Analysis of a nanoparticle-enriched fraction of plasma reveals miRNA candidates for Down syndrome pathogenesis

ALESSANDRO SALVI1, MARINKA VEZZOLI2, SARA BUSATTO1, LUCIA PAOLINI1,3, TERESA FARANDA1, EDOARDO ABENI1, MARIA CARACAUSt4, FRANCESCA ANTONAROS4, ALLISON PIOVESAN4, CHIARA LOCATELLI5, GUIDO COCCHI5,6, GUALTIERO ALVISI7, GIUSEPPINA DE PETRO1, DORIS RICOTTA1, PAOLO BERGESE1,3 and ANNALISA RADEGHIERI1,3

1Department of Molecular and Translational Medicine, University of Brescia; 2Unit of Biostatistics, Department of Molecular and Translational Medicine, University of Brescia, I-25123 Brescia; 3CSGI, Research Center for Colloids and Nanoscience, Sesto Fiorentino, I-50019 Florence; 4Department of Experimental, Diagnostic and Specialty Medicine (DIMES), Unit of Histology, Embryology and Applied Biology, University of Bologna; 5Neonatology Unit, St. Orsola-Malpighi Polyclinic; 6Department of Medical and Surgical Sciences (DIMEC), University of Bologna, I-40138 Bologna; 7Department of Molecular Medicine, University of Padua, I-35121 Padua, Italy

Received November 16, 2018; Accepted March 20, 2019

DOI: 10.3892/ijmm.2019.4158

Abstract. Down syndrome (DS) is caused by the presence of part or all of a third copy of chromosome 21. DS is associated with several phenotypes, including intellectual disability, congenital heart disease, childhood leukemia and immune defects. Specific microRNAs (miRNAs/miR) have been described to be associated with DS, although none of them so far have been unequivocally linked to the pathology. The present study focuses to the best of our knowledge for the first time on the miRNAs contained in nanosized RNA carriers circulating in the blood. Fractions enriched in nanosized RNA-carriers were separated from the plasma of young participants with DS and their non-trisomic siblings and miRNAs were extracted. A microarray-based analysis on a small cohort of samples led to the identification of the three most abundant miRNAs, namely miR-16-5p, miR-99b-5p and miR-144-3p. These miRNAs were then profiled for 15 pairs of DS and non-trisomic sibling couples by reverse transcription-quantitative polymerase chain reaction (RT-qPCR). Results identified a clear differential expression trend of these miRNAs in DS with respect to their non-trisomic siblings and gene ontology analysis pointed to their potential role in a number of typical DS features, including ‘neural crest development’, ‘neuronal cell body’ and certain forms of ‘leukemia’. Finally, these expression levels were associated with certain typical quantitative and qualitative clinical features of DS. These results contribute to the efforts in defining the DS-associated pathogenic mechanisms and emphasize the importance of properly stratifying the miRNA fluid vehicles in order to probe biomolecules that are otherwise hidden and/or not accessible to (standard) analysis.

Introduction

Down syndrome (DS) is the most frequent human chromosomal disorder with a frequency of 1/400 conceptions and 1/1,000 births worldwide (1,2). Since the initial discovery of Lejeune et al (3), it is known that the presence of full or partial chromosome 21 (Hsa21) in three copies (trisomy 21) in the cells of the affected subjects is responsible for the typical features of DS, in particular intellectual disability (ID), cardiovascular defects (4,5) and craniofacial dysmorphism. Importantly, a highly restricted ‘Down syndrome critical region’ of 34 kb on distal 21q22.13 appears to be specifically duplicated in all individuals with DS (6,7). In addition, a recent study conducted on plasma and urine samples of individuals with DS demonstrated a systematic deviation of metabolites involved in central metabolic processes associated with mitochondrial metabolism including the Krebs cycle, glycolysis and oxidative phosphorylation in DS (8).

Despite intense efforts of the scientific community, the pathogenic mechanism linking chromosome 21 and DS is remains largely uncharacterized, and pharmacological therapies targeting genes located on Hsa21 have not been developed so far. Therefore, current research shifted to study the involvement of non-coding RNAs (e.g., microRNAs) in the process (7). microRNAs (miRNAs/miR) are small single stranded nucleic acids that regulate gene expression post-transcriptionally via binding to different mRNA targets, resulting in inhibition of mRNA translation (9,10). Increasing evidence indicates that miRNAs serve important roles in a large variety of biological...
processes including development, differentiation, proliferation and apoptosis (11). It is well documented that intracellularly produced miRNAs can be secreted in the extracellular milieu, mostly associated with Ago proteins, bound to lipoproteins or secreted in extracellular vesicles (EVs) (12-15). Due to their nature, the latter two may provide an enriched and preserved source of miRNAs (16-19), therefore they may be better associated with the disease state than total plasma analysis (20,21), allowing a deeper understanding of the ongoing pathological processes and complementing previous studies.

Intriguingly, several circulating miRNAs have been differentially retrieved from either placenta or plasma samples of euploid and subjects with DS (22,23). However, only a few of them (miR-99a, miR-125b, let-7c and miR-155) were mapped on Hsa21 (24). It should also be noted that in other studies no Hsa21-derived miRNAs were identified to be differentially expressed in association with DS (25,26). This variability may derive from the type of sample analyzed and the extraction procedure used to recover the nucleic acids.

