

ORIGINAL ARTICLE

Asthma and Lower Airway Disease

Asthma diagnosis using integrated analysis of eosinophil microRNAs

José M. Rodrigo-Muñoz^{1,2}  | José A. Cañas^{1,2}  | Beatriz Sastre^{1,2}  | Natalia Rego³ | Gonzalo Greif³  | Manuel Rial⁴  | Pablo Mínguez⁵  | Ignacio Mahillo-Fernández⁶  | Mar Fernández-Nieto^{2,4} | Inés Mora¹ | Pilar Barranco^{2,7}  | Santiago Quirce^{2,7}  | Joaquín Sastre^{2,4}  | Victoria del Pozo^{1,2} 

¹Department of Immunology, IIS-Fundación Jiménez Díaz, Madrid, Spain

²CIBER de Enfermedades Respiratorias (CIBERES), Madrid, Spain

³Institut Pasteur de Montevideo, Montevideo, Uruguay

⁴Department of Allergy, IIS-Fundación Jiménez Díaz, Madrid, Spain

⁵Department of Genetics, Bioinformatics Group, IIS-Fundación Jiménez Díaz-UAM, Madrid, Spain

⁶Epidemiology and Biostatistics Unit, IIS-Fundación Jiménez Díaz, Madrid, Spain

⁷Department of Allergy, Hospital La Paz-Institute for Health Research (IdiPAZ), Madrid, Spain

Correspondence: Victoria del Pozo, Immunology Department, IIS-Fundación Jiménez Díaz, Av. Reyes Católicos 2, 28040 Madrid, Spain (vpozo@fjd.es).

Funding information

This study was supported by Fondo de Investigación Sanitaria—FIS [PI15/00803, FI16/00036], Merck Health Foundation; CIBER de Enfermedades Respiratorias (CIBERES), a Carlos III Institute of Health initiative; Conchita Rábago Foundation (FCR); and FEDER funds.

Abstract

Background: Asthma is a syndrome characterized by airway inflammation and obstruction. Due to its heterogeneity, the difficulties in asthma diagnosis and treatment make the discovery of new biomarkers a focus of research. So, we determined the differential miRNA expression of eosinophils between healthy and asthmatic patients and to establish a differentially expressed miRNA profile detectable in sera for use as biomarker.

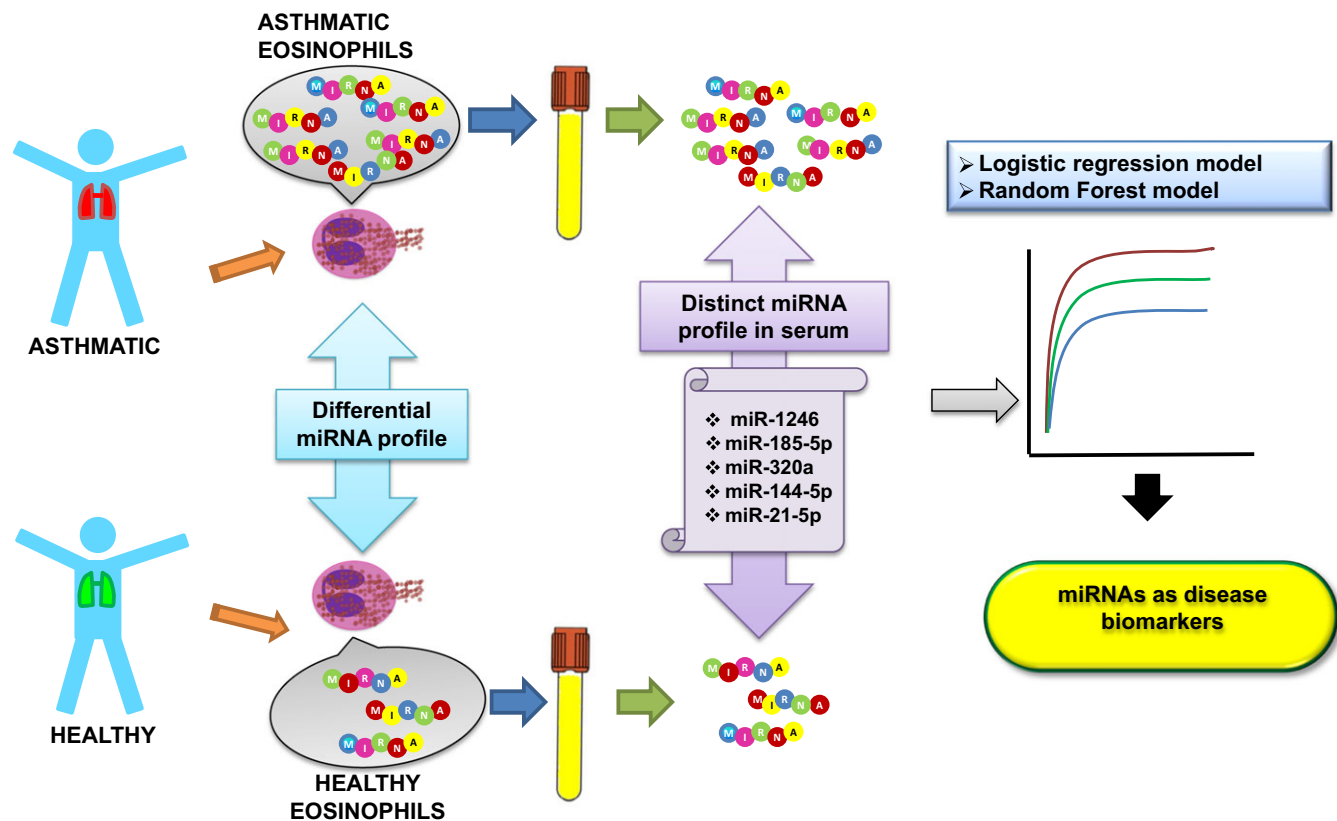
Methods: MicroRNAs from peripheral eosinophils from healthy and asthmatic subjects were isolated and analyzed by next-generation sequencing and confirmed by quantitative PCR in 29 asthmatics and 10 healthy individuals. The levels of serum miRNAs were performed by quantitative PCR in 138 asthmatics and 39 healthy subjects. Regression analysis and Random Forest models were performed.

Results: We found a set of miRNAs whose expression differs between eosinophils from asthmatics and healthy subjects. These miRNAs can classify asthmatics into two clusters that differed in the number of eosinophils and periostin concentration in serum. Some of these miRNAs were also confirmed in sera, as miR-185-5p which discriminates asthmatics from healthy subjects. Together with other two miRNAs, miR-185-5p allowed us to create a logistic regression model to discriminate better both conditions and a Random Forest model that can even sort the asthmatics into intermittent, mild persistent, moderate persistent, and severe persistent asthma.

Conclusion: Our data show that miRNAs profile in eosinophils can be used as asthma diagnosis biomarker in serum and that this profile is able to rank asthma severity.

KEYWORDS

asthma, biomarker, diagnosis, eosinophil, microRNA



GRAPHICAL ABSTRACT

Eosinophils from asthmatics present a different profile in microRNAs (miRNAs) compared to eosinophils from healthy subjects. As eosinophils obtention from patients is not a standardized method, we analyzed these miRNAs in serum showing that miRNAs profile expression in this bio-fluid can be used for asthma diagnosis and for severity classification.

1 | INTRODUCTION

Asthma is a heterogeneous chronic inflammatory disorder of the airways characterized by variable airflow obstruction and episodes of wheezing and breathlessness in association with airway hyperresponsiveness.¹ Asthma pathogenesis depends on numerous factors, including genetic predisposition and environmental agents (allergens, infections, and air pollutants).² It can be classified into different phenotypes and endotypes, which further complicates diagnosis in the absence of fully effective treatment.³ Asthma symptoms are accompanied by T helper (Th)2-driven inflammatory cell infiltrate of eosinophils and T lymphocytes.^{4,5}

Eosinophils are differentiated granulocytes that play a key role in asthma pathogenesis. In allergic or asthmatic processes, these cells are recruited to the airway from bone marrow via the bloodstream.⁶ Once in the airways, eosinophils release cationic proteins, cytokines, chemokines, and lipid mediators contained in their cytoplasmic granules.⁷ These molecules act on other cells and on eosinophils themselves and can contribute to perpetuation of inflammation and tissue damage.⁸ Previously, our group reported that eosinophils can release exosomes acting on different cells.⁹⁻¹¹ Exosomes contain a great number of proteins and microRNAs (miRNAs) that will be transferred

to recipient cells. Thus, eosinophils can carry miRNAs to other cells, principally via exosomes.¹²

MicroRNAs are single-stranded RNA sequences of 19-24 nucleotides in length that do not code for proteins. miRNAs act as regulators of gene expression by binding to the 3' untranslated region (UTR) of target mRNA, resulting in inhibition of protein translation or mRNA degradation.¹³ miRNAs have different expression profile depending on the cell type and disease status. Serum miRNAs are very stable and resistant to blood RNases, and some have been described as differentially expressed in several pathologies.¹⁴ All these features make miRNAs an emerging field of biomarker discovery.

The aim of our study was to search miRNAs which are differentially expressed between eosinophils from healthy and asthmatic individuals and to evaluate whether the expression signature of these miRNAs on serum can be used as asthma diagnosis biomarker and also to rank asthma severity.

2 | METHODS

More exhaustive descriptions about subject populations, reagents, eosinophil purification, miRNAs isolation, quantitative PCR,

bioinformatics analyses, and statistical analyses are provided in the section Appendix S1 (Methods).

2.1 | Patient selection and sampling

Asthmatics ($n = 173$) and healthy subjects ($n = 53$) were selected and recruited by Allergy Departments of IIS-Fundación Jiménez Díaz and Hospital La Paz. Subjects were properly informed and signed informed consent. The study was conducted following the principles of the Declaration of Helsinki. The study was approved by Fundación Jiménez Díaz Ethics Committee. Patients with asthma presented consistent history of asthma¹⁵ with greater than 12% improvement in FEV₁ 10 minutes following administration of inhaled terbutaline (500 µg) or had methacholine airway hyperresponsiveness (PC20 methacholine <16 mg/mL). Inhaled corticosteroids were withdrawn at least 2 weeks before sample obtaining. Nobody received oral corticosteroids (at least 6 months before the study), leukotriene receptor antagonists, aspirin, or any cyclooxygenase inhibitor.

Spanish guidelines on asthma management (GEMA) were applied for disease severity classification.¹⁶ Nonasthmatic healthy individuals with similar gender, age, race, and geographic region were selected as controls. The clinical and epidemiologic characteristics of patients and controls are collected in Tables S1-S3.

2.2 | Eosinophil's purification

Eosinophils were purified from peripheral blood of healthy and asthmatic donors as previously described.¹⁷ Serum was obtained by blood clotting and centrifugation.

2.3 | Eosinophil and serum miRNAs isolation and cDNA synthesis

Total eosinophils's RNA was extracted with miRNeasy Mini Kit (Exiqon; Qiagen, Hilden, Germany) and serum miRNAs with miRCURY RNA Isolation Kit—Biofluids (Qiagen). Both were retrotranscribed to cDNA using Universal cDNA Synthesis Kit II (Qiagen) following manufacturer's instructions.

2.4 | Illumina sequencing and expression analysis

Eosinophil miRNAs were precipitated with GenElute-LPA carrier (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany). Short RNA libraries were prepared with NEBNext® Small RNA Library Prep Set for Illumina® (Multiplex Compatible) (New England Biolabs, Ipswich, MA). Library quality control was performed with As suggested we accept the change to High Sensitivity DNA ChIP Kit (Agilent, Santa Clara, CA). Double-stranded DNA broad range kit (Invitrogen, Thermo Fisher Scientific Massachusetts, MA) was used for quantification in Qubit. Single-read sequencing for 50 cycles was run in an Illumina MiSeq sequencer (Illumina, San Diego, CA).

2.5 | Eosinophil and serum miRNA quantitative PCR (qPCR)

Differential expressed miRNAs found in next-generation sequencing (NGS) were validated in a different eosinophils cohort by qPCR using ExiLent SYBR® Green Master Mix (Exiqon; Qiagen) following manufacturer's instructions.

Eosinophil miRNAs and serum miRNAs samples C_t values were normalized to hsa-miR-103a-3p and hsa-miR-191-5p, respectively. Eosinophil and serum miRNAs relative expression was calculated using $2^{-\Delta\Delta C_t}$ method¹⁸; more details are explained in the online supplement.

2.6 | Periostin determination by ELISA

Periostin serum levels were measured using Human Periostin/OSF-2 DuoSet ELISA Kit (R&D Systems, Minneapolis, MN), following manufacturer's recommendations.

2.7 | Bioinformatic analysis of differentially expressed miRNAs

Multivariate analyses of differentially expressed miRNAs were performed with ClustVis¹⁹ by principal component analysis (PCA) and unsupervised hierarchical clustering heatmaps.

2.8 | Statistical analyses

Data were compared with unpaired, two-tailed Student's *t* test, or Mann-Whitney *U* test and ANOVA with Bonferroni post hoc test or Kruskal-Wallis with Dunn for multiple comparisons. $P < 0.05$ was considered significant. Correlations were estimated by Spearman's *R*. miRNAs performance was evaluated by receiver operating curve (ROC) analysis. Logistic regression models and Random Forest models²⁰ were developed for diagnosis.

Statistical calculations and graphs were performed with GraphPad Prism 6 (GraphPad Software Inc., San Diego, CA). ROC analysis, multivariate logistic regression, and Random Forest models were performed using R (<https://www.r-project.org/>).

3 | RESULTS

3.1 | Differential eosinophil miRNA detection between eosinophils from healthy and asthmatic subjects

To identify miRNA differences in eosinophils from asthmatic and healthy subjects, we analyzed the miRNAome by NGS in six samples of eosinophils from asthmatics and four samples from healthy individuals. Clinical and epidemiological characteristics of participating subjects can be consulted in Table S1. After NGS, 539 mature miRNAs were observed in eosinophils and 24 of them were differentially expressed between both types of eosinophils (expressed as Log₂ fold

change; Table 1). Of these, 14 were upregulated in eosinophils from asthmatics and the other 10 were downregulated.

Cluster and heatmap analysis were performed using the 24 differentially expressed miRNAs data, revealing 2 major clusters that segregate in accordance with disease status (Healthy and Asthma, Figure 1A,B).

3.2 | Eosinophil miRNA validation by qPCR

To validate the previous data, we studied these 24 miRNAs in a different population (29 asthmatic patients and 10 healthy donors) by qPCR. Clinical characteristics are recovered in Table S2. After qPCR validation, 14 miRNAs presented differential expression between both conditions (Table 1).

Out of 14 eosinophil miRNAs, nine were upregulated in asthma and five were downregulated (Table 1).

Principal component analysis was applied to reduce the dimensionality of the dataset to search for the most relevant combination. PCA showed the formation of two principal clusters (Figure 2A), according to disease status.

