Egyptian Journal of Medical Human Genetics

RESEARCH Open Access

Non-randomness distribution of micro-RNAs on human chromosomes

Fariba Boroumand, Iraj Saadat and Mostafa Saadat*

Abstract

Background: Micro-RNA (miRNA) is one of the non-coding RNAs that exist in human genome. miRNAs play an important role in the expression of target genes. Several studies have indicated that organization of human genome is not random. In order to investigate the distribution of miRNAs on human chromosomes, the present study was carried out.

Results: Using the data from miRBase database, we found 1913 loci coding for miRNAs (MIRs). Human chromosome bands 1p36, 1q22, 1q24, 2q13, 2q35, 3p21, 6p21, 7q22, 8p23, 8q24, 9q22, 9q34, 11q12-q13, 12q13, 14q32, 16p13, 16q24, 17p13, 17q11, 17q21, 17q25, 19p13, 19q13, 20q13, 21p11, 22q13, and Xq26-q28 were significantly bearing higher number of MIRs. The 14q32 and 19q13 with 4.11 and 3.59 MIRs per mega-base pair, respectively, were the most MIR-richest human chromosomal bands. The number of MIRs on chromosomal bands significantly decreased as a function of distance from telomere ($r = -0.949$, df = 5, $P = 0.001$).

Conclusions: Our current data suggest that MIRs are not randomly distributed on human genomes.

Keywords: Genome, miRNA, 14q32, 19q13

Background

Multiple non-coding RNAs such as micro-RNAs (miRNAs), small nucleolar RNAs (snoRNAs), long non-coding RNAs (lncRNAs), and circular RNAs (circRNAs) exist in human genome [1, 2]. The size of the mature miRNAs is ~ 22 nucleotides in length that is produced by multiple splicing from a ~ 75 nucleotides primary transcribed precursor [1]. The majority of miRNA (MIR) sequences are found in introns of non-coding or coding transcripts although some MIRs are encoded within exons. The main function of MIRs is post-transcriptional gene regulation. It has been suggested that each MIR is predicted to have multiple potential target mRNAs and a single gene can be modulated by several MIRs [3].

Control of gene expression by MIRs plays an important role in multiple cellular pathways, such as cell proliferation and differentiation, apoptosis, control of cell cycle, migration, invasion, and many tissue-specific functions [1, 4, 5]. Many studies have demonstrated that MIRs play critical roles in cancers [4, 6]. Over-expression of MIR-197 and MIR-346 repressed the expression of their predicted target genes in vitro and in vivo. The MIR-197 and MIR-346 are contributed to carcinogenesis of follicular thyroid carcinomas [7]. MIR495 is among a group of MIRs which is significantly upregulated in breast cancer cell lines [8]. MIR9, MIR125A, and MIR125B are downregulated in primary neuroblastoma tumors [9].

Numerous studies indicated that organization of human genome is not random. Mutation rates show significant differences among regions of the mammalian genome [10–12]. In recent years, it has been shown that susceptible polymorphic genes involved in complex human diseases such as gastric cancer [13], breast cancer [14], schizophrenia [15], Parkinson's disease, and multiple sclerosis [16] are not randomly distributed on the chromosomes. Gene-rich bands and oncogenes are not randomly distributed on human chromosomes [17–25].

Calin and his colleagues have reported that human MIRs are commonly located at fragile sites and at genomic regions which are involved in cancers [26]. On the other hand, it has been shown that fragile sites and oncogenes are non-randomly located in light G bands of the human chromosomes [17]; similarly, human oncogenes show a specific chromosome territory [19]. Oncogenes are mainly located in telomeric regions [18]. Taken together, we suggested that

* Correspondence: saadat@shirazu.ac.ir
Department of Biology, College of Sciences, Shiraz University, Shiraz 71467-13565, Iran

MIR loci may be distributed non-randomly on the human chromosomes. Considering that to date the chromosomal distribution of MIRs in humans has not been investigated, the present study was carried out.

Methods

Based on the tools4mir.org, there are numerous databases concerning MIRs. These databases have different citations. Among these databases, TargetScans, MicroCosm Targets, and miRBase have 4440, 1753, and 1664 citations, respectively. Other databases have less than 500 citations. Data from TargetScans database was not available for download. The MicroCosm Target database contains computationally predicted targets of MIRs while the miRBase is a searchable database of published MIR sequences. Therefore, we chose the miRBase database for the present study.

