

RESEARCH ARTICLE

Role of the long non-coding RNA PVT1 in the dysregulation of the ceRNA-ceRNA network in human breast cancer

Federica Conte¹, Giulia Fiscon¹, Matteo Chiara², Teresa Colombo¹, Lorenzo Farina³, Paola Paci^{1*}

1 Institute for Systems Analysis and Computer Science “Antonio Ruberti”, National Research Council, Rome, Italy, **2** Department of Biosciences, University of Milan, Milan, Italy, **3** Department of Computer, Control and Management Engineering, “Sapienza” University, Rome, Italy

* paola.paci@iasi.cnr.it



OPEN ACCESS

Citation: Conte F, Fiscon G, Chiara M, Colombo T, Farina L, Paci P (2017) Role of the long non-coding RNA PVT1 in the dysregulation of the ceRNA-ceRNA network in human breast cancer. *PLoS ONE* 12(2): e0171661. doi:10.1371/journal.pone.0171661

Editor: Turgay Unver, Dokuz Eylul Universitesi, TURKEY

Received: September 21, 2016

Accepted: January 24, 2017

Published: February 10, 2017

Copyright: © 2017 Conte et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was financially support by The Epigenomics Flagship Project (Progetto Bandiera Epigenomica) EPIGEN funded by Italian Ministry of Education, University and Research (MIUR) and the National Research Council of Italy (CNR), URL: <http://www.epigen.it/>.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Recent findings have identified competing endogenous RNAs (ceRNAs) as the drivers in many disease conditions, including cancers. The ceRNAs indirectly regulate each other by reducing the amount of microRNAs (miRNAs) available to target messenger RNAs (mRNAs). The ceRNA interactions mediated by miRNAs are modulated by a titration mechanism, *i.e.* large changes in the ceRNA expression levels either overcome, or relieve, the miRNA repression on competing RNAs; similarly, a very large miRNA overexpression may abolish competition. The ceRNAs are also called miRNA “decoys” or miRNA “sponges” and encompass different RNAs competing with each other to attract miRNAs for interactions: mRNA, long non-coding RNAs (lncRNAs), pseudogenes, or circular RNAs. Recently, we developed a computational method for identifying ceRNA-ceRNA interactions in breast invasive carcinoma. We were interested in unveiling which lncRNAs could exert the ceRNA activity. We found a drastic rewiring in the cross-talks between ceRNAs from the physiological to the pathological condition. The main actor of this dysregulated lncRNA-associated ceRNA network was the lncRNA PVT1, which revealed a net binding preference towards the miR-200 family members in normal breast tissues. Despite its up-regulation in breast cancer tissues, mimicked by the miR-200 family members, PVT1 stops working as ceRNA in the cancerous state. The specific conditions required for a ceRNA landscape to occur are still far from being determined. Here, we emphasized the importance of the relative concentration of the ceRNAs, and their related miRNAs. In particular, we focused on the withdrawal in breast cancer tissues of the PVT1 ceRNA activity and performed a gene expression and sequence analysis of its multiple isoforms. We found that the PVT1 isoform harbouring the binding site for a representative miRNA of the miR-200 family shows a drastic decrease in its relative concentration with respect to the miRNA abundance in breast cancer tissues, providing a plausibility argument to the breakdown of the sponge program orchestrated by the oncogene PVT1.

Introduction

The last years have been marked by an increasing widespread interest in non-coding RNAs, emerging as key regulators of many cellular processes in both physiological and pathological states [1–4]. This class of RNA species appears really heterogeneous, including the intensively studied microRNAs (miRNAs)—small non-coding RNAs of 20–22 nucleotides long [5–7]—as well as the most recently acknowledged long non-coding RNAs (lncRNAs). lncRNAs are non-protein coding transcripts greater than 200 nucleotides in length and lacking of extended open reading frames [8–10]. As broadly suggested by several works [11–31], lncRNAs critically participate in transcriptional and post-transcriptional regulation, though the biological functions of the majority of them largely remain to be defined yet. Recent studies have shown that some lncRNAs may have a role linked to their secondary structure [32–40], whose specific substructures can function as guide or scaffold by binding chromatin-modifying protein complexes [34, 37, 41, 42]. Although lncRNAs have low sequence conservation [43, 44], increasing evidence indicates that also their primary structure (*i.e.* nucleotides sequence) could be instrumental for their implication in a wide variety of processes, including competition for miRNA binding [30, 45–61].

Competing endogenous RNAs (ceRNAs), also known as miRNA “decoy” or miRNA “sponges”, are RNA transcripts that compete for the binding to the same miRNA *via* the base-pairing with miRNA recognition/response elements (MREs) [62–70], subsequently enabling the reduction of the amount of miRNAs available to target messenger RNAs (mRNAs). Such a mechanism of regulation of miRNA activity was firstly discovered in plants and called “target mimicry” process [71].

The first experimental evidence of lncRNAs acting as miRNA decoys modulating the derepression of miRNA targets has been found in wide variety of human cancers and specifically concerns the functioning of pseudogenes (*i.e.* copies of real genes that originate from duplications or retro-transpositions) as competitors of their ancestral genes for miRNA binding [72]. They are not translated into functional proteins because their coding potential is corrupted by premature stop codons, deletions/insertions and frameshift mutations. Nevertheless, nucleotide sequences contained within pseudogenes are well preserved, suggesting that selective pressure to maintain these genetic elements exists, and that they may indeed have an important cellular role [72]. Moreover, pseudogenes are almost as numerous as coding genes and represent a significant proportion of the transcriptome [73]. They are perfect endogenous competitors of their ancestral genes, since they retain many of the miRNA binding sites.

lncRNAs functioning as ceRNAs can be also observed in: mouse and human myoblasts, where the large intergenic non-coding RNA (lincRNA) called linc-MD1 controls muscle differentiation by targeting miR-133 and miR-135 to regulate the expression of MAML1 and MEF2C [74]; human embryonic stem cells, where linc-RoR competes with the transcription factors NANOG, OCT4, SOX2 for binding to miR-145 regulating cell pluripotency and self-renewing [75]; human thyroid cancer, where the thyroid-specific lncRNA PTSC3 targets miR-574-5p [76]; human embryonic kidney 293 (HEK293) cells, where the lncRNA H19 modulates the let-7 miRNAs family availability causing precocious muscle differentiation [77].

Most recently, also the new-appreciate circular RNAs (circRNAs) appear to exert ceRNA activity [70, 78–80]. They are a class of non-coding RNAs derived mostly from a non-canonical form of alternative splicing, whereby the exon ends are joined to form a continuous loop [81–84]. In particular, the exonic circRNA CDR1 relieves the activity of miR-7 on its target impairing midbrain development in mammals [85] and the testis-specific circRNA Sry serves as a miR-138 sponge [80].

In our previous work [86], we developed a purely data-driven approach focused on the identification of lncRNAs acting as new putative ceRNAs in a large set of tumour and matched-normal samples (*i.e.* tissues that are adjacent to the tumour and taken from the same patient) of breast invasive carcinoma available from The Cancer Genome Atlas (TCGA) [87, 88]. By applying a multivariate statistical analysis refined by the requirement of a seed match enrichment, we built a network of miRNA-mediated sponge interactions (MMI-networks) in both physiological and pathological states and compared the two obtained MMI-networks. We found a marked rewiring in the ceRNA program between normal and pathological breast tissues. At the heart of this phenomenon is the lncRNA PVT1 that serves as miRNA sponge in normal tissues, but not in cancer. Moreover, it revealed, in normal MMI-network, a net binding preference towards the miR-200 family, which it antagonizes to regulate the expression of hundreds of mRNAs known to be related to the cancer development and progression (*e.g.* GATA3, CDH1, TP53, TP63, TP73, RUNX1, and RUNX3).

PVT1 is a large intergenic non-coding RNA that appears to be strongly conserved between mouse and human [89–94]. The PVT1 gene [95] spans across a genome interval of over 300 kb (*i.e.* bases 128806779–129113499 within the February 2009 human genome build GRCh37/hg19) on the forward strand of chromosome 8 [96]. Moreover, PVT1 lies in a recognized cancer risk locus that it shares with MYC and shows highly complex gene architecture. Indeed, its locus gives rise to over 20 different variants of the lncRNA according to the Ensembl annotations of the human genome (release 75) and also produces a cluster of six annotated microRNAs (*i.e.* miR-1204, miR-1205, miR-1206, miR-1207-5p, miR-1207-3p, and miR-1208) [96–101]. The last years have been the scene of increasing advancements in studying PVT1 role in tumour cells [28, 96, 102–107] and its overexpression appears as a frequent event in a wide variety of cancers [98, 103, 104, 108, 109]. In addition to a putative ceRNA activity, interesting ways of functioning of PVT1 have been suggested, such as the regulation of the protein stability of the well-known MYC oncogene through its secondary structure [99, 110–112].

The miR-200 family consists of five members: miR-200a, miR-200b, miR-200c, miR-141 and miR-429. On the basis of the similarities of their seed sequences (*i.e.* 6 nucleotides at positions 2–7 from the miRNA 5'-end [113]), the miR-200 family members can be clustered into two groups only differing for one nucleotide in the seed sequence: miR-200a/141 (AACACU) and miR-200b/200c/429 (AAUACU) [114, 115]. The miR-200 family is one of the most widely studied for its crucial role in cancer initiation, metastasis, diagnosis, and treatment. A large number of studies showed that the down-regulation of the miR-200 family members appears to promote the epithelial-mesenchymal transition, proving their suppressive effects on cancer cell proliferation, migration, and invasion [115–118]. However, Park *et al.* [119] experimentally demonstrated how the down-regulation of all members of the miR-200 family would result in mesenchymal cell lines, while a their up-regulation would appear characteristic of an epithelial phenotype.

In the dataset we analysed in [86], all members of the miR-200 family appear to be highly up-regulated in cancer tissues (from 4- to 8- folds) and this up-regulation is counteracted by a similar, even if not comparable, overexpression of PVT1 that in cancer tissues appears to increase of about two folds. This observation could in principle warrant the annihilation of the PVT1 sponge activity noted in cancer dataset. In fact, Salmena *et al.* [63] suggested that the breakdown of the ceRNA activity could be due to a titration mechanism, *i.e.* large changes in the ceRNA expression levels that either overcome, or relieve, the miRNA repression on competing ceRNAs; similarly, large changes in the miRNA expression allow miRNAs to escape the recruitment accomplished by ceRNAs.

Here, we are interested to analyse the specific conditions required for a ceRNA landscape to occur, betting on the titration mechanism as the main culprit. In particular, inspiring by our

amazing results of [86] and by the growing interest of the scientific community in the oncogenic role of the lncRNA PVT1, we focused on its activity as sponge modulator of the activity of the miR-200 family members on their targets and on the withdrawal of its decoy service in breast cancer tissues.

Materials and methods

Algorithm for identifying ceRNA-ceRNA interactions

The pipeline of the algorithm for searching putative ceRNAs and for building the MMI-network (Fig 1) in breast invasive carcinoma was presented in our previous work [86] and encompassed the following four steps: i. data collection and processing; ii. statistical analysis; iii. seed match analysis; iv. network building.

