

Relationship between Mutations in DNA Sequences Loci Coding Pre-miRNAs and Genes Related to Biogenesis of SncRNAs with MiRNA Expression in Endometrial Carcinoma Tissues

Massimiliano Chetta¹, Angela Cordella², Giuseppina Bruno¹, Maurizio Guida³, Alessandro Weisz^{1,4} and Maria Ravo¹

1. Department of Medicine and Surgery—Laboratory of Molecular Medicine and Genomics, University of Salerno, Baronissi 84081, Salerno, Italy

2. Foundation Institute for Hospitalization and Scientific Care (IRCCS) SDN, Naples 80143, Italy

3. Department of Obstetrics and Gynecology, “San Giovanni Di Dio Ruggi D’Aragona” University Hospital—Schola Medica Salernitana, Salerno 84131, Italy

4. Department of Molecular Pathology and Medical, “San Giovanni Di Dio Ruggi D’Aragona” University Hospital—Schola Medica Salernitana, Salerno 84131, Italy

Abstract: Endometrial cancer (EC) is the most common and lethal gynaecological cancer type in Europe and in North America. Frequently EC arises more in the corpus proper and manifests as round, polypoid expansile masses, but it may also originate in the lower uterine segment or spread in endometrium with necrosis and hemorrhage. The analysis was performed using a custom panel containing all DNA sequences loci coding pre-miRNAs and genes related to biogenesis and regulation of sncRNAs in normal and tumor tissues extracted from 6 unrelated patients with endometrial carcinoma. The identified variations were correlated with mature miRNAs differentially expressed in the same normal and tumor endometrial tissues. The comparison analysis confirmed the high degree of cellular and genetic intratumoral heterogeneity with a temporal and spatial miRNA expression distribution in association with genomic variants identified. The classification of specific DNA mutations, onto the loci identified, should be suitable to characterize possible instability genome regions and help classification of tumors to ameliorate the clinical management of patients affected by endometrial carcinoma.

Key words: Endometrial cancer, pre-miRNAs, sncRNA, miRNA expression.

1. Introduction

EC is one of the most common invasive and lethal gynecological cancer types in Europe and North America [1]. It is also considered the most common cause of death from cancer in women in Western Europe (1%-2% of all deaths from cancer), and the incidence is increasing. Median age of occurrence is 63 years and > 90% of affected women are older than

50. The majority of endometrial cancers occur after menopause, but up to 25% of cases may be premenopausal [2]. According to the FIGO (International Federation of Gynecology and Obstetrics) classification system 2006 [3], endometrial cancer or uterine sarcomas can be divided into four stages related to the region location (confined to the body of the uterus) or widespread in the body. Various evidences suggest that endometrium is subject to several autocrine and paracrine regulators, including many growth factors, cytokines, chemokines, proteases,

Corresponding author: Massimiliano Chetta, Ph.D., research fields: genetics and genomics. E-mail: mchetta@unisa.it.

extracellular matrix and potential regulators of their signaling pathways that include sncRNAs. Now, it is clearly established the functional role of sncRNAs and aberrant expression has been associated with inappropriate tissue regeneration, embryo implantation failure and other uterine abnormalities, including dysfunctional uterine bleeding, endometrial hyperplasia and carcinoma [4]. The present study has investigated occurrence of mutations in a custom panel containing all DNA sequence loci coding pre-miRNA and genes related to biogenesis and regulation of sncRNAs using a NGS approach. The possible role of DNA variants was corroborated by direct observation of differentially mature miRNAs expression characterized in the same normal and tumor endometrial tissue, which in order to associate somatic mutations with temporal and/or spatial miRNAs distribution.

2. Material and Methods

2.1 Patients

Patients were subjected to transvaginal ultrasound and office hysteroscopy in the Departments of Obstetrics and Gynecology of the University of Salerno and University of Naples Federico II. Endometrial biopsies from tumor (T) and adjacent normal (N) tissue specimens from 6 patients were selected for this study. Informed consent was obtained from all patients. Genomic DNA from endometrial biopsies was extracted using GenElute Mammalian Genomic DNA Miniprep Kit (Sigma-Aldrich), following the manufacturer's instructions. A custom HaloPlex panel was designed using Agilent's Sure Design tool (S1, supplementary materials).

2.2 Enlis Genome Analysis

Enlis Genome software platform is used for the analysis. The following filters are applied to identify variations: "IN single genome parameter" (tumor sample) and "NOT in single genome" (normal sample) and quality score ≥ 30 . The data are also filtered for the custom panel [5].

