Beyond genome wide association studies in celiac disease by exploring the non-coding genome
de Almeida, Rodrigo

PART II
Chapter 6

Circulating miRNAs as potential biomarkers of early celiac disease

Manuscript in preparation

Rodrigo Almeida[1], Sabine Lisa Vriezinga[2]*, Anna Stachurska[1]*, Urmo Vsa[3,4], Rutger Modderman[1], Vinod Kumar[1], M Luisa Mearin[2#], Cisca Wijmenga[1#], Sebo Withoff[1]

[1]University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands.
[2]Leiden University Medical Center, Department of Pediatrics, Leiden, the Netherlands.
[3]Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia,
[4]Estonian Genome Center, University of Tartu, Tartu, Estonia.

*These authors shared the second authorship.
#These authors contributed equally.

Abstract

Celiac disease (CeD) is an immune-related disease, which is diagnosed on the basis of symptoms, the detection of CeD-specific antibodies, and biopsy results. The only available treatment is a life-long, gluten-free diet. Despite the availability of these diagnostic options, as many as 7 out of 8 CeD patients are either not diagnosed or incorrectly diagnosed, highlighting the need to find relevant biomarkers. Circulating micro-RNA (miRNA) profiles have been shown to be disease- or even disease-stage-specific, in patients with cancer or gastro-intestinal disease. We therefore examined whether circulating miRNAs in serum samples of CeD patients can be used as biomarkers for CeD. In the PreventCD cohort, newborns at risk for CeD were challenged between 4-6 months of age with low levels of gluten to see if this would induce gluten tolerance. Plasma samples were taken from these babies at age 4, 6, 9, 12, 18, 24 and 36 months of age and annually thereafter until diagnosis and after approximately 6 months on a gluten-free diet. Using next generation sequencing, we profiled miRNAs in 105 serum samples from 12 CeD patients, 5 anti-gliadin positive patients who did not develop CeD, and 5 control subjects from the PreventCD cohort. Comparing data from patient samples taken at the time of diagnosis versus samples taken before the gluten introduction (at age 4 months), we found 45 miRNAs to be significantly differentially expressed (FDR < 0.05). Of these, six miRNAs showed a consistent behavior over different time points up to diagnosis, and seemed to return to normal levels when the patients started on a gluten-free diet. Our results suggest that circulating miRNAs could be promising biomarker candidates for early CeD.

Introduction

Celiac disease (CeD) is a complex autoimmune disease triggered by dietary gluten in genetically predisposed individuals. With a worldwide prevalence of about 1%, CeD is considered one of the most common, genetically-based food intolerances in the world (1). Thus far, CeD is the only immune-related disease caused by a well-known environmental factor (gluten). In fact, the only current treatment is a life-long, gluten-free diet (GFD). The diagnosis of CeD is reached by a combination of clinical and serological data, and the most definitive diagnosis is based on the histology of biopsies taken from the small intestine. Despite the availability of these diagnostic tools, as many as 7 out of 8 CeD patients are either not diagnosed or incorrectly diagnosed (2), indicating that additional biomarkers would be very useful, especially if they could be used to detect the slow progression into full-blown disease (3).

MicroRNAs (miRNA) are small non-coding RNAs (~22 nucleotides) that are involved in many cellular

processes. They fine-tune gene expression by targeting a protein-complex, the RNA-induced silencing complex (RISC), to the 3'-untranslated region (3'-UTR) of target messenger RNAs (mRNAs). Binding of the RISC complex can result in degradation or inhibition of translation of the target mRNA (4). Recently a lot of attention has been paid to miRNAs as potential biomarkers since they have shown surprising stability in circulating body fluids, such as serum and plasma (5, 6). Disease-specific circulating miRNA profiles have been described in various immune-mediated diseases, including multiple sclerosis (7), rheumatoid arthritis (8) and type 1 diabetes (9). Thus far, only three studies have shown miRNAs to be differentially expressed in biopsies from celiac patients (10–12), however, none of these studies focused on identifying specific circulating miRNAs in CeD.

In this study we used the Dutch cohort of the prospective PreventCD study (n = 105) to assess whether circulating miRNAs could be used as early biomarkers for CeD. PreventCD is an ongoing European study investigating if it is possible to induce tolerance to gluten by introducing small quantities of gluten to babies early in life (13). We compared samples obtained from babies at 4 months of age, (before they were introduced to gluten) with samples obtained from children at the time of CeD diagnosis (average age at diagnosis was 33 months). This analysis identified 45 miRNAs as significantly differentially expressed (FDR < 0.05) between these two groups. Of these candidates, six miRNAs showed a consistent behavior at different time points of sample collection.

