A peer-reviewed version of this preprint was published in PeerJ on 5 January 2018.

Hsu P, Chiou B, Huang C. 2018. On revealing the gene targets of Ebola virus microRNAs involved in the human skin microbiome. PeerJ 6:e4138 https://doi.org/10.7717/peerj.4138

Title

Bioinformatics approach reveals the gene targets of Ebola virus microRNAs involved in the human skin microbiome

Authors

Pei-Chun Hsu [1*†], Bin-Hao Chiou[2] and Chun-Ming Huang [2,3]

*Correspondence: d98945017@ntu.edu.tw

[1]Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, No. 1, Sec. 4, Roosevelt Rd., Taipei 10617, Taiwan

[2]Department of Systems Biology and Bioinformatics, National Central University, Jhongli, Taiwan

[3]Department of Medicine, Division of Dermatology, University of California, San Diego, 3525 John Hopkins Court, Rm. 276, San Diego, CA, 92121, USA

Abstract

The Ebola virus, a negative-sense single-stranded RNA virus, causes severe viral hemorrhagic fever and is highly lethal. Histopathology and immunopathologic study of Ebola virus have revealed that histopathologic changes in skin tissue were mainly various degrees of endothelial cell swelling and necrosis. The interactions of microbes within or on a host are a crucial aspect of the skin immune shield. The discovery of microRNAs in Ebola virus implies that immune escape, endothelial cell rupture and tissue dissolution during Ebola virus infection are all results of the action of Ebola virus miRNAs.

Keratinocytes obtained from normal skin and subsequently attached and

spread on the thrombospondin protein family may play a role in initiating cell-mediated immune responses in the skin. Several miRNAs have been observed to bind the 3' untranslated region of the thrombospondin mRNA, thereby controlling its stability and translational activity.
In this study, we first discover short RNA sequences that might act as miRNAs from *Propionibacterium acnes* by design a practical workflow of bioinformatics methods. Subsequently, we deciphered the common target gene. These RNA sequences tend to binding to the same thrombospondin protein. These RNA sequences tend to bind to the same protein, THSD4, emphasizing the potential importance of the synergistic binding of miRNAs from Ebola virus, *Propionibacterium acnes*, and humans to the target. By RNA expression validation, we prove the potential synergistic binding of the miRNA from Ebola virus, *Propionibacterium acnes* and human to the target.

Keyword: Ebola virus, THSD4, *Propionibacterium acnes*, microRNA, microbiome, bioinformatics

Background

1. Ebola Virus Infections and Host Responses

The Ebola virus (EBOV), a negative-sense single-stranded RNA virus, which are five species identified, including four that mainly affect in humans. Ebola virus could cause sporadic outbreaks of hemorrhagic fever with a high fatality rate in Central Africa and Southeast Asia.

EBOV is primarily transmitted from human to human by direct contact with blood or tissue of patients. The virus infects host cell by destroying the protein synthesis systems and immune defenses (Beer, Kurth et al. 1999) at an unexpected high replication rate. Host immune responses against Ebola virus and direct infection of monocytes and macrophages result in the

release of cytokines. Those immune responses associated with inflammation and fever led to the Ebola hemorrhagic fever (EHF) (Yang, Duckers et al. 2000).

Histopathologic and immunopathologic studies of EBOV had revealed that the infection of EBOV is unconventional. Histopathologic changes in the skin tissue were consisted mainly of various degrees of endothelial cell swelling and necrosis (Zaki, Shieh et al. 1999). For some cases, EBOV infections would cause moderate to severe clinical manifestations.

2. Bioinformatics analysis and prediction of potential EBOV microRNA

Recently, the genome-wide scanning of 2014 outbreak EBOV suggests that putative viral microRNAs (miRNAs) in EBOV imply the potential EVOB miRNA target genes. The interactions of miRNA and target genes during EBOV infection may regulate the endothelial cell rupture and tissue dissolution related with immune escape of host defense mechanism (Teng, Wang et al. 2015).