1. Data collection and processing

Collections of tumour and normal expression data from high-throughput RNA- and miRNA-sequencing of breast invasive carcinoma were downloaded from the TCGA data portal [87, 88]. High-throughput sequencing data for both RNAs and miRNAs correspond to level 3 data (*i.e.* normalized expression data) given in terms of FPKM (*i.e.* fragments per kilobase of exon per million fragments mapped). The analysis was restricted to 72

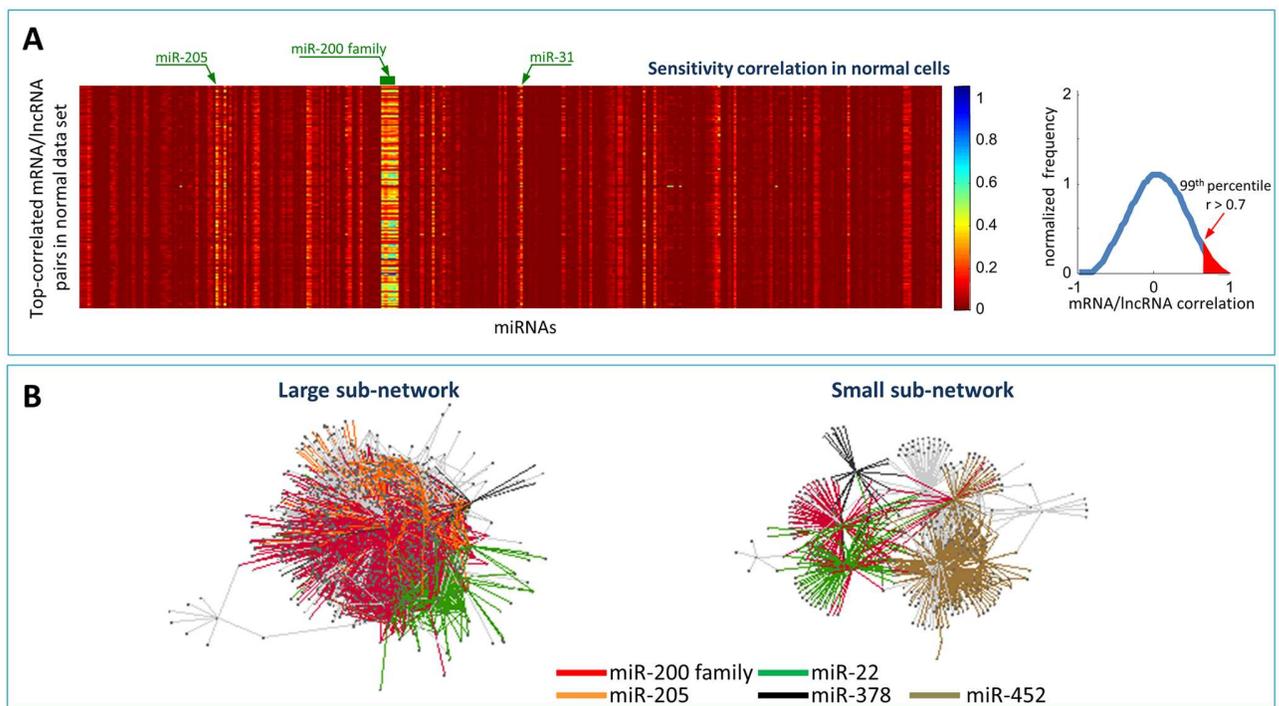


Fig 1. Sensitivity correlation and normal MMI-network. (A) Left: heat-map representing the sensitivity correlation, given in the Eq 1, for the top-correlated pairs ($N = 87398$) of mRNAs and lncRNAs (rows), previously identified in [86], in normal breast samples versus each miRNA (columns) that was expressed in the same tissues. Light vertical stripes point to a little pool of miRNAs that are responsible for the high correlation between all the top-correlated mRNA/lncRNA pairs. Colour key: red to blue scale corresponds to low to high sensitivity correlation. Right: the distribution of the Pearson correlations between mRNA and lncRNAs expression profiles. The pairs that are highlighted in red correspond to the top-correlated mRNA/lncRNA pairs: Pearson correlation values exceeding the 99th percentile of the overall correlation distribution (*i.e.* $\rho > 0.7$ in normal samples). (B) The normal MMI-network (1738 nodes and 32375 edges) built in [86] starting from the expression data of normal breast tissues. Nodes in this network represent both mRNAs and lncRNAs; edges represent miRNAs that are mediating their interactions. Each pair of linked nodes fulfils two requirements: i. sensitivity correlation > 0.3 and ii. one or more shared MREs, for each miRNA linking them. Colours correspond to different miRNAs.

doi:10.1371/journal.pone.0171661.g001

individuals for which the complete sets of tumour and matched-normal profiles—for both short and long RNA-seq data—were available. Entries with more than the 10% of missing values were filtered out; coding versus non-coding RNAs based on Entrez gene identifiers and human annotation obtained from NCBI [120] were separated. The analysis was limited to those mRNAs with an available 3' untranslated region (3'UTR) sequence at least equal to 500 nt in the curated UTRdb database [121]. All together, a total of 10492 mRNAs, 311 miRNAs, and 833 lncRNAs were analysed in [86].

2. Statistical analysis

The top-correlated mRNA/lncRNA pairs in normal and cancer data sets were selected by setting in both cases the correlation threshold to the 99th percentile of the corresponding overall correlation distribution (Fig 1A). Then, two regression models were built up: i. the expression profile of the mRNA is the dependent variable X and the expression profile of the miRNA is the explanatory variable Z ; ii. the expression profile of the lncRNA is the dependent variable Y and the expression profile of the miRNA is the explanatory variable Z . The X and Y variables correspond to the top-correlated mRNA/lncRNA pairs. To infer the role of Z in mediating $X - Y$ correlation, the partial correlation was computed:

$$\rho_{XY|Z} = \frac{\rho_{XY} - \rho_{XZ}\rho_{ZY}}{\sqrt{1 - \rho_{XZ}^2}\sqrt{1 - \rho_{ZY}^2}}$$

where $\rho_{X, Y}$ is the Pearson correlation. Then, the *sensitivity correlation* S was defined as:

$$S = \rho_{XY} - \rho_{XY|Z} \tag{1}$$

The XYZ triplets with $S > 0.3$, corresponding to a drop of about the 30% in the correlation between XY when Z is removed, were selected. Finally, these triplets were restricted to those enriched in binding sites of the shared miRNA (hypergeometric test p-value <0.01).

3. Seed match analysis

The minimal pairing requirement to predict a miRNA target recognition is a perfect match to positions 2 to 7 (6-mer miRNA seed) at the 5'-end of the mature miRNA sequence [122]. The miRNA seed sequences were obtained by mapping TCGA miRNA identifiers to miR-Base [123]. Complementary DNA (cDNA) sequences (*i.e.* without introns) for lncRNAs were obtained querying the Ensembl [124] data portal through its R/Bioconductor [125] interface provided by the package *biomaRt* and by using Entrez gene identifiers [126]. For each 3'UTR sequence included in the dataset analysed in [86], all the occurrences matching the reverse-complement of the 6-mer seed for the miRNAs analysed were recorded. Similarly, for each lncRNA included in the dataset analysed in [86] all the occurrences of short sites matching the reverse-complement of a miRNA seed in the entire transcript sequence were stored. The lists of coding and non-coding RNA identifiers used to retrieve corresponding sequences were built based on gene annotations obtained from the NCBI [120].

4. Network building

The MMI-network both in normal and cancer tissues was built by integrating the results of statistical analysis and seed match analysis. Nodes in the networks represent mRNAs and lncRNAs with highly correlated expression profiles while edges represent miRNAs mediating their interactions. Concretely, linked nodes are required to meet three conditions: i. matching high values of the Pearson correlation between their expression profiles ($\rho > 0.7$); ii. matching high values of the sensitivity correlation ($S > 0.3$); iii. sharing binding sites for miRNAs (6-mer miRNA seed match).

Raw data retrieval and processing

Data collection. Mapped read data (bam files) for the 72 patients (for which the complete sets of tumour and matched-normal profiles—for both short and long RNA-seq data—were available) analysed in [86] were downloaded from the TCGA [87] via controlled access (*i.e.* by using the TCGA dedicated software “gtdownload” to query via controlled access the restricted-access data repository). For each patient the relative two bam files corresponding to the breast tumour and normal sample are used as input for the Cufflinks software [127] in order to assemble transcripts and to estimate the relative abundances of these transcripts. As output formats the Cufflinks suite used FPKM tracking format. Then, we used Cuffmerge (a software included in Cufflinks) in order to merge together the 72 Cufflinks assemblies.

The PVT1 locus assembled by Cufflinks was compared with genome annotations for the same locus provided by Ensembl (release Homo sapiens GRCh37) by running the Cuffcompare utility and by careful inspection of the above assemblies and annotations on the UCSC genome browser.

Targeted reassembly of the PVT1 locus were performed using the Trinity software [128] with default parameters and digital normalization of the reads. To highlight possible differences between the healthy and tumour samples three independent assemblies were carried out, by using: a) all the reads mapping to the PVT1 locus from both cancer and normal tissue, b) only cancer reads, c) only reads from normal tissues. A UCSC genome browser track showing the main results of these analyses is available through this link: https://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=pantaleoM&hgS_otherUserSessionName=hg19_pone_S16_46501

While the overall agreement between the Trinity and Cufflinks assemblies is good, and both methods concur in recovering all of the PVT1 Refseq exons, we notice that some of exons predicted by Cufflinks are not supported by any of Trinity assemblies, and likely constitute false positives. Importantly the most prevalent isoforms reconstructed by Trinity are highly similar if not completely identical to the two most expressed PVT1 isoforms TCONS_147426 and TCONS_147501 predicted by Cufflinks, suggesting that these discrepancies in the assembly are not likely to play a major effect on isoforms abundance estimation. A differential alternative exon usage analysis (S1 Fig), performed by comparing the normalized reads counts distributions on the Refseq PVT1 exons, shows a striking pattern supporting the up-regulation of all the exons downstream of exon 5 in the tumour samples. This observation is highly consistent with our hypothesis that the up-regulation of PVT1 in tumour samples is mostly due to the up-regulation of isoforms of the gene devoid of the key exons exerting the sponge activity on miR-200 family members.

Statistical analysis. The PVT1 locus—assembled by the reference-based RNA-Seq transcriptome assembler Cufflinks using TCGA data of breast invasive carcinoma—is composed of 91 different isoforms (S2 Fig and S1 File). The FPKM normal and cancer data of these isoforms (S1 Table) were subjected to a pre-processing and filtering operation in order to reduce data noise and to select only the ones that show a statistically significant fold-changes between cancer and normal tissues (p -values of the Student’s t -test <0.05). Thus, the number of the PVT1 isoforms to analyse was trimmed to 17 isoforms.

Data classification. In order to classify the PVT1 isoforms on the basis of the FPKM data, we used the Principal Component Analysis (PCA) [129, 130]. PCA operates on a n -by- p data matrix X , whose rows correspond to the n observations and columns to the p variables. The representation of X in the principal component (PC) space is known as matrix of the principal component scores (S2 Table), whose rows correspond to observations and columns to the principal components (PCs). The transformation matrix from the old to the new coordinate

system is known as matrix of factors (S2 Table), whose rows correspond to variables and columns to components. In this study, the original variables are the 72 patients and the observations are the PVT1 isoforms' variations (*i.e.* the difference of the expression levels of the PVT1 isoforms between cancer and normal tissues).