2.3 MiRNA Expression

Identification of differentially expressed miRNAs was performed with the Bioconductor DESeq package available in iMir tool. A negative binomial distribution model was used to test differential expression in deep sequencing datasets. Data were filtered according to read count value (threshold: 10 reads) and those showing fold-change ≤ -1.5 or ≥ 1.5 ($|FC| \geq 1.5$), with a p-value ≤ 0.05 which were considered as differentially expressed [6].

2.4 Pre-miRNAs Structure Prediction

The pre-miRNAs structure prediction was obtained using Oligo Analyzer 3.1. Using the HAIRPIN function, the RNA secondary structures by mFold algorithm were predicted [7].

2.5 Dicer Cleavage Sites Prediction

The in silico Dicer cleavage sites prediction is a bioinformatics tool that use SVM and process the Dicer sites at 5 prime (5') arm of pre-miRNAs. The algorithm generates 14 nt long sliding structure, having cleavage site at center and binary pattern. The results generate several cleavage sites with different SVM score and site. The highest score is considered as most probable cleavage site (S2, supplementary materials) [8].

2.6 DROSHA Cleavage Sites Prediction

The in silico DROSHA cleavage sites prediction was performed using a Microprocessor SVM. The software was designed to predict the DROSHA processing site that determines one end of the mature microRNA in human genome (S3, supplementary materials) [9].

2.7 MiRNA Target Prediction

The in silico target prediction is performed using DIANA-microT-CDS. The analysis was performed using the default threshold miTGs score of 0.8, which provides an average of 350 targets per miRNA. The

mature miRNA sequences with and without variants are analyzed for each prediction [10].

3. Results

The presence of mutations in DNA sequences was investigated in tumor tissues and corresponding normal tissues, which obtained from 6 unrelated patients with endometrial carcinoma, using a targeted NGS approach. Identification of all variations in pre-miRNA DNA sequences loci by means of Enlis bioinformatics analysis (S4, supplementary materials). Different bioinformatics tools are used to analyze the possible effect of variations on secondary RNA stability (S5, supplementary materials) and/or possible Dicer and DROSHA cleavage sites alterations (S2-S3, supplementary materials) (Fig. 1). Furthermore, potential modifications in mature miRNA target sites were predicted when the detected variant was localized into the mature miRNA sequence (S6, Supplementary Materials). All miRNAs mutated are summarized in Table 1. Variations in genomic sequence were not found in the samples 7T and 14T

between normal and tumor.

The analysis of DNA sequences of genes involved in biogenesis and regulation of sncRNAs also showed a widespread distribution of mutations onto different loci except for the sample 1T and 14T which were not identified sequence variants (S7, supplementary materials).

The sample 3T contains the majority number of variants spread in different genes: 75 variations in 33 different genes, 2 nonsense new transversions, 6 synonymous variants, 37 intronic variants distributed in 17 genes and finally 20 new variants spread in 3 prime (3') UTR sequence of 11 different genes. Within these variants, the presence of several homozygote polymorphisms was underlined and could be associated to LOH (Loss of heterozygote) regions. In the sample 4T, a frame shift mutation was identified, caused by a deletion of Adenine in ATM (ataxia-telangiectasia mutated) gene and a new mutation in 3' UTR sequence of AR gene (androgen receptor). In the sample 7T, it was reported a new intron transversion in SMARCA4 (SWI/SNF related,