MicroRNAs (miRNAs) are tiny piece regulatory non-coding RNAs, about 18 to 23 nucleotides in length, that are processed from long precursor transcripts (pre-miRNAs). Cellular miRNAs are highly conserved and occur naturally binding to matching pieces of messenger RNA on its 3'-untranslated regions (3'-UTR) of target gene thereby down regulating the production of the corresponding protein (Filipowicz, Bhattacharyya et al. 2008, Siomi and Siomi 2010). The most rigorous discipline of miRNAs is their binding to mRNAs at nucleotide positions 2-7, representing the "seed region" or "seed sequence". Although the base pairing of a miRNA and its target mRNA is not a perfect match, the seed sequence must be perfectly complementary (Hibio, Hino et al. 2012).

3. The interactions of microbes and opportunistic pathogens

The opportunistic pathogens such as *Propionibacterium acnes* (*P. acnes*), a gram-positive anaerobic predominant bacterium in the skin; either within or on a host, is a crucial aspect of the skin's immune shield (Schommer and Gallo 2013). Microbial colonization on the skin, also known as the skin microbiome, strengthens the skin's defense against potentially pathogenic organisms. Keratinocytes obtained from normal skin attached and spread on thrombospondin (TSP, also referred to as THSD in the literature) protein families may play a role in initiating cell-mediated immune responses in the skin by releasing cytokines and compelling the TSP protein family (Varani, Nickoloff et al. 1988) to facilitate the movement of immune competent cells (Baker, Ovigne et al. 2003). TSP promotes keratinocyte attachment and spreading, this molecule may play an important role in maintaining the normal growth of the basal cell layer. To date, several miRNAs have been observed to bind the 3'UTR of the TSP mRNA, as well as controlling its stability and translational activity (Bartel 2009).
The human genome contains more than 1,000 miRNA genes that have been either predicted or experimentally observed to play critical roles in normal cellular functions, such as maintaining homeostasis, and regulating or modulating viral and cellular gene expression. Thus, miRNAs constitute one of the most abundant classes of gene-regulatory molecules in animals.

4. A new approach to discover EBOV infectious mechanism

Currently, no specific therapy demonstrated to be effective for the treatment of EHF is available. Doctors now can only provide supportive therapies to conquer the infection but fatality rate is still high. EBOV can spread through direct skin contact. Recent studies have proposed that EBOV infection is result of the action of EBOV miRNAs (Liang, Zhou et al. 2014). As previously mentioned, microbes in the skin microbiome, such as *P. acnes*, potentially play a vital role in protecting the host. In this paper, we present a bioinformatics approach based on sequence alignment that

detects some short RNA fragments in *P. acnes,* which are similar to viral miRNAs in EBOV, as well as provide a possible solution for determining the mechanism of EBOV infection.

Methods

1. Identify short RNA fragments in *P. acnes*

The whole genome of *Propionibacterium acnes* KPA171202 (Bruggemann, Henne et al. 2004) was retrieved from the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/nuccore/NC_006085.1). The predicted mature miRNAs in EBOV are used as query sequence.

The Basic Local Alignment Search Tool (BLAST) Two sequence program was employed to search against the entire *P. acnes* genome at the NCBI (http://blast.ncbi.nlm.nih.gov/). The specialized BLAST program, called "Align Two Sequences" (bl2seq) with the following parameter setting—word length: 11; E-value cutoff: 10—was selected to perform pairwise local alignments of whole genome sequences. The RNA sequences used as query in BLAST were as follow: EBV-miR-1-5p,5′-AAAAAGUCCAUAAUGCUGGGGA-3′; EBV-miR-1-3p,5′-GCCACCAUAGGACUUUUUCAAU-3′; EBV-miR-2-3p,5′-UUAUCCUUCUUGAAUCCUGAGA-3′.

