MiR-10b, miR-335, and miR-21 are classes of microRNAs (miRNAs) that are overexpressed in breast cancer. Thus, in our study we aimed to test the hypothesis that miRNAs may have direct interactions with proteins and the possibility to inhibit/activate the functional site of proteins and enzymes. For this purpose, we choose three miRNAs involved in breast cancer to study interactions between some proteins and genes, including *BRCA1* and *PTEN*, by processing the docking and matching tools using the Hex8 and HADDOCK server. Mathematically, the hidden Markov models were created by using MATLAB script according to the algorithm in order to study and validate the interactions and bonds between proteins and miRNAs. The main results demonstrate the ability of miR-10b, miR-335, and miR-21 to create direct interactions with 3D protein structures. Furthermore, these results may lead to another pathway of research, i.e. the direct interaction between proteins and their sub-units, to highlight the data obtained previously and demonstrate that proteins may directly interact with ncRNA instead of mRNA. Moreover, our study suggests developing research on different pathways of association proteins-miRNAs as a part of epigenetic extra-nuclear regulation. Taken together, our study provides the first evidence of direct interactions between miRNAs and proteins.

Key words: miR-10b, miR-335, miR-21, breast cancer, gene interactions.

A computational approach to the study of interactions between proteins and miR10-b, miR-335, and miR-21 involved in breast cancer

Rahma Ait Hammou, Yassine Kasmi, Moulay Mustapha Ennaji

Team of Virology, Oncology, and Medical Biotechnologies, Laboratory of Virology, Microbiology, Quality, and Biotechnologies/ETB, Faculty of Sciences and Techniques, University Hassan II of Casablanca, Mohammedia, Morocco

Introduction

Breast cancer (BC) is the most common type of malignancy affecting females worldwide, with about 2.1 million new diagnoses worldwide and 626,679 deaths estimated each year. The incidence of this malignancy has increased by about 20% over the past five years. It is also the main cause of mortality among females, affecting one in four women suffering from malignancy [1]. miRNAs are short non-coding RNAs that regulate gene output at the post-transcriptional level by targeting degenerate elements primarily in 3'untranslated regions (3'UTRs) of mRNAs and are involved in many biological processes, such as cell proliferation, cell cycle regulation, differentiation, apoptosis, and development. Individual miRNAs can regulate networks of hundreds of genes. Therefore, their dysregulation can lead to a variety of diseases including cancer. Individual miRNAs can regulate networks of hundreds of genes, but for the majority of miRNAs few, if any, targets are known. Misexpression of miRNAs is also a major contributor to cancer progression, thus there is a critical need to validate miRNA targets in high throughput to understand the contribution of miRNAs to tumourigenesis. In human BC, it has been shown that they can act either as tumour suppressors or oncogenes [2]. In particular, miR-10b was proven to play a prominent role in regulation of tumour invasion and metastasis. This miRNA expressed in metastatic BC cell lines is able to generate metastases when growing as a primary tumour in mice [3]. miR-335 was demonstrated and reported to play important roles in tumour development, including roles in tumour cell proliferation, survival, migration, invasion, and colonisation. Furthermore, it has been shown that this miR acts as a BC suppressor. However, the mechanisms by which miR-335 affects these biological processes in BC remain unclear [4].

miR-21 is significantly elevated in the majority of human tumours. It plays an important role in tumour formation and development, and it has been shown to be a key regulator of oncogenic processes. Iorio *et al.* reported that miR-21 is overexpressed in BC tissue and may be a useful marker of BC [5].

Two large classes of regulatory molecules working post-transcriptionally are microRNAs (miRNAs) and RNA-binding proteins (RBPs). Although various mechanisms are responsible for the dysregulation of miRNAs in cancer, the miRNA biogenesis pathway is almost controlled by cancer cells to promote dysregulated expression. Furthermore, many human miRNA genes are located at fragile sites or in other genomic regions that are subjects of mutation, deletion, amplification, or translocation in cancer [6, 7]. Recently, cancer regulation by miRNAs had been extensively studied and understood to such an extent that several miRNAs are nowadays classified as "oncomiRs" or,

A computational approach to the study of interactions between proteins and miR10-b, miR-335, and miR-21 involved in breast cancer

221

conversely, as tumour suppressor miRNAs [8]. As for miR-NAs, the involvement of some RBPs in oncogenesis and tumour progression has been shown in several types of cancer [9]. Depending on the protein, the mRNA, and the biological context, the regular actions of RBPs may be positive (activators) or negative (suppressors) [10].