Material and Methods
PreventCD Cohort

PreventCD is a multicenter European study, focusing on ~1,000 children who are at high risk of developing CeD. The subjects come from eight countries (the Netherlands, Italy, Poland, Spain, Germany, Croatia, Hungary and Israel) and were born between 2007 and 2010. The PreventCD study hypothesized that it is possible to induce tolerance to gluten by introducing small quantities of gluten to babies between the age of 4 to 6 months (13). Serum samples were taken at 4 (before intervention), 6 (after intervention), 9, 12, 18, 24 and 36 months after birth and annually thereafter. Additional samples were collected around the time of CeD diagnosis in patients, and also during standard care check-ups after the start of a GFD for a subset of patients. Children developing high titers of CeD-associated antibodies or with a clinical suspicion of CeD, were offered small bowel biopsies to make a definitive diagnosis (14). For the current study, we investigated 105 samples, from three different groups (Table 1), from which 65 samples were obtained from 12 patients with CeD, 20 samples from 5 individuals who displayed transient high levels of anti-gliadin antibodies (AGA) at

age 6 months, directly after the gluten intervention and did not develop CeD during the course of the study, and 20 samples from 5 controls (individuals who did not develop CeD during the course of the study) (Table 1).

RNA isolation, library preparation and sequencing

Total RNA was isolated from 250 l of serum according to the mirVana PARIS kit manual (Ambion, Carlsbad, CA, USA). Next generation sequencing libraries were constructed using the TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, CA, USA) following the manufacturer's protocol. The libraries were checked by Experion microfluidic capillary electrophoresis (Bio-Rad, Hercules, CA, USA) and pooled (20 libraries per lane) before deep-sequencing on the Illumina HiSeq2500 (Illumina, San Diego, CA, USA).

Next generation sequencing data analysis

Adapter trimming of raw reads was performed using CLCBIO Genomic Workbench 7.0 (CLC Inc, Aarhus, Denmark). MiRNA expression analysis was performed using miRanalyzer 0.3 using human genome version 19 (hg19) and miRBase v20 as the reference (15). Only miRNAs with five or more reads on average between all samples were included for further analysis. Differential expression in samples at age 4 months (all groups; n = 20) versus diagnostic samples (CeD group; n = 11) was performed using the Wald test of the DESEq2 R-package (16). The resulting P-values for differentially expressed miRNAs were adjusted for multiple testing using the Benjamini and Hochberg correction for false discovery rate (FDR) (17), and only miRNAs with FDR-corrected P-values < 0.05 were considered to be significantly differentially expressed.

Results

Differential expression analysis

To investigate differentially expressed miRNAs, we decided to compare samples taken at age 4 months, with samples from patients at the time of their CeD diagnosis. These analyses identified 45 circulating miRNAs to be significantly differentially expressed (FDR < 0.05) between the two groups. Of these, 40 miRNAs were downregulated at the time of diagnosis (Table 2).

Table 1. Description of the samples studied

Phenotype	M3	M9	M12	M18/ M24	Diagnosis	On GFD for 6 months	Total samples
CeD (n = 12)	10	12	10	10	11	9	65
AGA and not CeD (n = 5)	5	5	5	5	n.a.	n.a.	20
Control (n = 5)	5	5	5	5	n.a.	n.a.	20

M3, M9 = age 3 and 9 months, etc.; CeD celiac disease; AGA anti-gliadin antibodies; n.a. not applicable