2. Target gene prediction

The target genes of RNA fragments were predicted by TargetScan online software (release 6.0) (http://www.targetscan.org/) (Agarwal, Bell et al. 2015). MiRNAs regulate protein coding gene expressions by binding to the 3′ UTR, TargetScan is specifically designed for predicting such interactions, and thus, was employed to predict the binding of the seed regions of miRNAs. We employed customized TargetScanHuman software (verson5.2) (http://www.targetscan.org/vert_50/seedmatch.html) to search for the predicted target genes of small RNA fragments. This customized software

is specially developed to predict biological targets of RNA fragments by searching for the presence of 7-8 nucleotides that match the seed region of each RNA fragments. For instances, a short nucleotide sequence, nucleotides 1-7 from the 5′end of *P. acnes* RNA fragments, referred to as the seed sequence. The RNA fragments applied for these analyses are shown in Table 1 and came from three sources: EBOV, *P. acnes* and human. As a result, the three groups of gene lists were separated by using EBOV miRNA, P. acnes short RNA sequences and Human microRNA mir-1248 that was found from miRBase as query. We cross-match these three gene lists to find the common target gene.

3. Analysis of mRNA Expression of THSD4

To explore the regulation of gene expression mediated by miRNA, it is necessary to identify the target genes and regulating efficacies of each miRNA. In order to further demonstrate the regulatory role THSD4, Human embryonic kidney cell-line 293T (HEK293T) cells (purchased from American Type Culture Collection, ATCC company) were transfected with short RNA sequences listed in Figure 2. With the use of the siTran transfection reagent (Origene) according to the manufacturers protocol and incubated for 48 hr. The expression level of THSD4 mRNA is quantified by real-time quantitative PCR (RT-qPCR). The fold change of mRNA was compute by the change of threshold cycle (Ct) and represented by $2^{-\triangle\triangle Ct}$ value.

Results and Discussion

1. Prediction of P. acnes small RNA fragments from EBOV miRNAs

With the explosion of sequence and biological information available to researchers, the field of bioinformatics, or more properly, computational biology, is playing an increasingly large role in the study of fundamental biomedical problems. The challenges that biologists have faced are the vast amount of data produced by experiments and potentially reveal previously unknown relationships with respect to the function of genes and proteins. Provide an easy way for biologists to integrate information from

different biological database by themselves than by programmers. We create a workflow of retrieving entries from databases and annotating collection of publicly available sequences, shown in figure 1.

MiRNAs are thought to be found mostly in eukaryotes. Recently, miRNAs have also been employed by viruses to regulate the expression of their own genes, the host's genes, or both (Grice and Segre 2011). EBOV miRNAs can be generated through cellular miRNA processing machinery. Three mature miRNAs are identifiable from EBOV (Liang, Zhou et al. 2014).

By mapping the clean reads of EBOV microRNA to the *P. acens* and human genome, 16 small RNA fragments from *P. acnes* and one human miRNA were detected in this study. We obtained only sequences with 100% similarity, as well as several mismatches. The lengths of the identified small RNA sequences were approximately 13 to 11 nt each. The RNA sequence listed in Table 1 represent newly identified small RNA sequences identified. The RNA sequence CCUUCUU of PA-mir-2-3p-5 is identical to the seed region of the human miRNA family, miR-1248.