Positive actions in miRNA biogenesis result in them behaving as general enhancers of miRNA functions. Thus, this protein dysregulation is often observed in numerous types of human tumour cells. For example, overexpression of DICER1 in acute myeloid leukaemia cells [11], downregulation of miRNA biogenesis components in hepatocellular carcinoma [12], and DICER downregulation are used to predict poor prognosis in chronic lymphocytic leukaemia [13] and many other types of cancer [14].

In the second class of RBPs, which contrasts with miRNA function, the activity of miRNAs is hindered by RBPs, as shown in the study of p53 inducing the RBP RBM38, which limits the accessibility of miRNA sites on 3'UTRs of its target gene transcripts. Hence, an important proportion of wild type p53 tumours, like breast and prostate cancer, suggest an active mechanism to silence RBM38 in those tumours because it possesses DNA methylation of the RBM's promoter region [15].

The functional relationship between miRNAs and RBPs can also be carried out via the regulation of RBP expression by specific miRNAs. Obviously, it results in a modulation of the effects achieved by RBPs and RBP-controlled genes, and this may be considered as an additional way for the miRNAs to control gene expression and then to contribute to the diversity of cancer cell characterisation.

Furthermore, due to the chemical nature of ribonucleic acid, miRNAs are regulated by RBPs which have an obvious effect on miRNAs precursor pathways in both nuclear and cytoplasmic processes in all cell types, DROSHA and DICER [16]. In defined conditions such as cancer, some other RBPs act specifically on selected miRNAs.

Also in this study we investigate interactions between miRNAs and other proteins that play main and important role in the cell cycle, and some are associated with BC. Among those proteins and genes, we cite: BRCA1 – since its cloning in 1994, numerous studies have established its role in diverse cellular and biochemical processes, such as DNA damage repair, cell cycle control, and transcriptional regulation as well as ubiquitination. BRCA is also well-known as a tumour suppressor implicated in familial breast and ovarian cancer [17]. In vivo, BRCA1 exists as a heterodimer with the BARD1 tumour suppressor, and both proteins harbour a phosphate-binding BRCT domain. Thereby, the main role of BARD1 is an essential contribution to the tumour suppression activity of the BRCA1/BARD1 heterodimer [18]. Furthermore, a study by Ganesan et al. suggests that BRCA1 and BARD may form direct or indirect interactions with Xi-specific transcript (XIST) RNA, and they demonstrated that loss of BRCA1 can lead to a specific disruption of several characteristic features of inactive X chromosome (Xi), including the normal localisation of XIST RNA, and that this interaction may be part of a more general role of BRCA1 in the regulation of chromatin structure [19].

CHEK2 plays a key role in the cellular response to DNA damage, chromosomal stability, and regulation of mitosis. CHEK2 mutation along with a family history of BC imparts a 28–37% lifetime risk of breast carcinoma [20]. PTEN is a crucial tumour-suppressor gene, which may negatively regulate cell proliferation, renewal, and differentiation [21]. This gene had been shown to be dysregulated in many malignant tumours, especially BC [22]. Furthermore, down the expression of PTEN is associated with several outcomes in many diseases [23]. Computational tools allow us to understand better the mechanism of our bodies and accelerate research and innovation. Thus, tertiary structure analysis of miRNA–target interactions provides a more detailed understanding of siRNA/miRNA function, which could open new avenues for developing more accurate algorithms for target prediction. Current computational methods to identify mRNA targets of miRNAs use rules based on primary and secondary structure information. Furthermore, hidden Markov and Bayesian models coupled to Monte Carlo methods provide the possibility to get more information about the statements and the relations into complex systems using a stochastic model.

Here, in this study, the proteins and genes were selected from the literature as important targets, like miRNA target site prediction and miRNA-mRNA duplex prediction, and molecular docking between a variety of 40 proteins and three miRNAs: miR-10b, miR-21, and miR-335, were conducted to elucidate the protein-assisted molecular interaction between miRNAs and these proteins computationally.

