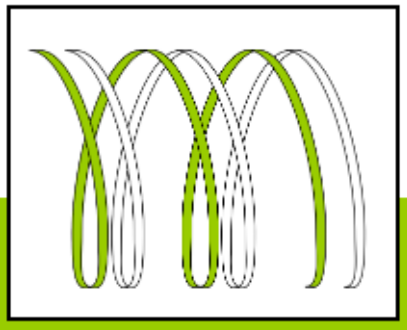
EN

ФГБУН Институт проблем передачи информации им. А.А. Харкевича Российской академии наук
(ИППИ РАН)

9-ая Московская конференция по вычислительной молекулярной биологии MCCMB'19
Москва, 27-30 Июля 2019



MCCMB 2019

Сбоник трудов

ISBN 978-5-901158-31-9

Москва
ИППИ РАН
2019

Сведения об издании

A

up

Eldar Abdullaev *, Peter Arndt*

Modelling Segmental Duplications in the Human Genome

Anahita Amiri , Gianfilippo Coppola , Soraya Scuderi , Feinan Wu , Tanmoy Roychowdhury , Flora Vaccarino , ***Alexej Abyzov***

Transcriptome and epigenome landscape of human cortical development modeled in organoids

Darya Afenteva *, Mikhail Orlov , Maria Poptsova*

Joint mechanical, thermodynamic and electrostatic machine-learning model for prediction of nucleosome positions

Dmitry Afonnikov *, Mikhail Genaev , Nikolay Shmakov , Zakhar Mustafin , Alexey Mukhin , Dmitrii Konstantinov , Alexey Doroshkov , Sergey Lashin*

Analysis of out of reference transcripts from RNA-seq libraries in crops.

Sergey Aganezov *, Fritz Sedlazeck , Sara Goodwin , Gayatri Arun , Isac Lee , Sam Kovaka , Michael Kirsche , Rachel Sherman , Robert Wappel , Melissa Kramer , Karen Kostroff , David Spector , Winston Timp , Michael Schatz , Richard McCombie*

Reconstructing haplotype-specific cancer genome karyotypes with multiple sequencing technologies

Dana Aisina *, Anatoliy Ivashchenko*

Prediction of clusters of miRNA binding sites in mRNA candidate genes of luminal A and B subtypes breast cancer

Aigul Akimniyazova *, Alexandra Kim*

Prediction of clusters of miRNA binding sites in mRNA candidate genes of luminal A and B subtypes breast cancer

Aisina D.E., Ivashchenko A.T.*

SRI of Biology and Biotechnology Problems, al-Farabi KazNU, Department of biotechnology, Almaty, Kazakhstan, a_ivashchenko@mail.ru

Breast cancer (BC) occupies one of the first places among all cancers in the world. These statistics demonstrate an intense, steady increase in the incidence and mortality from BC among women [1]. Establishing the interaction of miRNAs with mRNA genes involved in the development of BC (candidate genes) is one of the promising areas of research. miRNAs are found in tumors, blood, and may be potential biomarkers of BC [2]. The establishment of a correlation between the expression of miRNA and various BC subtypes is devoted to several publications [3].

The nucleotide (nt) sequences of candidate genes of BC subtypes were downloaded from GenBank (http://www.ncbi.nlm.nih.gov). These candidate genes are specific for the development of luminal A and B subtype of BC. The nucleotide sequences of 2565 miRNAs were taken from miRBase and 3707 miRNAs from the publication [4]. RPKM value [5] given in the Human Protein Atlas data (https://www.proteinatlas.org). Human Protein Atlas data were used as a quantitative measure of transcript expression in breast. The miRNAs binding sites (BS) in 5'UTRs, CDSs and 3'UTRs of several genes were predicted using the MirTarget program [6]. This program defines the following features of miRNA binding to mRNA: a) the start of the initiation of miRNA binding to mRNAs; b) the localization of miRNA BS in 5'UTRs, CDSs and 3'UTRs of the mRNAs; c) the free energy of interaction miRNA and the mRNA (ΔG, kJ/mole); d) the schemes of nucleotide interactions between miRNAs and mRNAs. The ratio ΔG/ΔGm (%) was determined, where ΔGm equals the free energy of miRNA binding with its fully complementary nucleotide sequence. Eighteen miRNA BS with overlapping nucleotide sequences were identified in 5'UTR mRNA of *FOXA1* gene (Table 1). Twenty BS formed a cluster with the length of 52 nt, from 99 nt to 150 nt. The total length of all 20 BS is 447 nt, which is longer than 5'UTR with length of 312 nt. All miRNA BS were located in the first half of 5'UTR.