Results and discussion

Inspiring by our previous study presented in [86], in this manuscript we have investigated the specific conditions required for a ceRNA interaction network to occur. In particular, we thoroughly studied the intriguing phenomenon of the breakdown of the PVT1 functioning as sponge of the miR-200 family members in the breast invasive carcinoma by analysing the expression data of its multiple isoforms (S1 Table). The starting point of the present analysis, which complements the results obtained in [86], is represented by the investigation of the sensitivity correlation behaviour (Fig 1A), formerly inspected in [86] and whose mathematical expression is reported in Eq 1. This enables measuring the contribution of miRNAs in mediating the ceRNAs cross-talk and provides compelling clues on the nature of the ceRNA interactions, *i.e.* indirect (direct) interaction meaning that the ceRNAs communication is (is not) arbitrated by one or more microRNAs. As already mentioned in [86], in physiological conditions the value of the sensitivity correlation is almost zero, *i.e.* the Pearson correlation is equal to the partial correlation, leading to the expected conclusion that, in normal breast tissues, the majority of the miRNAs is not arbitrating the cross-talk between long non-coding RNAs and coding RNAs. Thus, the observed high correlations between the expression profiles of the top-correlated lncRNA/mRNA pairs could be presumably ascribable to a common transcriptional regulatory mechanism, rather than to a post-transcriptional regulation program orchestrated by shared miRNAs. Nevertheless, a small pool of miRNAs appears as responsible of the vertical light stripes that unexpectedly stand out from the prevailing red colour of the background of Fig 1A. Hence, these miRNAs can be reasonably envisaged as the mediators of the interactions between all the highly correlated pairs in the normal breast samples. Among them, there are all members of the miR-200 family, whose importance in breast cancer is well-known and is related to the epithelial-mesenchymal transition. This pattern completely disappears in cancer [86] to give way to the activation of a different ceRNA landscape. This “on/off” switch from normal to cancer, and *vice-versa*, leads to the inference of a marked rewiring in the ceRNA program between normal and pathological breast tissue that confers an interesting character to ceRNAs as potential oncosuppressive, or oncogenic, protagonists in cancer.

Using the sensitivity correlation and the results of the seed match analysis, summarised in Materials and Methods section, the MMI-network was built in both the physiological and pathological condition of human breast cancer dataset analysed in [86]. Nodes of these network are lncRNAs and mRNAs that are competing for miRNA binding and links are the “bone of contention” miRNAs (Fig 1B). The lncRNA PVT1 with its 2169 edges represents the first hub (*i.e.* the node with the largest number of links or the highest degree in the network) in the normal-MMI-network. It is connected to 753 different mRNAs (~ 50% of total mRNAs in the network) and the miR-200 family members are arbitrating over the 80% of these interactions (Fig 2A). Moreover, PVT1 has as nearest neighbours some of the well-known cancer genes (Fig 2B) and is connected to 753 different mRNAs representing more than the 50% of all the mRNAs in the whole normal MMI-network (Fig 2C).

By analysing the expression profiles of PVT1, obtained from the dataset studied in [86], over all the patients, we found that it is up-regulated in breast cancer tissues both as mean value (Fig 3A) and individually on each patient, regardless of the breast cancer subtypes (Fig 3B). This up-regulation is counteracted by a similarly, but even more significant,

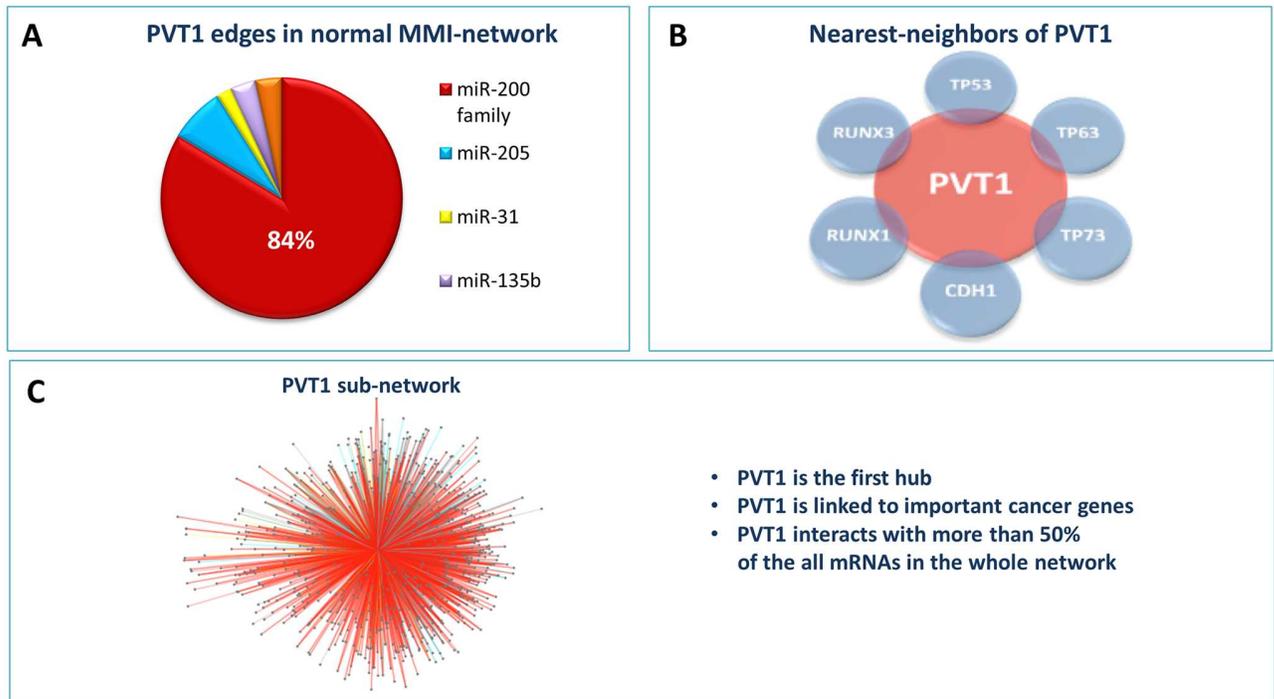


Fig 2. PVT1: the first hub of the normal MMI-network. (A) The percentage of the miRNAs sponged by PVT1 with respect to all of its links. More than the 80% corresponds to the miR-200 family members. (B) Some nearest neighbours of PVT1 that are well-known cancer genes as the members of the p53 family or the members of the RUNX family, as well as E-cadherin. (C) The sponge interactions sub-network of PVT1. It consists of 2169 edges and 753 nodes, more than the 50% of all the mRNAs in the whole normal MMI-network.

doi:10.1371/journal.pone.0171661.g002

overexpression of the miR-200 family members (see Fig 3A and 3C for the representative case of the miR-200b). The question, then, arises: if PVT1 and the miR-200 family are both up-regulated in cancer, why PVT1 stops working as sponge in cancer?

The analysis of the PVT1 genomic locus showed the existence of multiple isoforms (Fig 4 and S2 Fig) representing all the possible configurations: hosting the binding site for some (e.g. Iso6 or Iso7 in Fig 4) or all members of the miR-200 family (e.g. Iso1 in Fig 4); missing the binding site (e.g. Iso11 and Iso12 in Fig 4). This consideration together with the observed synchronised up-regulation of the PVT1 gene and the miR-200 family members encouraged us to hypothesize different scenarios that could be in principle compatible with the ceasing of the PVT1 sponge activity in breast cancer tissues. From one hand, the absence in two PVT1 isoforms of the exon where the MREs for the all members of the miR-200 family reside could lead to support the hypothesis of a preferential expression in cancer tissues of these two isoforms, thus justifying the lack of the miRNA/target interaction with a consequent breakdown of the PVT1 ceRNA activity (i.e. the exon skipping mechanism). From the other hand, the observation of a simultaneous up-regulation of the PVT1 gene and the miR-200 family members could lead to support the alternative hypothesis of different relative concentrations between each isoform and the miR-200 family members. According to that, a substantial decrease in cancer tissues of the relative variation of the isoform harbouring the binding site for one or more members of the miR-200 family could be due to a huge increase of the miR-200 family associated with a moderate growth in cancer of the expression levels of this PVT1 isoform. This situation, completely different from what occurs in normal tissues where the miRNA/target concentrations are comparable, could give reason of the PVT1 cease-activity as ceRNA in cancer (i.e. a titration mechanism).

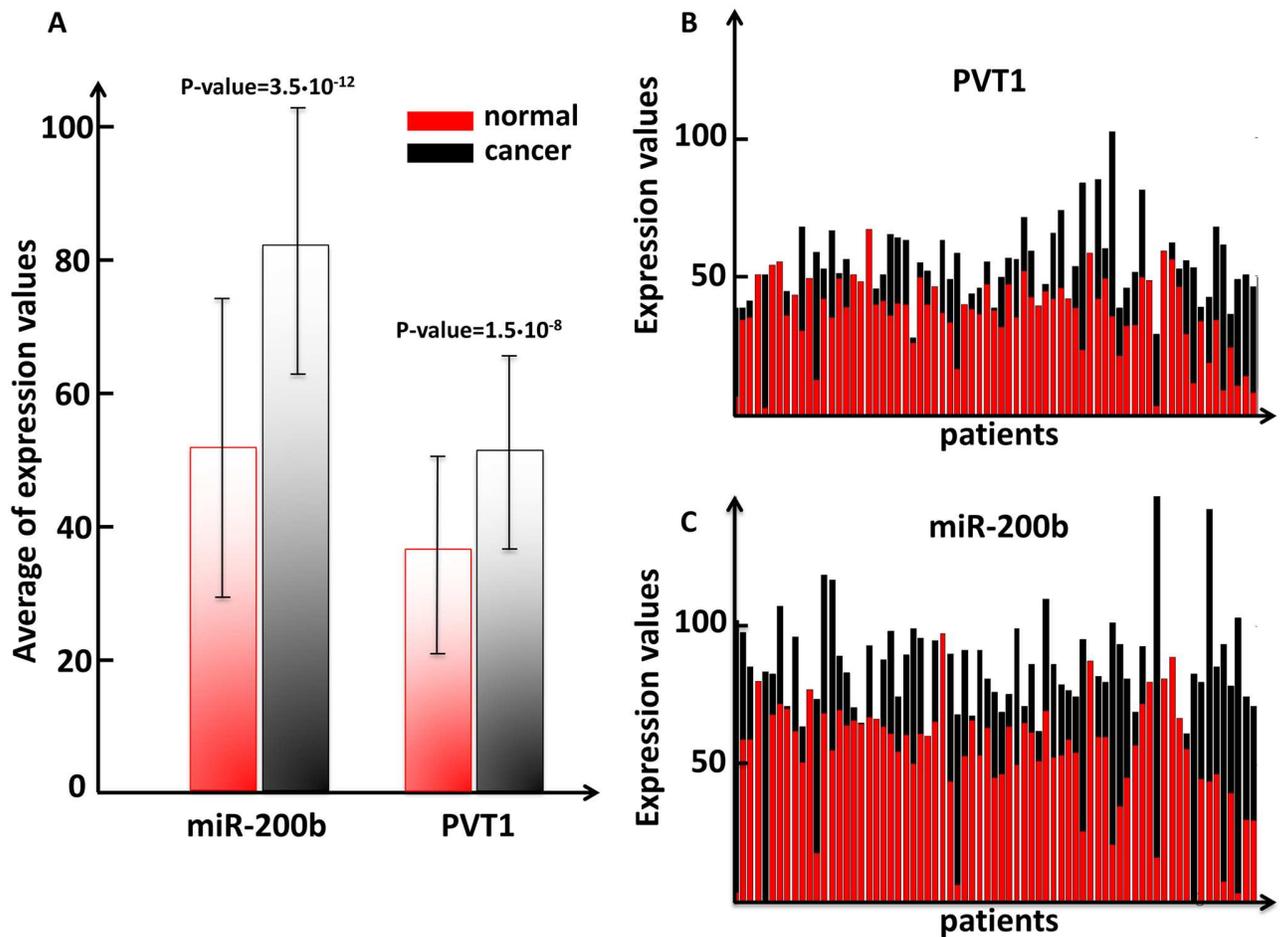


Fig 3. PVT1 and miR-200b expression levels in human breast cancer tissues. The main actors of the normal MMI-network are the miR-200 family members and the long non-coding PVT1. Both of them appear to be up-regulated in breast cancer tissues with respect to normal breast tissues. (A) The mean and the standard deviation for the long non-coding PVT1 and for the miR-200b, one member of the miR-200 family, in normal (red boxes) and cancer samples (black boxes). In figure the p-values resulting from the statistical hypothesis Student's t-test are reported. (B-C) Level 3 (*i.e.* normalized expression data) IlluminaHiSeq expression data of PVT1 and the miR-200b for all patients given in terms of FPKM (*i.e.* fragments per kilobase of exon per million fragments mapped). Red boxes correspond to normal tissues while black boxes correspond to cancer tissues.

