families. Clustered and homologous miRNAs may have evolved from the same ancestral gene through complex duplication processes, perhaps even genome-wide duplication history (Guo et al., 2009; Heimberg et al., 2008; Hertel et al., 2006; Sun et al., 2013; Zhang et al., 2007). The potential functional relationships among these miRNAs have attracted considerable attention. Increasing number of reports indicates that many miRNA gene clusters and families have important roles in cancer development. For example, the miR-17-92 cluster has been shown to contain potential oncomiRs and contribute to the occurrence and development of multiple human cancers (Cho, 2007; Concepcion et al., 2012; Olive et al., 2010). Although the expression and evolutionary patterns of these miRNAs are widely studied, the evolutionary relationships between miRNA gene clusters and families remain largely unexamined. miRNA gene cluster is defined if two or more miRNAs have close physical distance (such as less than10 kb), and miRNA gene family is defined if two or more miRNAs have higher sequence similarity. According to the location distribution and sequence similarity, these miRNA groups may also have evolutionary and functional relationships. It is therefore quite necessary to assess the distribution and evolutionary patterns between these different miRNAs.

Abbreviations: miRNA, microRNA; ncRNA, non-coding RNA; chr, chromosome; pre-miRNA, precursors miRNA; NJ, neighbor-joining; MJ, median-joining; GO, gene ontology.

* Corresponding author.
E-mail addresses: gl8008@163.com, lguo@njmu.edu.cn (L. Guo), fengchen@njmu.edu.cn (F. Chen).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.10.037

In order to identify the functional and evolutionary relationships among these miRNAs, an integrated analysis of human miRNA gene clusters and families was performed. These results provide data regarding the potential genetic, evolutionary, and functional relationships between miRNA gene clusters and families.

2. Results

2.1. Overview of human gene clusters and families

One hundred and thirty-five human miRNA gene clusters were found to be involved in 422 pre-miRNA sequences (http://www.mirbase.org/cgi-bin/mirna_summary.pl?org=hsa&cluster=10000). These miRNAs were more prone to cluster on chromosomes (chr) 8, 17 and X than most miRNAs (Figs. 1A and B). Although the numbers of clusters and known miRNAs showed moderate relationships, they also showed different distributions. These miRNA clusters always had 2–8 members, and 68.49% of them were composed of two miRNA genes (Fig. 1B). However, the two special larger clusters (mir-379 cluster and mir-512-1 cluster), located on chromosomes 14 and 19, were found to involve 42 and 46 members, respectively.

We found that 62.22% of the clusters were involved in homologous miRNAs (Fig. 1C). Of these, 19.05% were composed of multicopy miRNA genes. The multicopy pre-miRNAs could yield to the same mature miRNAs, although they might have different sequences and might be located on different genomic regions. Generally, these multicopy clusters

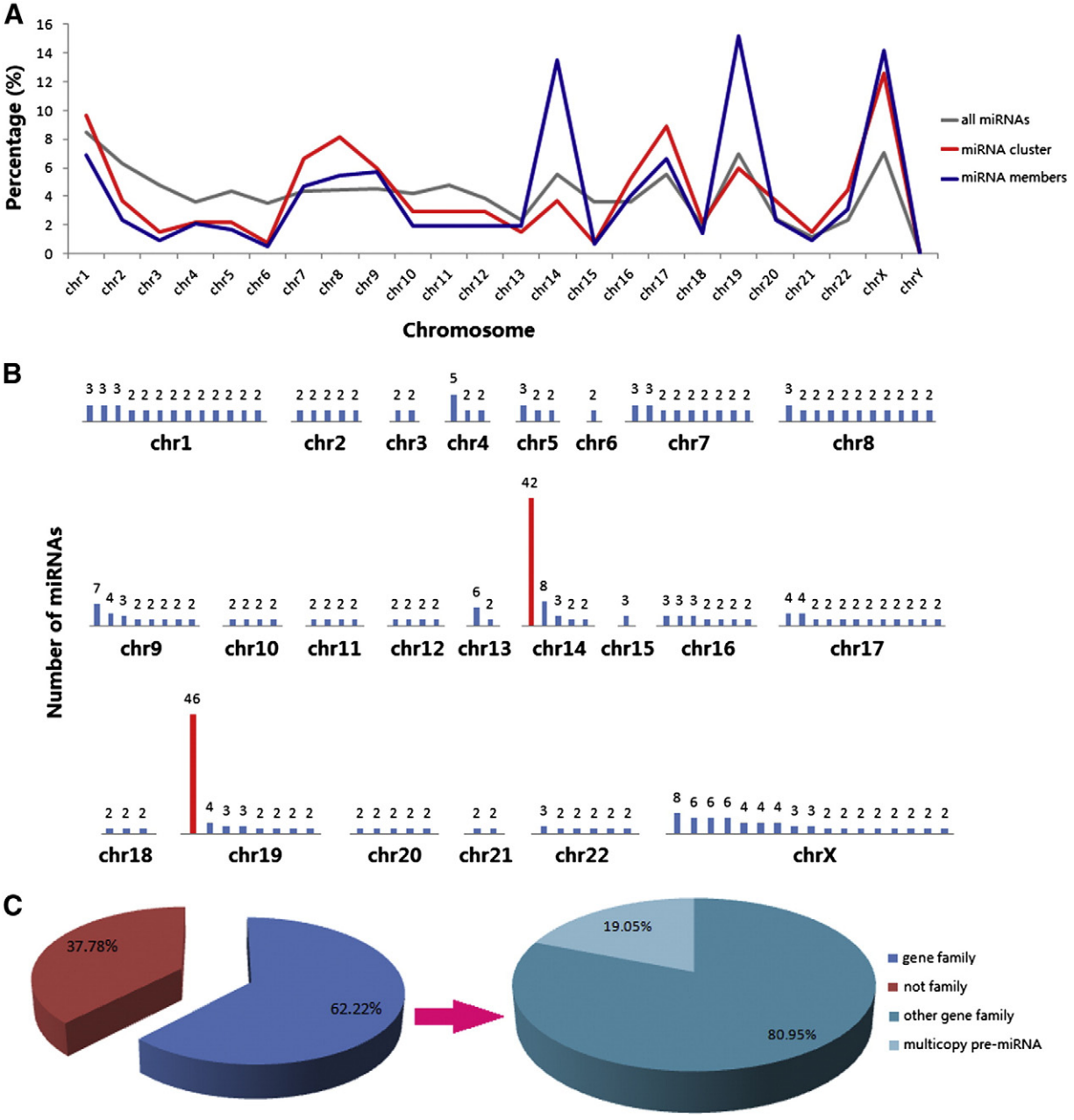

Fig. 1. Human miRNA gene clusters and related gene families. (A) Distributions of all known miRNAs (miRNAs), miRNA gene clusters (miRNA clusters), and members of those families (miRNA members) on human chromosomes. Different distributions are detected. (B) The details of the distributions of the clusters and the number of miRNAs on each chromosome. The larger miRNA gene clusters (containing at least 40 miRNA genes) are indicated with red bars. (C) Pie chart of clusters and related gene families. Out of all the gene clusters, 62.22% are found to contain homologous miRNAs (gene family), and members of 37.78% of the clusters are not in any gene families (not family). In the clusters with homologous miRNAs, 19.05% are found to contain multicopy pre-miRNAs (multicopy pre-miRNAs), but others are not found to be in any gene family (other gene family).