Material and methods

The workflow of the method used in this study is described in Figure 1.

Data set

The sequences of miR-10b, miR-335, and miR-21 were collected from the miRBase microRNA database [27]. The crystal structures of proteins were collected from the PDB database; their accession numbers are shown in Table 1. All the ligands and antibodies were removed from the PDB structure files as well as the water molecules.

2D and 3D structures prediction

To predict 2D structures of miR-10b, miR-335, and miR-21 we used RNA structure software [28] and ViennaRNA Web Services [29]. However, the prediction of the 3D structure was made using the software RNAComposer [30] and Sim-RNAweb [31]. Evaluations and benchmarks were obtained based on the ModeRNA server [32].

Interaction protein-miRNA

Many studies have developed the interactions between miRNA, mRNA, and promoter genes. However, to our knowledge, and because during our search of the database no studies demonstrated direct interactions between miRNAs and our target proteins, 3D crystal structure is not available in databases. In this context, we used chemoin-
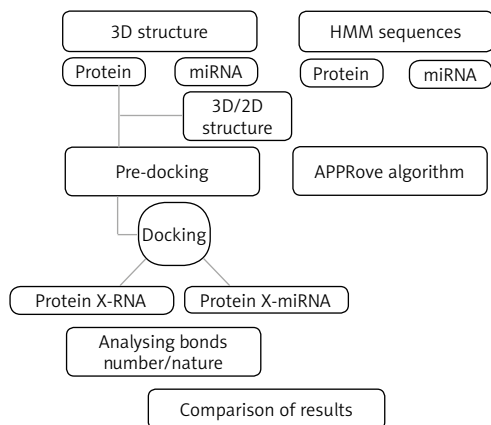
Fig. 1. Workflow of the analysis. We used computational methods to predict the possible interactions between miRNAs and proteins. The first method was based on docking and chemoinformatics approaches to reveal the strengths and natures of the bonds based on docking 3D structures. The second method was based on HMM applied to amino acids and nucleotide sequences revealing essentially the positioning of the bonds. Afterwards, we proceeded to the comparison of the results obtained from the two methods

formatics tools with the objective to study the possible interactions between proteins and miRNAs, based on molecular modelling using docking protein-RNA and matching protein-RNA.

After pre-docking preparation, the miRNAs and protein 3D structures were treated by NP Dock HADDOCK 3.2 [33] and Hex 8.0 to predict the eventual interactions using docking protein-RNA tools with default parameters. The docking process was repeated four times for each protein-miRNA couple, and the average of the free binding energy was calculated, taking into consideration the number of bonds revealed by Discovery Studio 4.0.

Hidden Markov model Bayesian protein-RNA

In order to confirm our results of docking we used the HMM prediction method to compare its results with the docking results. Thus, in the present study, we used the aPPRove algorithm based on HMM [34] to validate the

results using the HMM coupled to Bayesian models. The transition matrix was defined in order to study the transition probability among states. Whereas, for the probability from state α to β, the transition probability of leaving the state was defined according to the formula:

$$\frac{(F\,(\alpha,\,\beta) + \gamma\,(\alpha,\,\beta))}{(\sum (F\,(\alpha,\,\beta) + \gamma\,(\alpha,\,\beta))}$$

where η and γ are set as pseudo-counts used in order to determine A and T probabilities. We defined $\gamma\,(i,\,j)$ for all possibilities of i and j, where i and j state in the pair hidden Markov model. The Viterbi algorithm was used in order to validate the scores of transmission and emission from one state to another.

Results and discussion

2D and 3D structures

Tertiary structure-based computational methods are required for a more detailed analysis of the structural mechanisms of miRNA target recognition and post-transcriptional regulation. We used Monte Carlo (MC) methods, which are frequently applied to sample putative interaction geometries of proteins, including also possible conformational changes of the binding partners. Figure 2 demonstrates the score of prediction using HMM of amino acid interactions.