doi:10.1371/journal.pone.0171661.g003

To shed light on which of the two hypothesised mechanisms lies the origin of the PVT1 stoppage as sponge, we looked at the PVT1 abundance in terms of its isoforms and we found that in both normal (Fig 5A and S3 Table) and cancer tissues (Fig 5B and S3 Table) only two isoforms represent the biggest slices: the first largest slice—which corresponds to the 50% (48%) of the PVT1 total abundance in normal (cancer) breast samples—represents the isoform missing the binding site for the miR-200 family (TCONS_147501); the second largest slice—which corresponds to the 15% (17%) of the PVT1 total abundance in normal (cancer) breast samples—represents the isoform hosting the binding site for the miR-200b/200c/429 cluster (TCONS_147426). Overall, both in normal and cancer tissues the two isoforms TCONS_147501 and TCONS_147426 represent about the 65% of the total abundance of PVT1 (S3 Table). Moreover, PVT1 resulted up-regulated also in terms of its total isoforms abundance (Fig 5C), confirming the result obtained at gene level (Fig 3A).

The observation that both the isoforms, with and without the exons where the MREs of the miR-200 family members reside, resulted expressed in both cancer and normal breast tissues

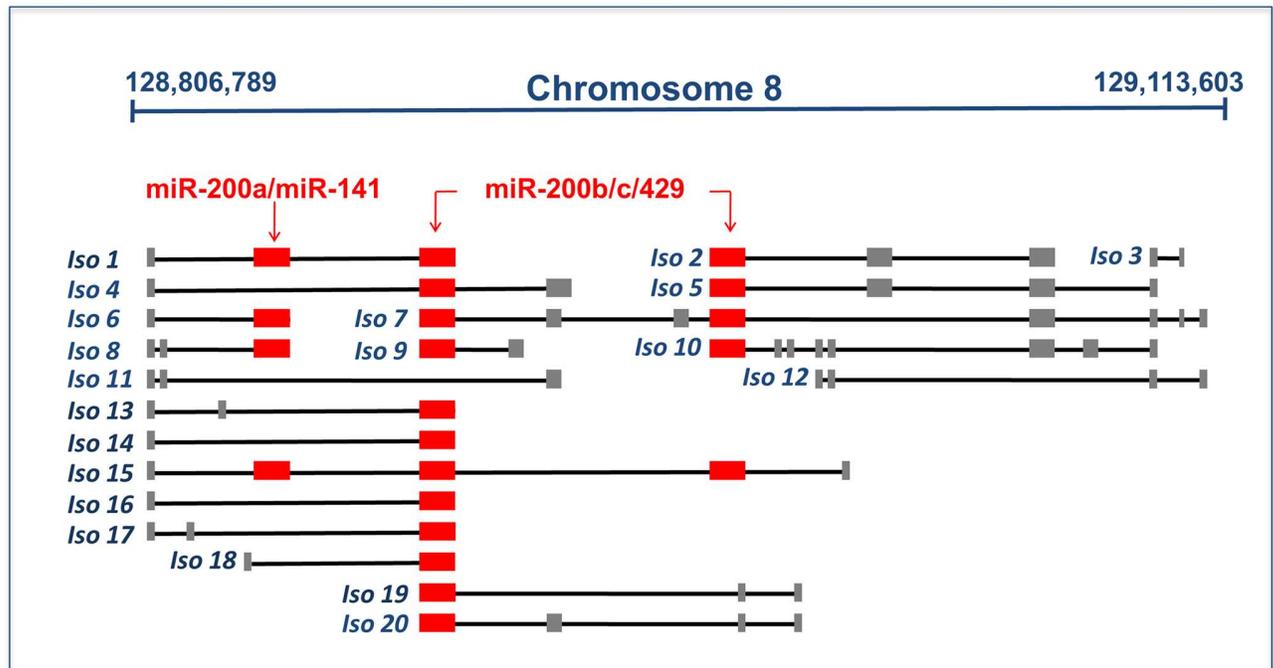


Fig 4. Sketch of the PVT1 locus in humans. Model of PVT1 genomic locus as reconstructed by Cufflinks (S1 File) spans across a genome interval of over 300 kb (*i.e.* bases 128,806,789-129,113,603 within the February 2009 human genome build GRCh37/hg19) on the forward strand of chromosome 8. The large PVT1 locus gives rise to 91 different variants (S2 Fig) according to raw RNA-seq data of TCGA for breast invasive carcinoma. The isoform names correspond to an increasing symbolic numbering and not to the actual nomenclature of the PVT1 variants. Lines represent introns and boxes (red and grey) represent exons. Red boxes correspond to the binding sites for the miR-200 family members. Note that some isoforms lack such binding sites (*e.g.* Iso11 and Iso12).

doi:10.1371/journal.pone.0171661.g004

undermine the truthfulness of the hypothesis rested on the exon skipping mechanism and corroborates the proposal based on the relative concentrations of the PVT1 isoforms and the miR-200 family members.

Thus, in order to sift through the validity of a titration mechanism, we performed the principal component analysis using the feature abundance levels of all the PVT1 isoforms across samples (S1 Table). The aim of PCA is to determine the principle axes of the abundance variation and to separate the isoforms according to this feature. This is achieved through a reduction of the space dimensionality that transforms a high-dimensional dataset—where the dimension of the space is equal to the linear independent variables (*i.e.* patients)—into a smaller-dimensional subspace—where the dimension of the space is equal to number of PCs that are able to explain the first 100% of the cumulative distribution of the explained variance of the data. The first step of this analysis is to draw a new axis representing the direction of maximum variation through the data (the first PC). Next, another axis is added orthogonal to the first and positioned to represent the next highest variation through the data (the second PC), and so on.

We found that two PCs are able to explain more than the 80% of the variance of the data (Fig 6A and S2 Table). In order to understand the meaning of these two PCs, we drew the score plot (Fig 6B and S2 Table) and found that the first PC is able to separate the contribution of the isoform missing the binding site for any members of the miR-200 family from the others, while the second PC is able to separate the contribution of the isoform hosting the binding site for the miR-200b/200c/429 cluster from the others.

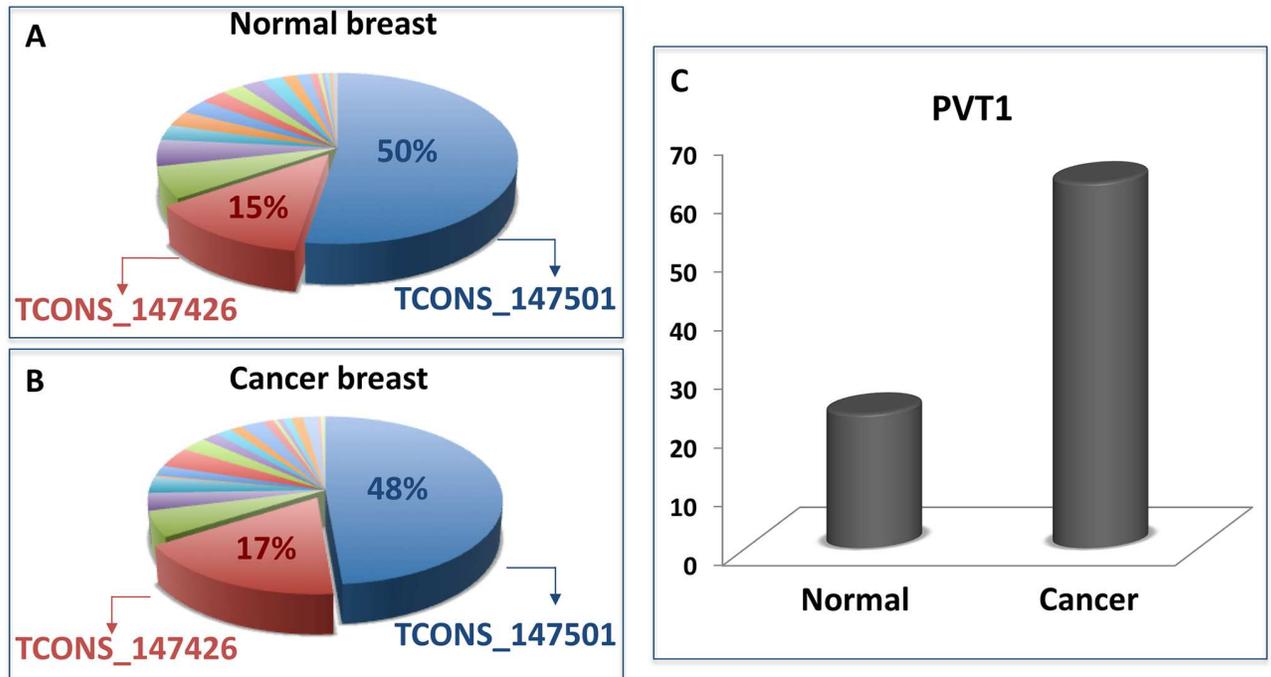


Fig 5. PVT1 isoforms in normal and cancer breast tissues. The PVT1 abundance in terms of its isoforms in normal (A) and cancer tissues (B) of breast invasive carcinoma. The percentage are calculated with respect to the total abundance of PVT1 in normal and cancer tissues (S3 Table). In both panels, the red slice corresponds to the isoform (TCONS_147426) with seed match for the miR-200b/200c/429 cluster and the blue slice corresponds to the isoform (TCONS_147501) lacking the binding site for any member of the miR-200 family. (C) The average of PVT1 abundance on all the isoforms both in normal and cancer tissues.

doi:10.1371/journal.pone.0171661.g005

This suggests the following argument of plausibility of the PCA analysis results: the first PC, which explain by alone about the 60% of the total variance of the analysed data (S2 Table), corresponds to the variation of the isoform that, missing the binding site, does not interact with the miR-200 family; while the second PC, explaining by alone about the 20% of the total variance of the analysed data (S2 Table), represents the variation of the isoform that, hosting the binding site for the miR-200b/200c/429 cluster, could be act as competitors of the targets of these miRNAs. Overall the variation between cancer and normal tissues of these two isoforms accounts for more than the 80% of the variance of the data (Fig 6A and S2 Table).

Studying the variation of each PVT1 isoform between normal and cancer breast tissues with respect to the variation of TCONS_147501, the results of PCA seems to be confirmed (Fig 7 and S4 Table): the isoform harbouring the binding site for the miR-200b/200c/429 cluster and the isoform missing the binding site for any member of the miR-200 family, are the only isoforms that change (Fig 7A).