The present study focused for the first time on a subset of miRNAs carried by circulating biogenic nanoparticles, namely EVs and high-density lipoproteins (HDLs). Biogenic nanoparticles formulations, also referred to as nanoparticle-enriched fraction (NEF), were obtained from the plasma of young subjects with DS and their miRNA expression profiles systematically compared with those of NEF from their non-trisomic siblings. Interestingly, three novel candidate miRNAs emerged from the present study (miR-16-5p, miR-144-3p and miR-99b-5p) that could help unraveling molecular aspects underlying the complex DS pathogenic phenotype.

Materials and methods

Subject enrolment. For all participants involved in the present study (approved by the competent Ethics Committee of Sant’Orsola-Malpighi Hospital in Bologna, Italy; approval no. 39/2013/U/Tess) written informed consent was obtained from the parents in the case of subjects with age below 18 years and from the subjects themselves if aged over 18 years, according to the approved protocol.

Patient enrolment was performed at the Unit of Neonatology of Sant’Orsola-Malpighi Hospital in Bologna, Italy from February 2014 to March 2016. All methods were performed in accordance with ethical principles for Medical Research involving human subjects of the Helsinki Declaration.

The exclusion criteria of the enrolment to the study were: Age <2 years old, distress at birth, strongly premature (EG <35 weeks) and severe central nervous system (CNS) disease at birth. The presence or absence of each clinical feature and comorbidity was assessed based on the agreed clinical judgment of at least two pediatricians with long-time experience in the follow-up of children with DS, aiming to assess if each subject fell in the relative top, medium or lower class of cognitive and global performance within the population.

Collection of blood samples. Blood samples (3 ml) were collected in EDTA-coated blood collection tubes and processed within 2 h from blood draw (27,28). All traceable identifiers were removed prior to analysis to protect patient confidentiality, all samples were analyzed anonymously (29). The sample was transferred in a new tube and centrifuged at 1,250 x g for 10 min at room temperature (centrifuge ALC 4235 A, rotor ALC T111). The plasma fraction was isolated and centrifuged for a second time at 800 x g for 30 min at room temperature (centrifuge ALC4214, rotor 6642) and the supernatant was transferred to new tubes without touching the pellet or the bottom of the tube, and divided in aliquots of 300 µl. All plasma samples were rapidly frozen and stored at -80°C until needed for subsequent analysis.

The exclusion criteria of plasma samples from the subsequent analysis were: Blood sample treatment after 2 h from the draw, or evident contamination of plasma samples by residual erythrocytes at the end of the treatment.

NEF separation from plasma and miRNA isolation. A total of 300 µl of plasma were used to obtain the NEF with Exosome Precipitation Solution (Serum/Plasma; Macherey-Nagel GmbH, Düren, Germany), according to manufacturer’s protocol. Briefly the plasma was centrifuged at 10,000 x g for 15 min at 4°C to pellet intact cells, cellular debris and bigger vesicles. Then the supernatant was mixed with ‘Exosome precipitation solution’, incubated for 30 min on ice and finally centrifuged at 600 x g for 5 min at 4°C. The supernatant was kept for further analyses. The pellet was then resuspended in 300 µl RNAase free water and miRNAs were extracted using ‘NucleoSpin miRNA Plasma’ kit (Macherey-Nagel GmbH), according to the manufacturer’s protocol.

For reverse transcription-quantitative polymerase chain reaction (RT-qPCR) analyses, 3 µl of a 4.16 nM solution of the synthetic miRNA cel-miR-39-3p from C. elegans (custom synthesized by Integrated DNA Technologies, Coralville, IO, USA) were added. miRNA fractions were eluted in 35 µl of nuclease-free water.

Western blotting. Proteins from NEF preparations obtained from C and DS plasma and associated supernatants were extracted with reducing SDS sample buffer (80 mM Tris, pH 6.8, 2% SDS, 7.5% glycerol, and 0.01% bromophenol blue) supplemented with 2% -mercaptoethanol for 5 min at 95°C and quantified by a Bradford’s assay. A total of 50 µg of the preparations were loaded onto a 10% acrylamide/bis-acrylamide gel,
electrophoretically separated by SDS-PAGE and transferred to polyvinylidene difluoride membranes (30,31). Membranes were then blocked at 37°C for 1 h in 5% non-fat milk, 10 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.1% Tween-20 and probed with the following primary antibodies (all at 1:500 dilution) at 4°C for 16 h: Mouse anti-Lamp1 (BD Biosciences, San Jose, CA, USA; cat. no. 610142), mouse anti-cluster of differentiation (CD)63 (EMD Millipore, Billerica, MA, USA; clone RFAC4, cat. no. CBL553), rabbit anti-Adam10 (Acris OriGene Technologies GmbH, Germany, clone 3C6, cat. no. AP08308PU-N), rabbit anti-TSG101 (Santa Cruz Biotechnology, Inc., Dallas, TX, USA; clone C-2, cat. no. sc-7964), mouse anti-Flotillin2 (BD Biosciences, San Jose, CA, USA; cat. no. 610383), mouse anti-CD81 (Santa Cruz Biotechnology, Inc.; clone B11, cat. no. sc-7637), mouse anti-GM130 (BD Biosciences, San Jose, CA, USA; clone 35/GM130, cat. no. 610822) (32), mouse anti-ApoAI (Thermo Fischer Scientific, Inc.; clone 19H20L19, cat. no. 701239), mouse anti-Ago2 (Origene Technologies, Inc., Rockville, MD, USA; cat. no. TA352430) and incubated in the presence of specific horseradish-peroxidase conjugated immunoglobulin (Ig)G, Rabbit anti-mouse (cat. no. A90-117P) or Goat anti-rabbit (cat. no. A120-101P) (Bethyl Laboratories, Inc, USA, 1:3,000 dilution) at room temperature for 1 h (33,34). Immunoreactive bands were identified using the Luminata classicco detection system (EMD Millipore). Images were acquired using a G:Box Chemi XT Imaging system and quantified using Gene Tools software version 4.01 (Syngene, Frederick, MD, USA) (35).