Table 2. miRNAs differentially expressed between samples at age 4 months (before introduction of gluten) and samples taken at diagnosis

miRNAs	Log2 Fold change*	P-value (FDR corrected)
hsa-miR-4492	2.212054666	0.015419594
hsa-miR-378c	1.930461868	0.030193637
hsa-miR-15b-5p	1.278274304	0.01593734
hsa-miR-423-5p	1.261131204	0.043997852
hsa-miR-451a	0.955359691	0.040747629
hsa-miR-92b-3p	-0.856153759	0.041523273
hsa-miR-148a-3p	-0.891209773	0.019461592
hsa-miR-24-3p	-1.044150043	0.031730371
hsa-miR-181b-5p	-1.054121953	0.040799068
hsa-miR-361-3p	-1.307793647	0.038591895
hsa-miR-125b-2-3p	-1.340487331	0.040799068
hsa-miR-10a-5p	-1.382000345	0.014877126
hsa-miR-224-5p	-1.399292079	0.045238586
hsa-miR-150-3p	-1.449080462	0.015321755
hsa-miR-30c-5p	-1.514766136	0.026242809
hsa-miR-369-3p	-1.541316796	0.007960103
hsa-miR-1246	-1.594246402	0.040799068
hsa-miR-143-3p	-1.625759603	0.007960103
hsa-let-7b-3p	-1.647221131	0.031429726
hsa-miR-30b-5p	-1.660679689	0.040747629
hsa-miR-532-3p	-1.828680674	0.039386762
hsa-miR-100-5p	-1.893921902	0.002084626
hsa-miR-181a-2-3p	-1.951225088	0.00121631
hsa-miR-424-5p	-2.089006741	0.009780786
hsa-miR-99a-5p	-2.105939577	0.000717664
hsa-miR-132-3p	-2.112259123	0.005662518
hsa-miR-125b-1-3p	-2.219817879	0.009780786
hsa-miR-199b-5p	-2.222779216	0.005018145
hsa-miR-141-3p	-2.275324557	0.00121631
hsa-miR-10b-5p	-2.278271598	0.00052596
hsa-miR-30a-3p	-2.411914266	0.002567547
hsa-miR-200b-3p	-2.463230018	0.011854189
hsa-miR-432-5p	-2.721485352	4.97E-05
hsa-let-7c	-2.845367592	0.00052596
hsa-miR-34c-5p	-2.849996946	0.002567547
hsa-miR-654-5p	-2.919403485	0.000514378
hsa-miR-429	-2.943591481	0.000459433
hsa-miR-4454	-3.162501201	0.000264582
hsa-miR-23b-3p	-3.254659295	2.02E-05
hsa-miR-125b-5p	-3.259661259	7.34E-06
hsa-miR-195-5p	-3.277969576	0.000184321
hsa-miR-200a-3p	-3.332939963	0.000103371
hsa-miR-483-3p	-3.340526834	0.000134987
hsa-miR-145-5p	-3.927977724	7.34E-06
hsa-miR-483-5p	-4.890224389	2.17E-13

Bold text indicates miRNAs that will recur in the analyses below. *Fold changes related to time of diagnosis.

CeD-specific miRNAs can be detected before a diagnosis is made

In order to discover CeD-related miRNAs that are detectable before the time of diagnosis, we visually inspected the expression pattern of 45 miRNAs separately, in each of the 12 CeD patients, 5 subjects who became anti-gliadin positive but did not developed CeD, and 5 controls (subjects who did not develop CeD during the course of the follow-up). We identified six interesting miRNAs (miR-100-5p, miR-30c-5p, miR-15b-5p, miR-92b-3p, miR-224-5p, miR-369-5p). Five gradually decrease in circulation until the time of diagnosis but showed an increase at 6 months after starting on a gluten-free diet (miR-30c-5p, miR-92b-3p, miR-100-5p, miR-224-5p and miR-369-5p). Conversely, miR-15b-5p slowly increased until the time of diagnosis and had decreased at 6 months after starting on a GFD (Fig. 1). In addition, except for miR-224-5p and miR-369-5p, the other four miRNAs showed a similar pattern across the time points between controls, AGA individuals and CeD patients (Fig. 1). We observed that these miRNAs started to change early after birth and in most cases before the age of diagnosis.

