Serological profiling reveals hsa-miR-451a as a possible biomarker of anaphylaxis

Wojciech Francuzik, … , Magda Babina, Margitta Worm

*JCI Insight*. 2022. [https://doi.org/10.1172/jci.insight.156669](https://doi.org/10.1172/jci.insight.156669).

**Clinical Medicine**  **In-Press Preview**  **Immunology**

**Graphical abstract**

---

Find the latest version:

[https://jci.me/156669/pdf](https://jci.me/156669/pdf)
Serological profiling reveals hsa-miR-451a as a possible biomarker of anaphylaxis

Wojciech Francuzik*,1, Kristijan Pažur*,1, Magdalena Dalke1, Sabine Dölle-Bierke1, Magda Babina#,1, and Margitta Worm#,1,✉

* These authors contributed equally to this work.

# Shared senior authorship

1 Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Division of Allergy and Immunology, Department of Dermatology, Venereology and Allergology, Charitéplatz 1, 10117 Berlin, Germany

✉Correspondence: Prof. Dr. med. Margitta Worm <margitta.worm@charite.de>, Division of Allergy and Immunology, Department of Dermatology, Venereology and Allergology, Charitéplatz 1, 10117 Berlin, Germany, +49 30 450 518092

Keywords: anaphylaxis; hymenoptera; biomarkers; food allergy; hsa-miR-451a; tryptase; point-of-care; apolipoprotein; hsa-miR-25-3p; PGF2

MW discloses consulting fees and honoraria from Viatris, Aimmne, ALK and Advisory Board Membership in DBV-Technologies. MB discloses consulting fees and honoraria from escient and consulting fees from Health Advances. Other authors disclose no conflict of interest.
Abstract

Background: There is a need to support the diagnosis of anaphylaxis by objective markers. miRNAs are promising non-coding RNA species that may serve as serological biomarkers, but their use in diagnosing anaphylaxis has not been systematically studied. We aimed to comprehensively investigate serum biomarker profiles (proteins, lipids, and miRNAs) to support the diagnosis of anaphylaxis.

Methods: Adult patients admitted to the emergency room with a diagnosis of anaphylaxis (< 3 hours) were included. Blood samples were taken upon emergency room arrival and one month later.

Results: Next-generation sequencing of 18 samples (6 anaphylaxis patients in acute and non-acute condition, respectively, and 6 healthy controls) identified hsa-miR-451a to be elevated during anaphylaxis and verified by RT-qPCR in the remaining cohort. The random forest classifier enabled us to classify anaphylaxis with high accuracy using a composite model. We identified tryptase, 9α,11β-PGF2, apolipoprotein A1, and hsa-miR-451a as serological biomarkers of anaphylaxis. These predictors qualified as serological biomarkers individually but performed better in combination.

Conclusion: Unexpectedly, hsa-miR-451a was identified as the most relevant biomarker in our data set. We were also able to distinguish between patients with a history of anaphylaxis and healthy individuals with higher accuracy than any other available model. Future studies will need to verify miRNA biomarker utility in real-life clinical settings.
Funding: This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of the clinical research unit (CRU339): Food allergy and tolerance (FOOD@) and a grant to MW (Wo541-16-2); project numbers (409525714, 428094283, 428447634, 264921598).
Introduction

Anaphylaxis is a type I hypersensitivity reaction that is potentially life-threatening if not adequately treated. Although the simultaneous onset of skin symptoms and hypotension (in patients with a history of allergic reactions) suggests an anaphylactic episode – reactions may present atypically as skin symptoms may not be present in up to 20% of patients (1, 2). To more confidently administer adrenaline as the first-line treatment in cases where the elicitor of anaphylaxis is unknown and the diagnosis uncertain - there is a high medical need for objective laboratory parameters supporting the diagnosis of anaphylaxis (preferably using easy to acquire biological samples i.e., blood).

Currently, total mast cell tryptase is the only laboratory parameter routinely used for the diagnosis of anaphylaxis, but its moderate sensitivity and specificity require improvement (3). Moreover, the values of tryptase during acute reaction need to be related to their baseline levels in the same individual due to substantial inter-individual variability even on baseline (4). In addition, the relatively late diagnostic window for serum tryptase between 30 – 120 minutes after the onset of first anaphylaxis symptoms is too late concerning the necessity of rapid treatment initiation. Therefore multiple efforts have been undertaken in identifying other potential biomarkers of anaphylaxis (5–8)

miRNAs are small non-coding RNAs that regulate gene expression usually by silencing transcription (9). Their use as serological biomarkers has been proposed in various diseases (10). Data on the role of miRNA in human anaphylaxis is largely unknown. A recent study in pediatric oral food provocations of 4 children reported differential
expression of several miRNAs during the reaction, indicating their potential to assist in the
diagnosis of anaphylaxis (11). Since several miRNAs were shown to play a role in acute
asthma via airway inflammation and hyperresponsiveness (12), we hypothesized that
miRNA secretion may also occur in hypersensitivity reactions.