Chromosomal distribution of loci encoding for MIRs was extracted from miRBase database (http://www.mirbase. org/). For evaluation of the non-random chromosomal distribution of MIRs, expected values for numbers of MIRs were estimated using relative lengths of each chromosome and then the differences between observed and expected values were evaluated by chi-square test. Spearman's correlation coefficient was used to investigate the association between the number of MIRs and distance from the telomere.

In order to investigate the non-random distribution of MIRs on each chromosomal band(s), the statistical method of Tai et al. was used [27]. The relative width of each band was measured using the International System for Chromosome Nomenclature based on 400 bands. The $P < 0.001$ was considered statistically significant.

Results

Using the miRBase database, there were 1913 loci for MIRs in human genome (Additional file 1: Table S1). There was a significant difference between the number of MIRs which is assigned to each chromosome and expected values based on their relative lengths ($\chi^2 = 501$, df = 22, $P < 0.001$). Chromosomes 14, 16, 17, 19, 22, and X have higher number and chromosomes 3, 4, 5, 6, 13, and 18 have lower number of MIR loci than the expected values.

In the next step, we determine which chromosomal bands bear higher number of MIRs. Figure 1 shows the distribution of these loci on human chromosomes. Table 1 shows the number of MIRs and relative length of selected chromosomal bands, as well as the calculated F statistics and its degrees of freedom (df). All of the comparisons were statistically significant ($P < 0.001$). The bands 1p36, 1q22, 1q24, 2q13, 2q35, 3p21, 6p21, 7q22, 8p23, 8q24, 9q22, 9q34, 11q12-q13, 12q13, 14q32,

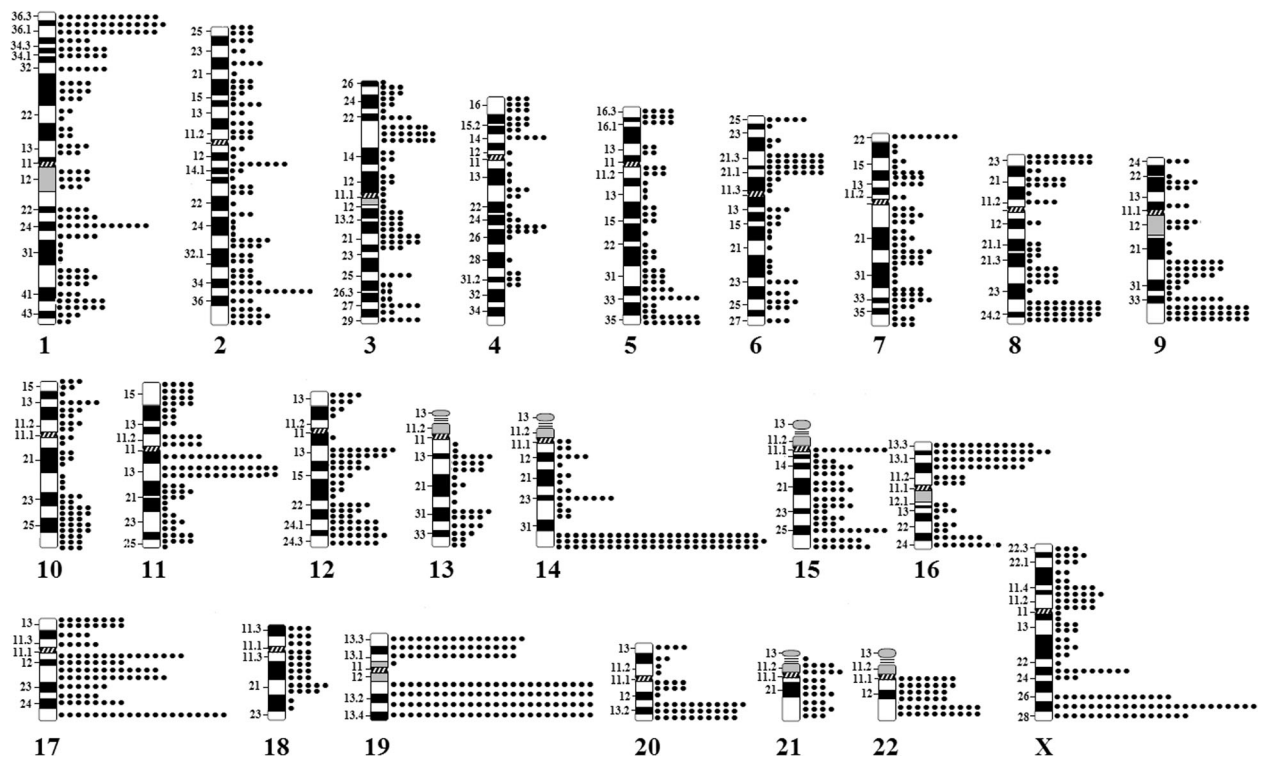

Fig. 1 Distribution of 1913 miRNAs (MIRs) on human chromosomes. International System for Chromosome Nomenclature based on 400 bands was used. Numbers on left side of the diagrams are the numbers of the bands. Each closed circle represents a MIR