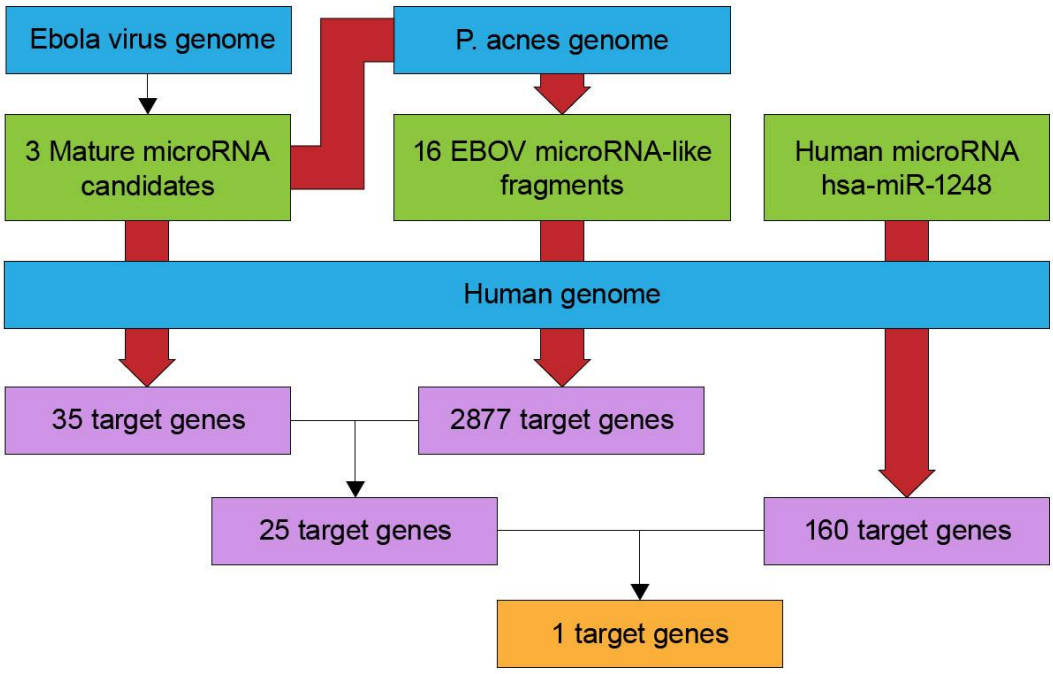

Figure 1. Flowchart of target genes prediction.

The BLAST Two sequence program was employed to predict short RNA sequences from P. acnes. All the target genes were predicted using the

TargetScan web service. Each red arrow represents a BLAST process.

2. Target gene prediction by using small RNA fragments

Total of 20 small RNA fragments, including one human mir-1248 miRNA, were used as query to retrieve TargetScanHuman database to scan the potential target genes for these small RNA fragments. TargetScan software scanned the 3′UTR of human genes that complementary between the first 2–7 nucleotides of small RNA fragment. We found that each RNA fragment could target to more than one gene. There are 2877 target genes identified from 16 P. acnes small RNA fragments, 35-target gene from 3 EBOV miRNA, and 160 target genes from human miRNA family miR-1248. Interestingly, three different groups of RNA fragments targeted to a unique gene THSD4.

Thrombospondins (TSP) are a family of multidomain extracellular matrix glycoproteins. The unique gene THSD4 is recently found as a new gene in NCBI database known as thrombospondin, type I, domain containing 4. However, the function of THSD4 has not clear.

Table 1. Short RNA sequences employed in this study.

The target genes of RNA fragments were predicted by TargetScanHuman online software. The short 7 nucleotides in RNA sequences, highlighted in bold, are referred to seed sequences.

RNA sequence id	Sources	RNA sequences(5' -> 3')	length(bp)
EBV-miR-1-5p	Ebola virus[12]	A**AAAAGUC**CAUAAUGCUGGGGA	22
EBV-miR-1-3p	Ebola virus[12]	G**CCACCAU**AGGACUUUUUCAAU	22
EBV-miR-2-3p	Ebola virus[12]	U**UAUCCUU**CUUGAAUCCUGAGA	22
PA-mir-1-5p-1	*P. acnes*	**GCAUUAU**GGACU	12
PA-mir-1-5p-2	*P. acnes*	**AUCAUGG**ACUUUUU	14
PA-mir-1-5p-3	*P. acnes*	**AAGUCCA**UAAU	11
PA-mir-1-5p-4	*P. acnes*	**CCCCAGC**AUUA	11
PA-mir-1-3p-1	*P. acnes*	**CCUAUGG**UGGC	11
PA-mir-1-3p-2	*P. acnes*	**GCCACCA**UAGG	11