Discussion

Circulating miRNAs have been proposed as biomarkers in several diseases, from cancer to autoimmune disease (18). Currently, it is not clear what the specific function is of circulating miRNAs, how they end up in circulation, nor what the cell of origin is. It could be that they are just waste products from dying cells, they could be secreted because specific cells want to get rid of them, or they could
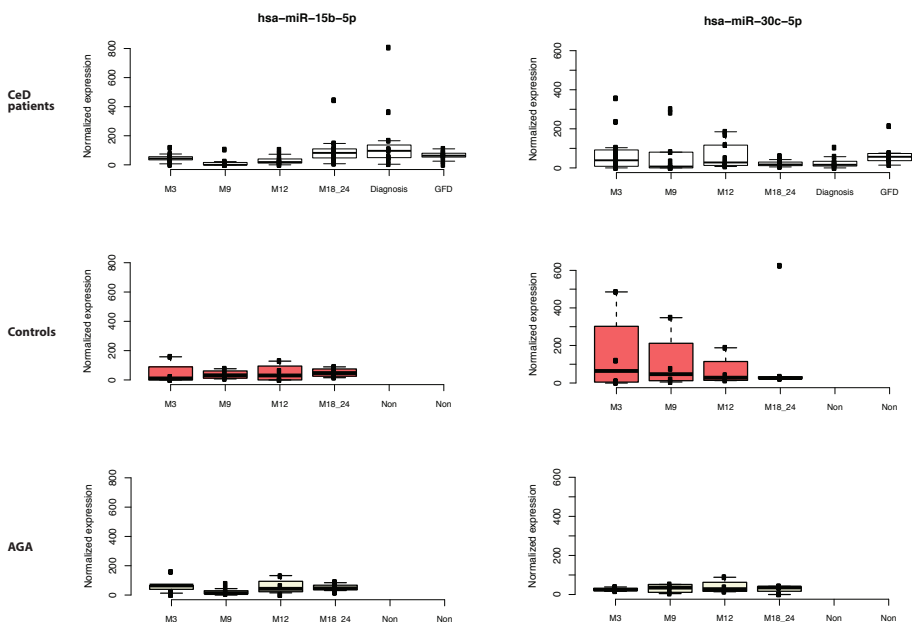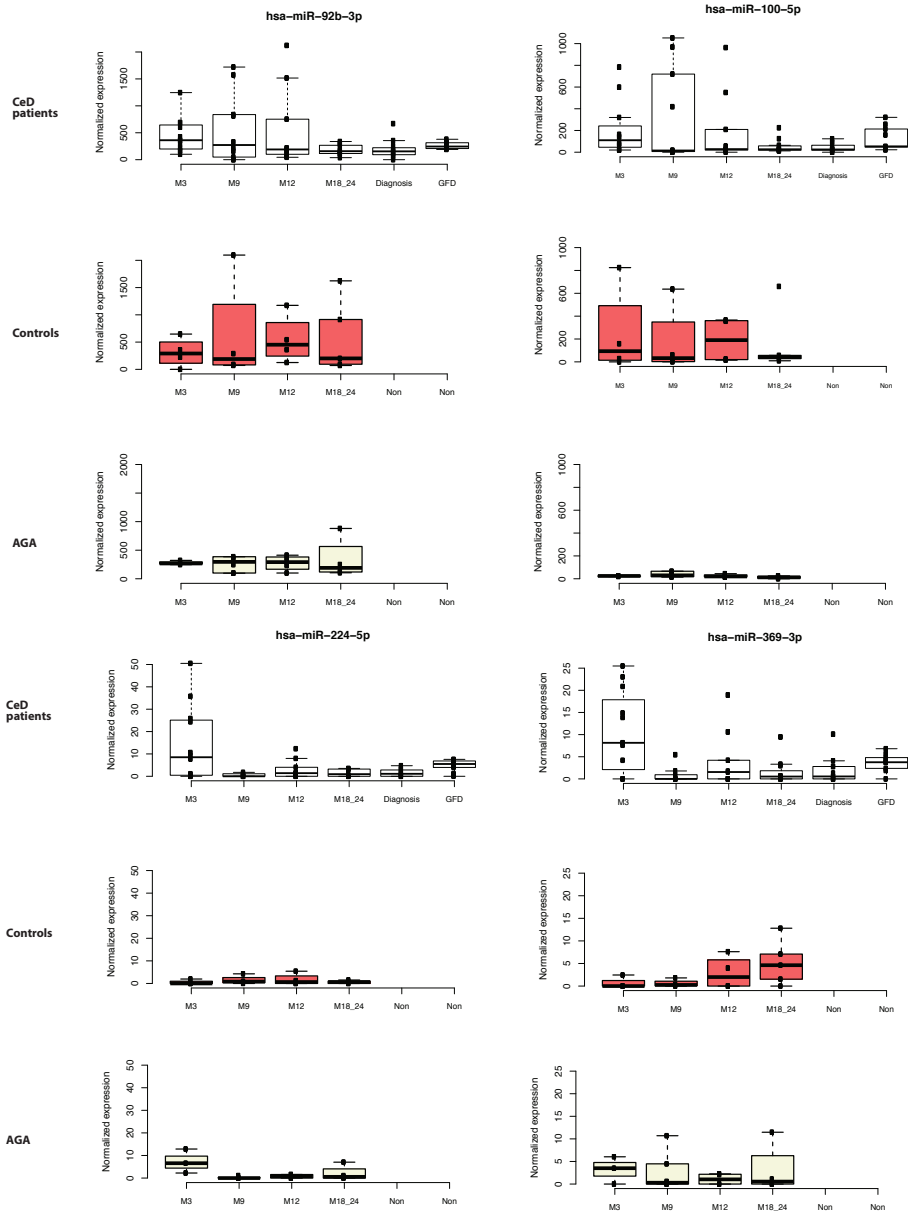
Fig. 1. Box plots illustrating six miRNAs with a consistent pattern across the different time points. CeD patients (in white), controls (in red) and individuals with positive AGA (in beige). M3: first sample collected at age 3 months; M9: sample collected at age 9 months; M12: sample collected at age 12 months; M18/24: sample collected at age 18 or 24 months.