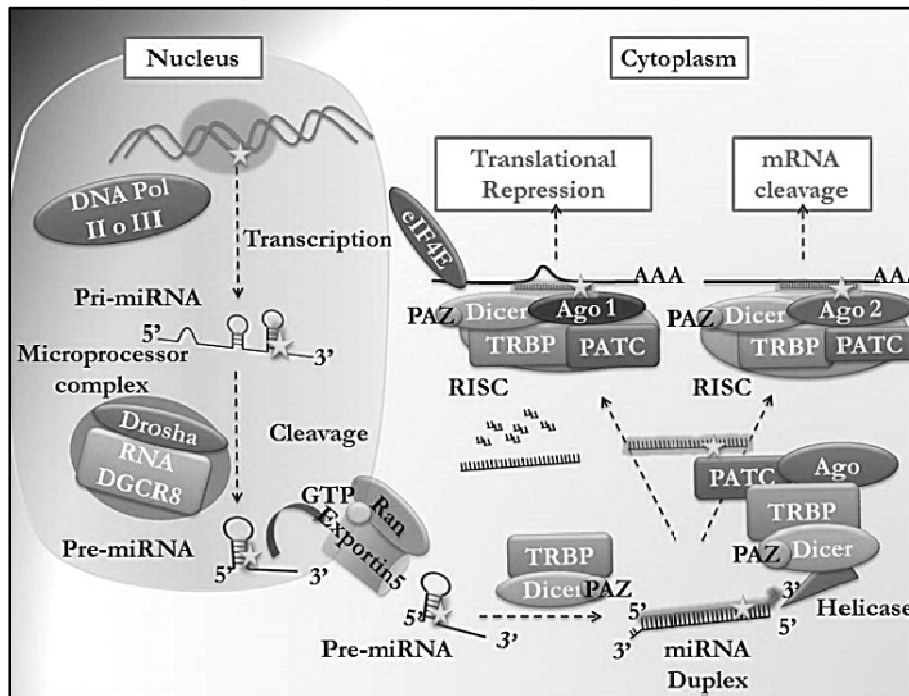


Fig. 1 The miRNA processing pathway, stars indicate same point mutation in miRNA biogenesis and its corresponding bioinformatics analysis.

Table 1 Results from the bioinformatic analysis (in italics were highlighted chromosome locus 14q32.2-14q32.31, pre-miRNAs that were mutated in more samples are indicated in bold).

	Predicted target sites	miRNA fold change
Sample 1T		
hsa-miR-941-1	No variation	> 3.48
Sample 3T		
hsa-miR-149	No variation	< -1.55
hsa-miR-182	No variation	> 630.44
hsa-miR-4477A	Target genes ranking variations	No detected
hsa-miR-612	No variation	No detected
(14q32.2-14q32.31)		
<i>hsa-miR-369</i>	No variation	< -4.11
<i>hsa-miR-345</i>	Increase affinity for UBAP2	No detected
<i>hsa-miR-487A</i>	No variation	No detected
<i>hsa-miR-539</i>	Variations of genes ranking	No detected
<i>hsa-miR-323B</i>	Increase number of target sites	No detected
<i>hsa-miR-382</i>	Decrease number of target sites	No detected
<i>hsa-miR-329-1</i>	Increase number of target sites	No detected
<i>hsa-miR-154</i>	No variation	< -10.94
<i>hsa-miR-548AW</i>	No variation	No detected
<i>hsa-miR-4481</i>	No variation	No detected
hsa-miR-608	No variation	No detected
<i>hsa-miR-412</i>	No variation	No detected
<i>hsa-miR-5189</i>	No variation	No detected
<i>hsa-miR-3615</i>	No variation	> 8.58
Sample 4T		
hsa-miR-4477A	Conservation of 13 target genes variations of genes ranking	No detected
Sample 15T		
hsa-miR-149	No variation	< -3.61
hsa-miR-612	No variation	No detected
hsa-miR-608	No variation	No detected
hsa-miR-941-1	No variation	> 5.29
hsa-miR-4432	No variation	No detected
hsa-miR-5579	87 more interacting target genes	No detected
hsa-miR-4751	No variation	No detected
hsa-miR-4752	No variation	No detected

matrix associated, actin dependent regulator of chromatin, subfamily a, member 4). Finally, in the sample 15T, a total of 18 variants distributed in 11 different genes: 1 new frameshift mutation, a homozygous polymorphic variant, 2 polymorphisms, 2 homozygous polymorphisms in ESR1 (Estrogen receptor alpha) and NBN (Nibrin), new heterozygote synonymous and finally 10 distributed in IVS regions

were identified (additional information on mutations is reported in S7, supplementary materials).

4. Discussion

SncRNAs are a class of regulatory molecules with the ability to control gene expression at the post-transcriptional level through degradation, repression or silencing. Each sncRNAs in particular

miRNAs has been predicted to have a broad range of target mRNA based on the degree of sequence homology. Changes in the sequences expression of even a single miRNA could have a detrimental impact on the outcome of different cellular activities [11]. The aim of study was the identification of DNA variants in pre-miRNA loci and genes related to biogenesis of sncRNAs, by comparing the sequences of DNA extracted from normal and tumor endometrial tissues. The variations identified were correlated with differentially mature miRNAs expression.