The gene targets of Ebola virus microRNAs

PA-mir-1-3p-3	*P. acnes*	**GUCCUAU**GGUG	11
PA-mir-1-3p-4	*P. acnes*	**AUAGGAC**UUUU	11
PA-mir-1-3p-5	*P. acnes*	**AGGACUU**UUUCGAU	14
PA-mir-1-3p-6	*P. acnes*	**UGAAAAA**GUCC	11
PA-mir-1-3p-7	*P. acnes*	**ACUUUUU**CAAU	11
PA-mir-2-3p-1	*P. acnes*	**UAUCCUU**CUUGA	12
PA-mir-2-3p-2	*P. acnes*	**AUCCUUC**UUGA	11
PA-mir-2-3p-3	*P. acnes*	**UCAAGAA**GGAU	11
PA-mir-2-3p-4	*P. acnes*	**UCCUUCU**UGAA	11
PA-mir-2-3p-5	*P. acnes*	**CCUUCUU**GAAU	11
hsa-miR-1248	Human[40]	A**CCUUCUU**GUAUAAGCACUGUGCUAAA	27

Note: P. acnes: Propionibacterium acnes KPA171202

3. Clustering and conserved segment of human microRNAs

MiRNAs in animals are found in diverse genomic locations. Most miRNAs are encoded in intergenic regions, but there are also many miRNAs that are hosted within the introns of pre-mRNAs or encoded within non-coding RNA genes. We first align RNA sequence from microbiome to human THSD4 gene. There are two conserved segments "AGAAGG" was found to located in THSD4 3' UTR region which show in Figure 2. This segment was also the target of seed region of human microRNA mir-1248 "UCUUCC". We next looked at the microbiome RNA sequence found in our study that PA-mir-2-3p-5 and EBV-miR-2-3p are with this segment as well. Moreover, a relaxed requirement of a 6-nt match to seed complementary 2-7 position of miRNA was announced to be retaining specificity (Lewis, Burge et al. 2005).

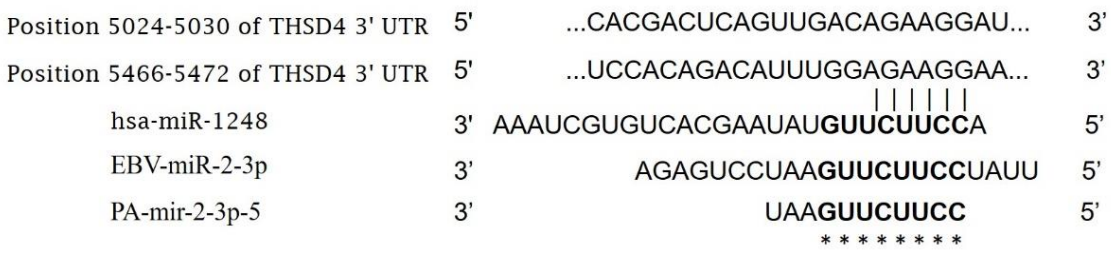

Figure 2. Alignment of orthologues RNA sequences to Human THSD4 DNA

sequence. The RNA sequence PA-mir-2-3p-5 came from P. acnes and EBV-miR-2-3p came from ebola virus.

4. Analysis of mRNA Expression of THSD4

To evaluate the importance of the 3′-terminal segments for the impact of translational control. HEK293T cells were transfected with short RNA sequences listed in Figure 2 with the use of the siTran transfection reagent (Origene). Real-time quantitative PCR was used to measure the expression level of THSD4 mRNA. We first note that cells transfected with PA-mir-2-3p-5 and EBV-miR-2-3p short RNA fragments show similar transcriptional level as transfected with human mir-1248 miRNA as demonstrate in Figure 3. In this experiment, we revealed that PA-mir-2-3p-5 might be a very short RNA sequence with only 11 nucleotides but it could also behave as a miRNA.