TABLE 1 Fold changes and statistical differences of eosinophil miRNAs between asthmatic patients and healthy individuals obtained by qPCR validation after NGS

miRNA	Next-generation sequencing Asthma (n = 6) vs Healthy (n = 4)			qPCR Asthma (n = 29) vs Healthy (n = 10)	
	Log ₂ fold	FDR	P-value	Fold (2 ^{-ΔΔC_t})	P-value
miR-1246	9.27	3.89E-036	7.22E-039	7.48	<0.0001
miR-1290	7.21	3.13E-013	1.16E-015	2.80	0.0185
miR-451a	-3.40	2.52E-006	1.40E-008	-11.23	<0.0001
miR-144-3p	-5.22	2.99E-006	2.22E-008	-9.22	0.0009
miR-5100	-3.04	3.90E-006	3.61E-008	-2.08	0.0268
miR-4521	2.29	0.00013	1.41E-006	2.19	0.0084
miR-144-5p	-3.59	0.00033	4.30E-006	-5.92	0.0012
miR320a	1.24	0.00056	8.24E-006	1.32	0.0106
miR-320b	1.56	0.0047	7.90E-005	1.56	0.0114
miR-185-5p	1.10	0.0060	0.00011	1.28	0.0115
miR-21-5p	0.94	0.0109	0.00022	1.66	0.0015
miR-146b-5p	1.20	0.013	0.00029	1.77	0.0098
miR-486-5p	-2.78	0.029	0.00087	-18.77	<0.0001
miR-629-5p	0.93	0.035	0.0013	1.56	0.0005
miR-664a-5p	1.09	0.028	0.00095	NA	NA
miR-1260b	-1.21	0.028	0.00095	NA	NA
miR-146a-5p	1.29	0.028	0.0011	NA	NA
miR-1-3p	10.7	0.035	0.0014	NA	NA
miR-4425	2.36	0.035	0.0014	NA	NA
miR-146b-3p	1.80	0.040	0.0017	NA	NA
miR-590-3p	-1.17	0.045	0.0020	NA	NA

FDR, false discovery rate; miRNA, microRNA; NA, not available.

Unsupervised hierarchical cluster analysis of miRNAs created two principal clusters corresponding to a subject's disease status: Asthma or Healthy (Figure 2B). Also, we observed two major subgroups inside the asthma cluster, and these could be related to different asthma classification (Figure 2A,B). Thus, we classified the samples into three groups: Healthy cluster (n = 10), Asthma 1 cluster (n = 11), and Asthma 2 cluster (n = 18) (Figure 2). The 14 studied miRNAs' expression differed between Healthy and Asthma 2 clusters (Figure 2C), seven were differentially expressed between Asthma 1 and Healthy clusters (Figure 2C) and eight between the Asthma 1 and Asthma 2 clusters (Figure 2C).

We also compared some clinical biomarkers of the subjects across the three clusters. Blood eosinophilia showed significant differences between Asthma 1 and Asthma 2 clusters (Figure 3A), and serum periostin levels showed significant differences between Healthy vs Asthma 2 cluster and between both Asthma clusters (Figure 3B). The rest of the evaluated parameters (FVC, FEV₁, FVC/FEV₁, FeNO, and IgE levels) did not show any differences between Asthma 1 and Asthma 2 subgroups (data not shown). The clusters' clinical characteristics are presented in Table S2.

Also, we observed a positive correlation between eosinophils number and miRNAs expression (expressed as 2^{-ΔC_t}): miR-1246 (Spearman's *R* = 0.57; *P* < 0.05), miR-146b-5p (Spearman's *R* = 0.54; *P* < 0.05), miR-451a (Spearman's *R* = 0.56; *P* < 0.05), and miR-1290 (Spearman's *R* = 0.49; *P* < 0.05; data not shown).

3.3 | miRNA serum expression as asthma disease biomarker

To evaluate whether the miRNA signature could be used as a biomarker, we analyzed if it was conserved in sera. We analyzed the 14 miRNAs in sera from 138 asthmatic patients and 39 healthy subjects (clinical characteristics in Table S3). Of the 14 miRNAs, miR-1246, miR-144-5p, miR-320a, miR-185-5p, and miR-21-5p were found as upregulated in sera from asthmatic patients compared to healthy controls. The rest were either undetected in sera or not differentially expressed (Figure 4).

Biomarker quality of each miRNA for asthma diagnosis was determined by ROC curve, setting an area under the curve (AUC) value over 0.7. Our results showed that miR-185-5p is the best asthma biomarker, with an AUC of 0.78, followed by miR-320a (AUC = 0.72) and miR-1246 (AUC = 0.70; Figure 5A and Table S4). MiR-144-5p and miR-21-5p AUC did not reach 0.7 (miR-144-5p AUC value = 0.68; miR-21-5p AUC value = 0.69; Table S4).

MiR-185-5p showed in our population a better performance than serum periostin, a previously described biomarker (AUC = 0.56; Table S4).

3.4 | Changes in serum miRNA expression are in accordance with disease severity

We next studied whether these miRNAs had a different expression in the asthmatic group depending on disease severity. As Figure S1

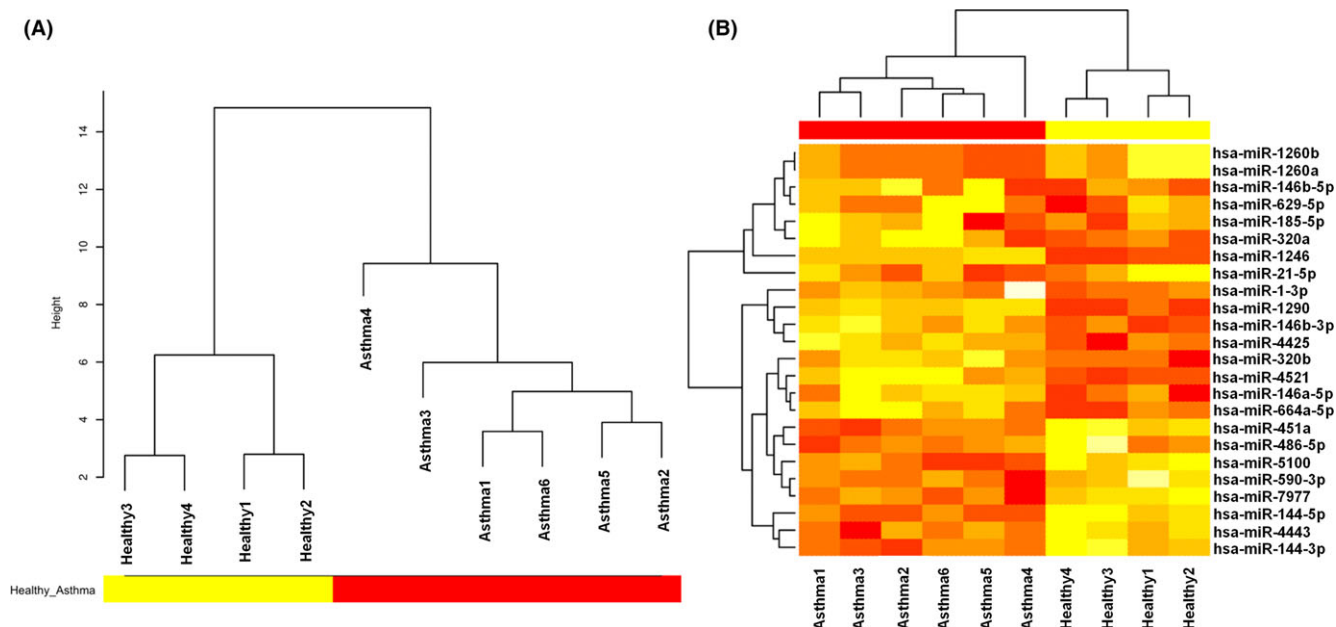


FIGURE 1 Differential eosinophil miRNAs expression between healthy and asthmatic subjects detected by NGS. A, Unsupervised hierarchical clustering based on 24 miRNA expression. B, Heatmap showing expression of 24 candidate eosinophil miRNAs after 2-way unsupervised hierarchical clustering. Healthy, $n = 4$; asthmatic, $n = 6$

shows, the expression of miR-185-5p may be related to severity, as it is higher in intermittent and mild asthmatics compared to controls, and even slightly higher in moderate and severe persistent asthmatics. Nevertheless, there are no statistical differences between the different asthma severity subgroups. When we analyzed the ROC curves, only miR-185-5p has AUC higher than 0.7 when each severity subgroup is compared to healthy individuals. For miR-320a and miR-1246, the AUC values obtained for each comparison are variable (Table S4).

We wanted to study whether the expression of this miRNA had any relationship with other indicators of severity. As shown in Figure S1, miR-185-5p expression is higher in patients with adult-onset asthma (>18 years old) compared to young-onset asthma, and it is also higher in patients who were admitted at least once in the intensive care unit (ICU) (Figure S1). Another miRNA that may be related to disease severity is miR-1246, which is highly upregulated in eosinophils from asthmatics compared to controls, and it is also upregulated in patients treated with antileukotrienes compared with nontreated asthmatics ($P < 0.01$; data not shown). We studied also if there were any differences in miRNAs expression when compared asthmatic patients with or without nasal polyps, sinusitis, rhinitis, or atopy, but no statistical differences were seen (data not shown).

We also determined the correlation between miRNA expression and clinical parameters. There was significant correlation between the expression of some miRNAs and FVC and FEV₁ (Figure S2). These results suggest that a higher expression of these miRNAs can be related to a worsening of the respiratory capacity. We did not find any correlation between the studied miRNAs and other

clinical parameters as eosinophil blood number or IgE levels (data not shown).

3.5 | Statistical models with miRNA relative expression for asthma diagnosis

For better discrimination of healthy vs asthma status, a multivariate logistic regression model was created using the cutoff values of miR-185-5p, miR-1246, and miR-144-5p expression defined by the Youden's index as explanatory variables (Table S4). The model combining these three miRNAs showed an AUC of 0.86, with a sensitivity of 0.89, and a specificity of 0.77 (Figure 5A,C). We also created two Random Forest models, a supervised machine learning method used for the classification of samples according to a set of variables, in this case, serum miRNAs expression. Our model is characterized by a 10-fold cross-validation using 90% of the samples as training dataset and the rest 10% for testing and by randomly resampling the dataset to keep similar number of samples. The first model was run to discriminate between healthy and asthma status, and the second was run to classify subjects into healthy or different asthmatic phenotypes (intermittent, mild persistent, moderate persistent, and severe persistent). Using miR-185-5p, miR-320a, and miR-320b in the first model, we are capable to classify subjects into asthma or healthy status with an AUC of 0.84 and a sensitivity and specificity of 0.75 (Figure 5B,C). On the second model, we found that expression of miR-320a, miR-185-5p, and miR-144-5p can classify subjects into one of the following groups; healthy, intermittent, mild persistent, moderate persistent, or severe persistent asthmatic, with an AUC of 0.75, a sensitivity of 0.5, and a specificity of 0.87 (Figure 5B,C).

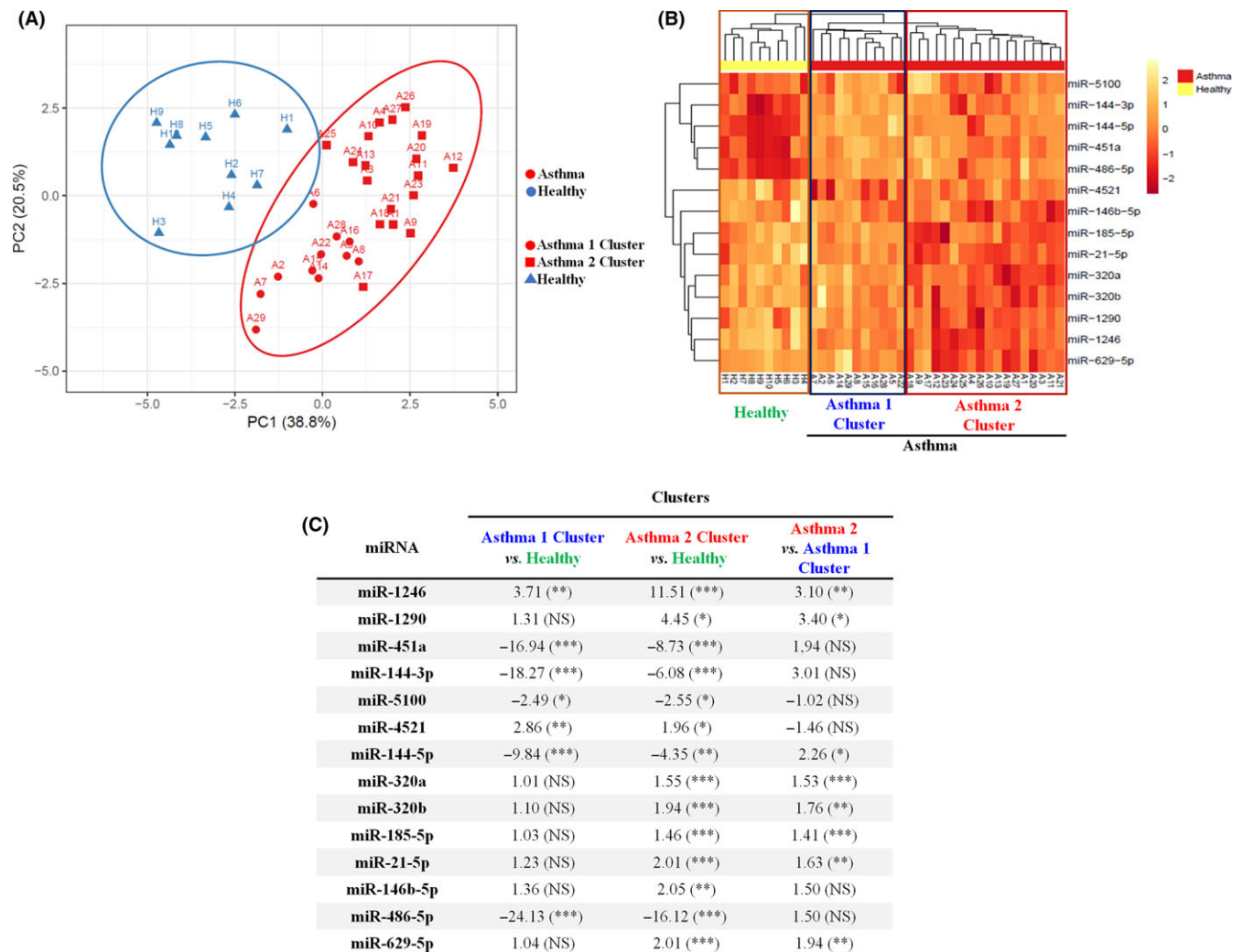


FIGURE 2 Validated eosinophil miRNAs allow discrimination between Healthy and Asthma status and show the formation of two asthma clusters with differences in miRNA expression. A, PCA with the data expression from 14 validated eosinophil miRNAs showed formation of 2 principal clusters: Asthma and Healthy. The Asthma group was also classified in the Asthma 1 cluster (red circles) and Asthma 2 cluster (red squares) following the subgroups created by the cluster analysis. B, Hierarchic clustering based on expression of 14 candidate eosinophil miRNAs in healthy ($n = 10$) and asthmatic subjects ($n = 29$). Asthma cluster present two principal clusters: Asthma 1 ($n = 11$) and Asthma 2 ($n = 18$). C, Fold changes and statistical differences among three hierarchical clusters of the study population. Results are expressed as mean $\pm$ SD. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

4 | DISCUSSION

Ours is the first study to demonstrate a distinct miRNA's profile in peripheral eosinophils between asthmatic and healthy subjects, showing that some of these miRNAs are also altered in serum.

In previous years, numerous studies have linked miRNAs to pulmonary diseases, such as asthma. These studies have been conducted in several cell types (eg, lymphocytes, airway epithelial and smooth muscle cells, and peripheral blood mononuclear cells) and in diverse biofluids, such as serum, bronchoalveolar lavage fluid, or exhaled breath condensate.²¹⁻²³ However, no studies have been performed with eosinophil miRNAs in asthma context.