Thus, we considered only these two isoforms (S2 File) and evaluated the ratio between the abundance of each one with respect to one representative member of the miR-200b/200c/429 cluster (*i.e.* miR-200b) in both normal and cancer tissues. For the TCONS_147501 isoform (missing the binding site) this ratio does not change between normal and breast cancer tissues, while in the case of the TCONS_147426 isoform (harbouring the binding site) this ratio shows a drastic decrease from normal to cancer tissues (Fig 7B). We speculate that the TCONS_147426 isoform acts as sponge regulator of the miR-200b in normal breast tissues, while the sponge mechanism is broken down in cancer tissues because this isoform shows a much lower concentration with respect to the miR-200b (Fig 8A). Informally speaking, such a

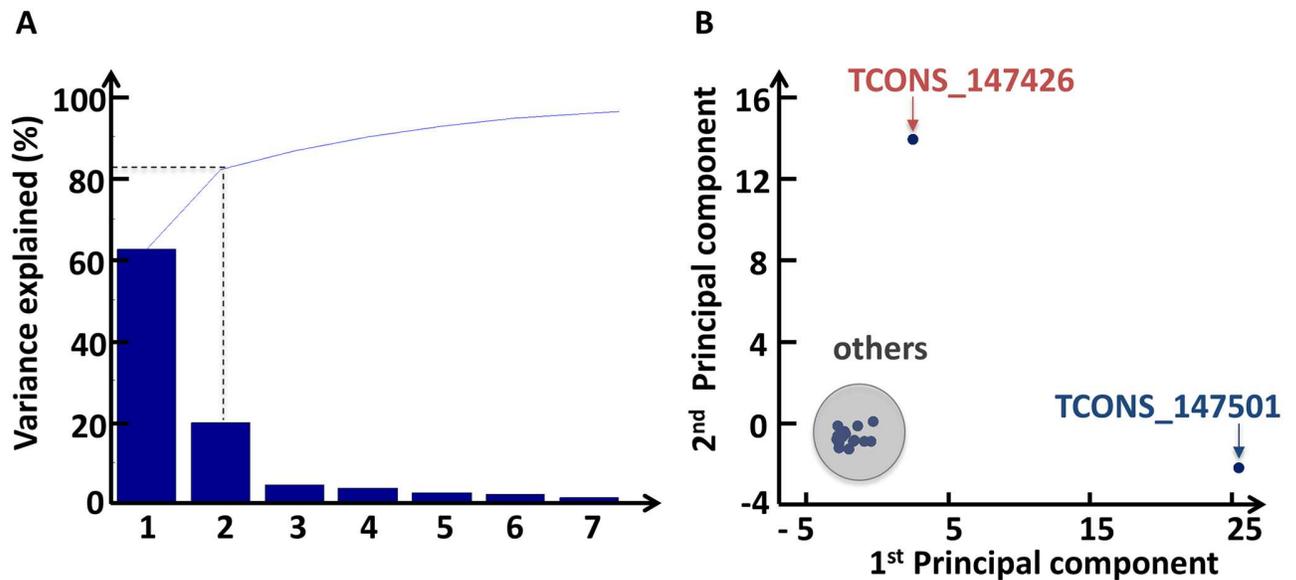


Fig 6. Principal component analysis. (A) The percent variability explained by each principal component (S2 Table). It is a type of chart, called Pareto chart, that contains both bars and a line graph, where individual values are represented in descending order by bars, and the line represents the cumulative total value. In particular, the y-axis represents the percentage of the data variance explained by each principal component, whereas the x-axis represents the principal components that are able to explain the first 100% of the cumulative distribution. The PCA is performed using the variations of all the isoforms between normal and cancer tissues. Two components explain more than the 80% of the variance of the data. (B) The scatter plot (score plot) of the projection of the original data (*i.e.* the variations of all the isoforms between normal and cancer tissues) onto the first two PCs; the x-axis contains the first PC while the y-axis contains the second PC (S2 Table). In this plot, it is possible to group isoforms in three classes: the isoform missing the binding site for the miR-200 family members (blue isoform, TCONS_147501), the isoform with the seed match for the miR-200b/200c/429 cluster (red isoform, TCONS_147426), and all the others. The first PC, which explains about the 60% of the variance in the original data, is able to separate the variation of the blue isoform from the others; the second PC, which explains about the 20% of the variance in the original data, is able to separate the variation of the red isoform from the others.

doi:10.1371/journal.pone.0171661.g006

sponge mechanism works as a real sponge: before saturation the sponge can hold more water, beyond saturation—there is too much water—the sponge can not hold more (Fig 8B).

So, our analysis supports the hypothesis that the “on/off” switch from normal to cancer state of the PVT1 sponge activity is mostly due to the variation of the relative concentration of PVT1 isoform hosting the binding site for the miR-200b/200c/429 cluster.

Conclusion

Starting from the results presented in [86]—where we analysed the complex interactions among mRNAs, long non-coding RNAs, and microRNAs in breast invasive carcinoma—here we investigated the mechanism underlying the marked rewiring of the sponge program between normal and cancer tissues. In particular, the analysis of the normal miRNA-mediated interactions network, built in [86], pointed out how the main actors of this rewiring were PVT1 and the miR-200 family members. Specifically, PVT1 emerged as a putative ceRNA modulating the activity of all members of the miR-200 family on their target mRNAs, which are well-known to be drastically involved in breast cancer morphogenesis and development. Interestingly, such a sponge mechanism resulted completely abolished in cancer tissues, although both PVT1 and the miR-200 family members appeared up-regulated in the pathological condition. Thus, processing the raw data from TCGA, which provided the abundance of the multiple isoforms generated by the PVT1 genomic locus, we tried to grasp the rationale behind the turning off of this sponge mechanism. In particular, the principal component

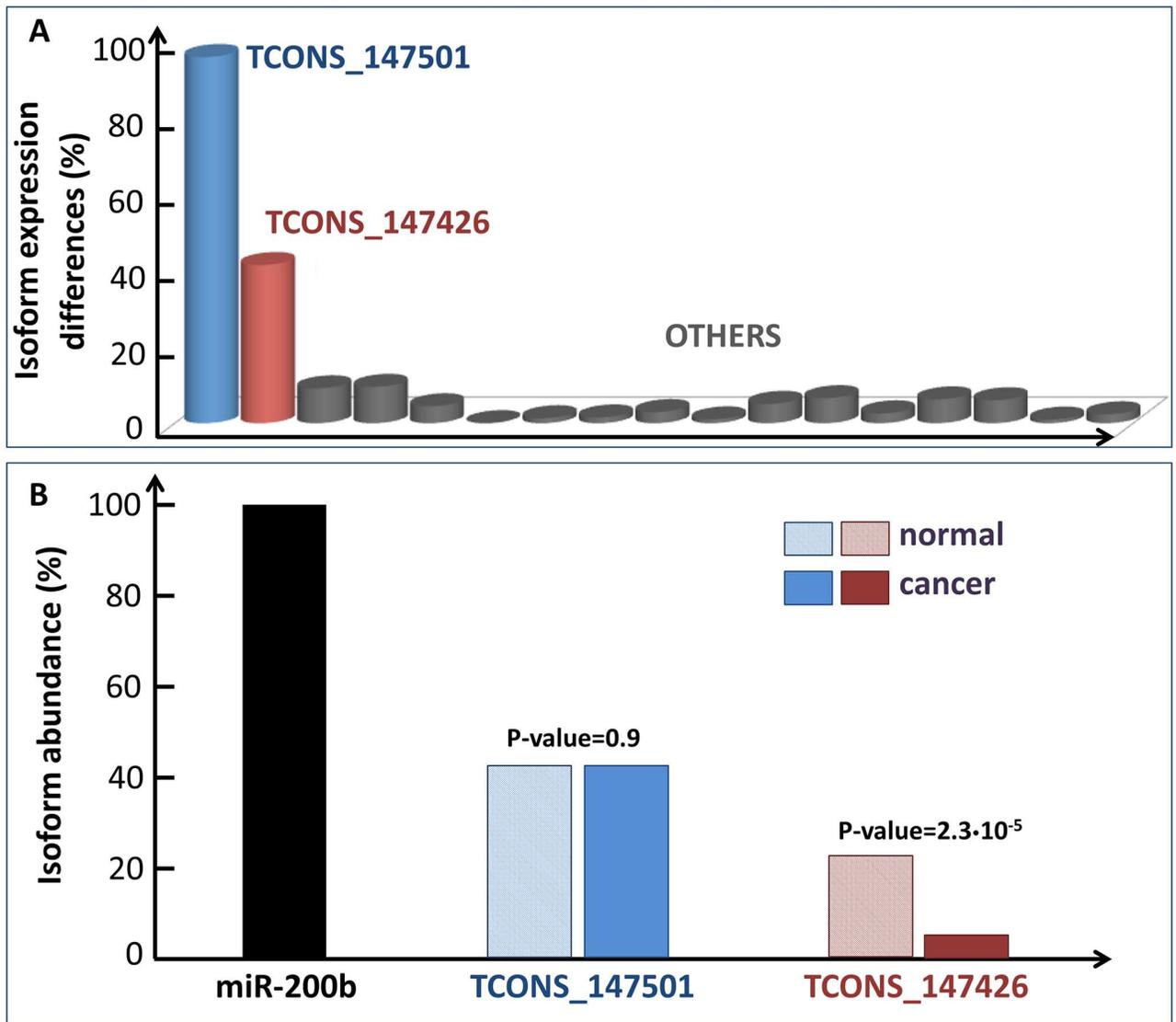


Fig 7. PVT1 isoforms variation. (A) The variations between cancer and normal tissues of all the PVT1 isoforms with respect to the variation of the blue isoform lacking the binding site for the miR-200 family members (S4 Table). The red and blue isoforms are the only isoforms that change. (B) The ratio between the abundance of the blue and red isoform with respect to the miR-200b in both normal (striped rectangle) and cancer tissues (full boxes). The p-values resulting from the statistical hypothesis Student's t-test are reported. The ratio between the blue isoform and the miR-200b does not change, while the ratio between the red isoform and the miR-200b shows a drastic fall in cancer tissues.

doi:10.1371/journal.pone.0171661.g007

analysis suggested that the variations between cancer and normal breast tissues of all PVT1 isoforms can be explained by only two principal components: one corresponding to the isoform harbouring the binding site for the miR-200b/200c/429 cluster and the other one representing the isoform missing the binding site for any member of the miR-200 family members. Moreover, comparing the relative expression levels of these two isoforms both in normal and cancer tissues with respect to the ones of the one representative member of the miR-200b/200c/429 cluster (*i.e.* miR-200b), we found a drastic drop, in the pathological condition, in the relative concentration of the PVT1 isoform hosting the binding site for the miR-200b. The drastic change observed in the sponge program, which is suggestive of a marked ceRNA rewiring that

