Human cytomegalovirus microRNAs target prediction by dynamic expression analysis

M. Trevisan[1], T. Sanavia[2], C. Albonetti[1], E. Lavezzo[1], B. Di Camillo[2], A. Sinigaglia[1], S. Toppo[3], G. Toffolo[2], C. Cobelli[2], G. Palù[1] and L. Barzon[1]

[1] *Department of Histology, Microbiology and Medical Biotechnologies, University of Padova, Padova, Italy*
[2] *Department of Information Engineering, University of Padova, Padova, Italy*
[3] *Department of Biological Chemistry, University of Padova, Padova, Italy*

***Abstract*— Human cytomegalovirus belongs to the viral family of herpesviruses and its infection modulates different microRNAs (miRNAs) involved in cellular differentiation and tumor development. miRNA/mRNA dynamic expression analysis was performed using microarray data. 68 miRNAs were selected as differentially expressed and they were clustered in 6 groups according to their temporal expression profile. The prediction of target genes on one of selected miRNAs was carried out using an approach based on different predictive algorithms and a correlation analysis on miRNA/mRNA expression data. Results, validated using real time PCR, confirmed the relationships between this miRNA and some host's genes during infection.**

***Keywords*—Cytomegalovirus, miRNA, expression analysis**

I. Introduction

Human cytomegalovirus (HCMV) is an ubiquitous virus that can establish lifelong latent infection and reactivate periodically. In contrast to primary infection, which is generally asymptomatic, reactivation of the virus, particularly in immuno-compromised host, often results in severe disease.

HCMV encodes its own microRNAs (miRNAs) that are expressed during acute lytic infection. miRNAs are a large class of 22 nucleotide non-coding RNAs that post-transcriptionally regulate gene expression by binding to the 3' UTR sequence of mRNAs, leading either to translation inhibition or to mRNA degradation. These flexible, small-sized and non-immunogenic molecules are the ideal tools for viruses to regulate both their own and host's gene expression during infection. Currently, the activity of only a few HCMV miRNAs have been characterized and involved in viral life cycle and evasion from host antiviral response.

II. Materials and Methods

A. *Analysis of HCMV miRNA and mRNA expression during lytic infection in human fibroblasts*

Embryonic lung human fibroblasts MRC-5 were seeded at a density of 105 cells/well in a 24 well plate. The cells were infected with Towne HCMV strain at MOI 2 PFU/cell; at 4, 8, 24, and 48 hours post infection (p.i.), cells were harvested and total RNA was extracted and processed for both miRNA and mRNA expression analysis by using Agilent microarrays. Also time series without infection (control case) are available for both miRNA and mRNA expression analysis. Each miRNA measurement has 2 or 4 replicates whereas for mRNA measurements, dye-swap experiment was performed, thus each time point has 4 replicates: 2 with common reference in Cy5 and 2 with common reference in Cy3. Loess normalization without background subtraction was applied to both miRNA and mRNA intensity signal.

B. *Selection of differentially expressed miRNAs and mRNAs*

Differentially expressed miRNAs/mRNAs were selected using the method described in [1], which calculates the area of the region bounded by the *treated* (infected) *minus control* (not infected) expression profile. The method assigns a p-value to each miRNA/mRNA by evaluating the significance of this area against the null hypothesis. Available replicates were used to derive the experimental error distribution at different intensity expression values and the null hypothesis distribution. To account for multiple testing, the significance level was corrected according to a false discovery rate (FDR), i.e. the number of false positives divided by the number of selected miRNAs/mRNAs, equal to 0.05.

C. *Clustering of miRNAs expression profiles*

In order to identify the main temporal expression patterns p.i., *treated minus control* expression profiles of selected miRNAs were clustered using K means clustering based on Pearson correlation. Average Linkage was applied and the number K of clusters was set to 7.

D. *Identification of miR-US25-2-5p targets*

HCMV-encoded miRNA miR-US-25-2 was shown to reduce viral replication and DNA synthesis of HCMV [2]. Pre-miR-US25-2 encodes two complementary miRNAs, miR-US25-2-3p and miR-US25-2-5p, which were successfully cloned [3]. In the present study, prediction of target genes of miR-US25-2-5p, selected as differentially expressed, was performed. Candidate target genes were first identified by a method that integrates three existing algorithms (RNAhybrid, Pita, Miranda), selecting only genes predicted by all the three implemented methods. Starting from these candidate genes, correlation analysis was performed using Pearson correlation between miR-US25-2-5p and candidate genes time series. One-tailed test for associations was applied and only significant negatively-correlated target genes were selected considering a threshold on p-values equal to 0.05. Results were experimentally validated by repeating the correlation analysis using miR-US25-2-5p time series obtained from quantitative real time PCR using TaqMan® MicroRNA Assay Kit (Applied Biosystems). 10 ng of the RNA extracted were reverse-transcribed to cDNA using specific primers for this miRNA.