Asthma is a heterogeneous and chronic respiratory illness. Due to its diversity, it is complicated to diagnose correctly.^{24,25} Over time, many investigators have searched for biomarkers, preferably

noninvasive ones such as miRNAs detected in easily available biofluids.²⁶

In this study, a set of 14 miRNAs was validated as differentially expressed between eosinophils from asthmatic and healthy individuals. Although several studies have found a miRNA signature that discriminates between asthma and healthy condition,^{22,27} this is the first time that it has been demonstrated that a specific miRNA signature from peripheral eosinophils is capable of discriminating between the two statuses. miRNAs differential expression is unlikely to be a mere epigenetic phenomenon of asthma as it has been demonstrated that miRNAs directly or indirectly affect the expression of multiple genes associated with inflammatory response; thus, miR-16 downregulates *TNF α* ; the miR-17~92 cluster modifies *PTEN*; miR-155 alters *PU.1*, *c-Maf*, and *SOCS1*; and miR-146a regulates *TRAF6* and *IRAK1*.^{28,29} Others miRNAs have been related to the pathology of

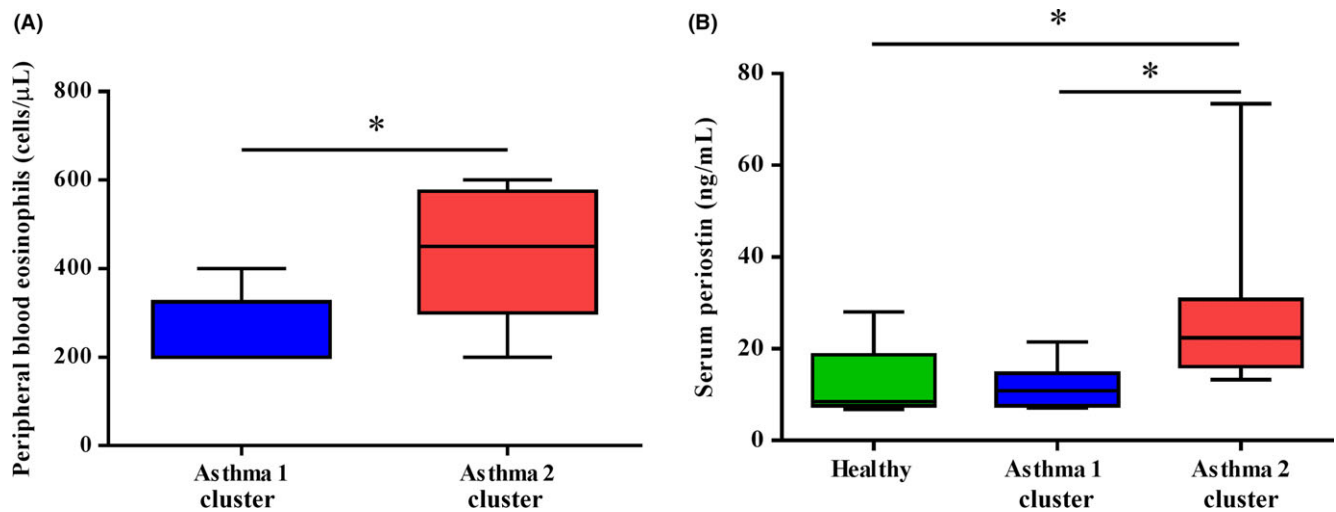


FIGURE 3 The two different asthma clusters present distinct levels of peripheral blood eosinophils and serum periostin levels. A, Peripheral blood eosinophils number. A significant increase in eosinophil number was observed in the Asthma 2 cluster compared to Asthma 1 cluster (* $P < 0.05$). B, Serum periostin levels were compared among the Healthy group and both Asthma clusters. Asthmatic subjects belonging to the Asthma 2 cluster showed the highest periostin levels compared to the Healthy and Asthma 1 clusters (* $P < 0.05$); no differences in periostin levels were found between Healthy and Asthma 1 clusters. Box plots represent medians and 25th and 75th percentiles and whiskers the 2.5th and 97.5th percentiles

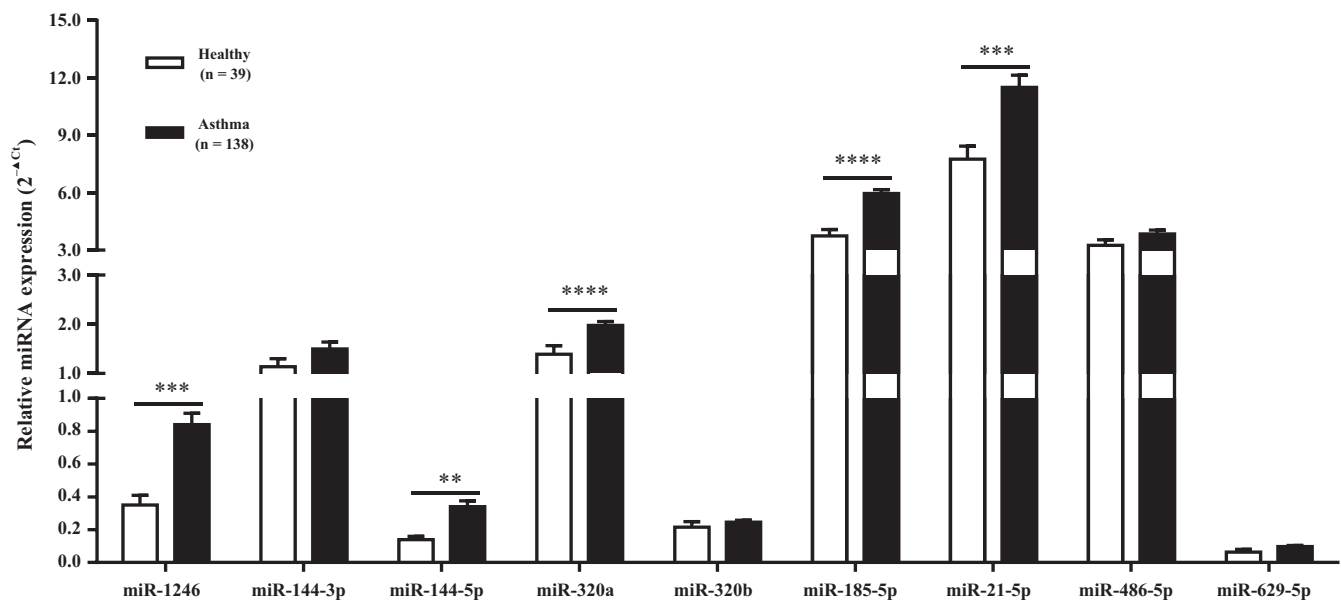


FIGURE 4 miRNA serum expression profiles differ between healthy and asthmatic subjects. Relative expression ($2^{-\Delta C_t}$) comparison between healthy ($n = 39$) and asthmatics ($n = 138$) for each miRNA in serum samples; error bars indicate mean $\pm$ SEM. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$

asthma and allergy as miR-21 that is upregulated in steroid-resistant asthma targeting *IL-12p35*; the miR-1248 targeting *IL-5*; and miR-15a inhibiting *VEGFA* in $CD4^+$ cells.³⁰ Furthermore, the miRNAs that we have found differentially expressed in eosinophils from asthmatics have predicted targets that may be related to asthma pathophysiology: *MAPK1*, *AKT1*, *CDC42*, *RHOA*, *PTEN*, *SMAD3*, *VEGFA*, *TGFBR1*, and *STAT3* using miRSystem (software for miRNA target determination, data not shown).

The profile of eosinophilic miRNAs is capable of discriminating between the two study conditions: asthma and healthy, as shown in the cluster analysis. The miRNA profile can also be used to distinguish clinical subgroups or phenotypes inside the asthmatic group. Thus, the population defined as Asthma 2 cluster presented higher percentage of peripheral eosinophils and serum periostin levels than Healthy and Asthma 1 clusters. Periostin is considered a biomarker of airway inflammation. Its serum concentration is usually higher in

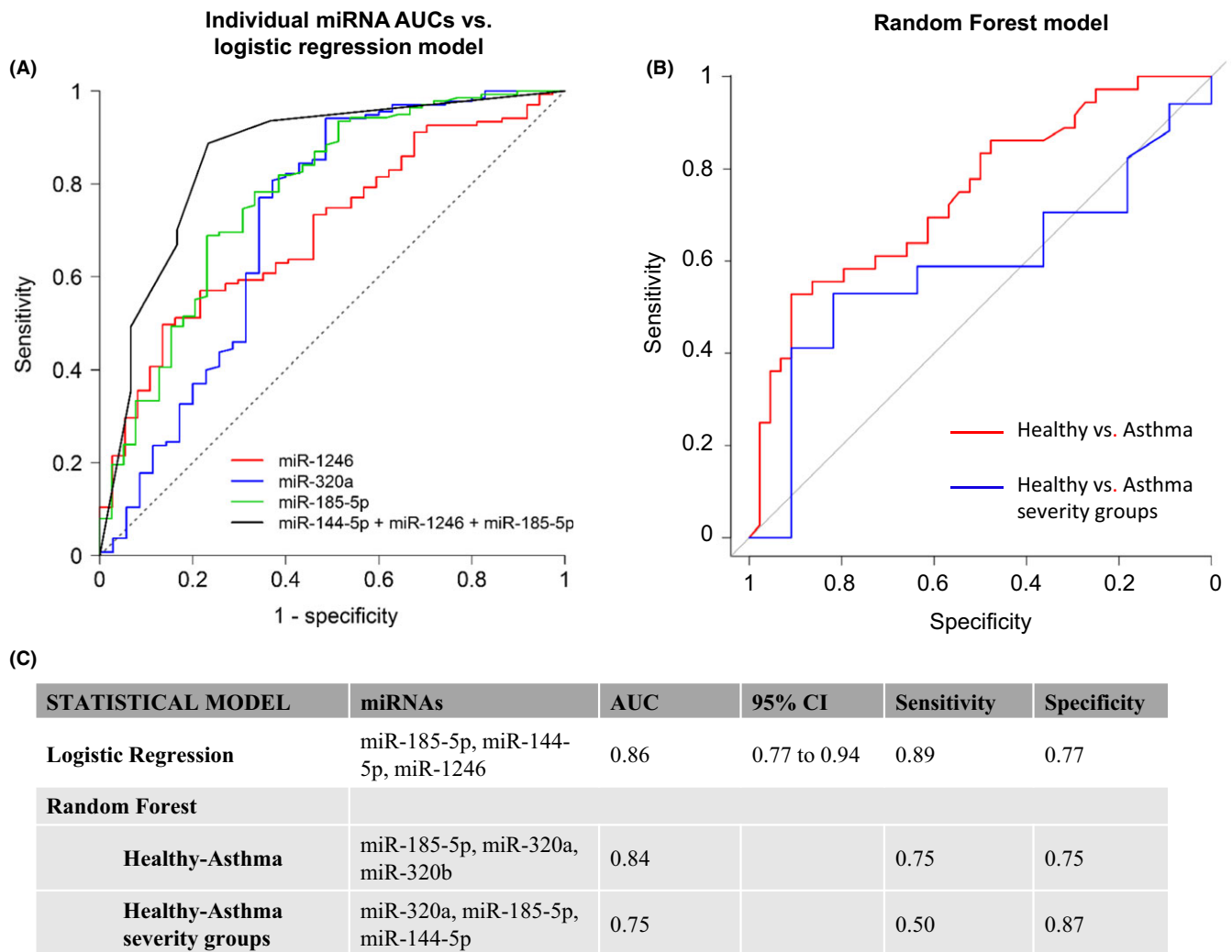


FIGURE 5 Statistical models created with serum miRNAs expression ($2^{-\Delta C_t}$) for asthma diagnosis. A, Receiver operating characteristic curve for the logistic regression model (using miR-185-5p, miR-1246, and miR-144-5p) and for individual miRNAs miR-185-5p, miR-320a, and miR-1246. B, Receiver operating characteristic curve for the Random Forest model of asthma diagnosis and for the Random Forest model for healthy and asthma severity groups classification. C, Comparison of the characteristics of the distinct statistical models created, logistic regression for asthma diagnosis and Random Forest models for asthma and subphenotypes asthma diagnosis

asthmatic patients with a T2 immune response associated with eosinophils.³¹⁻³³ In accordance with the results, our miRNA profile may be related to the Th2-high phenotype, as seen by the high level of peripheral eosinophils and serum periostin. miRNAs expression in eosinophils positively correlates with blood eosinophil counts. According to this, we suggest that this miRNA signature in eosinophils may also be used to distinguish between two different endotypes: Th2-high (greater severity as well as eosinophilia) and Th2-low (less severity and absence of eosinophilia).³⁴ Nevertheless, more studies must be performed to determine whether this signature is related to a full Th2 endotype characterization.

Among the discovered miRNAs, some have been found for the first time in this study (ie, miR-1246 and miR-1290), while others have already been described and associated with asthmatic pathology.³⁵⁻³⁷ For example, miR-21 is overexpressed in the asthmatic airways and miR-146b-5p is highly expressed in stimulated asthmatic airway smooth muscle cells.^{38,39}

The next step was to confirm that the specific eosinophilic miRNA signature associated with asthma can also be found in more accessible samples like serum. miRNAs are released from cells principally enclosed in exosomes, which reach recipient cells traveling through the body biofluids, such as serum.¹² Previously, we showed that eosinophils are able to release functional exosomes,⁹⁻¹¹ so we hypothesized that the 14 miRNAs altered in eosinophils from asthmatics compared to healthy subjects could be released inside exosomes. Thus, we may be able to detect these differentially expressed miRNAs in serum, a more easily obtained sample than blood eosinophils, being these eosinophilic miRNAs potential biomarkers. Conveniently we have validated a set of miRNAs in serum that are helpful for asthma diagnosis. It is possible that these miRNAs signature in serum may not be exclusively derived from eosinophils; in fact, some of these miRNAs have been previously described in other cell populations, such as miR-146a; miR-185; miR-144-3p; or miR-451a.³⁹⁻⁴³

Among the studied miRNAs, we can highlight miR-185-5p, which may be used to distinguish between asthmatic and healthy subjects with high values of sensitivity and specificity. Using statistic models that combine the discrimination power of the expression of several miRNAs, we can differentiate even better between healthy and asthmatic subjects, and it could be an alternative for asthma diagnosis.

Although similar studies have been performed, the advantage of our study is that the logistic regression model that we present has a sensitivity of 0.89 and a specificity of 0.77 for classifying patients into healthy or asthma using only three miRNAs, in contrast to Milger et al⁴⁴ that use 13 miRNAs or Panganiban et al²⁷ that used six miRNAs to differentiate healthy, asthmatic, and allergic rhinitis patients. Milger's study was developed in plasma from mice so the fact that they only found miR-320a from all our detected miRNAs may be related to the type of sample, the model (house dust mite induced), and the species used. The differences found between our study and the study developed by Panganiban et al can be due mainly to the fact that we searched specifically miRNAs differentially expressed in eosinophils. Nevertheless, converging with our results, they found differences in two miRNAs that we also found: miR-21 and miR-144, and therefore, our approach may be more specific for detecting miRNAs poorly represented.

One advantage of our study is that miR-185-5p alone may be used for asthma diagnosis (AUC of 0.78), instead of a big group of miRNAs, and could be more easily applied to the clinical routine as potential biomarker. Recently, and according to our data, McGeachie et al⁴⁵ have included miR-185-5p in their model for predicting future airways responsiveness.

As cutoff expression values from the individual miRNAs ROC curves are similar between each asthma severity subgroups, individual miRNAs expression is not suitable for asthma severity classification.¹⁶ For this purpose, we created a Random Forest model capable of classifying patients into healthy or each one of asthmatic severity using combined miRNAs expression in a trained algorithm. This model is characterized by 0.87 of sensitivity and 0.50 of specificity, which means that the algorithm classified correctly the 87% of positive samples in severity group and that 50% of negative samples that do not belong to a certain class are identified as such. Although having a high specificity (0.87), this model still has a low value of sensitivity (0.5) and more samples must be included in the model to improve sensitivity for a better clinical application.