Atomic force microscopy (AFM) imaging. AFM imaging was conducted as described (36,37). Briefly, NEP preparations were diluted 1:200 with deionized water. A total of 5-10 µl of samples were then spotted onto freshly cleaved mica sheets (Grade V-1, thickness 0.15 mm, size 10x10 mm). All mica substrates were dried at room temperature and analyzed using a Nanosurf NaioAFM (Nanosurf AG, Liestal, Switzerland), equipped with Tap190AI-G tips (Budget Sensor Solutions Bulgaria Ltd., Sofia, Bulgaria). Images were captured in tapping mode; the scan size ranged from 0.5-15 mm; the scan speed ranged from 0.6 to 1.5 sec x line.

miRNA expression by qPCR. Mature hsa-miR-16-5p, hsa-miR-99b-5p and hsa-miR-144-3p were amplified by a two-step Taq-Man RT-PCR analysis, using primers and probes obtained from Thermo Fisher Scientific, Inc., (hsa-miR-144-3p, cat.no.002676; hsa-miR-16-5p,cat.no.000391;hsa-miR-99b-5p, cat. no. 000436; cel-miR-39-3p, cat. no. 000200). cDNA was synthesized from 5 µl of the each RNA fraction in 15 µl reactions, using TaqMan MicroRNA Reverse Transcription kit (Thermo Fisher Scientific, Inc.). The reverse transcription reaction was performed by incubating the samples at 16°C for 30 min, followed by incubation at 42°C for 30 min and 85°C for 5 min. The RT-qPCR reaction (20 µl) contained 1.3 µl of reverse transcription product, 10 µl of Taq-Man 2X Universal PCR Master Mix and 1 µl of the appropriate TaqMan MicroRNA Assay (20X) specific for the miRNA targeted by the assay. The PCR mixtures were incubated at 95°C for 10 min, followed by 40 cycles at 95°C for 15 sec and 60°C for 60 sec. PCRs were performed in triplicate using the 7500 real time PCR system (Applied Biosystems; Thermo Fisher Scientific, Inc.). The expression of miRs was based on the 2^-∆∆Cq method (38), using cel-miR-39-3p as reference.

Bioinformatic analysis. For functional enrichment analysis of gene ontology terms and molecular signalling pathway, the following online software packages were used: Database for Annotation, Visualization and Integrated Discovery (DAVID) (39) and WebGestalt (40) and Kyoto Encyclopedia of Genes and Genomes analysis (KEGG; www.genome.jp/kegg/). The human disease database MalaCards (41) was used to search the diseases and the genes associated with DS. The miRNA targets were predicted by miRWALK (42). Considering that these procedures may have a certain false positive rate, the target genes of miRNA only considered the genes predicted by at least 7 procedures.

Statistical analysis. The descriptive statistics for all quantitative variables available in the dataset (minimum, maximum, first, second and third quartile, mean ± standard deviation, number and corresponding percentage of missing values) and the frequency distributions of 37 typical DS pathologies (presence/absence of the pathology-in absolute value and %-) have been computed.

The investigation on the linear association between the quantitative variables collected on 15 individuals with DS has been performed by calculating the Pearson correlation coefficient $r$.

For the graphical representation of the significant associations (correlation tests with $P<0.05$) between variables under inspection, bubble charts were used that introduce in a simple scatterplot a third dimension, namely the diameter of the bubble. This dimension is proportional to the categorical variable called ‘number of comorbidities’. In detail, for each individual with DS his/her number of comorbidities were summed, obtaining a quantitative score; then the first (Q1) and third (Q3) quartile were computed and the quantitative score were turned into the following three categories: i) First quartile (Q1=17.5), each subject with number of comorbidities $<17.5$ was classified as Low comorbidities; ii) second quartile (Q3=21), each subject with number of comorbidities $>$17.5 was classified as Medium comorbidities; and iii) each subject with number of comorbidities $>21$ was classified as High comorbidities.

Consequently, the scatterplots that highlight the significative association between quantitative variables report the dot circumference proportional to the number of assigned comorbidities: Small dots when subjects with DS are classified as ‘Low comorbidities’ (green); medium size dots for ‘Medium comorbidities’ (orange) and big dots when they are labeled as ‘High comorbidities’ (red).