The aim of this study was therefore to identify biomarkers predictive of anaphylaxis with a
special emphasis on miRNAs.

Results

Study cohort characteristics

We prospectively recruited 81 patients from the emergency room (ER) with a suspected
diagnosis of anaphylaxis. Out of this cohort 31 patients had to be excluded from the
analysis due to low severity, or prolonged time (>3 hours) from the onset of the first
symptoms till blood draw in the ER or due to not fulfilling the diagnostic criteria for
anaphylaxis. Second sampling during a diagnostic visit was performed in 24 out of 47
anaphylactic patients at the Allergy and Immunology Division (Figure 1A). 24 healthy
individuals without a history of hypersensitivity reactions were recruited as controls. The
mean age of patients and healthy controls was 48.77 (median = 52) and 46.26 (median =
40) years respectively and the sex distribution was comparable (Figure 1B-C). The majority
of patients were White (European 85.11%, five, 10.64% were of Middle East descent), and
two (4.26%) were Asian (Far East). The classification was made by the investigators. Table
S1 summarizes the studied individuals.
The majority of cases in our cohort were venom-induced anaphylaxis (VIA, n = 22) and food-induced anaphylaxis (FIA, n = 19), followed by drugs (DIA, n = 3), and three cases of idiopathic anaphylaxis. VIA did not show significantly more severe reactions, classified as Ring and Messmer (13) grade III, than FIA (p = .089). There were no significant differences in severity between sexes (data not shown).

Skin symptoms were most prevalent regardless of elicitor. Symptoms of VIA were more often associated with the cardiovascular system (p = .036) whereas there was a tendency of greater gastrointestinal and respiratory involvement in FIA than in VIA without reaching significance (p = .386 and p = .204 respectively). There were no cases of respiratory or cardiac arrest (Figure 1E).

Identification of 19 miRNA candidates via next-generation sequencing

The initial screening of 6 anaphylactic patients (3 VIA and 3 FIA) resulted in successful labeling and counting of 2656 miRNAs after removing entities with less than 100 counts. We observed differential expression of 19 miRNAs, with p values < .05 after FDR adjustment using Benjamini-Hochberg correction, between samples taken from the patients in anaphylaxis and their respective baselines. (Figure 2A). The samples from an acute reaction clustered together when unsupervised clustering was performed, separating them from healthy and baseline as indicated in the dendrogram (Figure 2A).

Principal component analysis showed some but no precise differentiation across the groups (regardless of the number of differentially expressed miRNAs, Figure 2B).
the differential expression analysis (with the DESeq2 package for R (14)) we selected 5 miRNAs as potential candidate biomarkers for further study: hsa-miR-451a, hsa-miR-143-3p, hsa-miR-486-5p, hsa-miR-25-3p, and hsa-miR-484, performing RT-qPCR validation on the sera of patients who provided baseline samples.

hsa-miR-451a and hsa-miR-486-5p showed a significant difference in \( \Delta_{ct} \) values between groups and it was highly enhanced in anaphylaxis samples (Figure 2D). Notably, a control group of patients with atopic dermatitis, and no medical history of anaphylaxis, showed similar levels of hsa-miR-451a to healthy controls (Figure S6). Conversely, hsa-miR-25-3p and hsa-miR-484 showed significantly higher \( \Delta_{ct} \) in both anaphylaxis and baseline samples when compared to healthy controls, but they did not differentiate between anaphylaxis and baseline samples of the same patients. hsa-miR-143-3p did not show significant differences across the groups (Figure S1).

**Patterns of serological biomarkers in anaphylaxis - a collective view**

We observed elevated tryptase in anaphylaxis compared to baseline and healthy control samples (Figure 3). Moreover, we were able to confirm our recently identified markers 11\( \beta \)-prostaglandin F2\( \alpha \) (PGF2, increase) and apolipoprotein A1 (ApoA1, decrease), in this independent cohort (6, 8). In addition, we identified arachidonic acid (AA) to be significantly decreased in our anaphylaxis samples when compared to healthy controls (Figure 3D). No significant differences were observed for other potentially relevant proteins including cysteinyl leukotriene (cys-LT; Figure 3E), CCL13, CCL27, Apolipoprotein
E (ApoE), CCL17, eosinophil cationic protein (ECP), or chitinase-3-like protein 1 (YKL-40), as shown in Figure S1A – F.

