Prediction of microRNA target genes for diagnosis and prognosis of asthma development

Shara Atambayeva*, *Saltanat* Orazova., *Gulzira* Yernazarova, *Svetlana* Turasheva and *Arailym* Bekenkali

Al-Farabi Kazakh National University, Almaty, 050040, Kazakhstan

Abstract. Given the intricacy of diagnosis, there's an urgent necessity to establish objective and quantifiable biological indicators to enhance the accuracy of diagnosing asthma in its early stages. MicroRNAs emerge as promising candidates for asthma diagnosis. By leveraging the correlations between established target genes and microRNAs, a methodology for early asthma diagnosis can be formulated. A comprehensive database containing microRNAs and target genes implicated in asthma development has been compiled. The dynamics of microRNA interaction with target genes crucial for diagnosing and predicting the progression of asthma have been investigated. Specific microRNA binding sites with target genes involved in asthma development have been pinpointed. Using the DIANA-microTCDS program, features of microRNA interaction with target genes linked to asthma development have been analyzed. Notably, binding sites with significant occurrences have been identified in genes such as *IL13, VEGFA, ADRB2, ALDH2, IL10*, and *DPP10*. The conservativity of microRNA binding sites to mRNA genes has been demonstrated.

1 Introduction

Asthma stands as one of the prevalent ailments in contemporary society. Extensive literature, both domestic and international, underscores a persistent and concerning trend: a steady rise in asthma incidence, unfavorable patterns in its progression, and a global uptick in patient mortality rates.

While the precise cause of asthma remains elusive, researchers have identified various risk factors, emphasizing the significant interplay between genetic predisposition and environmental influences. Genetic factors are recognized contributors, with estimates of asthma heritability ranging from 35% to 95%. Large-scale genetic investigations have unveiled numerous genetic variants linked to heightened asthma susceptibility. Furthermore, epigenetic modifications affecting how genetic information is expressed have been implicated in asthma development.

*Corresponding author: Shara.Atambaeva@kaznu.kz

Respiratory infections, particularly viral infections during early childhood, significantly elevate the risk of asthma onset, particularly in cases of severe symptomatology [1].

MicroRNAs represent the main mechanism of post-transcriptional regulation of genes in eukaryotes. As modulators of the immune response, microRNAs regulate the informational RNA (mRNA) of target genes and, thus, play a crucial role in the development and pathogenesis of asthma [2].

Functional analysis and a plethora of corroborative studies have underscored the role of microRNAs in regulating various signalling pathways. Several microRNAs have been identified as key players in modulating biological processes throughout the course of asthma development. Through bioinformatic analysis, microRNA-mediated regulation of several signalling pathways in asthma has been elucidated. These pathways encompass matrix metalloproteinase activity, inflammatory responses, transforming growth factor signalling, as well as pivotal biological processes like apoptosis and inflammation [3]. MicroRNAs also appear to play a role in the pathogenesis of asthma as a regulatory layer between genetic and environmental factors and allergic inflammation induced by the immune system and programming 19 of type 2 T helper response cells (Th2). Thus, therapeutic microRNA modulation may make it possible to regulate or suppress allergic inflammation [4].

In recent studies, elevated serum levels of several microRNAs including miR16, miR-21, miR-125b, miR-126, miR-145, miR-146a, miR-148a, miR-221, miR223, miR-338, miR-485-3p have been demonstrated in asthma populations in general found reduced miR-18a expression, miR-126, let-7e, miR-155, miR-224, and activation of miR-498, miR-187, miR-874, miR-143, miR-886-3p in nasal biopsies from patients with asthma without differences between allergic and nonallergic phenotypes. miR-296-5p, miR-16-5p, miR-203, and miR-30d-5p were found to correlate with bronchial hyperresponsiveness [5]. The significance of diverse microRNA expression in asthma is controversial. The role of microRNAs in asthma pathomechanism can be analyzed in three aspects: regulation of smooth muscle cells, epithelium and immune system [6].