Docking and protein interactions

In order to study eventual interactions between miRNAs and proteins or genes, molecular docking with default parameters was carried out. Thereby, we tested by docking the possible interactions between our miRNAs and *BRCA1*, *PTEN*, *BRCA1/BARD1*, *BRCT*, *CHEK2*, *BP1 BRCT7/8-BACH1*, PI3K-γ cancer, *BRCA1*/BRCT, E-cadherin, Mre11, HER2-, CDK1/CKS1, E1, E2, and E3 enzymes and ATM. As the literature has reported, a pocket of miRNAs implicates strong binding interactions at the molecular level between AGO protein and various miRNAs. Our work was mainly focused on the hypothesis that miRNAs can make covalent and non-covalent linkages with oncoproteins and suppressor
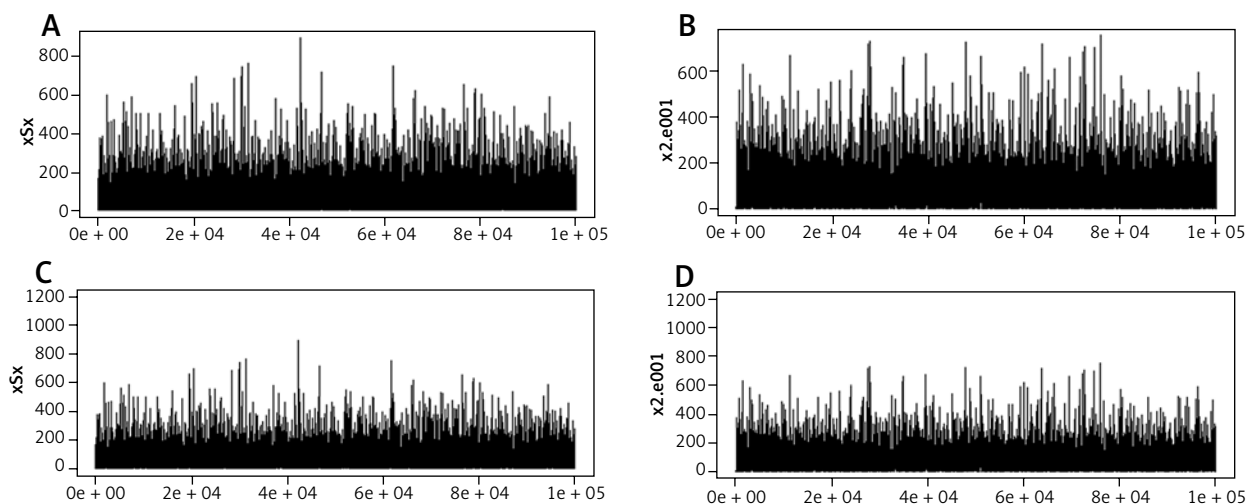
Fig. 2. Monte Carlo score vs. steps of simulation (refinement of prot-na_10809 with MC)
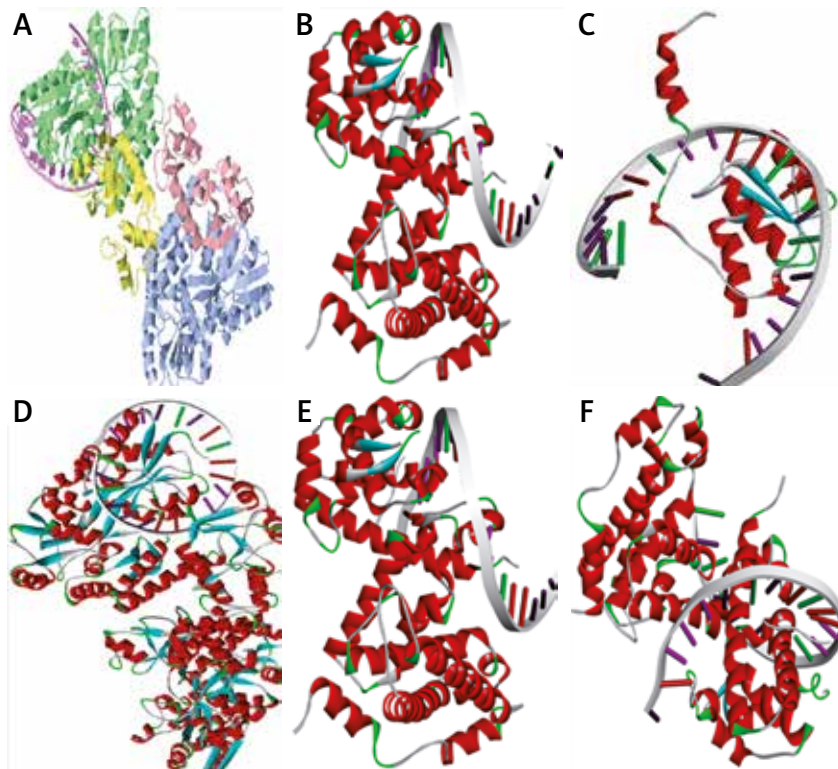
Fig. 3. Interaction between miR-10b and **A)** E6, **B)** P27 (KIP1)/cyclin A/CDK2 com, **C)** Mre11, **D)** Skp1-Skp2-Cks1, **E)** E-cadherin, and **F)** *BRCA1*