We have also observed that some of our miRNAs are inversely correlated with FVC and FEV₁ values, implicating that when the expression of those miRNAs increases there may be a worse outcome and an increase in severity due to biological mechanisms that are still unknown. MiR-185-5p is the main miRNA that may be associated with asthma severity, as it presents higher expression in patients that have been admitted in the ICU and in patients who developed an adult-onset asthma (>18 years). Classically, both subgroups have a more severe phenotype, requiring higher doses of corticosteroids. Sometimes this treatment is ineffective. As previously described, miR-185-5p is related to corticosteroid sensitivity, and it seems to restore sensitivity in CEM-C1 cells which are steroid-

resistant by suppressing the mammalian target of rapamycin complex signaling pathway (mTORC).⁴⁶ Moreover, this miRNA has been implicated in cell growth through targeting CDC42 and RHOA directly and VEGFA indirectly,⁴¹ and it plays a role in calcium signaling by targeting the proteins NFAT and CaMKII δ in myocardial cells during cardiac hypertrophy.⁴⁷ These features make miR-185-5p a miRNA that may perform similar roles in the physiopathology of asthma, possibly being related to muscle cells hypertrophy, hyperplasia, cell contractility, and resistance to glucocorticoids; nevertheless, these functions have to be studied in more detail. In light of these results, miR-185-5p expression may be a good biomarker of disease severity, and thus, a better treatment may be chosen for patients with upregulated miR-185-5p.

To our knowledge, this is the first study to show differential miRNA expression profile in eosinophils associated with asthma pathology. Additionally, some of these miRNAs were located in serum and can be integrated in statistical models for asthma diagnosis.

ACKNOWLEDGMENTS

The authors recognize Oliver Shaw, English editor of IIS-FJD, for his revision and editing in English.

CONFLICTS OF INTEREST

VdP has been a consultant/speaker for AstraZeneca. The rest of authors declare that they have no conflicts of interest.

AUTHOR CONTRIBUTION

J.M.R.-M. and J.A.C. designed and planned the experiments, participated in the analysis of the data, and wrote significant sections of the manuscript. B.S., N.R., G.G., P.M., I.M.-F., and I.M. acquired the data or analyzed and interpreted the data and contributed to edit the manuscript. M.F.-N., M.R., P.B., S.Q., J.S., and MF performed patients' recruitment and diagnosis and helped to draft the manuscript. V.d.P. conceived the study, analyzed and interpreted the data, and drafted the article. All authors have read and approved the final manuscript.

ORCID

José M. Rodrigo-Muñoz  <http://orcid.org/0000-0003-3873-4563>
 José A. Cañas  <http://orcid.org/0000-0002-0460-2813>
 Beatriz Sastre  <http://orcid.org/0000-0002-2394-6849>
 Gonzalo Greif  <http://orcid.org/0000-0002-6769-3249>
 Manuel Rial  <http://orcid.org/0000-0003-4706-6608>
 Pablo Mínguez  <http://orcid.org/0000-0003-4099-9421>
 Ignacio Mahillo-Fernández  <http://orcid.org/0000-0003-0558-6541>
 Pilar Barranco  <http://orcid.org/0000-0002-4083-9930>
 Santiago Quirce  <http://orcid.org/0000-0001-8086-0921>
 Joaquín Sastre  <http://orcid.org/0000-0003-4689-6837>
 Victoria del Pozo  <http://orcid.org/0000-0001-6228-1969>

REFERENCES

- Lambrecht BN, Hammad H. The immunology of asthma. *Nat Immunol.* 2014;16:45-56.
- Umetzu DT, McIntire JJ, Akbari O, Macaubas C, DeKruyff RH. Asthma: an epidemic of dysregulated immunity. *Nat Immunol.* 2002;3:715-720.
- Fitzpatrick AM, Moore WC. Severe asthma phenotypes — How should they guide evaluation and treatment? *J Allergy Clin Immunol Pract.* 2017;5:901-908.
- Holgate ST. Innate and adaptive immune responses in asthma. *Nat Med.* 2012;18:673-683.
- Tan L, Gódor D, Bratt J, Kenyon NJ, Louie S. Benralizumab: a unique IL-5 inhibitor for severe asthma. *J Asthma Allergy.* 2016;9:71-81.
- Johansson MW. Eosinophil activation status in separate compartments and association with asthma. *Front Med.* 2017;4:75.
- Rosenberg HF, Dyer KD, Foster PS. Eosinophils: changing perspectives in health and disease. *Nat Rev Immunol.* 2012;13:9-22.
- Melo RCN, Liu L, Xenakis JJ, Spencer LA. Eosinophil-derived cytokines in health and disease: unraveling novel mechanisms of selective secretion. *Allergy.* 2013;68:274-284.
- Mazzeo C, Cañas JA, Zafra MP, et al. Exosome secretion by eosinophils: a possible role in asthma pathogenesis. *J Allergy Clin Immunol.* 2015;135:1603-1613.
- Cañas JA, Sastre B, Mazzeo C, et al. Exosomes from eosinophils autoregulate and promote eosinophil functions. *J Leukoc Biol.* 2017;101:1191-1199.
- Cañas JA, Sastre B, Rodrigo-Muñoz JM, et al. Eosinophil-derived exosomes contribute to asthma remodelling by activating structural lung cells. *Clin Exp Allergy.* 2018;141:AB72.
- Gallo A, Tandon M, Alevizos I, Illei GG. The majority of microRNAs detectable in serum and saliva is concentrated in exosomes. In: Afarinkia K, editor. *PLoS One.* 2012;7:e30679.
- Xiao C, Rajewsky K. MicroRNA control in the immune system: basic principles. *Cell.* 2009;136:26-36.
- Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci.* 2008;105:10513-10518.
- Standards for the diagnosis and care of patients with chronic obstructive pulmonary disease (COPD) and asthma. This official statement of the American Thoracic Society was adopted by the ATS Board of Directors, November 1986. *Am Rev Respir Dis.* 1987;136:225-244.
- Plaza Moral V. Comité Ejecutivo de GEMA. [GEMA(4.0). Guidelines for Asthma Management]. *Arch Bronconeumol.* 2015;51(Suppl 51):2-54.
- López E, Zafra MP, Sastre B, et al. Suppressors of cytokine signaling 3 expression in eosinophils: regulation by PGE₂ and Th2 cytokines. *Clin Dev Immunol.* 2011;2011:917015.
- Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔCT} method. *Methods.* 2001;25:402-408.
- Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res.* 2015;43:W566-W570.
- Liaw A, Wiener M. Classification and Regression by randomForest. https://www.researchgate.net/profile/Andy_Liaw/publication/228451484_Classification_and_Regression_by_RandomForest/links/53fb24cc0cf20a45497047ab/Classification-and-Regression-by-Random-Forest.pdf. Date last updated: April 3 2018. Accessed April 3, 2018.
- Kho AT, Sharma S, Davis JS, et al. Circulating MicroRNAs: association with lung function in asthma. In: Mersha TB, editor. *PLoS One.* 2016;11:e0157998.
- Levänen B, Bhakta NR, Torregrosa Paredes P, et al. Altered microRNA profiles in bronchoalveolar lavage fluid exosomes in asthmatic patients. *J Allergy Clin Immunol.* 2013;131:894-903.
- Perry MM, Baker JE, Gibeon DS, Adcock IM, Chung KF. Airway smooth muscle hyperproliferation is regulated by microRNA-221 in severe asthma. *Am J Respir Cell Mol Biol.* 2014;50:7-17.
- Meyer N, Dallinga JW, Nuss SJ, et al. Defining adult asthma endotypes by clinical features and patterns of volatile organic compounds in exhaled air. *Respir Res.* 2014;15:136.
- Kupczyk M, Dahlén B, Sterk PJ, et al. Stability of phenotypes defined by physiological variables and biomarkers in adults with asthma. *Allergy.* 2014;69:1198-1204.
- Etheridge A, Lee I, Hood L, Galas D, Wang K. Extracellular microRNA: a new source of biomarkers. *Mutat Res Mol Mech Mutagen.* 2011;717:85-90.
- Panganiban RP, Wang Y, Howrylak J, et al. Circulating microRNAs as biomarkers in patients with allergic rhinitis and asthma. *J Allergy Clin Immunol.* 2016;137:1423-1432.
- Mehta A, Baltimore D. MicroRNAs as regulatory elements in immune system logic. *Nat Rev Immunol.* 2016;28(16):279-294.
- Lindsay MA. microRNAs and the immune response. *Trends Immunol.* 2008;29:343-351.
- Sastre B, Cañas JA, Rodrigo-Muñoz JM, del Pozo V. Novel modulators of asthma and allergy: exosomes and microRNAs. *Front Immunol.* 2017;8:826.
- Jia G, Erickson RW, Choy DF, et al. Periostin is a systemic biomarker of eosinophilic airway inflammation in asthmatic patients. *J Allergy Clin Immunol.* 2012;130:647-654.
- Johansson MW, Evans MD, Crisafi GM, Holweg CTJ, Matthews JG, Jarjour NN. Serum periostin is associated with type 2 immunity in severe asthma. *J Allergy Clin Immunol.* 2016;137:1904-1907.
- Bobolea I, Barranco P, Del Pozo V, et al. Sputum periostin in patients with different severe asthma phenotypes. *Allergy.* 2015;70:540-546.
- Fahy JV. Type 2 inflammation in asthma—present in most, absent in many. *Nat Rev Immunol.* 2015;15:57-65.
- Cushing L, Jiang Z, Kuang P. The roles of MicroRNAs and protein components of the MicroRNA pathway in lung development and diseases. *Am J Resp Cell Mol Biol.* 2015;52:397-408.
- Brown D, Rahman M, Nana-Sinkam SP. MicroRNAs in respiratory disease. A clinician's overview. *Ann Am Thorac Soc.* 2014;11:1277-1285.
- Heffler E, Allegra A, Pioggia G, Picardi G, Musolino C, Gangemi S. MicroRNAs profiling in asthma: potential biomarkers and therapeutic targets. *Am J Respir Cell Mol Biol.* 2017;57:642-650.
- Lee HY, Lee HY, Choi JY, et al. Inhibition of MicroRNA-21 by an antagomir ameliorates allergic inflammation in a mouse model of asthma. *Exp Lung Res.* 2017;43:109-119.
- Comer BS, Camoretti-Mercado B, Kogut PC, Halayko AJ, Solway J, Gerthoffer WT. MicroRNA-146a and microRNA-146b expression and anti-inflammatory function in human airway smooth muscle. *Am J Physiol Lung Cell Mol Physiol.* 2014;307:L727-L734. 2
- Lu TX, Munitz A, Rothenberg ME. MicroRNA-21 is up-regulated in allergic airway inflammation and regulates IL-12p35 expression. *J Immunol.* 2009;154:4994-5002.
- Tang H, Wang Z, Liu X, et al. LRR4 inhibits glioma cell growth and invasion through a miR-185-dependent pathway. *Curr Cancer Drug Targets.* 2012;12:1032-1042.
- Zhao Y, Xie Z, Lin J, Liu P. MiR-144-3p inhibits cell proliferation and induces apoptosis in multiple myeloma by targeting c-Met. *Am J Transl Res.* 2017;15:2437-2446.
- Fukumoto I, Kinoshita T, Hanazawa T, et al. Identification of tumour suppressive microRNA-451a in hypopharyngeal squamous cell carcinoma based on microRNA expression signature. *Br J Cancer.* 2014;15:386-394.

44. Milger K, Götschke J, Krause L, et al. Identification of a plasma miRNA biomarker signature for allergic asthma: a translational approach. *Allergy*. 2017;72:1962-1971.
45. McGeachie MJ, Davis JS, Kho AT, et al. Asthma remission: predicting future airways responsiveness using an miRNA network. *J Allergy Clin Immunol*. 2017;140:598-600.
46. Chen P, Shen T, Wang H, et al. MicroRNA-185-5p restores glucocorticoid sensitivity by suppressing the mammalian target of rapamycin complex (mTORC) signaling pathway to enhance glucocorticoid receptor autoregulation. *Leuk Lymphoma*. 2017;58:2657-2667.
47. Kim JO, Song DW, Kwon EJ, et al. miR-185 plays an anti-hypertrophic role in the heart via multiple targets in the calcium-signaling pathways. *PLoS One*. 2015;13:e0122509.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Rodrigo-Muñoz JM, Cañas JA, Sastre B, et al. Asthma diagnosis using integrated analysis of eosinophil microRNAs. *Allergy*. 2019;74:507–517.
<https://doi.org/10.1111/all.13570>

Online Supplement

Asthma diagnosis using integrated analysis of eosinophil microRNAs

Authors: José M. Rodrigo-Muñoz, MSc^{1,2,8}, José A. Cañas, MSc^{1,2,8} Beatriz Sastre, PhD^{1,2}, Natalia Rego, PhD³, Gonzalo Greif, PhD³, Manuel Rial, MD⁴, Pablo Mínguez, PhD⁵, Ignacio Mahíllo-Fernández, PhD⁶, Mar Fernández-Nieto, MD^{2,4}, Inés Mora, LT¹, Pilar Barranco, MD, PhD^{2,7}, Santiago Quirce, MD, PhD^{2,7}, Joaquín Sastre, MD, PhD^{2,4}, Victoria del Pozo, PhD^{1,2}.

¹Department of Immunology, IIS-Fundación Jiménez Díaz, Madrid, Spain; ²CIBER de Enfermedades Respiratorias (CIBERES); ³Institut Pasteur de Montevideo, Montevideo, Uruguay; ⁴Department of Allergy, IIS-Fundación Jiménez Díaz; ⁵Department of Genetics, Bioinformatics Group, IIS-Fundacion Jimenez Diaz-UAM; ⁶ Epidemiology and Biostatistics Unit, IIS-Fundación Jiménez Díaz; ⁷Department of Allergy, Hospital La Paz-Institute for Health Research (IdiPAZ), Madrid, Spain. ⁸These authors contributed equally to this work.

Corresponding author:

Victoria del Pozo, PhD
Immunology Dept.
IIS-Fundación Jiménez Díaz
Av. Reyes Católicos 2
28040 Madrid SPAIN
Tel.: +34-91-5504891
E-mail: vpozo@fjd.es

SHORT TITLE: Asthma diagnosis by eosinophil microRNAs

KEY WORDS: asthma, biomarker, diagnosis, eosinophil, microRNA.

Online Supplement

METHODS

Patient selection and sampling

For the miRNA study, asthmatics ($n=173$) and healthy subjects ($n=53$) were selected and recruited from the Allergy Departments of IIS-Fundación Jiménez Díaz and Hospital La Paz-Institute for Health Research, Madrid. All subjects were properly informed and signed informed consent. The study was conducted in accordance with the principles set forth in the Declaration of Helsinki. The study was approved by the Fundación Jiménez Díaz Ethics Committee (CEIC). Asthma patients had a consistent history of asthma (1) and all showed a greater than 12% improvement in FEV₁ 10 minutes following administration of inhaled terbutaline (500µg) or had methacholine airway hyperresponsiveness (PC20 methacholine<16 mg/mL). None had a history of respiratory tract infections for at least the 6-week period preceding the study. We included both atopic and non-atopic asthmatic subjects because there were no differences in the parameters assessed between the two sets of patients. For patients who were receiving inhaled corticosteroids, the drugs were withdrawn for at least 2 weeks before the samples were obtained. No patient was receiving oral corticosteroids (for at least 6 months before the study), leukotriene receptor antagonists, aspirin, or any other cyclooxygenase inhibitor.

For serum analyses, asthmatic patients were subdivided by disease severity into subjects with intermittent, mild persistent, moderate persistent and severe asthma following the Spanish guidelines on asthma management (GEMA) (2).

The clinical and epidemiologic characteristics of patients and controls are collected in tables S1, S2, and S3 in Online Supplementary Material.