be secreted as signaling entities that act in other cells once they are taken up. To date, circulating miRNAs have not yet been interrogated in CeD, although there is a need for biomarkers that can predict CeD development. In this chapter we present the first profile of circulating miRNAs in a unique, prospective cohort of babies at high genetic risk of developing CeD. Because we only had a limited number of samples available, we compared all

the samples taken at age 4 months for all the groups for which we had samples available with the diagnostic samples of 12 subjects who developed CeD; we investigated if we could identify miRNAs that were significantly differentially present at the time of diagnosis. This comparison resulted in a panel of 45 differentially expressed circulating miRNAs, of which 40 were downregulated in circulation at the time of diagnosis and six showed a consistent pattern that suggested they could be detected before the reported time of diagnosis.

Interestingly, in our list of downregulated miRNAs, several are known to play a role in the immune system. For instance, miR-150 was found to be downregulated in CeD patients at diagnosis (> 2 fold change) and it was also downregulated in the serum of patients critically ill with sepsis (19). MiR-150s target CXCR4 and c-Myb, which are both involved in activating the immune response (20). This suggests that downregulation of miR-150 may permit overexpression of these two targets and therefore promote activation of the immune response in CeD patients. Recently, it was suggested that miR-150 can be selectively packaged into microvesicles and actively secreted by human blood cells (21). This packaging feature may permit delivery of this miRNA to recipient cells so that they can regulate their respective target genes (22). However, it is still being debated whether circulating miRNAs are functional signaling molecules or only waste products since Turchinovich et al. has commented that the concentration of miRNAs in serum is so low that is likely to be below the threshold for mediating any significant physiological effect in vivo (23). It is also not clear which cells are the donors and which cells are the recipients of secreted exosomes.

If the differentially expressed miRNAs are truly involved in active CeD, one would expect them to revert to the healthy state once the patient starts a gluten-free diet. To check whether this was the case, we visually inspected the profiles of the 45 miRNAs in our 12 patients, and in the controls and AGA individuals. This revealed six miRNAs (miR-15b-5p, miR-30c-5p, miR-100-5p, miR-92b-3p, miR-224-5p and miR-369-3p) with a constant expression pattern over the successive time points up to diagnosis, with a reversal to the healthy level when the patients had been on a GFD for 6 months. In addition, these miRNAs could be detected before the reported time of diagnosis, suggesting that they could be used as early biomarkers in CeD development. As an example, miR-92b-3p was downregulated in our CeD patients at the time of diagnosis and was also found to be downregulated in the circulation of patients with active multiple sclerosis (24). This miRNA is part of the miR-17-92 cluster, which is involved in the proliferation and activation of naive CD4+ T cells (25). We found miR-92 to be consistently downregulated until the time of diagnosis, whereas its level return to

normal after starting a GFD. Our data suggest that the change in this miRNA can be detected well before the time of diagnosis.