For each of the qualitative variables reported (Table SI; considering also the neurological severity), the median values of the miRNAs expression levels in correspondence of subgroups of subjects defined by its categories were computed. Then, the non-parametric Wilcoxon signed-rank test (or Kruskal Wallis test when the subgroups are $>2$) was computed in order to compare the medians obtained (Table SII). These non-parametric tests do not require any assumptions on the shape.
distributions of the three miRNAs and permit analysis of the possible association between clinical variables and miRNAs expression. The unique significant associations were visualized by means of a boxplot for each category (Low, Medium and High comorbidities).

**Heatmap generation.** The data matrix with the investigated miRNAs was visualized by means of a heatmap. Starting from the standardized matrix (each miRNA has mean 0 and standard deviation 1), a heatmap was generated to visualize the miRNA expression (low values are represented with blue while high values with red). In this graph, similar values were placed near each other according to the clustering algorithm used in the analysis, thereby realizing one dendrogram appended on the left y-axis.

**k-medoids cluster analysis.** Starting from the data matrix containing only miRNA expression levels, we use the k-medoids cluster analysis generating $k=2$ groups. For evaluating the distance between subjects the Manhattan distance was used, which is based on absolute value distance. This choice should provide more robust results, whereas Euclidean distance would be influenced by unusual values in the data matrix.

Inside each group, subjects exhibit a high degree of similarity based on miRNA expression while subjects belonging to different groups are as dissimilar as possible. The $k$-medoids cluster analysis (known also with the acronym PAM-Partitioning Around Medoids) searched for $k$ representative subjects (in the present study $k=2$) among all the subjects in the dataset. These representative subjects are called medoids (43) and the algorithm assigns each patient of the dataset to nearest medoid.

By means of silhouettes a graphical representation of cluster analysis was provided where the entire clustering was displayed by plotting the silhouette into a diagram highlighting the cluster quality. Denoting Cluster 1 with $C_1$ and Cluster 2 with $C_2$, silhouette for the subject $i$ contained in $C_1$ was:

$$s(i) = \frac{b(i) - a(i)}{\max[a(i), b(i)]}$$

where: $a(i) =$average dissimilarity of $I$ to all subjects of $C_1$ and $b(i) =$average dissimilarity of $I$ to all subjects of $C_2$.

From the ratio above, it is evident that each silhouette belongs to the interval $-1 \leq s(i) \leq +1$. This means that when $s(i)$ is close to 1 the within dissimilarities $a(i)$ is smaller than the between dissimilarities $b(i)$ and $i$ is well classified. When $s(i)$
is about 0, it is not clear where $i$ should be assigned (cluster $C_1$ or $C_2$) and finally when $s(i)$ is close to -1, $a(i)$ is larger than $b(i)$ and this means that $i$ has been misclassified.

All statistical analyses were performed using R version 3.4.4.

**Results**

**NEF separation, characterization and miRNA extraction.**

NEF samples were obtained with a commercial polymer-based precipitation (PBP) kit as described in the Materials and methods section. PBP is a simple, timesaving and cheap EV concentration method (44) characterized by low specificity. Formulations obtained by PBP consist in a heterogeneous mix of nanosized particles that overlap EVs in size and morphology, including HDLs.

NEF morphological properties were determined by AFM (45). The preparation was adsorbed on a freshly cleaved mica surface and imaged under ambient conditions. Samples were composed by spherical nano-objects with a size ranging from 10s to a few 100s of nm, which are visible despite the presence of background material, very like the polymeric matrix on which the precipitation kit is based (44,46) (Fig. 1A). The resulting fraction is mainly composed by small EVs and HDLs, as can be deduced by the enrichment of specific EV markers (47), including Lamp1, CD63, Adam 10, TSG101, Ago2 and CD81, as compared with the supernatant fractions by western blot analysis (Fig. 1B). On the other hand, NEF is devoid of both GM130s and Flotillin, excluding the presence of intracellular contaminants and implying that that large EVs are not efficiently enriched by this protocol. On the other hand Apo-A1, a biomarker characteristic of HDL lipoproteins (48),

![Figure 2. Expression levels of miRNAs in NEF from the couples of siblings.](image-url)

(A) mature miR-16-5p, (B) miR-144-3p and (C) miR-99b-5p in each of the 15 siblings couples obtained by quantitative polymerase chain reaction. Average expression levels of (D) miR-16-5p, (E) miR-144-3p and (F) miR-99b-5p in individuals with DS and in control subjects. miR, microRNA; DS, Down syndrome; C, healthy control.
miRNAs are detectable in supernatant fractions and in a lesser extent also in NEFs (Fig. 1B). Overall, the above results indicate that the final NEF is specifically enriched with nanosized RNA-carriers, namely small EVs and HDLs.

**NEF miRNA extraction and expression analysis.** miRNAs were subsequently extracted from NEF using commercial columns. First, Agilent miR microarray analysis aimed to search for NEF miRNA content and was performed on samples derived from 4 couples of siblings: 4 subjects with DS and their 4 correspondent healthy siblings. This allowed identification of the most abundant miRNAs in the NEF recognized at least by 3 probes and expressed in all the 8 subjects included in the analysis (Table SIII). Consequently, the expression levels of the three most abundant miRNAs (hsa-miR-16-5p, hsa-miR-144-3p and hsa-miR-99b-5p) were further assessed in a larger cohort of DS and C by qPCR (15 couples of siblings; n=30).