Eosinophil's purification and serum extraction

Eosinophils were purified from the peripheral blood of healthy controls and asthmatic donors using a two-step procedure as previously described (3). We used a negative immunoselection by means of a magnetic beads separation technique, as described in the manufacturer's protocol (EasySep® StemCell Technologies, Vancouver, Canada). Eosinophil purity was measured by flow cytometry (SIGLEC8⁺CD16⁻ cells) and routinely it was greater than 96%.

Afterwards, purified eosinophils were placed in 700 µL of QIAzol (Qiagen, Hilden, Germany) and frozen at -80°C at least 24 hours.

To obtain serum samples, peripheral blood was recovered in tubes without anticoagulant and it was allowed to clot for 30 min. After, it was centrifuged at 2095g for 10 min at 4°C. Then, serum was collected and stored at -80°C.

Eosinophil and serum miRNA isolation

Total RNA from eosinophils was extracted using the miRNeasy Mini Kit (Qiagen) according to manufacturer's instructions. Two hundred microliters of serum from each sample were used as input for the miRCURY RNA Isolation Kit-Biofluids (Exiqon, Qiagen), following the instructions of the manufacturer. Synthetic spike-in miRNAs were added during extraction to control a correct isolation of miRNAs (RNA spike-in kit UniRT, Exiqon, Qiagen).

Illumina sequencing and expression analysis

Eosinophil miRNA samples were precipitated with GenElute-LPA carrier (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany) and they were sent to the Unit of Bioinformatics of Institut Pasteur of Montevideo (Uruguay) for massive parallel

sequencing and bioinformatic analysis. Starting from 2 µg of total RNA, short RNA libraries were prepared with the NEBNext[®] Small RNA Library Prep Set for Illumina[®] (Multiplex Compatible) (New England Biolabs, Ipswich, MA, USA), following the manufacturer's instructions. After purification and size selection by polyacrylamide gel electrophoresis, library quality control was performed with the High Sensitivity DNA chip kit for Bioanalyzer (Agilent, Santa Clara, CA, USA). Then, dsDNA broad range kit (Invitrogen, Thermo Fisher Scientific) was used for quantification in QuBit. Single-read sequencing for 50 cycles was run in an Illumina MiSeq sequencer (Illumina, San Diego, CA, USA).

Several quality controls and an analysis of size distribution were performed on the raw sequences. CUTADAPT (4) was used to remove the adapter sequence 5' - AGATCGGAAGAGCACACGTCT from the 3' end of the reads and a minimum read length of 17 nucleotides was established. Final reads were aligned against reference human genome hg38 (UCSC Genome Browser, <https://genome.ucsc.edu>) using the bioinformatic tool Bowtie (5), in a global alignment mode and -v2 -a -best -strata parameters. MiRNA annotation was performed following the miRBase Release 21 database (6). MiRNAs expression profiles and analysis of differential expression was conducted in the R/Bioconductor environment (<https://www.bioconductor.org>), specifically using featureCounts function (7) and edgeR package (8). FeatureCounts was run with the following options: strandSpecific = 1, allowMultiOverlap = TRUE, minReadOverlap = 16, and countMultiMappingReads = TRUE. Before searching those mature sequences differentially expressed between conditions, the count table was filtered out for those matures not expressed above 1 count per million (cpm) of mapped reads in at least five samples. Normalisation and common dispersion estimation were done using edgeR default parameters. Tagwise dispersion values were used in an exact

test with quantile-adjusted conditional maximum likelihood (qCML) methods to assess for differentially expressed matures, before multiple testing corrections with a 5% False Discovery Rate (FDR).

cDNA synthesis for eosinophil and serum miRNAs

Eosinophil miRNA samples were retrotranscribed into cDNA using Universal cDNA Synthesis kit II (Exiqon, Qiagen) according to manufacturer instructions. Briefly, 20 ng of total RNA were used in each reaction and were retrotranscribed into cDNA in a final volume of 10 μ L. In the case of sera samples, 2 μ L of isolated miRNAs were used for retrotranscription. UniSp6 RNA spike-in was used as control primer. The cDNA samples were stored at -20°C until used.

Eosinophil and serum miRNA quantitative PCR (qPCR)

Differential expressed miRNAs found in next generation sequencing (NGS) were validated in a different cohort by qPCR using ExiLent SYBR[®] Green master mix (Exiqon, Qiagen) according to manufacturer's instructions. Eosinophil miRNAs samples were obtained from a mild and moderate persistent asthmatic population.

Hsa-miR-103a-3p, hsa-miR-191-5p, RNU6B, hsa-miR-629-5p, hsa-miR-146b-3p, hsa-miR-144-5p, hsa-miR-590-3p, hsa-miR-1246, hsa-miR-1260a, hsa-miR-7977, hsa-miR-4443, hsa-miR-1260b, hsa-miR-4521, hsa-miR-5100, hsa-miR-664a-5p, hsa-miR-4425, hsa-miR-1290, hsa-miR-486-5p, hsa-miR-21-5p, hsa-miR-1-3p, hsa-miR-146b-5p, hsa-miR-146a-5p, hsa-miR-451a, hsa-miR-144-3p, hsa-miR-320a, hsa-miR-320b, hsa-miR-185-5p, Spike-in-Sp2, Sp4, Sp5 and Sp6 primers were purchased from Exiqon (Quiagen). Hsa-miR-103a-3p, hsa-miR-191-5p, and RNU6B were tested as endogenous

miRNAs. MiR-451a and miR-23a-3p were measured for calculation of serum hemolytic index (no hemolysis if [miR-23a-3p – miR-451a] < 7).

Previously, eosinophil and sera cDNA were diluted 1:80 and 1:40 respectively with nuclease-free water and 4 µL was used in each reaction. Each sample was run in triplicate. A two-step program was used as follows: 45 cycles of 95°C for 10 seconds and 60°C for 1 min. The reactions were carried out in a Light Cycler 96 thermocycler (Roche, Basel, Switzerland). Cycle threshold (Ct) data was taken in the last step of the melting cycle in a continuous way. Results were analysed with LightCycler® 96 SW 1.1 (Roche) software.

For the selection of endogenous controls possible candidates were tested (miR-191-5p, miR-103a-3p and RNU6B), and the ones with more homogeneous gene expression between patients were selected for each type of sample. Eosinophil miRNA sample Ct values were normalised to hsa-miR-103a-3p, and hsa-miR-191-5p was used as endogenous control for serum miRNAs. Their expression was calculated for each sample by the C_t value. Eosinophil miRNAs were examined for their relative expression, using the $2^{-\Delta\Delta C_t}$ method, and serum miRNA expression was calculated by using the $2^{-\Delta C_t}$ method (9), where:

$$\Delta\Delta C_t = \Delta C_{t_{Asthma}} - \Delta C_{t_{Healthy}} \text{ and } \Delta C_t = \Delta C_{t_{miRNA}} - \Delta C_{t_{endogenous control}}$$

Periostin determination by ELISA

Periostin levels were measured in serum samples by ELISA using the Human Periostin/OSF-2 DuoSet ELISA kit (R&D Systems, Minneapolis, MN, USA) according to the manufacturer's recommendations. Previously, serum samples were diluted 1:100 in 1X PBS with 1% BSA. Standard range was from 62.5 pg/mL to 4 ng/mL.

Bioinformatic analysis of differentially expressed miRNAs

Multivariate analyses of differentially expressed miRNAs were performed with ClustVis bioinformatic tool (10) by using principal component analysis (PCA) and unsupervised hierarchical clustering heatmap analysis. The two clusters for healthy and asthma were calculated with the Principal Component Analysis (PCA) for reducing the dimensionality of the multivariate data from the miRNAs expression. The PCA method was performed and our multivariate data set was linearly transformed into a set of uncorrelated variables, ordered in descending manner by the variance explained. For our specific miRNAs data for healthy and asthma, unit variance scaling was applied to rows and singular-value decomposition (SVD) with imputation was used to calculate principal components. The software represents prediction ellipses that are such that with probability 0.95, a new observation from the same group will fall inside the ellipse.

Statistical analyses

Comparisons between groups were performed using unpaired, two-tailed Student *t* test for Gaussian samples and Mann-Whitney U-test for non-Gaussian samples. ANOVA with Bonferroni *post-hoc* test was performed for comparisons between more than two groups of Gaussian samples, and Kruskal-Wallis with Dunn *post-hoc* test was applied for non-Gaussian ones. A *P*-value of less than 0.05 was considered as significant. In addition, Spearman's correlation was used for comparison and estimation of correlations between miRNA expression levels and clinical characteristics.

Receiver operating curve (ROC) analysis was performed for evaluation of specificity and sensitivity values of miRNAs expression levels for distinguishing between healthy and asthmatic patients. The Youden index was used to set the optimal cut-off point and was performed according to the mathematical formula:

$$\text{Youden index (J)} = \frac{\text{True positives}}{\text{True positives} + \text{false negatives}} + \frac{\text{True negatives}}{\text{True negatives} + \text{false positives}} - 1$$

A logistic regression model was performed using as predictive variables the cut-off values of individual miRNAs expression defined by the Youden index in the ROC curves, being those over 4.5 for miR-185-5p, over 0.5 for miR-1246, and over 0.17 for miR-144-5p to differentiate between healthy and asthma condition. The classification of the subject is made based on the probabilities estimated by the model, taking a cut-off value of 0.5. Two Random Forest models (11) were elaborated; briefly, Random Forest is a supervised machine learning method used for the classification of samples according to a set of variables. It works building a set of decision trees using random variables extracted from a training dataset. Each of the trees makes a classification of samples. All the decision trees generated are then put together to select the best classifiers and applied to a test dataset to generate a new classification. The new re-classification is then evaluated with the real classes in order to estimate the prediction power of the model. Our performance of Random Forest includes two extra features: i) a 10-fold cross validation approach, where the random forest is applied 10 times, each of the runs generates randomly a train (90% of the data) and test (10% of the data) datasets; and ii) a classes balance approach, where in each of the 10 Random Forest runs, the dataset is randomly re-sampled in order to keep similar number of samples per class to allow the algorithm learn from all the classes. The final accuracy, sensitivity and specificity are the mean of the 10 values generated by our program. The final AUC also includes the information from the 10 runs. The approach described was applied to the same datasets in order to build two models able to discriminate between: i) healthy controls and asthmatics, and ii) healthy controls and asthmatic phenotypes (intermittent, mild persistent, moderate persistent and severe persistent). In this second multiclass classification model, a sample is classified exactly in one of the five groups, being those

healthy, intermittent, mild persistent, moderate persistent and severe persistent. To feed the Random Forest models, we used the expression of miRNAs differentially regulated in eosinophils from asthma vs. healthy individuals, measured in serum from all samples. In both approaches, a first approximation was performed in order to select the three miRNAs with more weight for discrimination as reported by two importance parameters: mean decrease accuracy (MDA) and mean decrease Gini (MDG). After this, the Random Forest algorithm was run again only considering the selected miRNAs.-All statistical calculations and graphs were performed with GraphPad Prism 6 (GraphPad Software Inc., San Diego, CA, USA). ROC analysis, multivariate logistic regression model with qualitative predictors, and Random Forest Model using 10 cross-fold validations were performed using R (<https://www.r-project.org/>).

216
217

218 REFERENCES:

- 219 1. Standards for the diagnosis and care of patients with chronic obstructive
220 pulmonary disease (COPD) and asthma. This official statement of the American
221 Thoracic Society was adopted by the ATS Board of Directors, November 1986.
222 *Am Rev Respir Dis* 1987;**136**:225–244.
- 223 2. Plaza Moral V, Comité Ejecutivo de GEMA. [GEMA(4.0). Guidelines for
224 Asthma Management]. *Arch Bronconeumol* 2015;**51**Suppl 1: 2–54.
- 225 3. López E, Zafra MP, Sastre B, Gámez C, Fernández-Nieto M, Sastre J, *et al.*
226 Suppressors of cytokine signaling 3 expression in eosinophils: regulation by
227 PGE₂ and Th2 cytokines. *Clin Dev Immunol* 2011;2011 917015.
- 228 4. Martin M. Cutadapt removes adapter sequences from high-throughput
229 sequencing reads. *EMBnet.journal* 2011;**17**:10.
- 230 5. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient
231 alignment of short DNA sequences to the human genome. *Genome Biol*
232 2009;**10**:R25.
- 233 6. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and
234 deep-sequencing data. *Nucleic Acids Res* 2011;**39**:D152–157.
- 235 7. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program
236 for assigning sequence reads to genomic features. *Bioinformatics* 2014;**30**:923–
237 930.
- 238 8. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for

differential expression analysis of digital gene expression data. *Bioinformatics* 2010;**26**:139–140.

9. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 2001;**25**:402–408.

10. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res* 2015;**43**:W566–W570.

11. Liaw A, Wiener M. Classification and Regression by randomForest. At https://www.researchgate.net/profile/Andy_Liaw/publication/228451484_Classification_and_Regression_by_RandomForest/links/53fb24cc0cf20a45497047ab/Classification-and-Regression-by-RandomForest.pdf.

SUPPLEMENTARY FIGURES AND LEGENDS

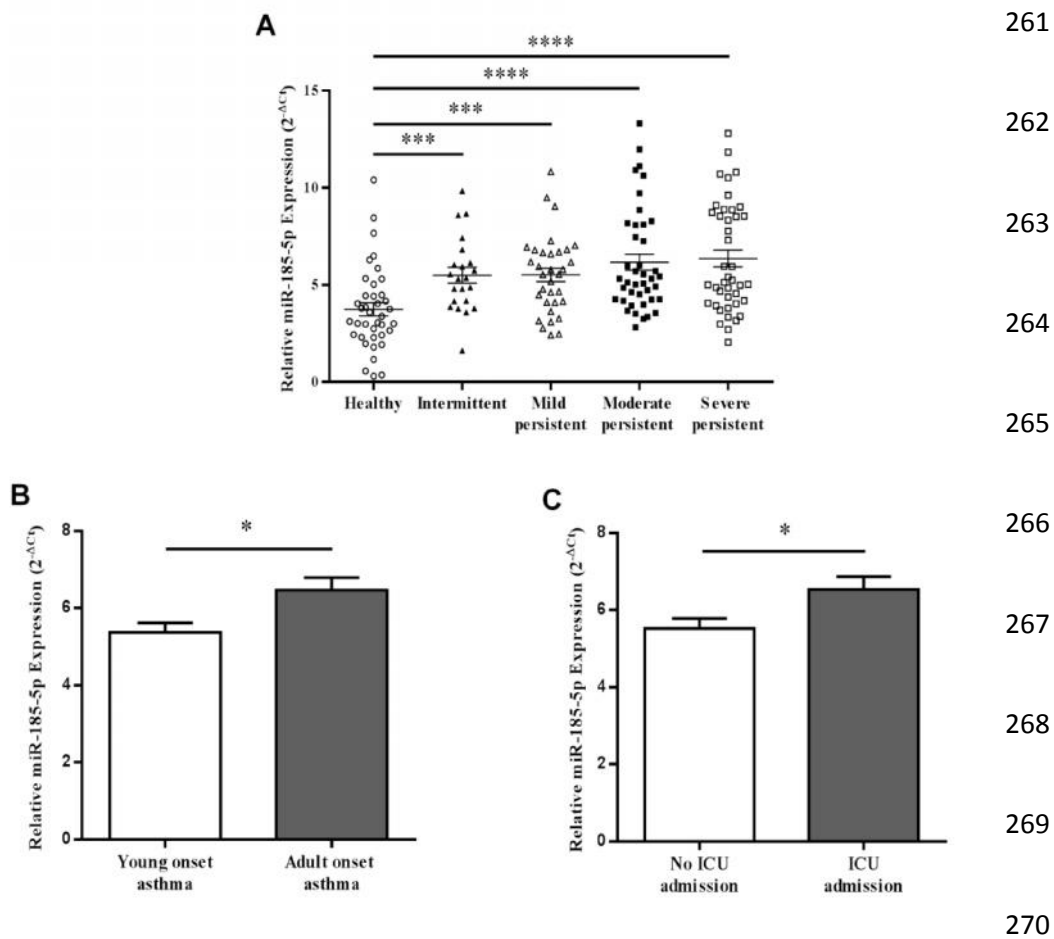


Figure S1. MiR-185-5p expression and asthma severity. A, Relative expression ($2^{-\Delta C_t}$) comparison between healthy ($n=39$) and intermittent ($n=22$), mild persistent ($n=33$), moderate persistent ($n=41$), and severe persistent ($n=42$) for miR-185-5p in serum samples. Results are expressed as dot plots with mean $\pm$ SEM. B, MiR-185-5p relative expression ($2^{-\Delta C_t}$) is higher in adult-onset asthmatics compared to young-onset asthmatics. C, MiR-185-5p relative expression ($2^{-\Delta C_t}$) is higher in the asthmatics who were admitted to the ICU at least once compared to those with no record of ICU admission. Results are expressed as mean $\pm$ SEM.