Table 1 Human chromosomal bands which bear higher numbers of miRNAs (MIRs)

Bands	Relative length	Number of MIRs	*F*	df
1p36	0.0089	37	2.19	74, 3754
1q22	0.0005	4	4.31	8, 3820
1q24	0.0024	11	2.37	22, 3806
2q13	0.0011	7	3.23	14, 3814
2q35	0.0020	10	2.61	20, 3808
3p21	0.0035	20	3.01	40, 3788
6p21	0.0051	28	2.91	56, 3772
7q22	0.0030	13	2.24	26, 3802
8p23	0.0041	16	2.03	32, 3796
8q24	0.0092	35	2.00	70, 3758
9q22	0.0039	20	2.71	40, 3788
9q34	0.0035	30	4.50	60, 3768
11q12	0.0025	11	2.28	22, 3806
11q13	0.0045	28	3.31	56, 3772
12q13	0.0038	15	2.08	30,3798
14q32	0.0057	73	6.86	146, 3682
16p13	0.0054	38	3.73	76, 3752
16q24	0.0020	8	2.07	16, 3812
17p13	0.0035	16	2.40	32, 3796
17q11	0.0027	15	2.90	30, 3798
17q21	0.0040	23	3.04	46, 3782
17q25	0.0034	20	3.14	40, 3788
19p13	0.0064	46	3.80	92, 3736
19q13	0.0087	96	6.05	192, 3636
20q13	0.0069	31	2.37	62, 3766
21p11	0.0016	9	2.91	18, 3810
22q11	0.0034	14	2.16	28, 3800
22q13	0.0044	20	2.38	40, 3788
Xq26-q28	0.0086	54	3.34	108, 3720

Note: Non-random distribution of MIRs was investigated using the statistical method of Tai et al. (see Ref. [26])

16p13, 16q24, 17p13, 17q11, 17q21, 17q25, 19p13, 19q13, 20q13, 21p11, 22q13, and Xq26-q28 were MIR-rich regions. The 14q32 and 19q13 with 4.11 and 3.59 MIRs per mega-base pair (Mbp), respectively, were the most MIR-richest human chromosomal bands.

There was a significant negative association between the number of MIRs and their distance from telomere (Fig. 2). This means that majority of MIRs are located near the telomeres of the chromosomes. We know that p and q arms of the human chromosomes have different lengths; therefore, the abovementioned association might be a quasi-association. In order to address this point, we analyzed the MIR distribution from telomere to position 21 Mbp. Although these regions account for approximately 30% of the human genome, 45.1% of the MIRs were located within 21 Mbp from telomeres. Number of MIRs located within 0–3, 3–6, 6–9, 9–12, 12–15, 15–18, and 18–21 Mbp regions, were 189, 167, 150, 93, 115, 77, and 72, respectively. Statistical analysis revealed that there was a very strong negative correlation between distance from telomeres and number of MIRs ($r = -0.949$, df = 5, $P = 0.001$).

Discussion

In the first step of the present study, we found that some human chromosomes (14, 16, 17, 19, 22, and X) are bearing higher and some other chromosomes (3, 4, 5, 6, 13, and 18) are bearing lower numbers of MIRs. Similar findings regarding the number of functional genes in human chromosomes have been reported previously. Chromosomes 1, 11, 12, 16, 17, 19, 20, and 22 have higher number of active genes [25]. Human chromosome 19 has the highest MIR density in comparison with the other human chromosomes. It should be noted that several studies have indicated that human chromosome 19 has some unusual characteristics compared to other human chromosomes including the highest gene density [25, 28, 29], high expression levels [30], and high density of minisatellites [31].