In particular, in sample 1T, a new heterozygous transition in hsa-miR-941-1 was detected. The hsa-miR-941-1 is located in the first IVS of the DNAJC5 gene (DnaJ (Hsp40) homolog, subfamily C and member 5) onto Chr20q13.33 within a miRNA cluster. The mature miR-941 was highly expressed in pluripotent cells, preferentially targets genes in hedgehog and insulin signaling pathways suggesting a central roles in cellular differentiation [11]. The same variant was found in sample 15T. The bioinformatics analysis revealed that this variant doesn't modify minimum Gibbs free energy. Consequently, the predicted RNA secondary structures, whereas, increase PHD score for Dicer cleavage site in a loss of site in nucleotide position 30th. This result was compatible with an increase RPM in sample 1T and 15T and FC of 3.48 and 5.29 respectively in the tumor as previously described [5].

Several mutations are identified in sample 3T. A polymorphism rs2674474, which also present in sample 4T, was detected into the seed region of hsa-miR-4477A. This polymorphism is not reported in the literature with a specific role in tumor differentiation. The polymorphism induced an increase of ΔG (Gibbs free energy) and an alteration of Dicer sites PHD score, but there is no evidence of expression difference between tumor and normal endometrial tissue. The predicted target sites for mature miRNA hsa-miR-4477A analysis showed a conservation of only 13 genes target site and entirely

variation of genes ranking. The possible involvement of hsa-miR-4477A was supported by the presence of the same variant in sample 4T. The other two polymorphisms rs2292832 and rs550894 are located on Chr2 and Chr11 respectively. The first one involved miR-149 that was found frequent hypermethylated and down regulated in endometrial cancer lines [12], instead, the second polymorphism concerned miR-612 was discovered as suppressor in EMT (epithelial-mesenchymal transition) and metastasis in hepatocellular carcinoma [13]. The polymorphism in 3' UTR region of hsa-miR-149 induced a decrease of free energy suggesting a subsequently thermodynamically more stable secondary structure and decrease of relative mature miRNA concentration. The expression analysis shows a negative FC of -1.55 for hsa-miR-149-5p. The polymorphism in 3' UTR region of hsa-miR-612 produced an increase of ΔG , and a different location of cleavage sites for DROSHA. No differential expression was found for hsa-miR-612 compatible with more late stage tumor development and related to transformation of cells into migratory and invasive. Also on Chr7, a new heterozygous transition C > T was detected in hsa-miR-182. Aberrant hsa-miR-182 expression has been implicated in melanoma metastasis through down-regulation of FOXO₃ (forkhead box O₃) and significantly up-regulated in endometrial cancer [14]. The variant increase the stability of secondary structure and modify the Dicer PHD cleavage score. Despite an increase of RNA stability and a loss Dicer site, it observed a considerable increase in hsa-miR-182-5p expression with a FC of 630.4, suggesting a central role of hsa-miR-182-5p on tumor development.

The homozygote transition identified in hsa-miR-548AW induced a decrease of ΔG and in addition, reduced the number of predicted Dicer cleavage sites and modified the position of DROSHA cleavage site. The mature miR-548AW was a potential biomarker and/or therapeutic target for metastatic prostate cancer

or other cancer types [15]. This evidence well matched with the absence of mature miRNA in normal and tumor tissue analyzed, underling a possible involvement in different cancer stage.

On Chr10, other two homozygote polymorphisms were identified: a transition in 3' UTR of hsa-miR-4481, located into IVS2 of CAMK1D (calcium/calmodulin-dependent protein kinase ID) and transversion C > G into the mature miRNA of hsa-miR-608 which also placed in IVS3 of SEMA4G (sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4G) gene. The first one decreased the stability of secondary RNA structure, instead, the second increased RNA stability and altered DROSHA cleavage site position. Both variants introduced a completely loss of Dicer cleavage sites. No differential expression for hsa-miR-4481 and hsa-miR-608 was observed. The same homozygote rs4919510 polymorphism within hsa-miR-608 was found in sample 15T, confirming the absence of expression, but possible pathogenic role in tumor development. Furthermore, in the sample 3T, a specific region encompassing approximately 40 kb was found on Chr14 (14q32.2-14q32.31). This region was associated with differential expression of estrogen-related gamma-receptor (ESRRG) and was also described down regulated in desmoplastic medulloblastoma. Moreover, the 14q32 region was positively correlated with transcriptional regulation of several miRNA cluster within the same region. In particular, the same miRNAs is down regulated in desmoplastic medulloblastoma: hsa-miR-369, hsa-miR-382, hsa-miR-487A, hsa-miR-539, hsa-miR-329-1, hsa-miR-154 and hsa-miR-323B. In the same region, another undescribed heterozygote transition C > T in hsa-miR-345 was identified. The mature miR-345 was previously reported as a potential prognostic biomarker, in whole blood, for metastatic colorectal cancer. The same region was also involved in very aggressive uterine carcinosarcoma