Table 1 - Characteristics of miRNAs interaction in mRNA of BC subtype luminal A and B

Gene, RPKM	miRNA	Start of BS, nt	ΔG, kJ/mole	ΔG/ΔGm	Length, nt
FOXA1	ID00297.5p-miR	99	-123	89	24
	ID02106.3p-miR	110	-123	89	23
	ID00252.5p-miR	111	-140	94	24
	ID02769.5p-miR	112	-127	92	22
	ID00296.3p-miR	115	-140	89	25
	ID01099.5p-miR	116	-108	100	17
	ID00071.3p-miR (2)	118 ÷ 121	-117 ÷ -121	93 ÷ 97	20
	ID01190.5p-miR	118	-108	100	17
	ID02457.3p-miR	118	-108	100	17
10.2	ID02595.5p-miR	118	-115	92	20
	ID01403.5p-miR	120	-123	91	23
	ID01702.3p-miR	120	-140	89	25
	miR-3960	120	-115	92	20
	ID03367.5p-miR (2)	121 ÷ 122	-117	93	20
	ID01641.3p-miR	122	-134	90	24
	ID00457.3p-miR	124	-123	91	22
	ID00061.3p-miR	127	-129	94	22
	ID02499.3p-miR (2)	127 ÷ 130	-119 ÷ -121	92 ÷ 93	21
*ITGB1**	ID02187.5p-miR	91	-127	92	23
	miR-4787-5p	92	-123	92	22
63.6	ID00457.3p-miR	95	-123	91	22
	ID02770.5p-miR	98	-117	93	20
	ID01184.3p-miR	101	-117	93	20
*SMAD3***	miR-4690-5p	2066	-115	92	22
	miR-3620-5p (2)	2069 ÷ 2074	-117 ÷ -115	87 ÷ 89	22
	ID02822.5p-miR	2070	-127	91	23
14.0	ID00978.5p-miR	2072	-119	90	22
	miR-6089 (2)	2073 ÷ 2078	-132 ÷ -136	89 ÷ 91	24
	ID01382.3p-miR	2075	-113	93	20

Note. There shown: Genes without * - miRNA BS are in the 5'UTR, genes with * - miRNA BS are in the CDS, ** - miRNA BS are in the 3'UTR; in parentheses indicates the number of BS; ÷ - the change of the parameter in the interval.

Since the cluster length is 52 nt, only two miRNAs can be contacted simultaneously, and other miRNAs will not affect the expression of the *FOXA1* gene. The formation of a cluster of BS for the *FOXA1* gene in 5'UTR indicates a greater ability of this gene for compaction, which causes competition among miRNA for the BS. Despite the fact that ID01099.5p-miR, ID01190.5p-miR and ID02457.3p-miR are fully complementary to mRNA gene, they had a free energy interaction of -108 kJ/mole, which is significantly less than for other miRNAs. At equal concentrations of all miRNAs, ID00252.5p-miR, ID00296.3p-miR and ID01702.3p-miR had ΔG values equal to -140 kJ/mole having the advantage in binding to the