Based on human chromosomal distribution of 186 MIRs, it has been reported that the majority of MIRs are located in cancer-associated genomic regions or at fragile sites, indicating the non-random distribution of MIRs [26]. In the present study, we investigated the distribution of large numbers of MIRs (1913 loci) at cytogenetic level and we found that MIRs were distributed in a non-random way on human chromosomes, which is consistent with the results of Calin and his colleagues [26]. In the present study and in the study of Calin and his colleagues, 1913 and 198 MIR loci were studied, respectively. It is self-evident that by increasing the number of MIR loci, the statistical power of the comparison will increase too.

It has been reported that MIR clusters are located on some human chromosomes, such as human chromosomes 14 and 19 [32]. It is well established that physically adjacent MIRs are exclusively or preferentially expressed in a tissue-specific manner [33, 34]. For example, the MIR clusters of chromosomes 14 and 19 are expressed in the human placenta [34–37]. The chromosome 19 MIR cluster is exclusively found in primates while the MIR located on human chromosome 14 appears to be conserved among eutherian species [37].

Our study furthermore demonstrated as evidence for non-random distribution of genetic materials across the human chromosomes. It should be noted that chromosome 19q13 is bearing polymorphic genes associated with the risk of breast [14] and gastric cancers [13], as

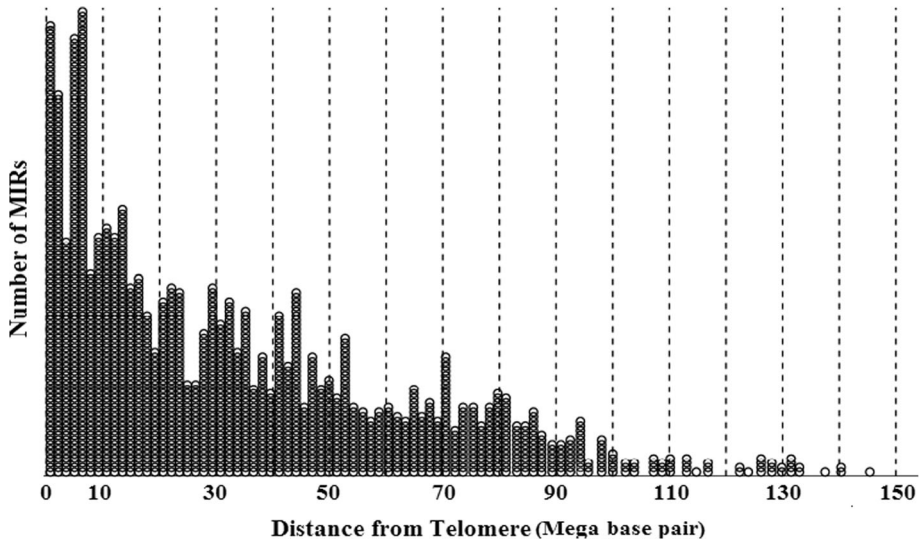

Fig. 2 Distribution of miRNAs (MIRs) on each arms of chromosomes in relation to their distance from telomeres

well as late-onset Alzheimer's disease [38] and multiple sclerosis [16].

As mentioned in the "Results" section, there is a very strong negative correlation between distance from telomeres and the number of MIRs. This means that the number of MIRs significantly decreased as a function of distance from telomere. This is a novel finding about chromosomal distribution of MIRs in human genomes, which has not been reported previously. However, similar distribution was reported for oncogenes on human chromosomes [18, 19].

Accordingly, our present data suggest that MIRs are not randomly distributed on human genome. Considering that MIRs are involved in the regulation of target genes, and the chromosome segments bearing a greater number of MIRs have been associated with several human complex diseases, further studies are needed to explain the biological significance of the non-random distribution of MIRs on human chromosomes. Attempts to confirm non-random distribution of functional genes on human genome/chromosomes may lead to new direction in etiology of human multifactorial traits and after that may lead to development of a novel tool for mass screening of complex diseases.

Conclusions

In the present study, we investigated the chromosomal distribution of 1913 loci coding for MIRs. Results indicated that some segments of human chromosomes are bearing higher numbers of MIRs. The 14q32 and 19q13 were the MIR-richest human chromosomal bands. The number of MIRs on chromosomal bands significantly decreased as a function of distance from telomere.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s43042-019-0041-2.

Additional file 1: Table S1. Human MicroRNAs.

Abbreviations

df: Degree of freedom; Mbp: Mega-base pair; MIR: miRNA; miRNA: Micro-RNA

Acknowledgements

This study was supported by Shiraz University.

Authors' contributions

IS and MS designed the research. FB built the database. FB and MS performed the statistical analyses. IS and MS interpreted the results. All authors read and approved the final manuscript.