As previously discussed, microRNAs are extensively recognized as biomarkers for various diseases and conditions, largely owing to their biochemical nature and associative attributes. Key features of disease biomarkers include their distinctive expression patterns in specific disease states and their amenability to easy detection and resistance to degradation. MicroRNAs fulfil these criteria admirably due to their inherent stability and disease-associated expression in biological fluids. They exhibit resilience to degradation as they are encapsulated within vesicles and bind to proteins in tissues, thus playing a pivotal role in post-transcriptional regulation [7].

2 Materials and methods

In this study, human genes that are involved in the development of asthma were taken as materials. During the search, the database produced 2096 target genes, all of which were tested separately. The association of the gene with this disease was identified, and a database of genes involved in the development of asthma was created. The search for genes was carried out in the DisGeNET database (https://www.disgenet.org//) [8]. Characteristics of the interaction of microRNAs with target genes associated with the development of asthma, obtained using the DIANA-microTCDS program [9, 10].

The DIANA-microT-CDS (http://diana.cslab.ece.ntua.gr/) is based on several parameters of individual calculation for each microRNA, taking into account conservative and non-conservative recognition of microRNAs. The program DIANAmicroT-CDS web

server v4.0 is based on the prediction of the binding sites of microRNA target genes, taking into account the value of the interaction energy. DIANA-microT-CDS identifies microRNA targets in both 3'UTR and CDS. This version of the program was developed due to the fact that data on the functioning of binding sites located in CDS are accumulating. DIANA-microT-CDS is based on three algorithms: DIANA-microT-CDS 3.0, PicTar and TargetScan. This program uses a dynamic programming algorithm to calculate the value of a parameter based on the interaction of microRNAs with target genes. Each computer program predicts binding sites based on certain principles: complete complementarity of the seed region, thermodynamic stability of the microRNA complex: mRNA complex, evolutionary conservativeness, search sites only 3'UTR or in all areas, etc [9, 10].

3 Results and Discussion

The study identified 100 genes involved in the development of asthma, and these genes are targets for 248 microRNAs. The DIANA-microT-CDS program was employed to predict the binding sites of microRNAs to target genes. Through analysis of the identified binding sites, 65 effective associations between microRNAs and target genes have been proposed. Among the genes implicated in asthma development, single binding sites were found in the 3'UTR of *CCL5, TNF, TSLP, ORMDL3, HLA-DQA1, PDE4D, IKZF3, CDHR3, EDN1, ARG1, NPSR1, PTEN, BGLAP, SPRR2B, CD14, BCL2,* and *AGER*. Conversely, multiple microRNA binding sites were identified in the 3'UTR of other genes. Notably, genes such as *IL13, VEGFA, ADRB2, ALDH2, IL10,* and *DPP10* exhibited a substantial number of binding sites with high conservation (Table 1).

Table 2. Characteristics of the interaction of microRNA with mRNA genes involved in the development of asthma disease.

Target genes	microRNA	Region	Binding Type	Transcript position	Score	Conservation
IL13	hsa-let-7b-5p	3'UTR	8mer	551-573	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau, canFam2, dasNov2, loxAfr3, echTel1, monDom5*
	hsa-let-7i-5p	3'UTR	8mer	548-573	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1, monDo5*
	hsa-miR-4478	3'UTR	8mer	122-144	0.93	*panTro2, rheMac2, rn4, mm9, canFam2, echTel1, monDom5*
ADRB2	hsa-let-7a-5p	3'UTR	8mer	284-306	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau, canFam2, dasNov2, loxAfr3, echTel1, monDom5*
	hsa-let-7b-5p	3'UTR	8mer	284-306	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1, monDom5*
	hsa-let-7c-5p	3'UTR	8mer	291-306	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1, monDom5*