In addition, we also like to know the potential synergistic effect among human and microbiome RNA fragments. The cells were co-transfected with human mir-1248 miRNA plus PA-mir-2-3p-5; human mir-1248 miRNA plus EBV-miR-2-3p; and PA-mir-2-3p-5 plus EBV-miR-2-3p. Interestingly, we found that dual transfection of the cell shows higher expression fold change than transfection alone. These findings support our previous hypothesis that short RNA fragments with a 6-nt match to seed complementary segment could function as a microRNA. The 3′UTR of mRNA tend to contain several binding sites for one miRNA, emphasizing the potential importance of synergistic binding of the miRNA to the target. Although thrombospondins were known to be a astrocyte-secreted glycoproteins that might influences central nervous system (CNS) synaptogenesis (Wang, Guo et al. 2012). However, thrombospondins, type I, domain containing 4 (*THSD4*) is a new gene in NCBI gene database. The function of THSD4 has not defined.

Recent work has proposed that EBOV infection might due to the action of the EBOV miRNAs (Liang, Zhou et al. 2014). As mentioned above, *P. acnes*, potentially play an important role to protect the host. We present a

bioinformatics approach that detects the short RNA structure in P. acnes based on sequence alignment, which may act as miRNA and protect human from evasion of EBOV by regulation the expression their common target gene THSD4.

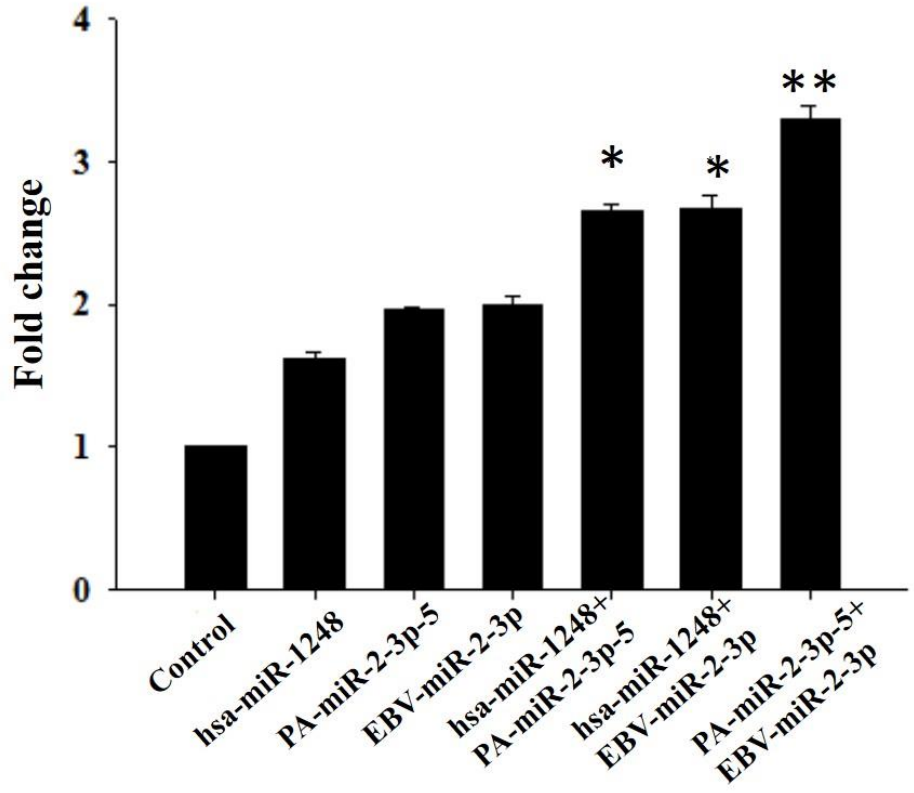

Figure 3. Relatived expression of THSD4 mRNA. Bar graph representing the fold changes of mRNA levels with ± SEM (n=3). * * . P< 0.001.