To better understand the biomarker profiles during anaphylaxis we stratified the patients according to the elicitor and grade of reaction. Ring and Messmer grade III reactions mostly showed greater differences in the measured biomarkers when compared to healthy controls and baseline (Figure S2A). Biomarker levels were not altered by adrenaline therapy, except for AA, which was lower in patients who did not receive adrenaline (Figure S2B).

hsa-miR-451a was higher in patients undergoing anaphylaxis due to insect venom when compared to food-induced anaphylaxis (Figure S2C). Baseline levels of hsa-miR-25-3p were elevated in VIA both on baseline and during anaphylaxis when compared to FIA. YKL-40 and AA showed the expected trends in VIA but not in FIA, despite comparable levels in the baseline of both groups (Figure S2C). Levels of hsa-miR-25-3p were elevated on baseline and anaphylaxis in VIA when compared to FIA.

Cross-sectional time analysis of the biomarker levels in various time intervals from the studied cohort indicated that tryptase and PGF2 levels peaked (and ApoA1 bottomed) between 60 – 100 minutes after anaphylaxis onset. The hsa-miR451a showed a sideways trend in time, with the highest value observed at around 60 minutes. AA levels post anaphylaxis were lower in cases when the time interval from anaphylaxis onset to blood draw was extended. The highest levels of hsa-miR-25-3p were observed 120 minutes after the onset of anaphylaxis symptoms (Figure S3A – F).
Random forest accurately classified samples using a set of biomarkers

To assess the diagnostic value of the distinct biomarkers, we performed a classification of samples using the random forest algorithm. In the initial model, all candidates were supplied as predictors and the importance of each predictor was calculated, indicating hsa-miR-451a as the most important predictor (Figure 4A). To minimize the set of predictors, we performed recursive feature elimination and arrived at a model with 4 predictors having the highest accuracy (Figure 4C). Hyperparameter tweaking optimally indicated 2 randomly selected variables for each split (Figure S4B). We performed 10 runs of 10-fold repeated cross-validation using the caret package for R (15) and arrived at 0.93 area under the receiver operating characteristic curve for the top 4 predictors (out of bag error (OOB) = 12%, Figure S4A). This was 0.15 higher when compared to a model where random forest used 3 predictors for classification: tryptase, PGF2, and ApoA1. Thus, the addition of hsa-miR-451a increased the precision to detect an acute reaction.

To validate the model, we used the remaining 30% of randomly pre-selected observations (n = 9 during anaphylaxis; n = 8 on baseline). The final model correctly classified 16 cases (94.12%). Only one sample originating from the baseline group was incorrectly assigned as anaphylaxis (Figure 4B).

Enrichment analysis of the differentially expressed miRNAs

We used two computational tools (miEAA2.0 (16), MIENTURNET (17)) and FANTOM5 mammalian expression miRNA atlas (18) to predict the pathomechanistic relevance of
selected miRNAs in anaphylaxis. The miEAA2.0 analysis yielded 163 enriched subcategories (out of over 13,000) in 7 categories out of 19 used. Figure 5A summarizes the first 25 most significantly enriched subcategories.

Most of the enriched subcategories came from Gene Ontology (miRWalk) category, indicating enrichment in protein transport, carbohydrate metabolism, oxidative stress, inflammation through the TNF-α signaling pathway, and cell activation with the PI3 kinase pathway. Regarding the localization of miRNAs based on the RNALocate2.0 database (19), we saw enrichment in the exosome subcategory, as expected.

Concurrent analysis was performed using the MIENTURNET online tool (17) with 19 differentially expressed miRNAs (2 minimum miRNA per gene target and FDR < .05) using the miRTarBase database (20). Target prediction indicated highly significant involvement of trans-Golgi network proteins (Golgin A8 A and B, vacuolar protein sorting-associated protein) as well as signaling (Protein Phosphatase 1 Regulatory Subunit 37), carbohydrate metabolic processes (Hexokinase 2), and transcription regulatory processes (Zinc Finger Protein 264, RNA Binding Motif Protein 27) (Figure S5). Target genes identified in miRTarBase overlapped with the pathways identified by miEAA2.0 analysis.

Reactome (21) pathway analysis with the predicted miRNA target genes identified enrichment in apoptosis, IL-4 and IL-13 signaling, activation of metalloproteases, and multiple tyrosine kinases signaling (Figure 5B).
The cellular origin of the 12 most differentially expressed miRNAs was determined based on the data from the FANTOM5 Project (18) providing expression levels of known miRNAs in 118 different cell types. No expression of hsa-miR-3178 and hsa-miR-10400-5p was detected. Neutrophils, monocytes, and B-cells were the primary source of the most differentially expressed miRNAs and hsa-miR-451a was primarily expressed in neutrophils.

miRNAs identified patients prone to anaphylaxis

Parameters that distinguish patients at the risk of developing anaphylaxis remain unknown but are highly desirable for risk assessment and the implementation of prophylactic measures or even for retrospective diagnosis.