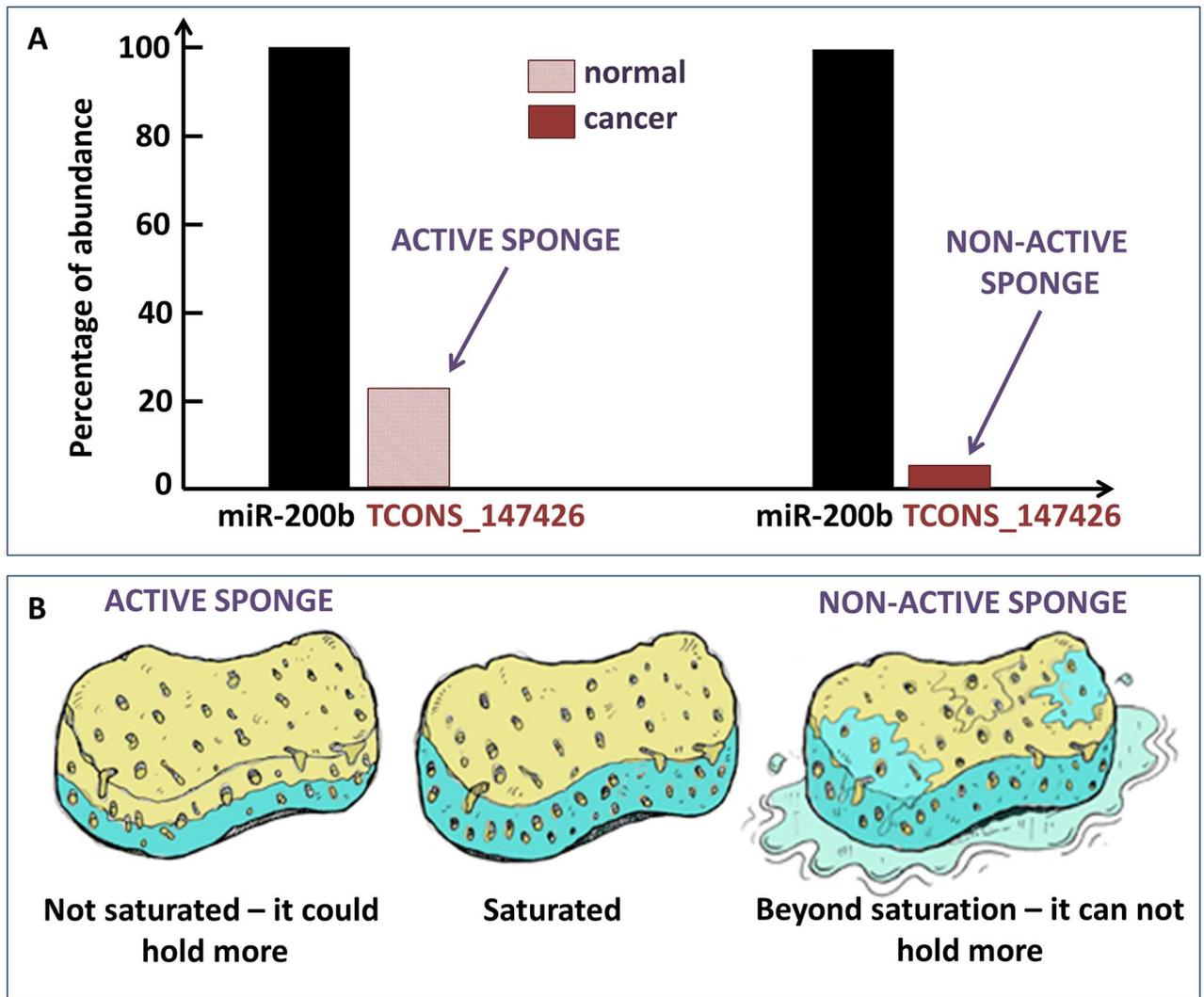


Fig 8. Sponge mechanism. (A) The ratio between the abundance of the red isoform (harbouring the binding site for the miR-200b/200c/429 cluster) over the abundance of the miR-200b in both normal (striped rectangle) and cancer tissues (full boxes). (B) Our hypothesis for the breakdown of the sponge mechanism in breast cancer orchestrated by PVT1 and the miR-200 family members. We speculate that in the normal tissues only the red isoform of PVT1 gene acts as sponge regulator of the miR-200 family members. In cancer tissues it stops working as sponge since its concentration is much lower than the concentration of the miR-200 family members (here is reported only the case of miR-200b). It is like in the case of a real sponge: before saturation the sponge can hold more water, beyond saturation—there is too much water—the sponge can't hold more.

doi:10.1371/journal.pone.0171661.g008

characterizes the cancer state, could support the testable hypothesis of a titration mechanism regarding the two main isoforms of PVT1 and the miR-200 family members.

Supporting information

S1 Fig. Differential alternative PVT1 usage analysis. This figure shows the results of a differential alternative exon usage analysis, performed by comparing the normalized reads counts distributions on the Refseq PVT1 exons. It shows a striking pattern supporting the up-regulation of all the exons downstream of exon 5 in the tumour samples. This observation is highly consistent with our hypothesis that the up-regulation of PVT1 in tumour samples is mostly

due to the up-regulation of isoforms of the gene devoid of the key exons exerting the sponge activity on miR-200 family members.

(PNG)

S2 Fig. Visualisation of PVT1 genomic locus in human. This figure shows the 91 PVT1 isoforms (*i.e.* bases 128,806,789-129,113,603 within the February 2009 human genome build GRCh37/hg19) visualised within the UCSC Genome browser (<https://genome.ucsc.edu/>) and assembled by the reference-based RNA-Seq transcriptome assembler Cufflinks by using the TCGA breast invasive carcinoma dataset.

(PNG)

S1 File. Reconstruction of PVT1 genomic locus in human. This file contains PVT1 gene annotations in GTF (Gene Transfer Format) format provided by the reference-based RNA-Seq transcriptome assembler Cufflinks. This file is a simple tab-delimited text file for describing genomic features and it can be uploaded to a genome browser such as the UCSC Genome browser (<https://genome.ucsc.edu/>) in order to obtain the [S2 Fig](#).

(GTF)

S2 File. Sequences of two PVT1 isoforms. This file contains the full genome sequences (in FASTA format) of the two PVT1 isoforms that mostly change between normal and cancer tissues: TCONS_147501 (missing the binding site for miR-200 family members) and TCONS_147426 (harbouring the binding site for the miR-200b/200c/429 cluster).

(FA)

S1 Table. PVT1 isoforms expression levels. This table reports the FPKM values of PVT1 isoforms across normal and cancer breast tissues in separate and accordingly named sheets.

(XLSX)

S2 Table. Principal Component Analysis. This table reports the results of the principal component analysis, in separate and accordingly named sheets: first sheet) the eigenvalues of the covariance matrix of the n -by- p data matrix X , whose rows correspond to observations (*i.e.* isoforms' variations that are the difference of the expression levels of the PVT1 isoforms between cancer and normal tissues) and columns to variables (*i.e.* patients), the variance accounted for by each component, and the cumulative function; second sheet) the matrix of the principal component scores, whose rows correspond to observations and columns to components; third sheet) the matrix of factors, whose rows correspond to variables and columns to components.

(XLSX)

S3 Table. PVT1 isoform abundance. This table reports the percentage of PVT1 abundance, showed in [Fig 5](#), in terms of its isoforms both in normal and cancer tissues of TCGA breast invasive carcinoma.

(XLSX)

S4 Table. PVT1 isoforms variation. This table reports the variations between cancer and normal tissues of the expression levels of all the PVT1 isoforms with respect to the variation of the TCONS_147501 isoform lacking the binding site for the miR-200 family members, showed in [Fig 7A](#).

(XLSX)

Acknowledgments

Paola Paci says: "Thank you Jacopo, my son, my life. You were born during the revision process of this paper and you inspired it."

The authors thank Prof. G. Macino for encouragement, fresh biological insights and for inspiring discussions on the whole analysis presented in this study.

The authors thank Leandro Castellano—Department of Surgery and Cancer, Imperial College London, Imperial Centre for Translational and Experimental Medicine (ICTEM)—for providing the RNA-seq raw data.

The results shown in this paper are in part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>.

Author Contributions

Conceptualization: PP.

Data curation: TC MC.

Formal analysis: PP FC GF.

Funding acquisition: PP.

Investigation: PP FC GF.

Methodology: PP.

Project administration: PP.

Resources: TC MC.

Software: PP LF FC.

Supervision: PP.

Validation: PP LF.

Visualization: PP FC GF MC.

Writing – original draft: PP FC GF.

Writing – review & editing: PP FC GF MC.

References

1. Mattick JS. The central role of RNA in human development and cognition. *FEBS letters*. 2011; 585(11):1600–1616. doi: [10.1016/j.febslet.2011.05.001](https://doi.org/10.1016/j.febslet.2011.05.001) PMID: [21557942](https://pubmed.ncbi.nlm.nih.gov/21557942/)
2. Esteller M. Non-coding RNAs in human disease. *Nature Reviews Genetics*. 2011; 12(12):861–874. doi: [10.1038/nrg3074](https://doi.org/10.1038/nrg3074) PMID: [22094949](https://pubmed.ncbi.nlm.nih.gov/22094949/)
3. Knowling S, Morris KV. Non-coding RNA and antisense RNA. Nature's trash or treasure? *Biochimie*. 2011; 93(11):1922–1927. doi: [10.1016/j.biochi.2011.07.031](https://doi.org/10.1016/j.biochi.2011.07.031) PMID: [21843589](https://pubmed.ncbi.nlm.nih.gov/21843589/)
4. Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 2007; 447(7146):799–816. doi: [10.1038/nature05874](https://doi.org/10.1038/nature05874) PMID: [17571346](https://pubmed.ncbi.nlm.nih.gov/17571346/)
5. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *cell*. 2004; 116(2):281–297. doi: [10.1016/S0092-8674\(04\)00045-5](https://doi.org/10.1016/S0092-8674(04)00045-5) PMID: [14744438](https://pubmed.ncbi.nlm.nih.gov/14744438/)
6. Bartel DP. MicroRNAs: target recognition and regulatory functions. *cell*. 2009; 136(2):215–233.
7. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Reviews Genetics*. 2008; 9(2):102–114. doi: [10.1038/nrg2290](https://doi.org/10.1038/nrg2290) PMID: [18197166](https://pubmed.ncbi.nlm.nih.gov/18197166/)
8. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nature Reviews Genetics*. 2009; 10(3):155–159. doi: [10.1038/nrg2521](https://doi.org/10.1038/nrg2521) PMID: [19188922](https://pubmed.ncbi.nlm.nih.gov/19188922/)
9. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell*. 2009; 136(4):629–641. doi: [10.1016/j.cell.2009.02.006](https://doi.org/10.1016/j.cell.2009.02.006) PMID: [19239885](https://pubmed.ncbi.nlm.nih.gov/19239885/)