It is difficult to formulate definitive conclusions on the control and AGA groups. Firstly, we had only five samples available for each group. The difficulty with circulating miRNA profiles is that they display much more sample-to-sample heterogeneity than tissue- or cell-type-specific profiles. The differences between the CeD group and the control group were not very strong, but it is possible that these subjects will develop CeD later in life even though this diagnosis was not reached at the mean age of 2.8 years old. The same holds true for the AGA group. Although the AGA group appears to show the most deviation with respect to the six miRNAs prioritized visually, it is possible for these subjects that they too will develop CeD at a later age. The already planed addition of extra controls and AGA subjects might shed more light on these issues.

It is possible that the miRNA profiles in the GFD samples will change even more towards normality. It is presently unclear whether 6 months is long enough to allow for CeD patients to re-establish a healthy gut and it has been suggested that this might actually take longer (26). Moreover, some of the patients may knowingly or unknowingly not adhere strictly to the diet so preventing full intestinal recovery. From this perspective, it would be worthwhile to follow-up our preliminary data, which suggests that miRNAs could be used as markers for adherence to a GFD.

The necessity of finding novel biomarkers for early CeD development remains. The current diagnostic options, such as CeD-associated antibodies, seem to be present after the development of the disease. The patients in this study were diagnosed between 2,6 and 5 years of age, which is a rather wide range. The hope is to find CeD-associated miRNAs that rise slowly, so that dietary or other preventive measures can be taken before full-blown intestinal inflammation erupts.

In conclusion, in this subgroup of the PreventCD cohort, we found 45 miRNAs that are differentially expressed in the circulation of CeD patients very early in life. This is a preliminary analysis and more samples of the rest of the PreventCD cohort will be analyzed in the near future. We hope to validate our findings in order to find novel biomarkers for early (or even better predictive of) CeD.

Acknowledgements

References

1. Guandalini, S. and Assiri, A. (2014) Celiac disease: a review. JAMA Pediatr., 168, 272–8.

2. DeGaetani, M., Tennyson, C.A., Lebwohl, B., Lewis, S.K, Abu Daya, H., Arguelles-Grande, C., Bhagat, G. and Green, P.H.R. (2013) Villous atrophy and negative celiac serology: a diagnostic and therapeutic dilemma. Am. J. Gastroenterol., 108, 647–53.

3. Vives-Pi, M., Takasawa, S., Pujol-Autonell, I., Planas, R., Cabre, E., Ojanguren, I., Montraveta, M., Santos, A.L. and Ruiz-Ortiz, E. (2013) Biomarkers for diagnosis and monitoring of celiac disease. J. Clin. Gastroenterol., 47, 308–13.

4. Bartel, D.P. (2009) MicroRNAs: target recognition and regulatory functions. Cell, 136, 215–33.

5. Arroyo, J.D., Chevillet, J.R., Kroh, E.M., Ruf, I.K, Pritchard, C.C., Gibson, D.F., Mitchell, P.S., Bennett, C.F., Pogosova-Agadjanyan, E.L., Stirewalt, D.L., et al. (2011) Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. Proc. Natl. Acad. Sci. U. S. A., 108, 5003–8.

6. Vickers, K.C., Palmisano, B.T., Shoucri, B.M., Shamburek, R.D. and Remaley, A.T. (2011) MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. Nat. Cell Biol., 13, 423–33.

7. Siegel, S.R., Mackenzie, J., Chaplin, G., Jablonski, N.G. and Griffiths, L. (2012) Circulating microRNAs involved in multiple sclerosis. Mol. Biol. Rep., 39, 6219–25.

8. Murata, K, Furu, M., Yoshitomi, H., Ishikawa, M., Shibuya, H., Hashimoto, M., Imura, Y., Fujii, T., Ito, H., Mimori, T., et al. (2013) Comprehensive microRNA analysis identifies miR-24 and miR-125a-5p as plasma biomarkers for rheumatoid arthritis. PLoS One, 8, e69118.

9. Nielsen, L.B., Wang, C., Srensen, K, Bang-Berthelsen, C.H., Hansen, L., Andersen, M.-L.M., Hougaard, P., Juul, A., Zhang, C.-Y., Pociot, F., et al. (2012) Circulating levels of microRNA from children with newly diagnosed type 1 diabetes and healthy controls: evidence that miR-25 associates to residual beta-cell function and glycaemic control during disease progression. Exp. Diabetes Res., 2012, 896362.

10. Magni, S., Comani, G. and Elli, L. (2014) miRNAs Affect the Expression of Innate and Adaptive Immunity Proteins in Celiac Disease. Am. J. Gastroenterol., 109, 1662–74.