In the preceding data (Figure 6), we fortuitously observed that in addition to the steep increase in hsa-miR-451a during anaphylaxis, other entities were enhanced in baseline measurements vis-à-vis healthy controls. Since this provided a strong hint that miRNAs may serve in the identification of patients at risk, we employed the same procedure as above to examine this possibility. This was performed, by designing a random forest model. Indeed, the combination of 5 predictors (hsa-miR-484, hsa-miR-25-3p, hsa-miR-451a, ApoE, YKL-40) resulted in a random forest model showing good accuracy to differentiate between patients prone to anaphylaxis (i.e. baseline samples, from patients that had anaphylaxis ~4 weeks prior to sampling - B) and healthy controls (HC). The model was greatly improved against the reference model (tryptase, PGF2, ApoA1) with an area under ROC curve of 0.82 versus 0.46 respectively (0.36 difference, Figure 6D).
Though the predictive capacity of the model was somewhat limited due to the lower number of observations used for training (n = 22, OOB = 18.18%), the model still correctly assigned 66.67% of cases (Figure 6B) on the 30% of previously unseen data. While more data is needed, this is arguably the first successful step towards the detection of a biomarker panel to assist in the identification of patients prone to anaphylaxis.

Discussion

This study finds hsa-miR-451a as a novel and robust biomarker of anaphylaxis which may improve the diagnostic accuracy, especially when combined with other available biomarkers. In our prospective clinical observational study, we determined a higher expression of hsa-miR-451a in serum of patients with anaphylaxis when compared to healthy controls and non-acute sera and were able to model its diagnostic capacity using a random forest classifier. Importantly, we achieved a high classification accuracy when the model was tested on previously unseen real-life data.

The studied cohort adequately represented anaphylaxis cases in Europe. Insect venom was the most prevalent elicitor of anaphylaxis in our cohort, corresponding to the previously published data for Europe (22). Although US epidemiological data indicate medication followed by food as the primary triggers of anaphylaxis (23), food anaphylaxis is most prevalent in children (24) who were not included in this study. In accordance with our data, the clinical presentation of anaphylaxis with cardiovascular symptoms is known to be more prevalent in VIA than in FIA (1, 25) and the frequency of grade IV cases (Ring &
Messmer scale) is known to be low (26). We did not include cases with Ring & Messmer
grade I, as they did not meet the definition of anaphylaxis by FAAN/NIAID (27). In line with
previous reports, we observed more severe cases in VIA (1). One of the limitations of this
study was the lack of anaphylaxis due to drugs (we had only one paired case). Therefore,
this may be of interest for future studies.

Nevertheless, acquiring samples from patients in anaphylaxis is difficult as the main
concern is to provide immediate therapy to a patient with a potentially lethal allergic
reaction. Therefore, in some cases, a low amount or incorrect biomaterial was sampled
upon the ER visit. Fortunately, a substantial number of patients complied with a follow-up
visit to our Allergy and Immunology Division ~4 weeks post-reaction, so that patient-
matched serum samples during anaphylaxis and baseline were available for the study. This
was not possible with 48% of patients, yet we were able to use unpaired samples for the
random forest modeling, which allowed us to achieve good model accuracy. It needs to be
pointed out that in real-life scenarios, patient-matched baseline sera are unavailable in
most situations, so that biomarker models, that do not require the matched non-acute
serum from patients will best serve in the clinic.

This study did not only explore a new molecular category, rarely investigated in the context
of anaphylaxis, i.e., miRNAs, but it also comprehensively assessed them in regard to other
high confidence biomarkers to grade their utility. The selection of protein and lipid
biomarkers was made based on literature research and our findings.
Tryptase has been extensively studied as a biomarker of anaphylaxis. The gold standard involves the measurement of serum total tryptase during an acute phase followed by a baseline measurement (≥24 h) (28). Although tryptase levels correlate with the severity of anaphylaxis (29) and this is also visible in our data, two similarly designed studies conducted in emergency departments failed to confirm a predictive role of tryptase or histamine in the diagnosis of anaphylaxis (30, 31). Nonetheless, tryptase is the best-studied entity already in clinical use despite its limitations. However, we previously identified new candidates, demonstrating the predictive power of PGF2, cys-LTs, ApoA1, and ApoE in the diagnosis of anaphylaxis (6, 8). Notably, PGF2 and ApoA1 were not only reproduced in our current, broader study using an independent cohort, but they turned out to be in the top 4 classifiers of the random forest, and thereby, part of the best biomarker composite model to predict anaphylaxis. AA was selected because it is a precursor for PGF2 synthesis (32). Upon anaphylaxis, levels of AA were expected to decrease as it is metabolized to prostaglandins, leukotrienes, and thromboxanes, and this was indeed observed. The other candidates included were mainly based on their utility in other atopic or mast cell driven diseases. For example, CCL27 (CTAK) is secreted in the skin and seemed to be associated with urticaria (33). Although skin symptoms were highly prevalent during anaphylaxis in our cohort, there were only a few samples with an increase in CCL27 in serum. CCL13 (MCP-4) has been implicated in inflammatory processes in asthma (34, 35), CCL17 (TARC) is a biomarker of atopic dermatitis severity with a reported case of transient increase after food-induced anaphylaxis (36). YKL-40 has been implicated in atopic diseases (allergic rhinitis (37), asthma (38, 39), and food allergy (40)). Raised ECP was described in food
anaphylaxis (41) and food provocation (42). However, none of them qualified as a potential biomarker of anaphylaxis in our study.