10. Chang HY. Genome Regulation by Long Non-Coding RNAs. *Blood*. 2013; 122(21):SCI–29.
11. Qureshi IA, Mattick JS, Mehler MF. Long non-coding RNAs in nervous system function and disease. *Brain research*. 2010; 1338:20–35. doi: [10.1016/j.brainres.2010.03.110](https://doi.org/10.1016/j.brainres.2010.03.110) PMID: [20380817](https://pubmed.ncbi.nlm.nih.gov/20380817/)
12. Nagano T, Fraser P. No-nonsense functions for long noncoding RNAs. *Cell*. 2011; 145(2):178–181. doi: [10.1016/j.cell.2011.03.014](https://doi.org/10.1016/j.cell.2011.03.014) PMID: [21496640](https://pubmed.ncbi.nlm.nih.gov/21496640/)
13. Clark MB, Mattick JS. Long noncoding RNAs in cell biology. *Seminars in cell & Developmental biology*. 2011; 22(4):366–376. doi: [10.1016/j.semcdb.2011.01.001](https://doi.org/10.1016/j.semcdb.2011.01.001) PMID: [21256239](https://pubmed.ncbi.nlm.nih.gov/21256239/)
14. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. *Molecular cell*. 2011; 43(6):904–914. doi: [10.1016/j.molcel.2011.08.018](https://doi.org/10.1016/j.molcel.2011.08.018) PMID: [21925379](https://pubmed.ncbi.nlm.nih.gov/21925379/)
15. Gibb EA, Brown CJ, Lam WL, et al. The functional role of long non-coding RNA in human carcinomas. *Mol Cancer*. 2011; 10(1):38–55. doi: [10.1186/1476-4598-10-38](https://doi.org/10.1186/1476-4598-10-38) PMID: [21489289](https://pubmed.ncbi.nlm.nih.gov/21489289/)
16. Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. *Cancer discovery*. 2011; 1(5):391–407. doi: [10.1158/2159-8290.CD-11-0209](https://doi.org/10.1158/2159-8290.CD-11-0209) PMID: [22096659](https://pubmed.ncbi.nlm.nih.gov/22096659/)
17. Moran VA, Perera RJ, Khalil AM. Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. *Nucleic acids research*. 2012; 40(14):6391–6400. doi: [10.1093/nar/gks296](https://doi.org/10.1093/nar/gks296) PMID: [22492512](https://pubmed.ncbi.nlm.nih.gov/22492512/)
18. Tano K, Akimitsu N. Long non-coding RNAs in cancer progression. *Frontiers in genetics*. 2012; 3. doi: [10.3389/fgene.2012.00219](https://doi.org/10.3389/fgene.2012.00219) PMID: [23109937](https://pubmed.ncbi.nlm.nih.gov/23109937/)
19. Tang JY, Lee JC, Chang YT, Hou MF, Huang HW, Liaw CC, et al. Long Noncoding RNAs-Related Diseases, Cancers, and Drugs. *The Scientific World Journal*. 2013; 2013. doi: [10.1155/2013/943539](https://doi.org/10.1155/2013/943539)
20. Li X, Wu Z, Fu X, Han W. Long Noncoding RNAs: Insights from Biological Features and Functions to Diseases. *Medicinal Research Reviews*. 2013; 33(3):517–553. doi: [10.1002/med.21254](https://doi.org/10.1002/med.21254) PMID: [22318902](https://pubmed.ncbi.nlm.nih.gov/22318902/)
21. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nature Reviews Genetics*. 2014; 15(1):7–21. doi: [10.1038/nrg3606](https://doi.org/10.1038/nrg3606) PMID: [24296535](https://pubmed.ncbi.nlm.nih.gov/24296535/)
22. Dey BK, Mueller AC, Dutta A. Long non-coding rnas as emerging regulators of differentiation, development, and disease. *Transcription*. 2014; 5(4). doi: [10.4161/21541272.2014.944014](https://doi.org/10.4161/21541272.2014.944014) PMID: [25483404](https://pubmed.ncbi.nlm.nih.gov/25483404/)
23. Yang G, Lu X, Yuan L. LncRNA: A link between RNA and cancer. *Biochimica et Biophysica Acta—Gene Regulatory Mechanisms*. 2014; 1839(11):1097–1109. doi: [10.1016/j.bbaggm.2014.08.012](https://doi.org/10.1016/j.bbaggm.2014.08.012) PMID: [25159663](https://pubmed.ncbi.nlm.nih.gov/25159663/)
24. Li X, Wu Z, Fu X, Han W. LncRNAs: Insights into their function and mechanics in underlying disorders. *Mutation Research—Reviews in Mutation Research*. 2014; 762:1–21. doi: [10.1016/j.mrrev.2014.04.002](https://doi.org/10.1016/j.mrrev.2014.04.002) PMID: [25485593](https://pubmed.ncbi.nlm.nih.gov/25485593/)
25. Morlando M, Ballarino M, Fatica A, Bozzoni I. The Role of Long Noncoding RNAs in the Epigenetic Control of Gene Expression. *ChemMedChem*. 2014; 9(3):505–510. doi: [10.1002/cmdc.201300569](https://doi.org/10.1002/cmdc.201300569) PMID: [24488863](https://pubmed.ncbi.nlm.nih.gov/24488863/)
26. Hansji H, Leung EY, Baguley BC, Finlay GJ, Askarian-Amiri ME. Keeping abreast with long non-coding RNAs in mammary gland development and breast cancer. *Frontiers in Genetics*. 2014; 5(OCT). doi: [10.3389/fgene.2014.00379](https://doi.org/10.3389/fgene.2014.00379) PMID: [25400658](https://pubmed.ncbi.nlm.nih.gov/25400658/)
27. Sosińska P, Mikula-Pietrasik J, Ksiazek K. The double-edged sword of long non-coding RNA: The role of human brain-specific BC200 RNA in translational control, neurodegenerative diseases, and cancer. *Mutation Research—Reviews in Mutation Research*. 2015; 766:58–67. doi: [10.1016/j.mrrev.2015.08.002](https://doi.org/10.1016/j.mrrev.2015.08.002) PMID: [26596549](https://pubmed.ncbi.nlm.nih.gov/26596549/)
28. Iden M, Fye S, Li K, Chowdhury T, Ramchandran R, Rader JS. The lncRNA PVT1 contributes to the cervical cancer phenotype and associates with poor patient prognosis. *PLoS ONE*. 2016; 11(5). doi: [10.1371/journal.pone.0156274](https://doi.org/10.1371/journal.pone.0156274) PMID: [27232880](https://pubmed.ncbi.nlm.nih.gov/27232880/)
29. Parasramka MA, Maji S, Matsuda A, Yan IK, Patel T. Long non-coding RNAs as novel targets for therapy in Hepatocellular Carcinoma. *Pharmacology & therapeutics*. 2016; 161:67–78. doi: [10.1016/j.pharmthera.2016.03.004](https://doi.org/10.1016/j.pharmthera.2016.03.004) PMID: [27013343](https://pubmed.ncbi.nlm.nih.gov/27013343/)
30. Shi Q, Yang X. Circulating MicroRNA and Long Noncoding RNA as Biomarkers of Cardiovascular Diseases. *Journal of Cellular Physiology*. 2016; 231(4):751–755. doi: [10.1002/jcp.25174](https://doi.org/10.1002/jcp.25174) PMID: [26308238](https://pubmed.ncbi.nlm.nih.gov/26308238/)
31. Liu FT, Zhu PQ, Luo HL, Zhang Y, Hao TF, Xia GF, et al. Long noncoding RNA ANRIL: A potential novel prognostic marker in cancer A meta-analysis. *Minerva Medica*. 2016; 107(2):77–83. PMID: [27176569](https://pubmed.ncbi.nlm.nih.gov/27176569/)
32. Zhang X, Rice K, Wang Y, Chen W, Zhong Y, Nakayama Y, et al. Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions. *Endocrinology*. 2010; 151(3):939–947. doi: [10.1210/en.2009-0657](https://doi.org/10.1210/en.2009-0657) PMID: [20032057](https://pubmed.ncbi.nlm.nih.gov/20032057/)

33. Liang JC, Bloom RJ, Smolke CD. Engineering biological systems with synthetic RNA molecules. *Molecular cell*. 2011; 43(6):915–926. doi: [10.1016/j.molcel.2011.08.023](https://doi.org/10.1016/j.molcel.2011.08.023) PMID: [21925380](https://pubmed.ncbi.nlm.nih.gov/21925380/)
34. Saxena A, Carninci P. Long non-coding RNA modifies chromatin. *Bioessays*. 2011; 33(11):830–839. doi: [10.1002/bies.201100084](https://doi.org/10.1002/bies.201100084) PMID: [21915889](https://pubmed.ncbi.nlm.nih.gov/21915889/)
35. Novikova IV, Hennelly SP, Sanbonmatsu KY. Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. *Nucleic acids research*. 2012; 40(11):5034–5051. doi: [10.1093/nar/gks071](https://doi.org/10.1093/nar/gks071) PMID: [22362738](https://pubmed.ncbi.nlm.nih.gov/22362738/)
36. Mortimer SA, Kidwell MA, Doudna JA. Insights into RNA structure and function from genome-wide studies. *Nature reviews Genetics*. 2014; 15(7):469–479. doi: [10.1038/nrg3681](https://doi.org/10.1038/nrg3681) PMID: [24821474](https://pubmed.ncbi.nlm.nih.gov/24821474/)
37. Mercer TR, Mattick JS. Structure and function of long noncoding RNAs in epigenetic regulation. *Nature structural & molecular biology*. 2013; 20(3):300–307. doi: [10.1038/nsmb.2480](https://doi.org/10.1038/nsmb.2480) PMID: [23463315](https://pubmed.ncbi.nlm.nih.gov/23463315/)
38. Somarowthu S, Legiewicz M, Chillón I, Marcia M, Liu F, Pyle AM. HOTAIR forms an intricate and modular secondary structure. *Molecular cell*. 2015; 58(2):353–361. doi: [10.1016/j.molcel.2015.03.006](https://doi.org/10.1016/j.molcel.2015.03.006) PMID: [25866246](https://pubmed.ncbi.nlm.nih.gov/25866246/)
39. Fiscon G, Paci P, Colombo T, Iannello G. A new procedure to analyze RNA Non-branching Structures. *BSP Current Bioinformatics*. 2015; 9(5):242–258. doi: [10.2174/1574893609666140820224651](https://doi.org/10.2174/1574893609666140820224651)
40. Fiscon G, Iannello G, Paci P. A Perspective on the Algorithms Predicting and Evaluating the RNA Secondary Structure. *Journal of Genetics and Genome Research*. 2016; 3(023):to appear.
41. Tsai MC, Manor O, Wan Y, Mosammamaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science*. 2010; 329(5992):689–693. doi: [10.1126/science.1192002](https://doi.org/10.1126/science.1192002) PMID: [20616235](https://pubmed.ncbi.nlm.nih.gov/20616235/)
42. Fiscon G, Paci P, Iannello G. MONSTER v1. 1: a tool to extract and search for RNA non-branching structures. *BMC genomics*. 2015; 16(6):1. doi: [10.1186/1471-2164-16-S6-S1](https://doi.org/10.1186/1471-2164-16-S6-S1) PMID: [26047478](https://pubmed.ncbi.nlm.nih.gov/26047478/)
43. Pang KC, Frith MC, Mattick JS. Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends in Genetics*. 2006; 22(1):1–5. doi: [10.1016/j.tig.2005.10.003](https://doi.org/10.1016/j.tig.2005.10.003) PMID: [16290135](https://pubmed.ncbi.nlm.nih.gov/16290135/)
44. Kutter C, Watt S, Stefflova K, Wilson MD, Goncalves A, Ponting CP, et al. Rapid turnover of long non-coding RNAs and the evolution of gene expression. *PLoS Genet*. 2012; 8(7):e1002841. doi: [10.1371/journal.pgen.1002841](https://doi.org/10.1371/journal.pgen.1002841) PMID: [22844254](https://pubmed.ncbi.nlm.nih.gov/22844254/)
45. Wang J, Liu X, Wu H, Ni P, Gu Z, Qiao Y, et al. CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic acids research*. 2010; 38(16):5366–5383. doi: [10.1093/nar/gkq285](https://doi.org/10.1093/nar/gkq285) PMID: [20423907](https://pubmed.ncbi.nlm.nih.gov/20423907/)
46. Tay Y, Kats L, Salmena L, Weiss D, Tan SM, Ala U, et al. Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell*. 2011; 147(2):344–357. doi: [10.1016/j.cell.2011.09.029](https://doi.org/10.1016/j.cell.2011.09.029) PMID: [22000013](https://pubmed.ncbi.nlm.nih.gov/22000013/)
47. Karreth FA, Tay Y, Perna D, Ala U, Tan SM, Rust AG, et al. In vivo identification of tumor-suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. *Cell*. 2011; 147(2):382–395. doi: [10.1016/j.cell.2011.09.032](https://doi.org/10.1016/j.cell.2011.09.032) PMID: [22000016](https://pubmed.ncbi.nlm.nih.gov/22000016/)
48. Wang L, Guo ZY, Zhang R, Xin B, Chen R, Zhao J, et al. Pseudogene OCT4-pg4 functions as a natural micro RNA sponge to regulate OCT4 expression by competing for miR-145 in hepatocellular carcinoma. *Carcinogenesis*. 2013; 34(8):1773–1781. doi: [10.1093/carcin/bgt139](https://doi.org/10.1093/carcin/bgt139) PMID: [23615404](https://pubmed.ncbi.nlm.nih.gov/23615404/)
49. Huarte M. The emerging role of lncRNAs in cancer. *Nature medicine*. 2015; 21(11):1253–1261. doi: [10.1038/nm.3981](https://doi.org/10.1038/nm.3981) PMID: [26540387](https://pubmed.ncbi.nlm.nih.gov/26540387/)
50. Marques AC, Tan J, Lee S, Kong L, Heger A, Ponting CP. Evidence for conserved post-transcriptional roles of unitary pseudogenes and for frequent bifunctionality of mRNAs. *Genome biology*. 2012; 13(11):1. doi: [10.1186/gb-2012-13-11-r102](https://doi.org/10.1186/gb-2012-13-11-r102)
51. Johnsson P, Ackley A, Vidarsdottir L, Lui WO, Corcoran M, Grandér D, et al. A pseudogene long-non-coding-RNA network regulates PTEN transcription and translation in human cells. *Nature structural & molecular biology*. 2013; 20(4):440–446. doi: [10.1038/nsmb.2516](https://doi.org/10.1038/nsmb.2516) PMID: [23435381](https://pubmed.ncbi.nlm.nih.gov/23435381/)
52. Liu Q, Huang J, Zhou N, Zhang Z, Zhang A, Lu Z, et al. LncRNA loc285194 is a p53-regulated tumor suppressor. *Nucleic acids research*. 2013; 41(9):4976–4987. doi: [10.1093/nar/gkt182](https://doi.org/10.1093/nar/gkt182) PMID: [23558749](https://pubmed.ncbi.nlm.nih.gov/23558749/)
53. Yu G, Yao W, Gumireddy K, Li A, Wang J, Xiao W, et al. Pseudogene PTENP1 functions as a competing endogenous RNA to suppress clear-cell renal cell carcinoma progression. *Molecular cancer therapeutics*. 2014; 13(12):3086–3097. doi: [10.1158/1535-7163.MCT-14-0245](https://doi.org/10.1158/1535-7163.MCT-14-0245) PMID: [25249556](https://pubmed.ncbi.nlm.nih.gov/25249556/)
54. Xie J, Guo B, Ding Z, Kang J, Deng X, Wu B, et al. Microarray analysis of lncRNAs and mRNAs co-expression network and lncRNA function as cerna in papillary thyroid carcinoma. *Journal of Biomaterials and Tissue Engineering*. 2015; 5(11):872–880. doi: [10.1166/jbt.2015.1389](https://doi.org/10.1166/jbt.2015.1389)