VEGFA	hsa-miR-16-5p	3'UTR	8mer	258-282	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1, monDom5, galGal3*
	hsa-miR-205-5p	3'UTR	9mer	137-156	0.99	*panTro2, rheMac2, oryCun2, canFam2, dasNov2, loxAfr3, monDom5, galGal3*
ALDH2	hsa-miR-498	3'UTR	8mer	1061-1076	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, loxAfr3, xenTro2, tetNig2, fr2, danRer6*
	hsa-miR-491-5p	3'UTR	8mer	18-42	0.97	*panTro2, rheMac2, rn4, oryCun2, bosTau4, canFam2, loxAfr3, xenTro2, tetNig 2, fr2*
	hsa-miR-5001-5p	3'UTR	8mer	1049-1075	0.96	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, loxAfr3, xenTro2, tetNig2, fr2, danRer6*
IL10	hsa-miR-202-3p	3'UTR	7mer	897-920	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, loxAfr3, echTel1, monDom5*
	hsa-miR-202-3p	3'UTR	7mer	889-908	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, loxAfr3, echTel1, monDom5*
DPP10	hsa-miR-361-5p	3'UTR	9mer	942-964	0.99	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1*
	hsa-miR-27a-3p	3'UTR	9mer	3341-3357	0.98	*panTro2, rheMac2, canFam2, dasNov2*
	hsa-miR-92a-3p	3'UTR	8mer	428-446	0.94	*panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1, monDom5, galGal3*
CCR4	hsa-miR-500a-5p	3'UTR	8mer	529-554	0.99	*panTro2, rheMac2*
	hsa-miR-4755-3p	3'UTR	8mer	153-180	0.97	*panTro2, rheMac2, loxAfr3*
	hsa-miR-4269	3'UTR	8mer	1795-1812	0.96	*panTro2, rheMac2, bosTau4*
	hsa-miR-8085	3'UTR	9mer	136-160	0.95	*panTro2, rheMac2*
	hsa-miR-149-5p	3'UTR	8mer	466-491	0.91	*panTro2, rheMac2*

IL13 is the target gene of three and more microRNAs, for example hsa-let-7b-5p, hsa-let-7i-5p, and hsa-miR-4478 that bind in the 3'UTR. The target prediction estimates between this gene and microRNA ranged from 0,93 to 0,99. The *ALDH2* gene is the target gene for hsa-miR-4498 (BSs 1061-1076 n), hsa-miR-491- 5p (BSs 18-42 n), and hsa-miR-5001-5p (BSs 1049-1075 n) microRNA. Target prediction estimates between this gene and microRNA ranged from 0,96 to 0,99. For the five microRNAs hsa-miR-500a-5p, hsa-miR-4755-3p, hsa-miR-4269, hsa-miR8085, hsa-miR-149-5p the CCR4 gene is the

target gene. The target prediction score between this gene and microRNA ranged from 0,91 to 0,99, and the highest score was for the relationship between the CCR4 gene and hsa-miR-500a-5p (BSs 529-554 36 n). The conservativity of microRNA binding sites to mRNA genes, were presented as an example of the genes in Table 1.

The binding sites of microRNAs with mRNAs of orthologous genes of 14 animal species were investigated to determine the specifics of microRNA interaction (Table 2).

Table 2. Logo diagrams demonstrating nucleotide sequence conservativity in regions of orthologous genes of mRNA microRNA binding sites of target genes.

Target gene, microRNA	Nucleotide sequence variability of BS microRNA mRNA target genes
IL13 hsa-let-7b-5p	 *Hsa, panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1, monDom5*
ADRB2 hsa-let-7a-5p	 *Hsa, panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1, monDom5*
VEGFA hsa-miR-16- 5p	 *Hsa, panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, dasNov2, loxAfr3, echTel1, monDom5, galGal3*
IL10 hsa-miR-202-3p	 *Hsa, panTro2, rheMac2, rn4, mm9, oryCun2, bosTau4, canFam2, loxAfr3, echTel1, monDom5*

WebLogo chart shows conservatism in microRNA binding sites in human genes *IL13, ADRB2, VEGFA, IL10,* and 14 animal species: *hsa – Homo sapiens (human), panTro2 - Pan troglodytes (chimp), rheMac2 - Macaca mulatta (rhesus), rn4 - Rattus norvegicus (rat), mm9 - Mus musculus (mouse), oryCun2 - Oryctolagus cuniculus (rabbit), bosTau4 - Bos taurus (cow), canFam2 - Canis lupus familiaris (dog), dasNov2 - Dasypus novemcinctus (armadillo), loxAfr3 - Loxodonta africana (elephant), echTel1 - Echinops telfairi (tenrec), monDom5 - Monodelphis domestica (opossum), galGal3 - Gallus gallus (chicken), xenTro2 - Xenopus tropicalis (X. tropicalis).*

The most robust associations, involving three or more microRNAs, were identified in *IL13, VEGFA, ADRB2, ALDH2, IL10,* and *DPP10* mRNA. These genes were selected as exemplars due to their involvement in regulating diverse cellular processes and active roles in modulating atopy, immune responses, and other interactions within the inflammatory cascade.