ApoA1 was discovered through a proteomics screen in our previous study (8). Since the strategy proved successful in biomarker exploration, we reapplied an analogous unbiased approach to exploit the untapped potential of miRNAs. The use of paired biosamples (during anaphylaxis and baseline) allowed us to achieve decent group discrimination in unsupervised hierarchical clustering, despite the relatively low number of sequenced samples. By using stringent differential expression analysis and state-of-the-art tools (14, 43), we were able to identify 19 significantly expressed miRNAs and tested extensively 5 promising candidates by RT-qPCR in the whole cohort. The NGS data could be reproduced on a larger cohort indicating the usefulness of this approach for the identification of serological biomarkers.

During modeling, we used a reference random forest model with 3 predictors that have previously shown predictive potential in diagnosing anaphylaxis. In our previous publication, a composite linear model using ApoA1 and PGF2 (8) was able to discriminate between anaphylaxis cases and healthy controls (but the results were not cross-validated). Therefore, we benchmarked our modeling to a presumably best performing composite model using ApoaA1, tryptase ad PGF2. The addition of hsa-miR-451a significantly improved the model accuracy and proved valid upon testing with previously unseen data. Notably, hsa-miR-451a was more important as a predictor for the random forest classifier than tryptase.
Regarding cellular origin, hsa-miR-451a was identified to be expressed by neutrophils and monocytes, based on the differentially expressed miRNAs in FANTOM5 (18). In addition, hsa-miR-451a is known to be expressed by erythrocytes (44–47), which were not included in the FANTOM5 cell atlas. Based on that, the increase in serum levels of hsa-miR-451a might be: 1) co-released into serum along with cellular activation and degranulation processes; 2) linked to the degradation or release from erythrocytes upon anaphylaxis.

Future studies will have to address the major cellular sources operative in anaphylaxis and mechanisms of release (direct versus effector cell driven, e.g., via mast cells or basophil activation). Although, we demonstrated that hsa-miR-451a was not increased in other atopic disease (atopic dermatitis) it would be crucial to verify its specificity when compared to other shock syndromes (i.e., myocardial infarction, sepsis and hypotensive shock).

An intriguing and unexpected outcome of this study was the differential miRNA profile in anaphylaxis prone patients (anaphylaxis patients sampled during baseline) versus healthy individuals. Not all individuals with an increased sIgE to a particular allergen respond to the allergen in question and the severity of the reaction and organ involvement are likewise unpredictable by sIgE levels or other diagnostic tests such as skin prick test or basophil activation test (48). Most importantly, the use of skin prick tests or sIgE failed to be predictive of reaction severity in food allergic children (49). Currently, the only usable predictors of future anaphylaxis severity are clinical symptoms of the previous reaction and the presence of the cofactors of anaphylaxis (50). The European Guidelines for the
Management of Anaphylaxis (PRACTALL) underline the need for an objective, predictive biomarker of future reaction severity (51). At the moment there are no point of care tests available for anaphylaxis and the first step towards achieving this goal is to confirm miRNAs value as a diagnostic tool in clinical setting. With the rapid development in nucleic acid tests seen during the COVID-19 pandemic (52) we can expect increasing use of similar technology for the diagnosis of other diseases.

It is therefore a major unmet clinical need to identify those patients who are at risk of severe hypersensitivity reactions based on serum biomarkers alone. This knowledge can be used to increase patients’ awareness, provide additional prophylactic measures, and even to help in the differential diagnosis of anaphylaxis retrospectively using serum. So far, no single or composite biomarkers are available to meet this aim. Here we find that our proposed composite biomarker (hsa-miR484, hsa-miR-25-3p, hsa-miR-451a, ApoE, and YKL-40) could distinguish between anaphylaxis prone and healthy with a higher accuracy than any other available model. Interestingly, the miRNAs that distinguish anaphylaxis prone from healthy individuals were different than the aforementioned hsa-miR-451a, the latter best suited to diagnose acute symptom precipitation. This was further emphasized by random forest analysis whereby the most significant predictors of the two independent questions (i.e. to diagnose acute anaphylaxis and to identify patients at risk) did barely overlap (compare Figure 4 with Figure 6). In fact, of all biomarkers measured, only miRNAs turned out to be reasonable predictors of anaphylaxis risk, highlighting the major potential of this class of biomolecules. Thus, the results can serve as a basis to implement miRNAs in
the analytic make-up to diagnose an acute anaphylaxis attack. Yet more, they also provide a rationale for future research to prospectively identify patients at risk based on serological miRNA profiles alone.