III. RESULTS

A. Expression of cellular miRNAs during HCMV infection

68 miRNAs were selected as differentially expressed p.i. Clustering analysis on miRNA temporal profiles identified 6 clusters of cellular miRNAs, involved in proliferation and cell differentiation and characterized by coherent changes in expression profile during the time course of infection (Figure 1). In particular, cluster 1 consists of miRNAs which were gradually downregulated and had low levels of expression at late time points p.i., such as miR-145, miR-143, miR-134, which recognize target involved in oncogenesis and are defined tumor suppressor miRNAs; cluster 2, including miRNAs which were markedly down-regulated since the early time points p.i. and during all the time course; cluster 3, including miRNAs transiently upregulated in the first 24 h p.i.; cluster 4 and 5, including miRNA up-regulated in late time points p.i. Cluster 4 consists of miRNAs that interfere with cell cycle arrest and apoptosis and are over-expressed in tumor tissues like the miRNA miR-17-92, which is regulated by c-Myc and binds the transcription factor E2F1, whereas miRNA miR-106b-25, belonging to this cluster, promotes entry into S phase by inhibiting CDKN1a/p21 the cyclin-dependent kinase. Cluster 5 is represented by miR-548 family, which has a large number of potential targets involved in many cellular processes, from cell proliferation to apoptosis and mitosis, suggesting a regulatory role related to the development of cancer. Finally, cluster 6 includes miRNA miR-192, showing a marked and progressive increase in expression levels since 24 h p.i.; its targets are p21 and other regulators of cell cycle arrest.

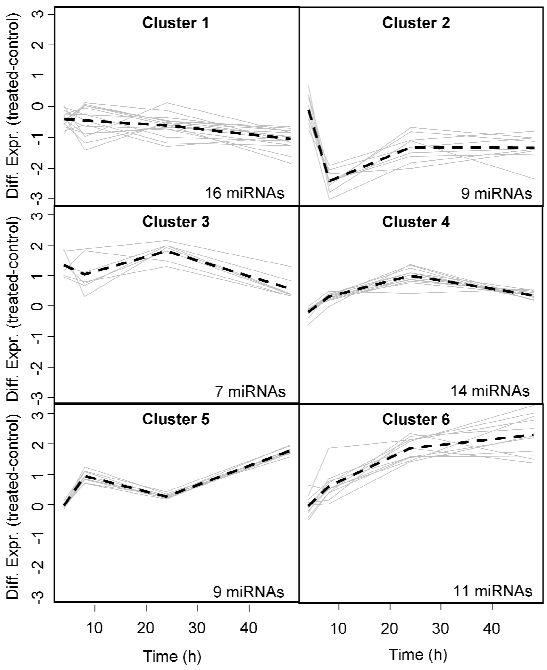

Fig. 1. Expression profile of miRNAs selected as differentially expressed, clustered in groups sharing the same temporal pattern.

B. Identification of miR-US25-2-5p targets

1034 genes were predicted by the three prediction algorithms for miR-US25-2-3p and 266 of these genes were selected as differentially expressed during post infection. For these 266 genes, correlation analysis between HCMV miR-US25-2-5p microarray expression profile and expression of predicted mRNA targets led to the selection of a set of 54 genes, which were down-regulated in HCMV infected fibroblasts and their expression profile was inversely related to HCMV miR-US25-2-5p expression profile. Using real time PCR expression profile of miR-US25-2-3p and repeating the correlation analysis, 49 genes were selected as target of this miRNA. Comparing the two list of target genes obtained from the two correlation analysis, 43 genes are in common.

Among these genes, *FAM127A* (family with sequence similarity 127, member A), and *FAM127B* (family with sequence similarity 127, member A), *SOD3* (extracellular superoxide dismutase) and *SGSM2* (small G protein signaling modulator 2) showed low p-values (less than 0.03) from the one-tailed test. *FAM127A* and *FAM127B* genes are transcribed during embryonic development but also in adult tissues and are down-regulated in tumors. *SOD3* encodes the extracellular superoxide dismutase, which has anti-inflammatory activity, whereas *SGSM2* is involved in vesicular transportation pathway.

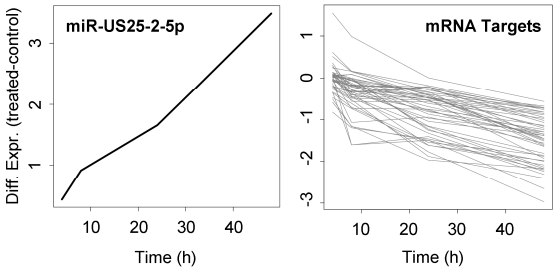

Fig. 2. Expression profile of miR-US25-2-5p and the related mRNA targets.

IV. CONCLUSION

An integrated analysis of viral and host miRNA, as well as host mRNA, in a time course experiment allowed the description of viral and host miRNA profiles during infection and the identification of host targets of viral miRNA. Target genes of a specific miRNA characterizing cytomegalovirus infection were identified using a metaconsensus approach based both on the application of three different prediction algorithms and a correlation analysis based on dynamic expression data. Correlation results obtained on microarray data are confirmed also by using the real time PCR expression profile of the selected miRNA. As future development, experimental validation of a pool of the resulting target genes will be carried out.

REFERENCES

[1] B. Di Camillo, G. Toffolo, S. K. Nair, L. J. Greenlund and C. Cobelli "Significance analysis of microarray transcript levels in time series experiments", *BMC Bioinformatics*, vol. 8, S10, March 2007.

[2] N. Stern-Ginossar, N. Elefant, A. Zimmerman, D. G. Wolf, N. Saleh, E. Horwitz *et al.* "Host immune system gene targeting by a viral miRNA", *Science*, vol. 317, pp. 376–381, July 2007.

[3] S. Pfeffer, A. Sewer, M. Lagos-Quintana, R. Sheridan, C. Sander *et al.* "Identification of microRNAs of the herpesvirus family", *Nat Methods*, vol. 2, pp. 269-276, April 2005.