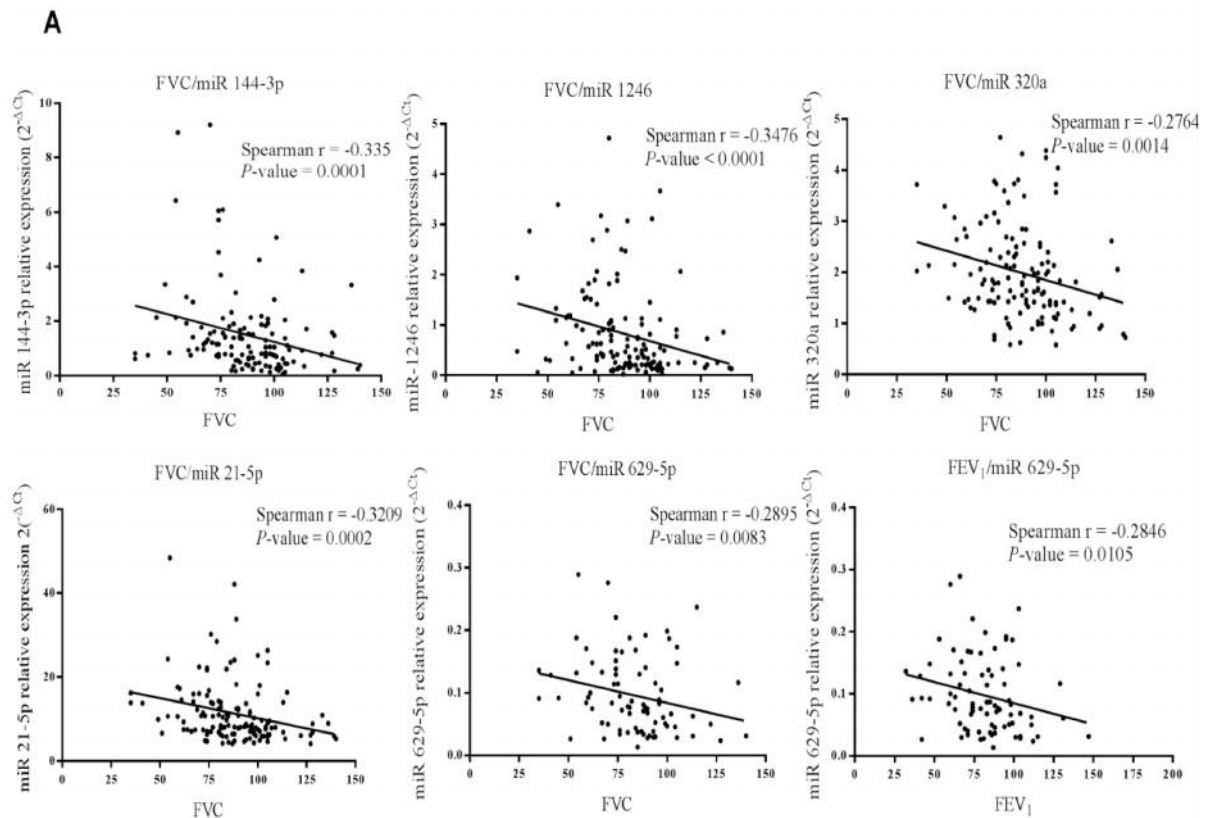


Figure S2. Correlation between clinical parameters and miRNA expression in asthmatic patients. MiR-144-3p, miR-1246, miR320a, miR-21-5p, and miR-629-5p expression (2^{-ΔCt}) are negatively correlated with FVC values in asthmatics. MiR-629-5p expression (2^{-ΔCt}) is negatively correlated with FEV₁ values in asthmatics.

Table S1. Clinical and epidemiological characteristics of asthmatic patients and healthy subjects recruited into the next generation sequencing eosinophil miRNA study.

	Asthma (n=6)	Healthy (n=4)
Age (years)*	33.17 (21-45)	34.20 (25-51)
Male (%)	0 (0.00)	3 (27.27)
Atopy (%)	3 (50.00)	3 (27.27)
Smoking history		
Current (%)	0 (0.00)	1 (9.09)
Ex-smoker (%)	2 (33.33)	3 (27.27)
Non-smoker (%)	4 (66.64)	7 (63.64)
FEV₁/FVC (%)†	78.66 (64.81-89.00)	NA
FeNO (ppb)*	49.60 (20.7-110.00)	NA
Total serum IgE (kU/L) †	50.00 (22.00-629.00)	NA
Eosinophils (%)*	3.48 (2.00-6.50)	<3.00
Eosinophils (per µL)*	275.00 (200.00-500.00)	<150.00

*Mean (range), †Median (range).NA (not available). FEV₁: forced expiratory volume in first second. FVC: forced vital capacity. FeNO: fractional exhaled nitric oxide.

Table S2. Clinical and epidemiological characteristics of the asthmatic patients and healthy subjects included in the qPCR miRNA validation study on eosinophils and classified by cluster.

	Healthy (n = 10)	Asthma (n = 29)	Asthma 1 Cluster (n = 11)	Asthma 2 Cluster (n = 18)	P-value
Age (years)*	34.20 (25-51)	39.38 (14-69)	37.36 (14-69)	40.61 (17-65)	NS
Male (%)	3 (27.27)	6 (20.69)	2 (18.18)	4 (22.22)	NS
Atopy (%)	3 (27.27)‡	20 (68.97)	6 (54.55)	14 (77.78)	§(P < 0.01)
Smoking history					
Current (%)	1 (9.09)	1 (3.45)	1 (9.09)	0 (0)	NS
Ex-smoker (%)	3 (27.27)	7 (24.14)	3 (27.27)	4 (22.22)	NS
Non-smoker (%)	7 (63.64)	21 (74.41)	7 (63.64)	14 (77.78)	NS
FEV₁/FVC (%)†	NA	75.89 (58.32-89)	77.00 (66-89)	75.66 (58.32-87.85)	NS
FeNO (ppb)*	NA	38.69 (7.20-110)	40.09 (7.20-84)	37.78 (11-110)	NS
Total serum IgE (IU/mL)†	NA	448.50 (22-6997)	50.95 (22-6997)	702.50 (268-864)	NS
Eosinophils (%)*	<3.00	5.13 (1.90-8.80)	3.17 (1.90-5.90)	5.80 (3.10-8.80)	** (P < 0.01)
Eosinophils (per µL)*	<150.00	368.42 (200-600)	250.00 (200-400)	409.09 (200-600)	** (P < 0.05)

*Mean (range), †Median (range). ‡P < 0.05 Healthy vs. Asthma. §P < 0.05 Healthy vs. Asthma 2, **P < 0.05 Asthma 1 cluster vs. Asthma 2 cluster. NA = (not available). NS = no signification. FEV₁: forced expiratory volume in first second. FVC: forced vital capacity. FeNO: fractional exhaled nitric oxide.

Table S3. Clinical and epidemiological characteristics of asthmatic patients and healthy subjects recruited for serum qPCR miRNA validation.

	Intermittent (n = 22)	Mild persistent (n = 33)	Moderate persistent (n = 41)	Severe persistent (n = 42)	Healthy (n = 39)	Asthmatics (n = 138)
Age (years)*	44.30 (26-86)	51.80 (27-82)	54.50 (16-86)	60.90 (29-88)	40.92 (21-79)	54.20 (16-88)
Atopy (%)	18 (81.81)	25 (75.75)	34 (82.92)	26 (61.90)	3 (7.69)	103 (75.63)
Smoking history						
Current (%)	2 (9.09)	6 (18.18)	5 (12.19)	1 (2.38)	2 (5.13)	13 (9.42)
Ex smoker (%)	3 (13.63)	6 (18.18)	12 (29.26)	9 (21.43)	0 (0.00)	30 (21.73)
Non-smoker (%)	17 (77.27)	21 (63.63)	24 (58.53)	32 (76.19)	16 (41.02)	94 (68.11)
FVC (%)†	99 (80-140)**‡	96 (61-136)**	84 (67-133)**	74 (35-139)	NA	87 (35-140)
FEV ₁ (%)†	97 (82-147)**	94 (60-129)**	84 (67-131)**	66 (32-98)	NA	84 (32-147)
Total serum IgE (kU/L)†	201 (70-1362)	249 (29-783)	132 (5-2000)	285 (4-1519)	NA	172 (4-2000)
Peripheral blood eosinophils (per µl)*	272.60 (3-470)	302.90 (98-433)	361.10 (5-1140)	440.90 (36-1100)	<350	365.2 (3-1140)
Broncodilators						
Long action (%)	3 (13.63)	19 (57.57)	41 (100.00)	42 (100.00)	0 (0)	105 (76.08)
B2 blockers (%)	10 (45.45)	15 (45.45)	28 (68.29)	34 (80.95)	0 (0)	87 (63.04)
Corticoid daily dose (µg)†	0 (0-400)**‡¶	320 (0-800)**‡	740 (160-1600)**	1600 (640-4000)	0.00	720 (0-4000)
Leukotriene receptor agonist (%)	2 (9.09)	2 (6.06)	17 (41.46)	30 (71.42)	0 (0)	51 (36.95)
Ipratropium bromide (%)	0 (0)	0 (0)	2 (4.87)	1 (2.38)	0 (0)	3 (2.17)
Tiotropium bromide (%)	0 (0)	0 (0)	1 (2.44)	5 (12.19)	0 (0)	6 (4.34)
Corticoids						
Oral (%)	1 (4.54)	29 (87.87)	41 (100)	42 (100)	0 (0)	113 (81.88)
Injected (%)	2 (9.09)	3 (9.09)	4 (9.756)	9 (21.43)	0 (0)	18/129 (13.04)
Systemic glucocorticoids (%)	2 (9.09)	8 (24.24)	9 (21.95)	25 (59.52)	0 (0)	44 (31.88)
ICU admissions (%)	6 (27.27)	18 (54.54)	13 (31.70)	23 (54.76)	0 (0)	60 (43.47)
>5 ICU admissions (%)	1 (4.54)	2 (6.06)	2 (4.87)	8 (19.04)	0 (0)	13 (9.42)
Periostin (ng/ml)*	129.3 (3.2-320.5)	109.69 (19.1-293.9)	145.8 (26-340.2)	150.6 (38.8-403.6)	115.3 (22.6-237.9) ^Δ	135.96 (3.2-403.6)

Results are shown as *Mean (range), †Median (range). FEV₁: forced expiratory volume in first second. FVC: forced vital capacity. ICU: intensive care unit. NA: not available ** $P < 0.05$ compared to severe persistent. ¶ $P < 0.05$ compared to mild persistent asthma. ‡ $P < 0.05$ compared to moderate persistent asthma. ^Δ $P < 0.05$ compared to moderate and severe persistent asthmatics.

Table S4. Statistic results for ROC curve analysis, between healthy and asthmatics, and between healthy and severity asthma groups.

	Biomarker	AUC	CUT-OFF ($2^{-\Delta Ct}$)	95% CI	P value	Sensitivity	Specificity
Asthma vs. Healthy	miR-185-5p	0.78	4.50	0.69 to 0.86	<0.001	0.688	0.769
	miR-320a	0.72	0.89	0.60 to 0.83	<0.001	0.941	0.514
	miR-1246	0.70	0.50	0.61 to 0.79	<0.001	0.496	0.865
	miR-144-5p	0.68	0.17	(0.59 - 0.77)	<0.001	0.587	0.774
	Periostin (ng/ml)	0.56	176.39	0.46 to 0.64	0.093	0.261	0.949
Intermittent vs. Healthy	miR-185-5p	0.76	3.59	0.64 to 0.89	0.002	0.955	0.538
	miR-320a	0.72	0.89	0.59 to 0.86	0.023	0.955	0.514
	miR-1246	0.68	0.236	0.53 to 0.82	0.011	0.773	0.541
Mild persistent vs. Healthy	miR-185-5p	0.76	4.07	0.65 to 0.87	<0.001	0.788	0.667
	miR-320a	0.71	1.11	0.57 to 0.84	0.041	0.938	0.571
	miR-1246	0.70	0.49	0.58 to 0.73	0.006	0.531	0.838
Moderate persistent vs. Healthy	miR-185-5p	0.79	4.18	0.69 to 0.89	<0.001	0.805	0.692
	miR-320a	0.68	0.89	0.56 to 0.81	0.068	0.951	0.514
	miR-1246	0.68	0.503	0.56 to 0.80	0.002	0.537	0.865
Severe persistent vs. Healthy	miR-185-5p	0.78	4.52	0.68 to 0.88	<0.001	0.690	0.769
	miR-320a	0.75	1.35	0.64 to 0.87	<0.001	0.825	0.657
	miR-1246	0.74	0.40	0.62 to 0.85	<0.001	0.625	0.784

Online Supplement

Asthma diagnosis using integrated analysis of eosinophil microRNAs

Authors: José M. Rodrigo-Muñoz, MSc^{1,2,8}, José A. Cañas, MSc^{1,2,8} Beatriz Sastre, PhD^{1,2}, Natalia Rego, PhD³, Gonzalo Greif, PhD³, Manuel Rial, MD⁴, Pablo Mínguez, PhD⁵, Ignacio Mahíllo-Fernández, PhD⁶, Mar Fernández-Nieto, MD^{2,4}, Inés Mora, LT¹, Pilar Barranco, MD, PhD^{2,7}, Santiago Quirce, MD, PhD^{2,7}, Joaquín Sastre, MD, PhD^{2,4}, Victoria del Pozo, PhD^{1,2}.

¹Department of Immunology, IIS-Fundación Jiménez Díaz, Madrid, Spain; ²CIBER de Enfermedades Respiratorias (CIBERES); ³Institut Pasteur de Montevideo, Montevideo, Uruguay; ⁴Department of Allergy, IIS-Fundación Jiménez Díaz; ⁵Department of Genetics, Bioinformatics Group, IIS-Fundacion Jimenez Diaz-UAM; ⁶ Epidemiology and Biostatistics Unit, IIS-Fundación Jiménez Díaz; ⁷Department of Allergy, Hospital La Paz-Institute for Health Research (IdiPAZ), Madrid, Spain. ⁸These authors contributed equally to this work.

Corresponding author:

Victoria del Pozo, PhD
Immunology Dept.
IIS-Fundación Jiménez Díaz
Av. Reyes Católicos 2
28040 Madrid SPAIN
Tel.: +34-91-5504891
E-mail: vpozo@fjd.es

SHORT TITLE: Asthma diagnosis by eosinophil microRNAs

KEY WORDS: asthma, biomarker, diagnosis, eosinophil, microRNA.