Methods

Inclusion criteria, sample collection and storage

We obtained serum samples from adult patients undergoing anaphylaxis who were admitted to the ER and provided written informed consent, as well as control (baseline) serum samples at least 1 month after anaphylaxis. The diagnosis of anaphylaxis was made according to the National Institute of Allergy and Infectious Disease/Food Allergy and Anaphylaxis Network criteria (27). We also acquired serum samples from healthy individuals who did not report a history of allergic diseases as well as a second control group consisting of atopic dermatitis patients without previous medical history of anaphylaxis. After collecting the whole blood, it was left undisturbed at room temperature for 20 to 30 minutes to allow blood to clot. Samples were then centrifuged in a pre-chilled centrifuge with a horizontal rotor (swing-out head) for 5 minutes at 1500 g at 4°C, aliquoted, and stored at -80°C.

Protein and lipid biomarker measurement

Human serum samples were analyzed using ELISA kits provided by R&D Systems, Minneapolis, USA (human CCL17, DY364; human CCL13, DY327; human CCL27, DY376;
human YKL-40, DC3L10), Cusabio Technology, Houston, Texas (Human Arachidonic Acid, CSB-E09040h; Human Eosinophilic Cationic Protein, CSB-E11729h), Cayman Chemical Company, Ann Arbor, USA (11β-Prostaglandin F2α, 516521; Cysteinyl Leukotriene, 500390-96), Abcam, Cambridge, UK (Human Apolipoprotein A1, ab108804), Raybiotech Life, Norcross, USA (Human ApoE, ELH-ApoE) following the manufacturer’s protocol. The total concentration of tryptase in serum was measured by ImmunoCAP (Thermo Fisher Scientific, Waltham, USA).

**Serum miRNA extraction and profiling by NGS sequencing**

Circulating miRNAs were isolated from 200 µl serum by the miRNeasy Serum/Plasma kit (Qiagen, 217184) and quantified with Agilent 2100 Bioanalyzer system. Library prep was performed using the QIASeq miRNA library kit (Qiagen, 331502) according to the manufacturer’s instructions. Briefly, a total of 5 µl RNA was used for miRNA library preparation. After ligation of the RNA 3’- and the RNA 5’ adapter, universal cDNA synthesis with UMI assignment, cDNA cleanup, library amplification, and library cleanup were performed. Library preparation quality control was performed using TapeStation (Agilent Technologies).

Sequencing was performed on Illumina miSeq Platform. Samples from 6 patients undergoing anaphylaxis were compared to their corresponding sera taken outside of an allergic reaction (i.e., baseline). The obtained reads were initially trimmed with cutadapt version 2.4 (53) using the first 10 bases of the Qiagen 3’ adapter sequence. Reads shorter than 10 bases after trimming were removed. Over 96% of reads in each sample contained
an adapter. Reads were aligned to the mirBase (release 22.1) human micro-RNA sequences (54) and counted using the Mirdeep2 package (55) (v. 2.0.1.2). To this end, reads were trimmed as described above, converted to the FASTA format and collapsed using the collapse_reads_md.pl program. Quantification was performed with the quantifier.pl program. Quality control was performed using fastQC v. 0.11.8 (56), RNA-Seqc, and dupradar (57).

Data availability

The sequencing data presented in this study have been deposited in the European Nucleotide Archive (ENA) under accession number PRJEB50710.

Differential expression analysis

Differential expression of miRNA sequencing data from 6 paired samples was analyzed using the DESeq2 package (14) for R (58), with log fold change shrinkage using the Ashr algorithm (59). Stringent false discovery rate adjustment was applied for the p values (60). A heatmap of differentially expressed miRNA was produced using the heatmaply package (61), with row wise-scaling and automatic clustering using euclidean distances.