proteins, which is not yet demonstrated by studies in vivo. According to the modelling of structures, we used chemical modelling tools, namely covalent and non-covalent docking and mathematical modelling by HMM, to illustrate and test this hypothesis. In general, miR10, miR335, and miR21 demonstrated their ability to create linkages with 3D protein structures, of which the long RNA or reference DNA sequences such as TATA box do not interact with the current protein. Thus, our study shows that miR-10b produces non-covalent bonds with allosteric sites of complex *BRCA1/BARD1*, which directly leads to an increase in or inhibition of their activity. Furthermore, we found that the possibility of interaction is favourable between miR-10b and oncoprotein E6, P27(KIP1)/cyclin A/CDK2 complex, Mre11 nuclease, which is essential for dsDNA break repair, Skp1-Skp2-Ck1, which plays a major role as mitotic regulatory protein, E-cadherin, which regulates cell growth, and *BRCA1*. The positioning of the interactions between the proteins and miR10 are located in different sites of the proteins according to the nature of the protein (Fig. 3).

On the other hand, we noticed that the strength of interactions was variable for each protein. Thereby, Table 1 shows the energy of binding of some studied proteins and target miRNAs. The interactions are significantly higher than for other proteins. We also found that these interactions are mostly covalent, with the presence of the non-covalent bonds enhancing the typical energy voices for stabilisation of complexes. It should be noticed that some proteins such as HE2 do not interact with the large RNAs and reference DNA. However, in the case of miR-10b and 335, these strong interactions elucidate the idea that the overexpression of miRNAs is not limited to the level of

mRNA but also to the proteomic level, which explains and provides another pathway of research.

Furthermore, studies focused on expression profiling have shown that the dysregulation of miRNA expression has various origins, which provide these biomarkers with their tissue-specific characteristics. Growing evidence suggests that miRNAs can be regulated extensively at the levels of promoter transcription methylation, miRNA processing, RNA editing, and miRNA-target interactions [24]. Thereby, transcriptional regulation by nuclear receptors is the primary level of control for miRNA expression. This regulation can be achieved via three mechanisms: direct binding to the promoter regions of miRNAs, indirect regulation of miRNA expression through NR target genes, and involvement in the regulation of miRNA biogenesis [25]. *BRCA1* regulates the expression of miRNAs, which may, in turn, regulate the expression of *BRCA1*. MiRNAs are known to down-regulate the expression of the protein in two different ways: by inhibiting the expression or degrading mRNA [26]. This mechanism can have the same effect in dysregulation of *BRCA1* function. These direct interactions allow us to conclude that the influence levels of miRNAs are not only genomic but can reach proteomic levels. Such direct interaction between miRNA and protein can influence the direct regulation of the action mechanisms in the biological system. The strength of the interactions between the crystallised structures and the miRNAs is very large compared to those observed with the same crystal structures and the DNA fragments obtained during the crystallisation; essentially those located at the promoter regions.