Conflict of Interest

The authors declare that they have no competing interests.

A vailability of data and materials

The target genes of EBOV-encoded miRNAs and miRNA sequence could be downloaded from http://link.springer.com/article/10.1007%2Fs11427-014-4759-2#SupplementaryMaterialhttp://link.springer.com/article/10.1007%2Fs11427

The gene targets of Ebola virus microRNAs

-014-4759-2 - SupplementaryMaterial

Author's contributions

Professor Huang offered the conception for design of the work. Bin-Hao Chiou finished the data analysis. The data collection, interpretation and draft the article were done by Pei-chun Hsu.

Acknowledgements

I wish to thank Dr. Eric Huang for thoughtful discussion and brilliant idea. I am grateful to Dr. Chao Kun-Mao for his encouragement, tolerance, and gentle direction. I would also like to thank my colleagues for their friend-ship and support.

We also appreciate kind assistance of Professor Kang in Graduate Institute of Toxicology of National Taiwan University College of Medicine for providing HEK293T cell.

References

Agarwal, V., et al. (2015). "Predicting effective microRNA target sites in mammalian mRNAs." Elife **4**.

Baker, B. S., et al. (2003). "Normal keratinocytes express Toll-like receptors (TLRs) 1, 2 and 5: modulation of TLR expression in chronic plaque psoriasis." Br J Dermatol **148**(4): 670-679.

Bartel, D. P. (2009). "MicroRNAs: target recognition and regulatory functions." Cell **136**(2): 215-233.

Beer, B., et al. (1999). "Characteristics of Filoviridae: Marburg and Ebola viruses." Naturwissenschaften **86**(1): 8-17.

Bruggemann, H., et al. (2004). "The complete genome sequence of Propionibacterium acnes, a commensal of human skin." Science **305**(5684): 671-673.

Filipowicz, W., et al. (2008). "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?" Nat Rev Genet **9**(2): 102-114.

Grice, E. A. and J. A. Segre (2011). "The skin microbiome." Nat Rev Microbiol **9**(4): 244-253.

Hibio, N., et al. (2012). "Stability of miRNA 5'terminal and seed regions is correlated

with experimentally observed miRNA-mediated silencing efficacy." Sci Rep **2**: 996.

Lewis, B. P., et al. (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell **120**(1): 15-20.

Liang, H., et al. (2014). "Identification of Ebola virus microRNAs and their putative pathological function." Sci China Life Sci **57**(10): 973-981.

Schommer, N. N. and R. L. Gallo (2013). "Structure and function of the human skin microbiome." Trends Microbiol **21**(12): 660-668.

Siomi, H. and M. C. Siomi (2010). "Posttranscriptional regulation of microRNA biogenesis in animals." Mol Cell **38**(3): 323-332.

Teng, Y., et al. (2015). "Systematic Genome-wide Screening and Prediction of microRNAs in EBOV During the 2014 Ebolavirus Outbreak." Sci Rep **5**: 9912.

Varani, J., et al. (1988). "Thrombospondin-induced adhesion of human keratinocytes." J Clin Invest **81**(5): 1537-1544.

Wang, B., et al. (2012). "Thrombospondins and synaptogenesis." Neural Regen Res **7**(22): 1737-1743.

Yang, Z. Y., et al. (2000). "Identification of the Ebola virus glycoprotein as the main viral determinant of vascular cell cytotoxicity and injury." Nat Med **6**(8): 886-889.

Zaki, S. R., et al. (1999). "A novel immunohistochemical assay for the detection of Ebola virus in skin: implications for diagnosis, spread, and surveillance of Ebola hemorrhagic fever. Commission de Lutte contre les Epidemies a Kikwit." J Infect Dis **179 Suppl 1**: S36-47.