with an extremely poor prognosis. The involvement of hsa-miR-369, hsa-miR-382, hsa-miR-487A, hsa-miR-539, hsa-miR-154 and hsa-miR-323A, located onto 14q32.2-14q32.31, was also corroborated by results obtained in recent publication of Kandoth [16]. Into the sample, a negative FC for hsa-miR-369-5p and hsa-miR-154-5p was observed but no differences were among hsa-miR-382, hsa-miR-487A, hsa-miR-539, hsa-miR-329-1, hsa-miR-345 and hsa-miR-412. In the hsa-miR-5189, they characterized the simultaneous presence of two homozygote polymorphisms transition (G > A and transversion C > G). Hsa-miR-5189 is located on Chr16 within the IVS1 region of ZFPM1 gene (zinc finger protein, FOG family member 1). These two polymorphisms produced a decrease of RNA secondary structure stability, remove completely the Dicer cleavage sites and diminish the probability score of DROSHA cleavage site. In sample 3T, no expression changing for hsa-miR-5189 was detected.

Lastly in sample 3T, it reported that the homozygote transversion G > C in 3' UTR region of hsa-miR-3615 on Chr17 also located in 5' UTR region of SLC9A3R1 (solute carrier family 9, subfamily A) gene. Hsa-miR-3615 was described as possible modulator of proteins production and/or activity resulting in haploinsufficiency and consequently, inability of the cell to maintain normal cellular function [17]. The polymorphism identified in hsa-miR-3615 decreases RNA secondary structure stability, which is compatible with the increase number of RPM and a FC of 8.58 and modify the predicted cleavage site score for Dicer and DROSHA.

In the sample 15T, in addition to hsa-miR-149, hsa-miR-608 and hsa-miR-612 previously discussed for sample 3T and hsa-miR-941-1 in sample 1T, other four homozygote polymorphisms were reported: rs243080 in hsa-miR-4432, rs11237828 in hsa-miR-5579 located in IVS1 of gene TENM4 (teneurin transmembrane protein 4) onto Chr11, rs8667 in hsa-miR-4751 located into exon 3 of gene

ATF5 (activating transcription factor 5) onto Chr19 and rs4112253 in hsa-miR-4752.

There is no specific correlation of these miRNA in endometrial cancer, although the in silico analysis shows a decrease RNA secondary structure stability in hsa-miR-4432 and hsa-miR-5579, and increase of stability in hsa-miR-4752. In hsa-miR-4751, it does not change the ΔG . Moreover, an alternative Dicer cleavage score in hsa-miR-4751 and a loss of sites in hsa-miR-4752 were calculated. The DIANA prediction results introduced more 87 interacting gene targets for mutant hsa-miR-5579. DNA sequence and mutational analysis of genes involved in biogenesis and regulation of sncRNAs, as expected, revealed a widespread distribution of mutations into different genes. Despite this evidence, common mutated genes were identified in different tumor sample.

In samples 3T and 15T, it was underlined the presence of mutation in DROSHA (drosha, ribonuclease type III)—a component of the microprocessor complex. All mutated genes revealed a possible genomic regions instability whose coexistence modify the cellular balance to tumor status. Moreover, the other specific variants confirmed the heterogeneity of endometrial carcinoma and possibly relationship with aggressive tumor progression and poor prognosis.

5. Conclusions

In conclusion, the samples analyzed corroborate the heterogeneity status of endometrial cancer, supporting the concept of tumor cell heterogeneity associated to hormonal and reproductive factors. The variants identified in pre-miRNAs loci and genes involved in regulation of sncRNA are consistent with a possible effect on miRNA maturation, with an increase or decrease of temporal expression regardless histologic type. NGS improved the heterozygous mutations sensitivity detection, proving a powerful tool for the characterization of specific regions of genomic instability in early stage of the disease that could have a direct role on tumor development. Furthermore, the

possible characterization of instability genome loci, in association with miRNA expression, could provide valuable resource to discover novel candidate biomarkers useful for clinical applications or development of novel targeted therapies.

Acknowledgements

Work supported by University of Salerno (FARB 2014 and 2015), Italian Association for Cancer Research (Grants IG-13176) and Genomix4Life S.r.l. GB is a Ph.D. student in “System Biology”, University of Salerno.