Validation of miRNA using quantitative RT-PCR

The differentially expressed miRNAs were further validated using reverse transcription quantitative real-time PCR (RT-qPCR). Briefly, total RNA was isolated from 200 µl serum by the miRNeasy Serum/Plasma kit (Qiagen, 217184) according to the manufacturer’s instructions. In addition, 3.5 µl miRNeasy serum/plasma spike-in control (Qiagen, 219610)
at $1.6 \times 10^8$ copies/μl was added to each sample. The total RNA was reverse transcribed using miRCURY LNA RNA kit (Qiagen, 339340) that generates universal cDNA templates for all miRNAs present in the sample. The synthetic spike-in (UniSp6, Qiagen, 339340) was added to each sample, and the reaction was performed in the GeneTouch thermal cycle (Bioer). Then, miRNA-specific quantification was performed using miRCURY LNA SYBR Green kit (Qiagen, 339347) according the manufacturer’s instructions. The expression of target miRNAs was normalized to the cel-miR-39-3p synthetic spike-in added during total RNA extraction.

Classification model

The supervised machine learning classification model was done using a random forest algorithm with the help of the caret package (15). All samples with complete observations (without missing values) were included in the training set and divided into anaphylaxis and baseline groups (the latter contained healthy controls and patients’ baseline samples). Features for the final random forest model were selected by a recursive feature elimination algorithm on the set of initially available biomarkers. Variable importance was derived from the initial random forest model, which included all predictors. Hypeparameter optimization was internally performed by the caret package to identify the optimal number of randomly drawn candidate variables at each decision tree split, based on the Out-of-Bag error estimate (62). The model fitness was calculated using 10 sets of 10-fold repeated cross-validation.
The functional miRNA set enrichment analysis was performed using the miEAA2.0 online tool (16). Briefly: 1) an ordered set of 500 most differentially expressed miRNAs based on the output from DESeq2 (adjusted p-values of the 6 screened anaphylaxis patients and their corresponding baseline samples) was provided for the miEAA2.0. 2) Subsequently the algorithm cross-referenced known and predicted miRNA-gene interaction targets restricting the significance level to .005 and the minimum required miRNA hits per subcategory of 5. We used FDR (Benjamini-Hochberg) p value adjustment for the whole set of analyses in all categories. 3) The resulting miRNA list for miRNA Set Enrichment Analysis (miRSEA) in all available 19 default gene-set categories including miRTarBase and miRWalk - for gene ontology, Cell type specific atlas, Immune Cells gene sets, Localization - RNALocate and pathways based on miRWalk database). 25 most significantly enriched sets were subsequently plotted to indicate the presence (or lack thereof) of 13 most significantly differentially expressed miRNAs in each of the gene sets.

Additionally, a set of 19 most differentially expressed miRNAs was provided for the MIENTURNET (17) online tool. After predicting the gene targets using mirTarBase 8.0 (20) the resulting gene sets were used in an enrichment analysis using the Reactome database (21) with 2 intersections.

Data from the FANTOM5 Project - promoter level mammalian expression atlas (18) were used to determine the cellular origin of most differentially expressed miRNAs. Relative
abundance was calculated as a ratio of each candidate miRNA expression in a specific cell type (tags per million) to the global expression level in all characterized cells.

Statistics

T Student’s test was used for comparing normally distributed values between unpaired observations, with Holm’s p value correction for multiple comparisons. Paired data was analyzed using paired T Student’s test with Holm’s correction where appropriate. Next-generation sequencing derived data were analyzed using Wald test with Benjamini-Hochberg FDR to correct for multiple comparisons. P values < .050 were considered significant.

Study approval

Institutional Review Board of the Charité Universitätsmedizin Berlin approved this study EA1/079/06. Written informed consent was acquired prior to participation in this study.

Author contributions

WF wrote the initial manuscript, analyzed the data and designed the model and the computational framework. KP wrote the initial manuscript, conceived and performed the experiments and analyzed the data. (WF and KP contributed equally to this work, names in alphabetical order). MD acquired the clinical samples, verified the analytical methods, performed the experiments and revised the manuscript critically. SDB contributed to the design and implementation of the research, to the analysis of the results and revised the
manuscript critically. MB conceived the analytical approach, developed the theoretical framework, supervised the findings of this work and wrote the manuscript. MW conceived of the presented study, supervised the clinical aspect of the study, developed the theoretical framework, aided in interpreting the results, and revised the manuscript critically.

Acknowledgements

We acknowledge the extraordinary assistance of Dennis Ernst and Dr. Davender Redhu in laboratory samples processing as well the support of the BIH Core Facility Genomics Team in our sequencing efforts. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as a part of a clinical research unit (CRU) 339 “Food Allergy and Tolerance” and a grant to MW (Wo541-16-2); project numbers (428094283 and 428447634, 264921598). The Supporting Institution had no influence on the study design, data, nor the contents of this manuscript.
References


15. Kuhn M. Caret: Classification and regression training. 2020:


44. Sun L et al. Red blood cells as potential repositories of MicroRNAs in the circulatory system. *Front Genet* 2020;11.