Funding

None

Availability of data and materials

The dataset analyzed during the current study is presented as a supplement file.

Ethics approval and consent to participate

Not applicable

Consent for publication

Consent to publish the data was obtained from all individual participants or their attendants included in the study.

Competing interests

The authors declare that they have no competing interests.

Received: 14 September 2019 Accepted: 10 December 2019
Published online: 19 December 2019

References

1. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297
2. Mattick JS, Makunin IV (2006) Non-coding RNA. Hum Mol Genet 15:R17–R29

3. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120:15–20
4. Peng Y, Croce CM (2016) The role of microRNAs in human cancer. Signal Transduct Target Ther 1:15004
5. Ehya F, Abdul Tehrani H, Garshasbi M, Nabavi SM (2017 Sep) Identification of miR-24 and miR-137 as novel candidate multiple sclerosis miRNA biomarkers using multi-staged data analysis protocol. Mol Biol Res Commun 6(3):127–140
6. Lee YS, Dutta A (2009) MicroRNAs in cancer. Annu Rev Pathol 4:199–227
7. Weber F, Teresi RE, Broelsch CE, Frilling A, Eng C (2006) A limited set of human MicroRNA is deregulated in follicular thyroid carcinoma. J Clin Endocrinol Metab 91:3584–3591
8. Hwang-Verslues WW, Chang PH, Wei PC, Yang CY, Huang CK, Kuo WH, Shew JY, Chang KJ, Lee EY, Lee WH (2011) miR-495 is up-regulated by E12/E47 in breast cancer stem cells, and promotes oncogenesis and hypoxia resistance via down-regulation of E-cadherin and REDD1. Oncogene 30:2463–2474
9. Laneve P, Di Marcotullio L, Gioia U, Fiori ME, Ferretti E, Gulino A, Bozzoni I, Caffarelli E (2007) The interplay between microRNAs and the neurotrophin receptor tropomyosin-related kinase C controls proliferation of human neuroblastoma cells. Proc Natl Acad Sci U S A 104:7957–7962
10. Wolfe KH, Sharp PM, Li WH (1989) Mutation rates differ among regions of the mammalian genome. Nature 337:283–285
11. Nachman MW, Crowell SL (2000) Estimate of the mutation rate per nucleotide in humans. Genetics 156:297–304
12. Smith NG, Webster MT, Ellegren H (2002) Deterministic mutation rate variation in the human genome. Genome Res 12:1350–1356
13. Mahjoub G, Saadat M (2018) Non-random distribution of gastric cancer susceptible loci on human chromosomes. EXCLI J 17:802–807
14. Saify K, Saadat M (2012) Non-random distribution of breast cancer susceptibility loci on human chromosomes. Breast Cancer Res Treat 136:315–318
15. Saadat M (2013) Chromosomal distribution of schizophrenia susceptibility loci. J Mol Neurosci 51:401–402
16. Saadat M (2014) Distributions of susceptibility loci of Parkinson's disease and multiple sclerosis on human chromosomes. EXCLI J 13:724–727
17. Hecht F (1988) Fragile sites, cancer chromosome breakpoints, and oncogenes all cluster in light G bands. Cancer Genet Cytogenet 31:17–24
18. Lima-de-Faria A, Mitelman F, Blomberg J, Pfeifer-Ohlsson S (1991) Telomeric location of retroviral oncogenes in humans. Hereditas 114:207–211
19. Lima-de-Faria A, Mitelman F (1988) The chromosome territory of human oncogenes. Biosci Rep 6:349–354
20. Mouchiroud D, D'Onofrio G, Aissani B, Macaya G, Gautier C, Bernardi G (1991) The distribution of genes in the human genome. Gene 100:181–187
21. Saccone S, Caccio S, Kusuda J, Andreozzi L, Bernardi G (1996) Identification of the gene richest bands in human chromosomes. Gene 174:85–94
22. Saccone S, De Sario A, Della Valle G, Bernardi G (1992) The highest gene concentrations in the human genome are in T-bands of metaphase chromosomes. Proc Nati Acad Sci USA 89:4913–4917