As presented in Fig. 2A-C, the expression levels of these three miRNAs were very different among the couples of siblings considered and in general the results demonstrated an increased average expression trend in DS NEF with respect to C [miR‑16‑5p: Mean Relative Quantification (RQ) of DS: 29.5±13.3, mean RQ of C: 11.9±3.3, Fig. 2D; miR‑144‑3p: Mean RQ of DS 66.6±18.7, mean RQ of C: 49.5±14.3, Fig. 2E; miR‑99b‑5p: Mean RQ of DS: 42.0±15.8, mean RQ of C: 27.7±10.6 Fig. 2F].

Since hsa-miR-16-5p, hsa-miR-144-3p and hsa-miR-99b-5p have all been previously identified as exosome-associated miRNAs and, in certain cases, associated with distinct pathologies (49-52), although never associated with the DS phenotype previously, it was decided to statistically correlate the expression levels of the three miRNAs with the available quantitative clinical data of the participants (quantitative variables, Table SIV) in order to have insight into their functional role in DS.

**Correlation analysis of miRNA expression with clinical features of subjects with DS.** In Table I the correlation matrix computed on the quantitative variables obtained from the subjects with DS was reported (correlation test with P<0.05). Bubble chart analysis (Fig. 3A-E) was subsequently performed only on the significant associations, introducing a third dimension, the diameter of the bubble. This dimension was proportional to the categorical variable named ‘number of comorbidities’, obtained as described in Material and methods section and reported in Table SV for each subject. An inverse significant linear association between the clinical feature ‘development babbling’ expressed in months and miR‑16‑5p and miR‑144‑3p expression levels was evident (Fig. 3A and B). From the bubble charts it is clear that these two miRNAs demonstrated high expression levels in subjects with DS with high number of comorbidities and beginning to babble earlier.

Concerning miR‑99b‑5p, it was observed that the median value of the expression level was significantly increased (Kruskal test: P=0.05) in individuals with DS presenting high comorbidities (Fig. 4 and Table SII) suggesting a potentially more relevant role of miR‑99b‑5p upregulation in this subgroup of subjects with DS. Instead, no significant differences were identified for miR‑144‑3p and miR‑16‑5p (Table SII).
Fig. 3C highlights a significant linear positive association between miR16-5p and miR-144-3p expression. The same consideration can be extended to miR16-5p versus miR-99b-5p and miR144-3p versus miR-99b-5p, respectively (Fig. 3D and E). Furthermore, individuals with DS with the highest expression levels of miRNAs exhibited a high number of DS-associated typical clinical features. It is interesting to outline that when the same correlations were calculated by Pearson correlation analysis on the healthy siblings (on the basis of the descriptive statistics reported in Table SVI), we obtained low values of $\rho$ (miR-16-5p vs. miR-144-3p: $\rho=0.1895$; miR-16-5p vs. miR-99b-5p: $\rho=0.5123$; miR-144-3p vs. miR-99b-5p: $\rho=0.6820$). This well agrees with the finding that the three selected miRNAs were upregulated in the DS cases in the present study, indicating a possible role in DS-associated typical clinical anomalies, probably mainly in subjects with 'high comorbidities'. As expected, by repeating the same analysis on all the 30 participants (DS+C), significant positive correlations were again obtained between miRNAs (Fig. 5A-C).
\( q \) was also computed on the expression level of each miRNA comparing couples of siblings (Fig. 5D-F). The only significant positive linear correlation was for miR-99b-5p (Fig. 5F). By computing \( q \) for couples of unrelated subjects (C subjects and subjects with DS were repeatedly randomized computing the correlation coefficients and corresponding test) it was noted that there was no correlation for miR-99b-5p (data and Figure provided upon request). This result demonstrates the same expression trend of miR-99b-5p among brothers, demonstrating the importance of the genetic background when analyzing miRs.

miRNA expression was visualized using heatmaps either on the entire sample (DS+C) (Fig. 6A), or separating C subjects and subjects with DS (Fig. 6B and C, respectively). The subjects with the highest miRNA expression (at the top three positions of the graph) belonged to the DS group. Furthermore, the C subjects at the top positions were siblings of the DS with highest miRNA expression. Focusing the attention on heatmaps in Fig. 6B and C, it was evident that miRNA expression levels were different between DS and C (in DS the color key is red-shifted compared with C); among the subjects with DS there was only one case exhibiting very low miRNA expression and notably this subject also presented low comorbidities; the miRNAs less expressed were miR-144-3p and miR-99b-5p in DS and C, respectively.

Using the Partitioning Around Medoids algorithm (PAM) 2 clusters of subjects were identified (Fig. 7A): The first one, called \( C_1 \), contained subjects with low median and mean expression values for each miRNA with a prevalence of C subjects (55% of individuals in \( C_1 \) were C which corresponded to 73% of healthy siblings in the study, see Table SVII for major details). Cluster 2, called \( C_2 \), contained a number of subjects with DS with increased miRNA expression (60% of individuals in \( C_2 \) were DS which corresponded to 40% of
subjects with Down syndrome in the study). Furthermore, certain subjects with DS in C were classified together with their siblings (DS10 and C10, DS12 and C12) confirming similar miRNA expression levels.