In conclusion, miRNAs play an increasingly conspicuous role in the regulation of epigenetic and post-transcriptom-

Table 1. Binding energy and RMSD values between proteins and miR-10b, miR-335, and miR-21

Proteins and genes		MiR-10b		MiR-335	MiR-21
		Docking HADDOCK	**RMSD**	**Docking HADDOCK**	**Docking HADDOCK**
1d5r	*PTEN*	−284.58	96.35	−309.94	−294.59
1jm7	*BRCA1/BARD1* RING-domain heterodimer	−	−	−	−
1jsu	P27(KIP1)/cyclin A/CDK2 complex	−350	50	383.89	−307.27
1l0b	*BRCA1* tandem-*BRCT* region	−305.18	56	−294.59	−297.92
1rgp	GTPase-activating domain	−351.24	61.32	−383.89	−370
1t2u	*BRCA1*	−299.38	90	−367.26	−263.9
2ass	Skp1-Skp2-Cks1	−433.73	101.29	−316.55	−401.3
2ing	*BRCT*	−329.69	76	383.89	−307.28
2wtd	Chk2	−288.84	48	−294.59	−297.93
3al3	BP1 BRCT7/8-BACH1	−400.39	53.79	−383.89	−376.3666667
3ene	PI3K-γ cancer	−394.68	47	−367.26	−392.0166667
3k0h	*BRCA1/BRCT*	−363.03	80	−316.55	−407.6666667
3lne	E-cadherin	−285.13	40	383.89	−307.29
3pom	Unliga + D19	−303.14	25	−294.59	−297.94
3qkr	Mre11	−290.01	45	−383.89	−423.3166667
3t1i	Mre11	−346.21	45	−367.26	−438.9666667
4hrl	HER2	−329.06	140	−316.55	−454.6166667
4yc6	*CDK1/CKS1*	−380.48	59	−367.26	−297.94
5l9t	E1, E2, and E2 *RING* E3	−271.27	65	−316.55	−423.3166667
5np0	ATM	−	−	−383.89	−438.9666667

RMSD – root-mean-square deviation

ic effects in cancerogenesis and oncological processes. miR-10b binds with oncogenic genes, explaining the increased expression of miR-10b. The results demonstrate covalent linkages between miR-335 and P53 and other suppressor proteins in the activating regions, whereas the interaction with the oncogenic progeny may be less likely. On the other hand, interactions with viral-oncogenic and oncogenic proteins are observed at the RNA-protein docking level. This may constitute a new pathway to other hypotheses of degradation and regulation of tumourigenesis mechanisms.

Acknowledgments

The authors declare no conflict of interest.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394-424.
2. O'Day E, Lal A. MicroRNAs and their target gene networks in breast cancer. Breast Cancer Res 2010; 12: 201.
3. Ma L Teruya-Feldstein J, Weinberg RA. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature 2007; 449: 682-868.
4. Dong Y, Liu Y, Jiang A, Li R, Yin M, Wang,Y. MicroRNA-335 suppresses the proliferation, migration, and invasion of breast cancer cells by targeting EphA4. Mol Cell Biochem 2018; 439: 95-104.
5. Iorio MV Croce, C.M. microRNA involvement in human cancer. Carcinogenesis 2012; 33: 1126-1133.
6. Calin GA, Sevignani C, Dumitru CD, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci 2004; 101: 2999-3004.
7. Zhang L, Huang J, Yang N, et al. microRNAs exhibit high frequency genomic alterations in human cancer. Proc Natl Acad Sci 2006; 103: 9136-9141.
8. Iorio MV, Croce CM. microRNA involvement in human cancer. Carcinogenesis 2012; 33: 1126-1133.
9. Kim MY, Hur J, Jeong S. Emerging roles of RNA and RNA-binding protein network in cancer cells. BMB Rep 2009; 42: 125-30.
10. Keene JD. RNA regulons: coordination of post-transcriptional events. Nat Rev Genet 2007; 8: 533-543.
11. Bai Y, Qiu GR, Zhou F, Gong LY, Gao F, Sun KL. Overexpression of DICER1 induced by the upregulation of GATA1 contributes to the proliferation and apoptosis of leukemia cells. Int J Oncol 2013; 42: 1317-1324.

A computational approach to the study of interactions between proteins and miR10-b, miR-335, and miR-21 involved in breast cancer

225

12. Kitagawa N, Ojima H, Shirakihara T, et al. Downregulation of the microRNA biogenesis components and its association with poor prognosis in hepatocellular carcinoma. Cancer Sci 2013; 104: 543-551.

13. Zhu DX, Fan L, Lu RN, et al. Downregulated Dicer expression predicts poor prognosis in chronic lymphocytic leukemia. Cancer Sci 2012; 103: 875-881.