23. Saccone S, Federico C, Solovei I, Croquette MF, Della Valle G, Bernardi G (1999) Identification of the gene-richest bands in human prometaphase chromosomes. Chromosom Res 7:379–386
24. Musio A, Mariani T, Vezzoni P, Frattini A (2002) Heterogeneous gene distribution reflects human genome complexity as detected at the cytogenetic level. Cytogenet Cell Genet 134:168–171
25. Rafiee L, Mohsenzadeh S, Saadat M (2008) Nonrandom gene distribution on human chromosomes. EXCLI J 7:151–153
26. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 101:2999–3004
27. Tai JJ, Hou CD, Wang-Wuu S, Wang CH, Leu SY, Wuu KD (1993) A method for testing the nonrandomness of chromosomal breakpoints. Cytogenet Cell Genet 63:147–150
28. Dehal P, Predki P, Olsen AS, Kobayashi A, Folta P, Lucas S, Land M, Terry A, Ecale Zhou CL, Rash S, Zhang Q, Gordon L, Kim J, Elkin C, Pollard MJ, Richardson P, Rokhsar D, Uberbacher E, Hawkins T, Branscomb E, Stubbs L (2001) Human chromosome 19 and related regions in mouse: conservative and lineage-specific evolution. Science 293:104–111
29. Grimwood J, Gordon LA, Olsen A, Terry A, Schmutz J, Lamerdin J, Hellsten U, Goodstein D, Couronne O, Tran-Gyamfi M, Aerts A, Altherr M, Ashworth L, Bajorek E, Black S, Branscomb E, Caenepeel S, Carrano A, Caoile C, Chan YM, Christensen M, Cleland CA, Copeland A, Dalin E, Dehal P, Denys M, Detter JC, Escobar J, Flowers D, Fotopulos D, Garcia C, Georgescu AM, Glavina T, Gomez M, Gonzales E, Groza M, Hammon N, Hawkins T, Haydu L, Ho I, Huang W, Israni S, Jett J, Kadner K, Kimball H, Kobayashi A, Larionov V, Leem SH, Lopez F, Lou Y, Lowry S, Malfatti S, Martinez D, McCready P, Medina C, Morgan J, Nelson K, Nolan M, Ovcharenko I, Pitluck S, Pollard M, Popkie AP, Predki P, Quan G, Ramirez L, Rash S, Retterer J, Rodriguez A, Rogers S, Salamov A, Salazar A, She X, Smith D, Slezak T, Solovyev V, Thayer N, Tice H, Tsai M, Ustaszewska A, Vo N, Wagner M, Wheeler J, Wu K, Xie G, Yang J, Dubchak I, Furey TS, DeJong P, Dickson M, Gordon D, Eichler EE, Pennacchio LA, Richardson P, Stubbs L, Rokhsar DS, Myers RM, Rubin EM, Lucas SM (2004) The DNA sequence and biology of human chromosome 19. Nature 428:529–535
30. Caron H, van Schaik B, van der Mee M, Baas F, Riggins G, van Sluis P, Hermus MC, van Asperen R, Boon K, Voûte PA, Heisterkamp S, van Kampen A, Versteeg R (2001) The human transcriptome map: clustering of highly expressed genes in chromosomal domains. Science 291:1289–1292
31. Wright FA, Lemon WJ, Zhao WD, Sears R, Zhuo D, Wang JP, Yang HY, Baer T, Stredney D, Spitzner J, Stutz A, Krahe R, Yuan B (2001) A draft annotation and overview of the human genome. Genome Biol 2:RESEARCH0025
32. Guo L, Zhao Y, Zhang H, Yang S, Chen F (2014) Integrated evolutionary analysis of human miRNA gene clusters and families implicates evolutionary relationships. Gene 534:24–32
33. Weber MJ (2005) New human and mouse microRNA genes found by homology search. FEBS J 272:59–73
34. Liang Y, Ridzon D, Wong L, Chen C (2007) Characterization of microRNA expression profiles in normal human tissues. BMC Genomics 8:166
35. Miura K, Miura S, Yamasaki K, Higashijima A, Kinoshita A, Yoshiura K-I, Masuzaki H (2010) Identification of pregnancy-associated microRNAs in maternal plasma. Clin Chem 56:1767–1771
36. Morales-Prieto D, Chaiwangyen W, Ospina-Prieto S, Schneider U, Herrmann J, Gruhn B, Markert U (2012) MicroRNA expression profiles of trophoblastic cells. Placenta 33:725–734
37. Morales-Prieto DM, Ospina-Prieto S, Chaiwangyen W, Schoenleben M, Markert UR (2013) Pregnancy-associated miRNA-clusters. J Reprod Immunol 97:51–61
38. Saadat M (2016) Distributions of susceptibility loci to late onset Alzheimer's disease on human chromosomes. EXCLI J 15:403–405

Publisher's Note