Figure 6. Heatmaps of the expression levels of miR-16-5p, miR-144-3p and miR-99b-5p. (A) Heatmaps on the entire sample DS+C, (B) on subjects with DS and (C) on healthy individuals. DS, Down syndrome; C, healthy individuals.
From the PAM algorithm it was also clear that miRNA levels were able to generate clusters of subjects that were well classified. In fact, the ‘within dissimilarities’ was smaller than the ‘between dissimilarities’ and 73% of individuals were well classified, as presented in Fig. 7B, where 22 subjects on 30 analyzed (on y-axis) exhibited a width silhouette $s(i)$ (on x-axis) higher than 0.4. Therefore, this representation confirmed the goodness of the classification induced by the miRNAs matrix (note that $s(i)$ reaches its maximum in correspondence of 1). Furthermore, in Fig. 7B each subject with DS is colored using the classification based on number of comorbidities: In $C_2$ cluster, subjects with DS with high comorbidities prevailed.

miRNAs gene target prediction and enrichment study. The main putative targets of the 16 most abundant miRNAs identified in the NEFs (Table SIII) were subsequently predicted using the online software miRWalk (6). The analysis in the present study identified 2,515 targeted genes, including 25 of the 536 genes located on Hsa21. Functional annotation analysis of the 2,515 genes was performed next. In a GO analysis conducted by DAVID (1,865 unique IDs identified), the candidate genes were analyzed in all categories including biological process (Table II), molecular function (Table III) and cellular component (Table IV). For the biological process category, the first 4 enriched terms were referred to the transcriptional process; the fifth biological term was ‘nervous system development’ with 53 predicted genes included in this term (P=0.0140). In the molecular function category, 1,010 genes were included in ‘protein binding’ term and the other enriched terms belonged essentially to ‘transcription process’ categories. In the cellular component category, the ‘nucleus’ was the predominant term with 620 associated-genes, the fourth term was ‘neuronal cell body’ with 54 genes (P=0.0080). Therefore, >50 candidate miRNAs target genes may have a role in ‘nervous system development’ or in the ‘neuronal cell body’ formation.

To characterize the predominant pathways, putative targets were searched by Kyoto Encyclopedia of Genes and Genomes
analysis, which revealed the term ‘chronic myeloid leukemia’ (P<0.01) with 23 target genes among the top 10 signal pathways (Table V). A disease association analysis of the predicted miRNA targets was also performed using Webgestalt demonstrating that terms as ‘leukemia T-cell’, ‘leukemia, myeloid, acute’, ‘chromosome aberration’, ‘Precursor T-Cell Lymphoblastic Leukemia-Lymphoma’, ‘leukemia myeloid’ ‘secondary leukemia’ were among the top 10 statistically significant (Table VI). In this context, the predicted target genes of miR-16-5p and 22 of miR-144-3p were identified to be differentially expressed in individuals with DS respect to controls (Ratio <0.8 and Ratio >1.2; in bold).

Furthermore, the miRNA enrichment analysis conducted by Webgestalt demonstrated that the major part (102 genes; P<0.01) of the predicted miRNA target genes were targeted by miR-144-3p, which recognizes the binding site ATACTGT (Table VII).

Among the predicted targets of miR-144-3p, two DS-associated genes, DYRK1A, encoding for dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1a and SIM1, encoding for the transcriptional factor, have been respectively reported at the 2nd and 26th positions in the Malacard gene list. DYRK1A and SIM1 gene transcripts were predicted to be

---

Table II. GO Top 10 significant terms belonging to the ‘biological process’ category.

<table>
<thead>
<tr>
<th>GO term (biological process)</th>
<th>Number of genes</th>
<th>%</th>
<th>P-value corrected P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive regulation of transcription. DNA-templated</td>
<td>102</td>
<td>5.5</td>
<td>1.6x10^-11</td>
</tr>
<tr>
<td>Positive regulation of transcription from RNA polymerase II promoter</td>
<td>163</td>
<td>8.7</td>
<td>2.80x10^-11</td>
</tr>
<tr>
<td>Negative regulation of transcription from RNA polymerase II promoter</td>
<td>126</td>
<td>6.8</td>
<td>2.50x10^-10</td>
</tr>
<tr>
<td>Transcription from RNA polymerase II promoter</td>
<td>96</td>
<td>5.1</td>
<td>1.40x10^-09</td>
</tr>
<tr>
<td><strong>Nervous system development</strong></td>
<td><strong>53</strong></td>
<td><strong>2.8</strong></td>
<td><strong>1.40x10^-08</strong></td>
</tr>
<tr>
<td>Positive regulation of transcription from RNA polymerase II promoter</td>
<td>73</td>
<td>3.9</td>
<td>4.90x10^-05</td>
</tr>
<tr>
<td>Steroid hormone mediated signaling pathway</td>
<td>17</td>
<td>0.9</td>
<td>9.30x10^-05</td>
</tr>
<tr>
<td>Dentate gyrus development</td>
<td>9</td>
<td>0.5</td>
<td>1.00x10^-04</td>
</tr>
<tr>
<td>Stem cell differentiation</td>
<td>11</td>
<td>0.6</td>
<td>1.60x10^-04</td>
</tr>
<tr>
<td>Transcription. DNA-templated</td>
<td>241</td>
<td>12.9</td>
<td>1.9x10^-4</td>
</tr>
</tbody>
</table>

**Bold font indicates typical Down syndrome features. GO, Gene Ontology.**

Table III. GO Top 10 significant terms belonging to ‘molecular function’ category.