To understand the characteristics of microRNA interaction with target genes, microRNA binding sites on mRNA orthologs from 14 animal species were investigated. This examination revealed evolutionary conservation of completely complementary microRNA binding sites, indicating the early emergence of gene expression regulation by microRNA molecules. The findings from this comparative analysis of microRNA binding sites in

mRNA orthologs justify the selection of an experimental animal model to validate the regulation of the studied gene targets by microRNA in subsequent experiments.

4 Conclusion

Using the DIANA-microT-CDS program, a database comprising 100 target genes implicated in asthma development was compiled, and binding sites of 248 microRNA to these target genes were predicted. Through analysis of the identified binding sites, 65 effective associations between microRNAs and target genes were proposed. The study revealed the association of 41 microRNAs with 31 genes involved in asthma development, with a target prediction accuracy of 0.999.

Further examination of the interaction characteristics between microRNAs and asthma target genes revealed that certain microRNAs possess binding sites in multiple genes. Predominantly, the binding sites of microRNAs are located in the 3'UTRs (untranslated regions) of mRNA transcripts of target genes. However, some genes exhibit microRNA binding sites in their 5'UTRs. Notably, the microRNA binding sites in 3'UTRs display evolutionary conservatism across orthologous mammalian genes.

References

1. S. Yanagisawa. Definition and diagnosis of asthma-COPD overlap (ACO). Allerg. Int., **67**, 172 (2018) https://doi.org/10.1016/j.alit.2018.01.002
2. L. Perez-de-Llano. Asthma-COPD overlap is not a homogeneous disorder: further supporting data. Respir. Res., **18**, 183 (2017) https://doi.org/0.1177/1753466618805662
3. J. Kim. Socioeconomic impact of asthma, chronic obstructive pulmonary disease and asthma-COPD overlap syndrome. J. Thorac. Dis., **9**, 1547 (2017) https://doi.org/10.21037/jtd.2017.05.07
4. M.P. Perron, P. Provost. Protein interactions and complexes in human microRNA biogenesis and function. Front. Biosci., **13**, 2537 (2008) https://doi.org/10.2741/2865
5. S. Griffiths-Jones. The microRNA Registry. Nucleic Acids Res., **32**, 109 (2004) https://doi.org/10.1093/nar/gkh023
6. A.L. Durham. Basic science: Epigenetic programming and the respiratory system. Breathe., **9**, 278 (2013) https://doi.org/10.1183/20734735.000413
7. A. Gurrola-Silva, J.G. Huerta-López. Asthma history. Alerg. Asma Inmunol. Ped., **22**, 77 (2013) https://doi.org/10.3390/medicina56090438
8. J. Piñero, J. Saüch, F. Sanz, L.I. Furlong. The DisGeNET cytoscape app: Exploring and visualizing disease genomics data. Comp. and Struct. Biotech. J., **19**, 2960 (2021) https://doi.org/10.1016/j.csbj.2021.05.015
9. M.D. Paraskevopoulou, G. Georgakilas, N. Kostoulas, I.S. Vlachos, T. Vergoulis, M. Reczko, C. Filippidis, T. Dalamagas, A.G. Hatzigeorgiou. DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. Nuc. Acids Res., **41**, 169 (2013) https://doi.org/10.1093/nar/gkt393
10. M. Reczko, M. Maragkakis, P. Alexiou, I. Grosse, A.G. Hatzigeorgiou. Functional microRNA targets in protein coding sequences. Bioinf., **28**, 771 (2012) https://doi.org/10.1093/bioinformatics/bts043