References

- [1] Dai, D., Bradford, A. P., and Prossnitz, E. R. 2014. “Endometrial Cancer: Molecular and Cellular Basis of Tumor Development, Novel Biomarkers and Therapeutic Agents, and Innovative Research Approaches.” *Obstet Gynecol Int*. 2014: 710405. doi: 10.1155/2014/710405.
- [2] Weiderpass, E., Antoine, J., Bray, F. I., Oh, J. K., and Arbyn, M. 2014. “Trends in Corpus Uteri Cancer Mortality in Member States of the European Union.” *Eur. J. Cancer* 50 (9): 1675-84.
- [3] FIGO (International Federation of Gynecology and Obstetrics). 2006. *The Results of Treatment in Gynecological Cancer*. 26th Annual report.
- [4] Press, J. Z., and Gotlieb, W. H. 2012. “Controversies in the Treatment of Early Stage Endometrial Carcinoma.” *Obstet Gynecol Int*. 2012: 578490.
- [5] Enlis, LLC. 2014. “Get Enlis Software.” Enlis Genomics. Accessed, February 8, 2014. <http://www.enlis.com/about.html>.
- [6] Ravo, M., Cordella, A., Rinaldi, A., Bruno, G., Alexandrova, E., and Saggese, P. et al. 2015. “Small Non-coding RNA Deregulation in Endometrial Carcinogenesis.” *Oncotarget* 6 (7): 4677-91.
- [7] Zuker, M. 2003. “Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction.” *Nucleic Acids Res*. 31 (13): 3406-15.
- [8] Ahmed, F., Kaundal, R., and Raghava, G. P. 2013. “PHDCleav: A SVM Based Method for Predicting Human Dicer Cleavage Sites Using Sequence and Secondary Structure of MiRNA Precursors.” *BMC Bioinformatics* 14 (Suppl. 14): S9. doi: 10.1186/1471-2105-14-S14-S9.
- [9] Helvik, S. A., Snøve, Jr, O., and Saetrom, P. 2007. “Reliable Prediction of Drosha Processing Sites Improves Micro-RNA Gene Prediction.” *Bioinformatics* 23 (2): 142-9.

- [10] Maragkakis, M., Alexiou, P., Papadopoulos, G. L., Reczko, M., Dalamagas, T., and Giannopoulos, G. et al. 2009. "Accurate Micro-RNA Target Prediction Correlates with Protein Repression Levels." *BMC Bioinformatics* 18 (10): 295.
- [11] Tsuruta, T., Kozaki, K., Uesugi, A., Furuta, M., Hirasawa, A., and Imoto, I. et al. 2011. "MiR-152 is a Tumor Suppressor Micro-RNA That is Silenced by DNA Hypermethylation in Endometrial Cancer." *Cancer Res.* 71 (20): 6450-62.
- [12] Gutierrez-Camino, A., Lopez-Lopez, E., Martin-Guerrero, I., Piñan, M. A., Garcia-Miguel, P., and Sanchez-Toledo, J. et al. 2014. "Noncoding RNA-Related Polymorphisms in Pediatric Acute Lymphoblastic Leukemia Susceptibility." *Pediatr Res.* 75 (6): 767-73.
- [13] Yan, J., Gumireddy, K., Li, A., and Huang, Q. 2013. "Regulation of Mesenchymal Phenotype by Micro-RNAs in Cancer." *Curr. Cancer Drug Targets* 13 (9): 930-4.
- [14] Watahiki, A., Wang, Y., Morris, J., Dennis, K., O'Dwyer, H. M., and Gleave, M. et al. 2011. "Micro-RNAs Associated with Metastatic Prostate Cancer." *PLOS One* 6 (9): e24950.
- [15] Devor, E. J., DE-Mik, J. N., Ramachandran, S., Goodheart, M. J., and Leslie, K. K. 2012. "Global Dysregulation of the Chromosome 14q32 Imprinted Region in Uterine Carcinosarcoma." *Exp. Ther. Med.* 3 (4): 677-682.
- [16] Kandath, C., Schultz, N., Cherniack, A. D., Akbani, R., Liu, Y., and Shen, H. et al. 2013. "Integrated Genomic Characterization of Endometrial Carcinoma." *Nature* 497 (7447): 67-73.
- [17] Romero, O. A., Setien, F., John, S., Gimenez-Xavier, P., Gomez-Lopez, G., and Pisano, D. et al. 2012. "The Tumor Suppressor and Chromatin-Remodelling Factor BRG1 Antagonizes Myc Activity and Promotes Cell Differentiation in Human Cancer." *EMBO Mol. Med.* 4 (7): 603-16.