<table>
<thead>
<tr>
<th>GO term (molecular function)</th>
<th>Number of genes</th>
<th>%</th>
<th>P-value corrected P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protein binding</td>
<td>1,010</td>
<td>54.2</td>
<td>1.6x10^-13</td>
</tr>
<tr>
<td>Transcriptional activator activity. RNA polymerase II core promoter proximal region sequence-specific binding</td>
<td>56</td>
<td>3.0</td>
<td>1.0x10^-9</td>
</tr>
<tr>
<td>Transcription factor activity. sequence-specific DNA binding</td>
<td>154</td>
<td>8.3</td>
<td>1.1x10^-9</td>
</tr>
<tr>
<td>Sequence-specific DNA binding</td>
<td>88</td>
<td>4.7</td>
<td>5.1x10^-7</td>
</tr>
<tr>
<td>RNA polymerase II core promoter proximal region sequence-specific DNA binding</td>
<td>66</td>
<td>3.5</td>
<td>6.9x10^-7</td>
</tr>
<tr>
<td>Protein serine/threonine kinase activity</td>
<td>66</td>
<td>3.5</td>
<td>5.2x10^-6</td>
</tr>
<tr>
<td>RNA polymerase II regulatory region sequence-specific DNA binding</td>
<td>41</td>
<td>2.2</td>
<td>2.9x10^-5</td>
</tr>
<tr>
<td>GDP binding</td>
<td>17</td>
<td>0.9</td>
<td>4.2x10^-5</td>
</tr>
<tr>
<td>Zinc ion binding</td>
<td>157</td>
<td>8.4</td>
<td>4.7x10^-5</td>
</tr>
<tr>
<td>Transcriptional repressor activity. RNA polymerase II core promoter proximal region sequence-specific binding</td>
<td>26</td>
<td>1.4</td>
<td>6.6x10^-5</td>
</tr>
</tbody>
</table>

**GO, Gene Ontology.**

**GO term (molecular function)**

**Number of genes**

**%**

**P-value corrected P-value**
targeted by miR-144-3p respectively by 8 and 6 prediction software programs (miRWalk). TargetScan (release 7.1, June 2016) predicted 2 conserved putative binding sites for *DYRK1A* and 1 for *SIM1* gene transcripts in the 3'UTR portions.

### Discussion

The present study compares the expression of miRNAs carried by biogenic nanoparticles separated from plasma samples of young individuals with DS and their healthy siblings. This is the first study that tackles noncoding RNAs encased in EVs and HDLs in association with DS pathogenesis, demonstrating specific miRNA-signatures (19). So far, only two studies have addressed an ‘exosome approach’ regarding this syndrome, although focusing on proteins rather than on miRNAs (54,55).

The adopted comprehensive approach of plasma NEF analysis was very informative since it allowed identification of novel miRNAs probably involved in DS pathogenetic mechanisms. This approach represents a first breakthrough in respect to previous findings about miRNA and DS and sets the proof of concept for further study. It will be interesting to indicate a sharp stratification of EV and HDL miRNAs, provided the availability of sufficient starting media to perform the very challenging separation of the two nanoparticle populations.
Gene target prediction on the over-represented miRNAs identified during the initial discovery phase by microarray analysis and subsequent GO analysis of the predicted targets, suggested the possible involvement of the identified miRNAs in 'nervous system development', 'neuronal cell body' and in the pathogenesis of certain DS-associated diseases like leukemia. These are interesting results since the nervous system is remarkably affected during development and aging in individuals with DS. Furthermore, it is well known that children with DS have a markedly increased risk for a subtype of myeloid leukemia (ML) (56) classified by the World Health Organization as ML-DS (Table VIII).

Expression profiling of the three most abundant miRNAs (miR-16-5p, miR-144-3p and miR-99b-5p) by RT-qPCR identified a trend of dysregulation in individuals with DS compared with C. The expression of these miRNAs has not been previously associated with DS and in general little data were published regarding the expression of such miRNAs in human plasma. miR-16-5p was identified in circulating exosomes of patients with multiple myeloma (57) as well as in circulating serum exosomes of patients with esophageal adenocarcinoma (49). miR-99b-5p is known as one of five most common miRNA in human plasma exosomes by RNA deep sequencing (58).

Concerning function, miR-16-5p was previously described in rat neurons as a negative regulator of dendritic complexity and mediates brain derived neurotrophic factor (BDNF)-induced dendritogenesis by regulating the translation of the BDNF mRNA itself, supporting the hypothesis that miR-16-5p may be important during neuronal development (59).

miR-144-3p has emerged as an important miRNA implicated in certain human CNS pathologies (60,61). In a mouse model, miR-144-3p targets a number of genes implicated in the control of neuronal plasticity-associated signaling cascades, including PTEN, Sprd1 and Notch1 (62).