49. Harleman L, Sie A. History, blood tests or skin prick testing? Is it possible to predict the severity of allergic reactions in children with IgE-mediated food allergy? *Arch Dis Child* 2015;100(6):594.1–598.


51. Muraro A et al. Precision medicine in allergic disease-food allergy, drug allergy, and anaphylaxis-PRACTALL document of the european academy of allergy and clinical


**Figure 1:** Studied cohort shows clinically relevant reactions. A) Flowchart illustrating the exclusion criteria of the study. B – C) Age and sex distribution in the studied cohort and their corresponding controls (ANA - acute anaphylaxis, HC - healthy control group, two-tailed Student's T test, n = 71). D) Distribution of elicitors and severity of anaphylaxis according to Ring and Messmer (VIA – venom-induced anaphylaxis, FIA – food-induced anaphylaxis). E) Frequency of symptoms in food and insect venom elicited anaphylaxis.
Figure 2: Screening and validation of miRNA candidate biomarkers show different expression in anaphylaxis when compared to healthy and baseline. A) Heatmap of 19 most differentially expressed miRNAs (using next generation sequencing) in 6 anaphylaxis samples (top row red) in comparison to the corresponding samples on baseline (top row yellow) applying unsupervised hierarchical clustering with euclidean distances. Additional 6 healthy control samples (top row brown) were provided for completeness. B) Principal component analysis of the sequenced cohort in corresponding colors (50 most differentially expressed miRNAs, data were scaled and centered). C) Volcano-plot illustrating differentially expressed miRNA upon anaphylaxis (n = 6) compared to baseline (n = 6). Fold-changes were adjusted using the ‘ashr’ algorithm. Red dots mark miRNAs with adjusted log-fold change > 1 and < -1. Red dashed line indicates adjusted p < 0.05 (Wald test with Benjamini-Hochberg FDR, n = 2656 miRNAs per sample). D – G) Quantification of selected miRNAs in serum using real-time qPCR on anaphylaxis patients’ sera in acute (ANA, n = 16), baseline (B, n = 16), and healthy controls (HC, n = 20). Comparisons between HC vs. B and HC vs. ANA - two tailed Student’s T test for unpaired data with Holm’s correction; comparisons between B vs. ANA - two tailed Student’s T test for paired data.
Figure 3: **Protein and lipid candidate serum biomarkers differentiate between anaphylaxis and baseline.**
A) ImmunoCAP measurement of tryptase in serum; B – E) ELISA measurement of selected biomarkers in healthy controls (HC), baseline (B), and acute anaphylaxis (ANA). Comparisons between HC (n = 24) vs. B (n = 24) and HC vs. ANA (n = 24) - two tailed Student’s T test for unpaired data with Holm’s correction; comparisons between B vs. ANA - two tailed Student’s T test for paired data.
Figure 4: Serological biomarkers allow predicting an acute anaphylactic episode using random forest classification model. A) Feature importance score using random forest algorithm with all potential biomarkers of anaphylaxis. B) Summary of sample selection for the algorithm training and test sets. C) Automatic selection of best features out of the 18 initial ones reported in panel A using recursive feature elimination (RFE) algorithm [15]. D) Receiver Operating Characteristic (ROC) in a random forest model with 4 top predictors (hsa-miR-451a, tryptase, PGF2, apolipoprotein A1 - red curve) compared to a model including tryptase, apolipoprotein A1 and PGF2 - black curve.
Figure 5: miRNA enrichment and localization analysis indicates the role of most differentially expressed miRNAs in protein transport, and cell signaling processes in immune cells. A) Enrichment analysis on the miRNA set produced by differential expression analysis using Enrichment Analysis and Annotation Package 2.0 (miEAA2.0). 1 most differentially expressed miRNAs are visualized according to gene set membership in 25 most relevant gene sets. (Q-value < 0.005 - p value adjusted using Bejamini-Hochberg FDR). B) Result of miRNA enrichment analysis using MIENTURNET (17). Functional enrichment using the Reactome pathways (21). C) Cell specific expression of miRNA in FANTOM5 mammalian shortRNA expression atlas (18)
Figure 6: **Serological biomarkers show potential to predict patients at risk of anaphylaxis.** A) Feature importance score using random forest algorithm with all potential biomarkers. B) Flowchart for the modeling procedure including model evaluation using repeated cross-validation. FPR - False Positive Rate, FNR - False Negative Rate. C) Automatic selection of best features out of the 17 initial ones reported in panel A using recursive feature elimination algorithm. Receiver Operating Characteristic area under the curve (AUC) in a random forest model with 5 predictors (hsa-miR-484, hsa-miR-25-3p, hsa-miR-451a, apolipoprotein E, YKL-40 - red curve) compared to a model including tryptase, apolipoprotein A1 and PGF2 - black curve.