11. Vaira, V., Roncoroni, L., Barisani, D., Gaudioso, G., Bosari, S., Bulfamante, G., Doneda, L., Conte, D., Tomba, C., Bardella, M.T., et al. (2014) microRNA profiles in coeliac patients distinguish different clinical phenotypes and are modulated by gliadin peptides in primary duodenal fibroblasts. Clin. Sci. (Lond)., 126, 417–23.

12. Capuano, M., Iaffaldano, L., Tinto, N., Montanaro, D., Capobianco, V., Izzo, V., Tucci, F., Troncone, G., Greco, L. and Sacchetti, L. (2011) MicroRNA-449a overexpression, reduced NOTCH1 signals and scarce goblet cells characterize the small intestine of celiac patients. PLoS One, 6, e29094.

13. Hogen Esch, C.E., Rosn, A., Auricchio, R., Romanos, J., Chmielewska, A., Putter, H., Ivarsson, A., Szajewska, H., Koning, F., Wijmenga, C., et al. (2010) The PreventCD Study design: towards new strategies for the prevention of coeliac disease. Eur. J. Gastroenterol. Hepatol., 22, 1424–30.

14. Vriezinga, S.L., Auricchio, R., Bravi, E., Castillejo, G., Chmielewska, A., Crespo Escobar, P., Kolaek, S., Koletzko, S., Korponay-Szabo,

I.R., Mummert, E., et al. (2014) Randomized Feeding Intervention in Infants at High Risk for Celiac Disease. N. Engl. J. Med., 371, 1304−1315.

15. Hackenberg, M., Rodrguez-Ezpeleta, N. and Aransay, A.M. (2011) miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. Nucleic Acids Res., 39, W132−8.

16. Anders, S. and Huber, W. (2013) Differential expression of RNA-Seq data at the gene level − the DESeq package. r-project.org, 1-24.

17. Benjamini, Y., and Y.H. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc, 57, 289−300.

18. Zeng, L., Cui, J., Wu, H. and Lu, Q. (2014) The emerging role of circulating microRNAs as biomarkers in autoimmune diseases. Autoimmunity, 23, 1−11.

19. Roderburg, C., Luedde, M., Vargas Cardenas, D., Vucur, M., Scholten, D., Frey, N., Koch, A., Trautwein, C., Tacke, F. and Luedde, T. (2013) Circulating microRNA-150 serum levels predict survival in patients with critical illness and sepsis. PLoS One, 8, e54612.

20. De Candia, P., Torri, A., Gorletta, T., Fedeli, M., Bulgheroni, E., Cheroni, C., Marabita, F., Crosti, M., Moro, M., Pariani, E., et al. (2013) Intracellular modulation, extracellular disposal and serum increase of MiR-150

mark lymphocyte activation. PLoS One, 8, e75348.

21. Zhang, Y., Liu, D., Chen, X., Li, J., Li, L., Bian, Z., Sun, F., Lu, J., Yin, Y., Cai, X., et al. (2010) Secreted monocytic miR-150 enhances targeted endothelial cell migration. Mol. Cell, 39, 133−44.

22. Valadi, H., Ekstrm, K, Bossios, A., Sjstrand, M., Lee, J.J. and Ltvall, J.O. (2007) Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat. Cell Biol., 9, 654−9.

23. Turchinovich, Samatov, T.R., Tonevitsky, G. and Burwinkel, B. (2013) Circulating miRNAs: cell-cell communication function? Front. Genet., 4, 119.

24. Gandhi, R., Healy, B., Gholipour, T., Egorova, S., Musallam, A., Hussain, M.S., Nejad, P., Patel, B., Hei, H., Khoury, S., et al. (2013) Circulating MicroRNAs as biomarkers for disease staging in multiple sclerosis. Ann. Neurol., 73, 729−40.

25. Liu, J., Wu, C.-P., Lu, B.-F. and Jiang, J.-T. (2013) Mechanism of T cell regulation by microRNAs. Cancer Biol. Med., 10, 131−7.

26. Ilus, T., Kaukinen, K., Virta, L.J., Huhtala, H., Mki, M., Kurppa, K, Heikkinen, M., Heikura, M., Hirsi, E., Jantunen, K, et al. (2014) Refractory coeliac disease in a country with a high prevalence of clinically-diagnosed coeliac disease. Aliment. Pharmacol. Ther., 39, 418−25.