Among the predicted targets of miR-144-3p, two DS-associated genes, DYRK1A and the transcription factor SIM1 were identified. While no experimental evidence about miR-144-3p and both DS-genes expression regulation has previously been demonstrated, it is known that DYRK1A expression can be regulated by miRNAs (63,64).

DYRK1A kinase produced by Hsa21 serves a pivotal role in the CNS and its overexpression in individuals with DS may be implicated in cognitive impairments, therefore it has been suggested as a putative therapeutic target for treating the cognitive deficiencies observed in DS (58,65,66). Deregulation of DYRK1A levels are also frequently observed in Alzheimer and Parkinson diseases (67-69).

A recent study about the transcriptome map analysis of DS vs. control human tissues (6) has demonstrated that in the differential transcriptome map of brain, induced pluripotent stem cells, blood and fibroblasts of DS and controls samples DYRK1A gene has expression ratios of 1.51, 1.36, 1.30 and 1.37, respectively. These ratios are consistent with a possible DYRK1A gene expression alteration in trisomy 21 samples due to the presence of the Hsa21 in three copies in the cells of

Table VII. miRNA enrichment analysis. Top 10 significant miRs targeting the gene list.

<table>
<thead>
<tr>
<th>ID (putative binding site and miR)</th>
<th>Number of genes</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATACTGT, MIR-144</td>
<td>102</td>
<td>0</td>
</tr>
<tr>
<td>CAGTGT, MIR-141, MIR-200A</td>
<td>124</td>
<td>0</td>
</tr>
<tr>
<td>GCACTTT, MIR-17-5P, MIR-20A, MIR-106A, MIR-106B, MIR-20B, MIR-519D</td>
<td>190</td>
<td>0</td>
</tr>
<tr>
<td>TGACTGTC, MIR-15A, MIR-16, MIR-15B, MIR-195, MIR-424, MIR-497</td>
<td>199</td>
<td>0</td>
</tr>
<tr>
<td>TGACTTT, MIR-519C, MIR-519B, MIR-519A</td>
<td>136</td>
<td>7.13x10^-11</td>
</tr>
<tr>
<td>AAGCACT, MIR-520F</td>
<td>85</td>
<td>1.33x10^-10</td>
</tr>
<tr>
<td>TTGCACT, MIR-130A, MIR-301, MIR-130B</td>
<td>119</td>
<td>8.25x10^-9</td>
</tr>
<tr>
<td>ATGCTGC, MIR-103, MIR-107</td>
<td>75</td>
<td>2.31x10^-8</td>
</tr>
</tbody>
</table>

Bold font indicates miRNA under study. FDR, false discovery rate; miR/miRNA, microRNA.

Table VIII. Down syndrome-associated diseases according to Malacards.

<table>
<thead>
<tr>
<th>Related disease ID</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leukemia, megakaryoblastic, with or without Down syndrome, somatic</td>
<td>12.0</td>
</tr>
<tr>
<td>Acute megakaryoblastic leukemia in Down syndrome</td>
<td>12.0</td>
</tr>
<tr>
<td>Acute megakaryoblastic leukemia without Down syndrome</td>
<td>11.9</td>
</tr>
<tr>
<td>Ayme-gripp syndrome</td>
<td>11.1</td>
</tr>
<tr>
<td>intellectual disability</td>
<td>10.4</td>
</tr>
<tr>
<td>Leukemia</td>
<td>10.4</td>
</tr>
<tr>
<td>Macroglossia</td>
<td>10.2</td>
</tr>
<tr>
<td>Dementia</td>
<td>10.1</td>
</tr>
<tr>
<td>Myeloid leukemia</td>
<td>10.1</td>
</tr>
<tr>
<td>Neuritis</td>
<td>10.1</td>
</tr>
</tbody>
</table>
The present study was partially supported by a BIOMANE grant (University of Brescia, Brescia, Italy) to AR, AS, GDP and PB. AR was also supported from Fondo per il Finanziamento delle Attività Base di Ricerca (Ministero dell’Istruzione, dell’Università e della Ricerca) and University of Brescia research fund (ex 60%). This study was partially supported by Centro Bresciano Down (CBD), Brescia, Italy to AR and DR.

Availability of data and materials

The analyzed datasets generated during the study are available from the corresponding author on reasonable request.

Authors’ contributions

AR and DR conceived the project. MC, FA, AP, CL and GC were involved in sample and medical record recruitment. AS, MV, AR wrote the manuscript. AR, LP, SB performed NEF separation and biophysical and biochemical characterization. AS, TF performed the qPCR experiments; AS, EA performed the GO analysis. GDP, GA and PB were involved in the study conception and design, analysis and interpretation of data, and critical revision of the manuscript. All authors read and approved the manuscript.

Ethics approval and consent to participate

The present study was approved by the competent Ethics Committee of Sant’Orsola-Malpighi Hospital in Bologna, approval number: 39/2013/U/Tess.

Patient consent for publication

Informed consent was obtained from the subjects for participation in this study.

Competing interests

No potential conflict of interest was reported by the authors.

References


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.