mRNA of *FOXA1* gene. The average free energy of miRNA binding, without three miRNAs with a length of 17 nt, was -126 kJ/mole, which is characteristic of miRNA binding in 5'UTR.The 5'UTR of mRNA *HMGA2* had 17 BS for 15 miRNAs. BS of these miRNAs were in a 95 nt cluster from 512 nt to 606 nt. The total length of BS was equal to 407 nt and it was 4.3 times longer than the cluster. miRNA BS were located in the first half of 5'UTR and had a ΔG value more than -125 kJ/mole. ID00296.3p-miR and ID00296.3p-miR had a ΔG equal to -142 kJ/mole and -146 kJ/mole, respectively. The *ITGB1* gene had no 5'UTR, but a cluster for five miRNA BS was located from 91 nt to 120 nt in the beginning of CDS with the length of 30 nt, which is 3.6 times less than the sum of the lengths of five miRNAs. For *HMGA2* gene there was a cluster for four BS from 1255 nt to 1295 nt located in the beginning of 3'UTR. The cluster length was equal to 41 nt with total length of BS comprising 98 nt in length. Apparently, the compaction of BS is due not only to the economy of gene length but also to the competition between miRNAs for interaction. For example, the cluster of eight BS with 3'UTR of mRNA *SMAD3* gene with the length of 35 nt was located from 2066 nt to 2101 nt. Therefore, only one miRNA can be bind in a cluster. At equal concentrations of all six miRNAs, ID02822.5p-miR and miR-6089 had free interaction energy of -127 kJ/mole to -136 kJ/mole will have an advantage in binding to cluster. The 3'UTR of mRNA *SOX4* gene had four miRNA BS organized in a cluster of 29 nt. ID01282.3p-miR and ID03445.3p-miR bound to mRNA with a ΔG equal to -125 kJ/mole and -127 kJ/mole, respectively.The mRNAs of the *TGFB1* gene had a cluster of BS for seven miRNAs with a length of 48 nt located from 2060 nt to 2107 nt. The length of 3'UTR was 146 nt with 10 miRNA BS equal to 230 nt, so the compacting of the BS was 4.8 times. Figure 1 shows the schemes of interaction of some miRNAs with mRNA of several candidate genes of the luminal A and B subtypes. The presented schemes clearly show the advantage of the MirTarget program in predicting the miRNA BS. For example, ID03367.5p-miR formed a non-canonical G-U pair in the mRNA *FOXA1* gene. But ID03367.5p-miR can bind to 19 nucleotide of mRNA and the free interaction energy was 93% of the maximum value. The ID02542.5p-miR interacted with 23 nucleotides of mRNA *FOXA1* gene, but had only one unpaired nucleotide. Such interaction between the miRNAs and their target genes is valid for the following pairs: ID00101.3p-miR and *HMGA2* gene, ID00849.3p-miR and *HMGA2* gene, miR-4507-3p and

SMAD3 gene, miR-937-5p and *TGFB1* gene, miR-937-5p and *SMAD3* gene, ID01403.5p-miR and *HMGA2* gene.

SMAD3; ID01020.5p-miR; 5'UTR; 194; -117; 100; 19 5'-GCGACCGCGGCAGGCCCCG-3' 3'-CGCUGGCGCCGUCCGGGGC-5'	*SMAD3*; miR-937-5p; 3'UTR; 2072; -102; 89; 20 5'-CCAGCCCC**G**CCC**C**GCC**C**C**G**C-3' 3'-GGUCGGGG**U**GGG**A**CUG**A**G**U**G-5
ANGPTL4; ID01593.5p-miR; CDS; 259; -134; 100; 23 5'-AGCGCUCAGGGCGGACCCGUGCA-3 3'-UCGCGAGUCCCGCCUGGGCACGU-5'	*FOXA1*; ID03332.3p-miR; CDS; 150; -134; 90; 24 5'-**A**GCGGA**A**GCGG**G**GGC**A**GCGGCGCC-3' 3'-**C**CGCCU**C**CGCC**U**CCG**C**CGCCGCGG-5'
Note: Gene; miRNA; the miRNA region; start of BS (nt); the free energy, ΔG (kJ/mole); the ΔG/ΔGm (%); length of miRNA (nt). The upper and lower nucleotide sequences of mRNA and miRNA, respectively. The nucleotides of non-canonical pairs G-U and A-C highlighted in bold type.	

Figure 1 - Schemes of miRNA interaction with mRNA of candidate genes of breast cancer luminal A and B subtypes.

1. Benson J.R., Jatoi I. (2012) The global breast cancer burden, *Future Onco*., **8:** 697-702.
2. Adhami M. et al. (2018) Candidate miRNAs in human breast cancer biomarkers: a systematic review, *Breast Cancer*, **25:** 198-205.
3. Telonis A.G. et al. (2015) Beyond the one-locus-one-miRNA paradigm: microRNA isoforms enable deeper insights into breast cancer heterogeneity, *Nucleic Acids Res*., **43:** 9158-9175.
4. Londin E. et al. (2015) Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs, *PNAS USA*, **112:** 1106-1115.
5. Mortazavi A. et al. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq, *Nature Methods*, **5:** 621-628.
6. Ivashchenko A.T. et al. (2016) Prediction of miRNA binding sites in mRNA, *Bioinformation*, **12:** 237-240.