Online Supplement

METHODS

Patient selection and sampling

For the miRNA study, asthmatics ($n=173$) and healthy subjects ($n=53$) were selected and recruited from the Allergy Departments of IIS-Fundación Jiménez Díaz and Hospital La Paz-Institute for Health Research, Madrid. All subjects were properly informed and signed informed consent. The study was conducted in accordance with the principles set forth in the Declaration of Helsinki. The study was approved by the Fundación Jiménez Díaz Ethics Committee (CEIC). Asthma patients had a consistent history of asthma (1) and all showed a greater than 12% improvement in FEV₁ 10 minutes following administration of inhaled terbutaline (500µg) or had methacholine airway hyperresponsiveness (PC20 methacholine<16 mg/mL). None had a history of respiratory tract infections for at least the 6-week period preceding the study. We included both atopic and non-atopic asthmatic subjects because there were no differences in the parameters assessed between the two sets of patients. For patients who were receiving inhaled corticosteroids, the drugs were withdrawn for at least 2 weeks before the samples were obtained. No patient was receiving oral corticosteroids (for at least 6 months before the study), leukotriene receptor antagonists, aspirin, or any other cyclooxygenase inhibitor.

For serum analyses, asthmatic patients were subdivided by disease severity into subjects with intermittent, mild persistent, moderate persistent and severe asthma following the Spanish guidelines on asthma management (GEMA) (2).

The clinical and epidemiologic characteristics of patients and controls are collected in tables S1, S2, and S3 in Online Supplementary Material.

Eosinophil's purification and serum extraction

Eosinophils were purified from the peripheral blood of healthy controls and asthmatic donors using a two-step procedure as previously described (3). We used a negative immunoselection by means of a magnetic beads separation technique, as described in the manufacturer's protocol (EasySep® StemCell Technologies, Vancouver, Canada). Eosinophil purity was measured by flow cytometry (SIGLEC8⁺CD16⁻ cells) and routinely it was greater than 96%.

Afterwards, purified eosinophils were placed in 700 µL of QIAzol (Qiagen, Hilden, Germany) and frozen at -80°C at least 24 hours.

To obtain serum samples, peripheral blood was recovered in tubes without anticoagulant and it was allowed to clot for 30 min. After, it was centrifuged at 2095g for 10 min at 4°C. Then, serum was collected and stored at -80°C.

Eosinophil and serum miRNA isolation

Total RNA from eosinophils was extracted using the miRNeasy Mini Kit (Qiagen) according to manufacturer's instructions. Two hundred microliters of serum from each sample were used as input for the miRCURY RNA Isolation Kit-Biofluids (Exiqon, Qiagen), following the instructions of the manufacturer. Synthetic spike-in miRNAs were added during extraction to control a correct isolation of miRNAs (RNA spike-in kit UniRT, Exiqon, Qiagen).

Illumina sequencing and expression analysis

Eosinophil miRNA samples were precipitated with GenElute-LPA carrier (Sigma-Aldrich, Merck KGaA, Darmstadt, Germany) and they were sent to the Unit of Bioinformatics of Institut Pasteur of Montevideo (Uruguay) for massive parallel

sequencing and bioinformatic analysis. Starting from 2 µg of total RNA, short RNA libraries were prepared with the NEBNext[®] Small RNA Library Prep Set for Illumina[®] (Multiplex Compatible) (New England Biolabs, Ipswich, MA, USA), following the manufacturer's instructions. After purification and size selection by polyacrylamide gel electrophoresis, library quality control was performed with the High Sensitivity DNA chip kit for Bioanalyzer (Agilent, Santa Clara, CA, USA). Then, dsDNA broad range kit (Invitrogen, Thermo Fisher Scientific) was used for quantification in QuBit. Single-read sequencing for 50 cycles was run in an Illumina MiSeq sequencer (Illumina, San Diego, CA, USA).

Several quality controls and an analysis of size distribution were performed on the raw sequences. CUTADAPT (4) was used to remove the adapter sequence 5' - AGATCGGAAGAGCACACGTCT from the 3' end of the reads and a minimum read length of 17 nucleotides was established. Final reads were aligned against reference human genome hg38 (UCSC Genome Browser, <https://genome.ucsc.edu>) using the bioinformatic tool Bowtie (5), in a global alignment mode and -v2 -a -best -strata parameters. MiRNA annotation was performed following the miRBase Release 21 database (6). MiRNAs expression profiles and analysis of differential expression was conducted in the R/Bioconductor environment (<https://www.bioconductor.org>), specifically using featureCounts function (7) and edgeR package (8). FeatureCounts was run with the following options: strandSpecific = 1, allowMultiOverlap = TRUE, minReadOverlap = 16, and countMultiMappingReads = TRUE. Before searching those mature sequences differentially expressed between conditions, the count table was filtered out for those matures not expressed above 1 count per million (cpm) of mapped reads in at least five samples. Normalisation and common dispersion estimation were done using edgeR default parameters. Tagwise dispersion values were used in an exact

test with quantile-adjusted conditional maximum likelihood (qCML) methods to assess for differentially expressed matures, before multiple testing corrections with a 5% False Discovery Rate (FDR).

cDNA synthesis for eosinophil and serum miRNAs

Eosinophil miRNA samples were retrotranscribed into cDNA using Universal cDNA Synthesis kit II (Exiqon, Qiagen) according to manufacturer instructions. Briefly, 20 ng of total RNA were used in each reaction and were retrotranscribed into cDNA in a final volume of 10 μ L. In the case of sera samples, 2 μ L of isolated miRNAs were used for retrotranscription. UniSp6 RNA spike-in was used as control primer. The cDNA samples were stored at -20°C until used.

Eosinophil and serum miRNA quantitative PCR (qPCR)

Differential expressed miRNAs found in next generation sequencing (NGS) were validated in a different cohort by qPCR using ExiLent SYBR[®] Green master mix (Exiqon, Qiagen) according to manufacturer's instructions. Eosinophil miRNAs samples were obtained from a mild and moderate persistent asthmatic population.

Hsa-miR-103a-3p, hsa-miR-191-5p, RNU6B, hsa-miR-629-5p, hsa-miR-146b-3p, hsa-miR-144-5p, hsa-miR-590-3p, hsa-miR-1246, hsa-miR-1260a, hsa-miR-7977, hsa-miR-4443, hsa-miR-1260b, hsa-miR-4521, hsa-miR-5100, hsa-miR-664a-5p, hsa-miR-4425, hsa-miR-1290, hsa-miR-486-5p, hsa-miR-21-5p, hsa-miR-1-3p, hsa-miR-146b-5p, hsa-miR-146a-5p, hsa-miR-451a, hsa-miR-144-3p, hsa-miR-320a, hsa-miR-320b, hsa-miR-185-5p, Spike-in-Sp2, Sp4, Sp5 and Sp6 primers were purchased from Exiqon (Quiagen). Hsa-miR-103a-3p, hsa-miR-191-5p, and RNU6B were tested as endogenous

miRNAs. MiR-451a and miR-23a-3p were measured for calculation of serum hemolytic index (no hemolysis if [miR-23a-3p – miR-451a] < 7).

Previously, eosinophil and sera cDNA were diluted 1:80 and 1:40 respectively with nuclease-free water and 4 µL was used in each reaction. Each sample was run in triplicate. A two-step program was used as follows: 45 cycles of 95°C for 10 seconds and 60°C for 1 min. The reactions were carried out in a Light Cycler 96 thermocycler (Roche, Basel, Switzerland). Cycle threshold (Ct) data was taken in the last step of the melting cycle in a continuous way. Results were analysed with LightCycler® 96 SW 1.1 (Roche) software.

For the selection of endogenous controls possible candidates were tested (miR-191-5p, miR-103a-3p and RNU6B), and the ones with more homogeneous gene expression between patients were selected for each type of sample. Eosinophil miRNA sample Ct values were normalised to hsa-miR-103a-3p, and hsa-miR-191-5p was used as endogenous control for serum miRNAs. Their expression was calculated for each sample by the C_t value. Eosinophil miRNAs were examined for their relative expression, using the $2^{-\Delta\Delta C_t}$ method, and serum miRNA expression was calculated by using the $2^{-\Delta C_t}$ method (9), where:

$$\Delta\Delta C_t = \Delta C_{t_{Asthma}} - \Delta C_{t_{Healthy}} \text{ and } \Delta C_t = \Delta C_{t_{miRNA}} - \Delta C_{t_{endogenous control}}$$

Periostin determination by ELISA

Periostin levels were measured in serum samples by ELISA using the Human Periostin/OSF-2 DuoSet ELISA kit (R&D Systems, Minneapolis, MN, USA) according to the manufacturer's recommendations. Previously, serum samples were diluted 1:100 in 1X PBS with 1% BSA. Standard range was from 62.5 pg/mL to 4 ng/mL.

Bioinformatic analysis of differentially expressed miRNAs

Multivariate analyses of differentially expressed miRNAs were performed with ClustVis bioinformatic tool (10) by using principal component analysis (PCA) and unsupervised hierarchical clustering heatmap analysis. The two clusters for healthy and asthma were calculated with the Principal Component Analysis (PCA) for reducing the dimensionality of the multivariate data from the miRNAs expression. The PCA method was performed and our multivariate data set was linearly transformed into a set of uncorrelated variables, ordered in descending manner by the variance explained. For our specific miRNAs data for healthy and asthma, unit variance scaling was applied to rows and singular-value decomposition (SVD) with imputation was used to calculate principal components. The software represents prediction ellipses that are such that with probability 0.95, a new observation from the same group will fall inside the ellipse.

Statistical analyses

Comparisons between groups were performed using unpaired, two-tailed Student *t* test for Gaussian samples and Mann-Whitney U-test for non-Gaussian samples. ANOVA with Bonferroni *post-hoc* test was performed for comparisons between more than two groups of Gaussian samples, and Kruskal-Wallis with Dunn *post-hoc* test was applied for non-Gaussian ones. A *P*-value of less than 0.05 was considered as significant. In addition, Spearman's correlation was used for comparison and estimation of correlations between miRNA expression levels and clinical characteristics.

Receiver operating curve (ROC) analysis was performed for evaluation of specificity and sensitivity values of miRNAs expression levels for distinguishing between healthy and asthmatic patients. The Youden index was used to set the optimal cut-off point and was performed according to the mathematical formula:

$$\text{Youden index (J)} = \frac{\text{True positives}}{\text{True positives} + \text{false negatives}} + \frac{\text{True negatives}}{\text{True negatives} + \text{false positives}} - 1$$

A logistic regression model was performed using as predictive variables the cut-off values of individual miRNAs expression defined by the Youden index in the ROC curves, being those over 4.5 for miR-185-5p, over 0.5 for miR-1246, and over 0.17 for miR-144-5p to differentiate between healthy and asthma condition. The classification of the subject is made based on the probabilities estimated by the model, taking a cut-off value of 0.5. Two Random Forest models (11) were elaborated; briefly, Random Forest is a supervised machine learning method used for the classification of samples according to a set of variables. It works building a set of decision trees using random variables extracted from a training dataset. Each of the trees makes a classification of samples. All the decision trees generated are then put together to select the best classifiers and applied to a test dataset to generate a new classification. The new re-classification is then evaluated with the real classes in order to estimate the prediction power of the model. Our performance of Random Forest includes two extra features: i) a 10-fold cross validation approach, where the random forest is applied 10 times, each of the runs generates randomly a train (90% of the data) and test (10% of the data) datasets; and ii) a classes balance approach, where in each of the 10 Random Forest runs, the dataset is randomly re-sampled in order to keep similar number of samples per class to allow the algorithm learn from all the classes. The final accuracy, sensitivity and specificity are the mean of the 10 values generated by our program. The final AUC also includes the information from the 10 runs. The approach described was applied to the same datasets in order to build two models able to discriminate between: i) healthy controls and asthmatics, and ii) healthy controls and asthmatic phenotypes (intermittent, mild persistent, moderate persistent and severe persistent). In this second multiclass classification model, a sample is classified exactly in one of the five groups, being those

healthy, intermittent, mild persistent, moderate persistent and severe persistent. To feed the Random Forest models, we used the expression of miRNAs differentially regulated in eosinophils from asthma vs. healthy individuals, measured in serum from all samples. In both approaches, a first approximation was performed in order to select the three miRNAs with more weight for discrimination as reported by two importance parameters: mean decrease accuracy (MDA) and mean decrease Gini (MDG). After this, the Random Forest algorithm was run again only considering the selected miRNAs.-All statistical calculations and graphs were performed with GraphPad Prism 6 (GraphPad Software Inc., San Diego, CA, USA). ROC analysis, multivariate logistic regression model with qualitative predictors, and Random Forest Model using 10 cross-fold validations were performed using R (<https://www.r-project.org/>).

216
217

218 REFERENCES:

- 219 1. Standards for the diagnosis and care of patients with chronic obstructive
220 pulmonary disease (COPD) and asthma. This official statement of the American
221 Thoracic Society was adopted by the ATS Board of Directors, November 1986.
222 *Am Rev Respir Dis* 1987;**136**:225–244.
- 223 2. Plaza Moral V, Comité Ejecutivo de GEMA. [GEMA(4.0). Guidelines for
224 Asthma Management]. *Arch Bronconeumol* 2015;**51**Suppl 1: 2–54.
- 225 3. López E, Zafra MP, Sastre B, Gámez C, Fernández-Nieto M, Sastre J, *et al.*
226 Suppressors of cytokine signaling 3 expression in eosinophils: regulation by
227 PGE₂ and Th2 cytokines. *Clin Dev Immunol* 2011;2011 917015.
- 228 4. Martin M. Cutadapt removes adapter sequences from high-throughput
229 sequencing reads. *EMBnet.journal* 2011;**17**:10.
- 230 5. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient
231 alignment of short DNA sequences to the human genome. *Genome Biol*
232 2009;**10**:R25.
- 233 6. Kozomara A, Griffiths-Jones S. miRBase: integrating microRNA annotation and
234 deep-sequencing data. *Nucleic Acids Res* 2011;**39**:D152–157.
- 235 7. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program
236 for assigning sequence reads to genomic features. *Bioinformatics* 2014;**30**:923–
237 930.
- 238 8. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for

differential expression analysis of digital gene expression data. *Bioinformatics* 2010;**26**:139–140.

9. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 2001;**25**:402–408.

10. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res* 2015;**43**:W566–W570.

11. Liaw A, Wiener M. Classification and Regression by randomForest. At https://www.researchgate.net/profile/Andy_Liaw/publication/228451484_Classification_and_Regression_by_RandomForest/links/53fb24cc0cf20a45497047ab/Classification-and-Regression-by-RandomForest.pdf.