14. Pellegrino L, Jacob J, Roca-Alonso L, Krell J, Castellano L, Frampton AE. Altered expression of the miRNA processing endoribonuclease Dicer has prognostic significance in human cancers. Expert Rev Anticancer Ther 2013; 13: 21-27.

15. Léveillé N, Elkon R, Davalos V, et al. Selective inhibition of microRNA accessibility by RBM38 is required for p53 activity. Nat Commun 2011; 2: 513.

16. Winter J, Jung S, Keller S, Gregory RI, Diederichs S. Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol 2009; 11: 228-234.

17. Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. Nat Rev Cancer 2007; 7: 937-489.

18. Billing D, Horiguchi M, Wu-Baer F, et al. The BRCT Domains of the BRCA1 and BARD1 Tumor Suppressors Differentially Regulate Homology Directed Repair and Stalled Fork Protection. Mol Cell 2018; 72: 127-139e8.

19. Ganesan S, Silver DP, Drapkin R, Greenberg R, Feunteun J, Livingston DM. Association of BRCA1 with the inactive X chromosome and XIST RNA. Philos Trans R Soc Lond B Biol Sci 2004: 359: 123-128.

20. Weischer M, Bojesen SE, Ellervik C, Tybjaerg-Hansen A, Nordestgaard BG. CHEK2*1100delC genotyping for clinical assessment of breast cancer risk: meta-analyses of 26,000 patient cases and 27,000 controls. J Clin Oncol 2008; 26: 542-548.

21. Groszer M, Erickson R, Scripture-Adams DD, et al. Negative regulation of neural stem/progenitor cell proliferation by the Pten tumor suppressor gene in vivo. Science 2001; 294: 2186-2189.

22. Li DM, Sun H. TEP1, encoded by a candidate tumor suppressor locus, is a novel protein tyrosine phosphatase regulated by transforming growth factor beta. Cancer Res 1997; 57: 2124-2129.

23. Chowdhury S, Ongchin M, Wan G, Sharratt E, Brattain MG, Rajput A. Restoration of PTEN activity decreases metastases in an orthotopic model of colon cancer. J Surg Res 2013; 184: 755-760.

24. Breving K, Esquela-Kerscher A. The complexities of microRNA regulation: mirandering around the rules. Int J Biochem Cell Biol 2010; 42: 1316-1329.

25. Zhihong Y, Wang L. Regulation of microRNA expression and function by nuclear receptor signalling. Cell Biosci 2011; 1: 31.

26. Nelson P, Kiriakidou M, Sharma A, Maniataki E, Mourelatos Z. The microRNA world: small is mighty. Trends Biochem Sci 2003; 28: 534-540.

27. Griffiths-Jones S, Grocock RJ, Van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006; 34: D140-D144.

28. Reuter JS, Mathews DH. RNA structure: software for RNA secondary structure prediction and analysis. BMC Bioinformatics 2010; 11: 129.

29. Hofacker IL. Vienna RNA secondary structure server. Nucleic Acids Res 2003; 31: 3429-3431.

30. Biesiada M, Pachulska-Wieczorek K, Adamiak RW, Purzycka KJ. RNA Composer and RNA 3D structure prediction for nanotechnology. Methods 2016; 103: 120-127.

31. Magnus M, Boniecki MJ, Dawson W, Bujnicki JM. SimRNA web: a web server for RNA 3D structure modeling with optional restraints. Nucleic Acids Res 2016; 44: W315-W319.

32. Rother M, Milanowska K, Puton T, Jeleniewicz J, Rother K, Bujnicki JM. ModeRNA server: an online tool for modeling RNA 3D structures. Bioinformatics 2011; 27: 2441-2442.

33. De Vries SJ, Van Dijk M, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. Nat Protoc 2010; 5: 883-897.

34. Chis T, Harrison P. Adapting Hidden Markov Models for Online Learning. Electron Notes Theor Comput Sci 2015; 318: 109-127.

Address for correspondence

Rahma Ait Hammou
Team of Virology, Oncology, and Medical Biotechnologies
Laboratory of Virology, Microbiology, Quality, and Biotechnologies
Faculty of Sciences and Techniques,
University Hassan II of Casablanca
Mohammedia, Grand Casablanca 20650, Morocco
e-mail: rahma.aith@gmail.com