SUPPLEMENTARY FIGURES AND LEGENDS

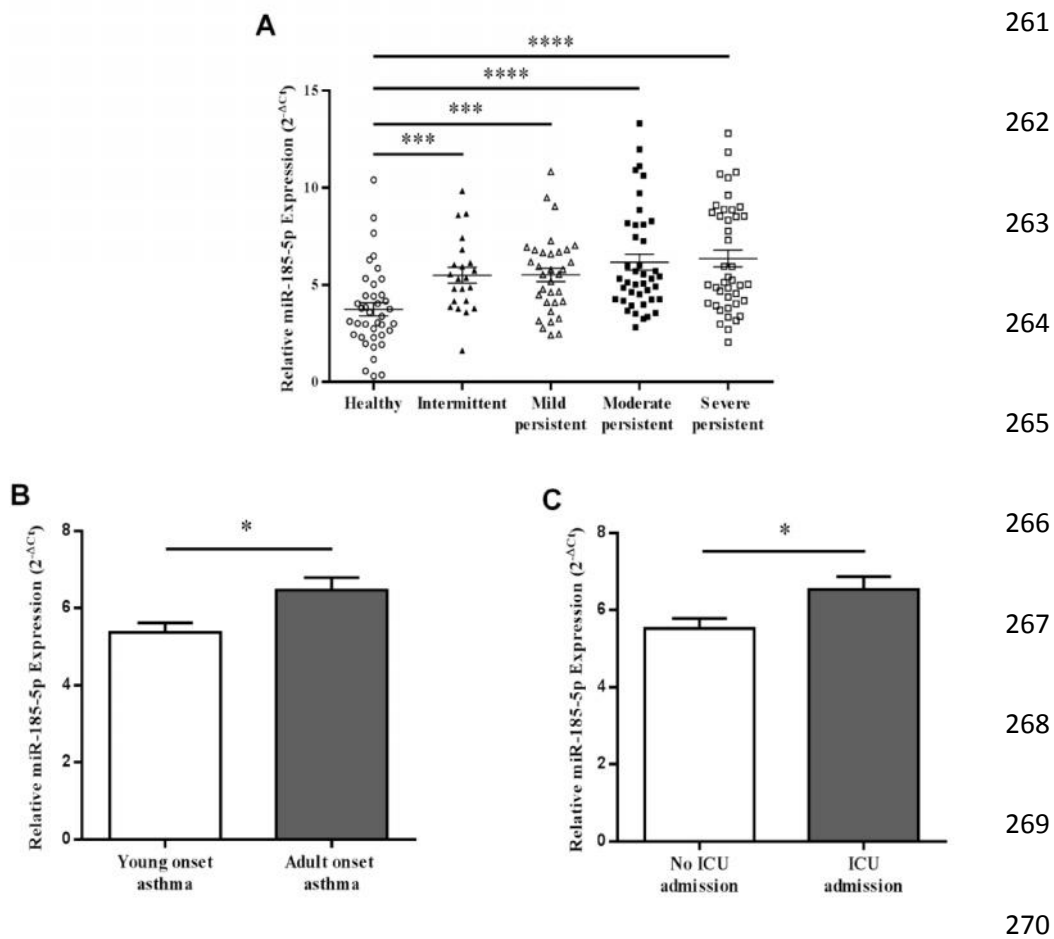


Figure S1. MiR-185-5p expression and asthma severity. A, Relative expression ($2^{-\Delta C_t}$) comparison between healthy ($n=39$) and intermittent ($n=22$), mild persistent ($n=33$), moderate persistent ($n=41$), and severe persistent ($n=42$) for miR-185-5p in serum samples. Results are expressed as dot plots with mean $\pm$ SEM. B, MiR-185-5p relative expression ($2^{-\Delta C_t}$) is higher in adult-onset asthmatics compared to young-onset asthmatics. C, MiR-185-5p relative expression ($2^{-\Delta C_t}$) is higher in the asthmatics who were admitted to the ICU at least once compared to those with no record of ICU admission. Results are expressed as mean $\pm$ SEM.

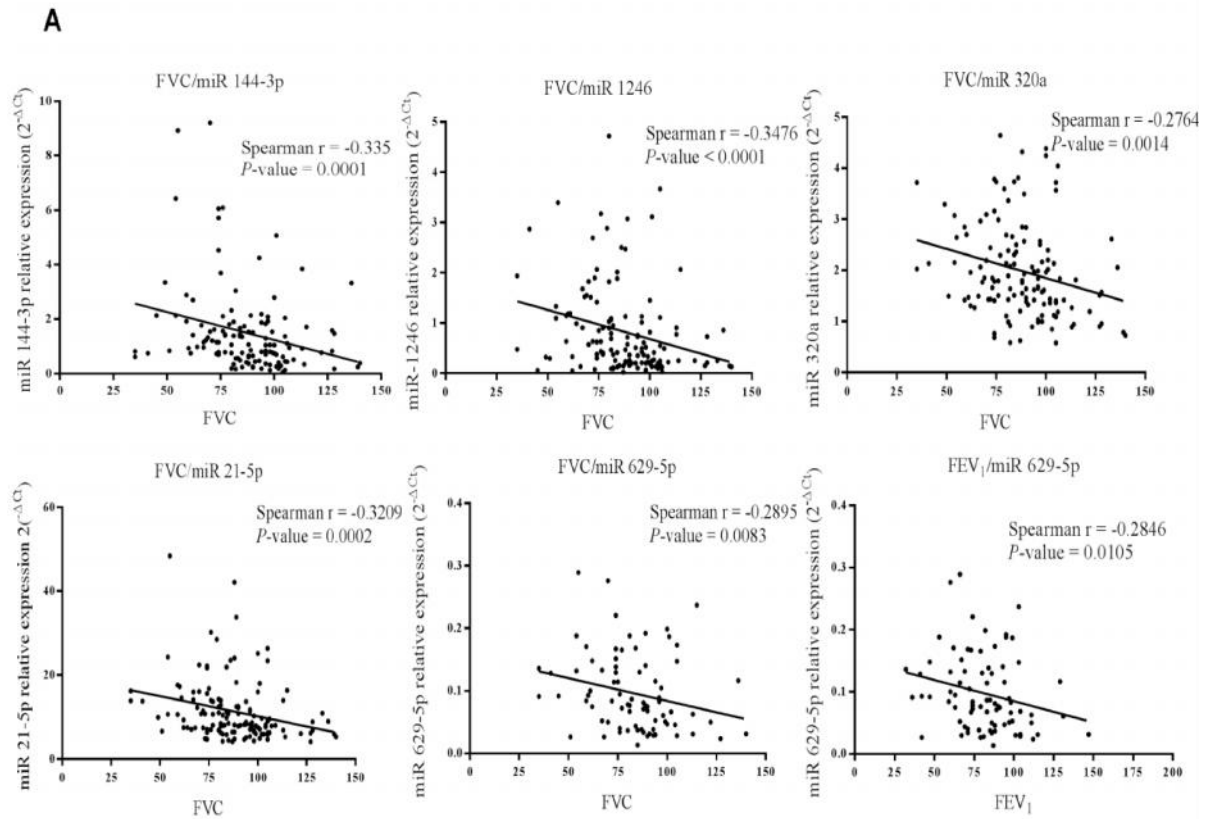


Figure S2. Correlation between clinical parameters and miRNA expression in asthmatic patients. MiR-144-3p, miR-1246, miR320a, miR-21-5p, and miR-629-5p expression ($2^{-\Delta C_t}$) are negatively correlated with FVC values in asthmatics. MiR-629-5p expression ($2^{-\Delta C_t}$) is negatively correlated with FEV₁ values in asthmatics.

287

288 **Table S1. Clinical and epidemiological characteristics of asthmatic patients and**
 289 **healthy subjects recruited into the next generation sequencing eosinophil**
 290 **miRNA study.**

	Asthma (n=6)	Healthy (n=4)
Age (years)*	33.17 (21-45)	34.20 (25-51)
Male (%)	0 (0.00)	3 (27.27)
Atopy (%)	3 (50.00)	3 (27.27)
Smoking history		
Current (%)	0 (0.00)	1 (9.09)
Ex-smoker (%)	2 (33.33)	3 (27.27)
Non-smoker (%)	4 (66.64)	7 (63.64)
FEV₁/FVC (%)†	78.66 (64.81-89.00)	NA
FeNO (ppb)*	49.60 (20.7-110.00)	NA
Total serum IgE (kU/L) †	50.00 (22.00-629.00)	NA
Eosinophils (%)*	3.48 (2.00-6.50)	<3.00
Eosinophils (per µL)*	275.00 (200.00-500.00)	<150.00

291 *Mean (range), †Median (range).NA (not available). FEV₁: forced expiratory volume in
 292 first second. FVC: forced vital capacity. FeNO: fractional exhaled nitric oxide.

Table S2. Clinical and epidemiological characteristics of the asthmatic patients and healthy subjects included in the qPCR miRNA validation study on eosinophils and classified by cluster.

	Healthy (n = 10)	Asthma (n = 29)	Asthma 1 Cluster (n = 11)	Asthma 2 Cluster (n = 18)	P-value
Age (years)*	34.20 (25-51)	39.38 (14-69)	37.36 (14-69)	40.61 (17-65)	NS
Male (%)	3 (27.27)	6 (20.69)	2 (18.18)	4 (22.22)	NS
Atopy (%)	3 (27.27)‡	20 (68.97)	6 (54.55)	14 (77.78)	§(P < 0.01)
Smoking history					
Current (%)	1 (9.09)	1 (3.45)	1 (9.09)	0 (0)	NS
Ex-smoker (%)	3 (27.27)	7 (24.14)	3 (27.27)	4 (22.22)	NS
Non-smoker (%)	7 (63.64)	21 (74.41)	7 (63.64)	14 (77.78)	NS
FEV₁/FVC (%)†	NA	75.89 (58.32-89)	77.00 (66-89)	75.66 (58.32-87.85)	NS
FeNO (ppb)*	NA	38.69 (7.20-110)	40.09 (7.20-84)	37.78 (11-110)	NS
Total serum IgE (IU/mL)†	NA	448.50 (22-6997)	50.95 (22-6997)	702.50 (268-864)	NS
Eosinophils (%)*	<3.00	5.13 (1.90-8.80)	3.17 (1.90-5.90)	5.80 (3.10-8.80)	** (P < 0.01)
Eosinophils (per µL)*	<150.00	368.42 (200-600)	250.00 (200-400)	409.09 (200-600)	** (P < 0.05)

*Mean (range), †Median (range). ‡P < 0.05 Healthy vs. Asthma. §P < 0.05 Healthy vs. Asthma 2, **P < 0.05 Asthma 1 cluster vs. Asthma 2 cluster. NA = (not available). NS = no signification. FEV₁: forced expiratory volume in first second. FVC: forced vital capacity. FeNO: fractional exhaled nitric oxide.

Table S3. Clinical and epidemiological characteristics of asthmatic patients and healthy subjects recruited for serum qPCR miRNA validation.

	Intermittent (n = 22)	Mild persistent (n = 33)	Moderate persistent (n = 41)	Severe persistent (n = 42)	Healthy (n = 39)	Asthmatics (n = 138)
Age (years)*	44.30 (26-86)	51.80 (27-82)	54.50 (16-86)	60.90 (29-88)	40.92 (21-79)	54.20 (16-88)
Atopy (%)	18 (81.81)	25 (75.75)	34 (82.92)	26 (61.90)	3 (7.69)	103 (75.63)
Smoking history						
Current (%)	2 (9.09)	6 (18.18)	5 (12.19)	1 (2.38)	2 (5.13)	13 (9.42)
Ex smoker (%)	3 (13.63)	6 (18.18)	12 (29.26)	9 (21.43)	0 (0.00)	30 (21.73)
Non-smoker (%)	17 (77.27)	21 (63.63)	24 (58.53)	32 (76.19)	16 (41.02)	94 (68.11)
FVC (%)†	99 (80-140)**‡	96 (61-136)**	84 (67-133)**	74 (35-139)	NA	87 (35-140)
FEV ₁ (%)†	97 (82-147)**	94 (60-129)**	84 (67-131)**	66 (32-98)	NA	84 (32-147)
Total serum IgE (kU/L)†	201 (70-1362)	249 (29-783)	132 (5-2000)	285 (4-1519)	NA	172 (4-2000)
Peripheral blood eosinophils (per µl)*	272.60 (3-470)	302.90 (98-433)	361.10 (5-1140)	440.90 (36-1100)	<350	365.2 (3-1140)
Broncodilators						
Long action (%)	3 (13.63)	19 (57.57)	41 (100.00)	42 (100.00)	0 (0)	105 (76.08)
B2 blockers (%)	10 (45.45)	15 (45.45)	28 (68.29)	34 (80.95)	0 (0)	87 (63.04)
Corticoid daily dose (µg)†	0 (0-400)**‡¶	320 (0-800)**‡	740 (160-1600)**	1600 (640-4000)	0.00	720 (0-4000)
Leukotriene receptor agonist (%)	2 (9.09)	2 (6.06)	17 (41.46)	30 (71.42)	0 (0)	51 (36.95)
Ipratropium bromide (%)	0 (0)	0 (0)	2 (4.87)	1 (2.38)	0 (0)	3 (2.17)
Tiotropium bromide (%)	0 (0)	0 (0)	1 (2.44)	5 (12.19)	0 (0)	6 (4.34)
Corticoids						
Oral (%)	1 (4.54)	29 (87.87)	41 (100)	42 (100)	0 (0)	113 (81.88)
Injected (%)	2 (9.09)	3 (9.09)	4 (9.56)	9 (21.43)	0 (0)	18/129 (13.04)
Systemic glucocorticoids (%)	2 (9.09)	8 (24.24)	9 (21.95)	25 (59.52)	0 (0)	44 (31.88)
ICU admissions (%)	6 (27.27)	18 (54.54)	13 (31.70)	23 (54.76)	0 (0)	60 (43.47)
>5 ICU admissions (%)	1 (4.54)	2 (6.06)	2 (4.87)	8 (19.04)	0 (0)	13 (9.42)
Periostin (ng/ml)*	129.3 (3.2-320.5)	109.69 (19.1-293.9)	145.8 (26-340.2)	150.6 (38.8-403.6)	115.3 (22.6-237.9) ^Δ	135.96 (3.2-403.6)

Results are shown as *Mean (range), †Median (range). FEV₁: forced expiratory volume in first second. FVC: forced vital capacity. ICU: intensive care unit. NA: not available ** $P < 0.05$ compared to severe persistent. ¶ $P < 0.05$ compared to mild persistent asthma. ‡ $P < 0.05$ compared to moderate persistent asthma. ^Δ $P < 0.05$ compared to moderate and severe persistent asthmatics.

Table S4. Statistic results for ROC curve analysis, between healthy and asthmatics, and between healthy and severity asthma groups.

	Biomarker	AUC	CUT-OFF ($2^{-\Delta Ct}$)	95% CI	P value	Sensitivity	Specificity
Asthma vs. Healthy	miR-185-5p	0.78	4.50	0.69 to 0.86	<0.001	0.688	0.769
	miR-320a	0.72	0.89	0.60 to 0.83	<0.001	0.941	0.514
	miR-1246	0.70	0.50	0.61 to 0.79	<0.001	0.496	0.865
	miR-144-5p	0.68	0.17	(0.59 - 0.77)	<0.001	0.587	0.774
	Periostin (ng/ml)	0.56	176.39	0.46 to 0.64	0.093	0.261	0.949
Intermittent vs. Healthy	miR-185-5p	0.76	3.59	0.64 to 0.89	0.002	0.955	0.538
	miR-320a	0.72	0.89	0.59 to 0.86	0.023	0.955	0.514
	miR-1246	0.68	0.236	0.53 to 0.82	0.011	0.773	0.541
Mild persistent vs. Healthy	miR-185-5p	0.76	4.07	0.65 to 0.87	<0.001	0.788	0.667
	miR-320a	0.71	1.11	0.57 to 0.84	0.041	0.938	0.571
	miR-1246	0.70	0.49	0.58 to 0.73	0.006	0.531	0.838
Moderate persistent vs. Healthy	miR-185-5p	0.79	4.18	0.69 to 0.89	<0.001	0.805	0.692
	miR-320a	0.68	0.89	0.56 to 0.81	0.068	0.951	0.514
	miR-1246	0.68	0.503	0.56 to 0.80	0.002	0.537	0.865
Severe persistent vs. Healthy	miR-185-5p	0.78	4.52	0.68 to 0.88	<0.001	0.690	0.769
	miR-320a	0.75	1.35	0.64 to 0.87	<0.001	0.825	0.657
	miR-1246	0.74	0.40	0.62 to 0.85	<0.001	0.625	0.784