55. Zhou X, Ye F, Yin C, Zhuang Y, Yue G, Zhang G. The interaction between MiR-141 and lncRNA-H19 in regulating cell proliferation and migration in gastric cancer. *Cellular Physiology and Biochemistry*. 2015; 36(4):1440–1452. doi: [10.1159/000430309](https://doi.org/10.1159/000430309) PMID: [26160158](https://pubmed.ncbi.nlm.nih.gov/26160158/)
56. Zheng L, Li X, Gu Y, Lv X, Xi T. The 3' UTR of the pseudogene CYP4Z2P promotes tumor angiogenesis in breast cancer by acting as a ceRNA for CYP4Z1. *Breast cancer research and treatment*. 2015; 150(1):105–118. doi: [10.1007/s10549-015-3298-2](https://doi.org/10.1007/s10549-015-3298-2) PMID: [25701119](https://pubmed.ncbi.nlm.nih.gov/25701119/)
57. Wang P, Ning S, Zhang Y, Li R, Ye J, Zhao Z, et al. Identification of lncRNA-associated competing triplets reveals global patterns and prognostic markers for cancer. *Nucleic Acids Research*. 2015; 43(7):3478–3489. doi: [10.1093/nar/gkv233](https://doi.org/10.1093/nar/gkv233) PMID: [25800746](https://pubmed.ncbi.nlm.nih.gov/25800746/)
58. Xiao H, Tang K, Liu P, Chen K, Hu J, Zeng J, et al. lncRNA MALAT1 functions as a competing endogenous RNA to regulate ZEB2 expression by sponging miR-200s in clear cell kidney carcinoma. *Oncotarget*. 2015; 6(35):38005–38015. doi: [10.18632/oncotarget.5357](https://doi.org/10.18632/oncotarget.5357) PMID: [26461224](https://pubmed.ncbi.nlm.nih.gov/26461224/)
59. Zhou M, Sun Y, Xu W, Zhang Z, Zhao H, Zhong Z, et al. Comprehensive analysis of lncRNA expression profiles reveals a novel lncRNA signature to discriminate nonequivalent outcomes in patients with ovarian cancer. *Oncotarget*. 2016;.
60. Tan JY, Marques AC. miRNA-mediated crosstalk between transcripts: The missing “linc”? *BioEssays*. 2016; 38(3):295–301. doi: [10.1002/bies.201500148](https://doi.org/10.1002/bies.201500148) PMID: [26835602](https://pubmed.ncbi.nlm.nih.gov/26835602/)
61. Larriba E, del Mazo J. Role of non-coding RNAs in the transgenerational epigenetic transmission of the effects of reprotoxicants. *International Journal of Molecular Sciences*. 2016; 17(4). doi: [10.3390/ijms17040452](https://doi.org/10.3390/ijms17040452) PMID: [27023531](https://pubmed.ncbi.nlm.nih.gov/27023531/)
62. Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nature methods*. 2007; 4(9):721–726. doi: [10.1038/nmeth1079](https://doi.org/10.1038/nmeth1079) PMID: [17694064](https://pubmed.ncbi.nlm.nih.gov/17694064/)
63. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell*. 2011; 146(3):353–358. doi: [10.1016/j.cell.2011.07.014](https://doi.org/10.1016/j.cell.2011.07.014) PMID: [21802130](https://pubmed.ncbi.nlm.nih.gov/21802130/)
64. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature*. 2014; 505(7483):344–352. doi: [10.1038/nature12986](https://doi.org/10.1038/nature12986) PMID: [24429633](https://pubmed.ncbi.nlm.nih.gov/24429633/)
65. Ergun S, Oztuzcu S. Oncocers: ceRNA-mediated cross-talk by sponging miRNAs in oncogenic pathways. *Tumor Biology*. 2015; 36(5):3129–3136. doi: [10.1007/s13277-015-3346-x](https://doi.org/10.1007/s13277-015-3346-x) PMID: [25809705](https://pubmed.ncbi.nlm.nih.gov/25809705/)
66. Qi X, Zhang DH, Wu N, Xiao JH, Wang X, Ma W. ceRNA in cancer: Possible functions and clinical implications. *Journal of Medical Genetics*. 2015; 52(10):710–718. doi: [10.1136/jmedgenet-2015-103334](https://doi.org/10.1136/jmedgenet-2015-103334) PMID: [26358722](https://pubmed.ncbi.nlm.nih.gov/26358722/)
67. Kagami H, Akutsu T, Maegawa S, Hosokawa H, Nacher JC. Determining associations between human diseases and non-coding RNAs with critical roles in network control. *Scientific Reports*. 2015; 5. doi: [10.1038/srep14577](https://doi.org/10.1038/srep14577)
68. Guo LL, Song CH, Wang P, Dai LP, Zhang JY, Wang KJ. Competing endogenous RNA networks and gastric cancer. *World Journal of Gastroenterology*. 2015; 21(41):11680–11687. doi: [10.3748/wjg.v21.i41.11680](https://doi.org/10.3748/wjg.v21.i41.11680) PMID: [26556995](https://pubmed.ncbi.nlm.nih.gov/26556995/)
69. Yang C, Wu D, Gao L, Liu X, Jin Y, Wang D, et al. Competing endogenous RNA networks in human cancer: Hypothesis, validation, and perspectives. *Oncotarget*. 2016; 7(12):13479–13490. doi: [10.18632/oncotarget.7266](https://doi.org/10.18632/oncotarget.7266) PMID: [26872371](https://pubmed.ncbi.nlm.nih.gov/26872371/)
70. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nature Reviews Genetics*. 2016; 17(5):272–283. doi: [10.1038/nrg.2016.20](https://doi.org/10.1038/nrg.2016.20) PMID: [27040487](https://pubmed.ncbi.nlm.nih.gov/27040487/)
71. Franco-Zorrilla JM, Valli A, Todesco M, Mateos I, Puga MI, Rubio-Somoza I, et al. Target mimicry provides a new mechanism for regulation of microRNA activity. *Nature genetics*. 2007; 39(8):1033–1037. doi: [10.1038/ng2079](https://doi.org/10.1038/ng2079) PMID: [17643101](https://pubmed.ncbi.nlm.nih.gov/17643101/)
72. Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*. 2010; 465(7301):1033–8. doi: [10.1038/nature09144](https://doi.org/10.1038/nature09144) PMID: [20577206](https://pubmed.ncbi.nlm.nih.gov/20577206/)
73. Harrison PM, Zheng D, Zhang Z, Carriero N, Gerstein M. Transcribed processed pseudogenes in the human genome: an intermediate form of expressed retrosequence lacking protein-coding ability. *Nucleic Acids Res*. 2005; 33(8):2374–83. doi: [10.1093/nar/gki531](https://doi.org/10.1093/nar/gki531) PMID: [15860774](https://pubmed.ncbi.nlm.nih.gov/15860774/)
74. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*. 2011; 147(2):358–369. doi: [10.1016/j.cell.2011.09.028](https://doi.org/10.1016/j.cell.2011.09.028) PMID: [22000014](https://pubmed.ncbi.nlm.nih.gov/22000014/)
75. Wang Y, Xu Z, Jiang J, Xu C, Kang J, Xiao L, et al. Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Developmental cell*. 2013; 25(1):69–80. doi: [10.1016/j.devcel.2013.03.002](https://doi.org/10.1016/j.devcel.2013.03.002) PMID: [23541921](https://pubmed.ncbi.nlm.nih.gov/23541921/)

76. Fan M, Li X, Jiang W, Huang Y, Li J, Wang Z. A long non-coding RNA, PTCSC3, as a tumor suppressor and a target of miRNAs in thyroid cancer cells. *Experimental and therapeutic medicine*. 2013; 5(4):1143–1146. doi: [10.3892/etm.2013.933](https://doi.org/10.3892/etm.2013.933) PMID: [23599737](https://pubmed.ncbi.nlm.nih.gov/23599737/)
77. Kallen AN, Zhou XB, Xu J, Qiao C, Ma J, Yan L, et al. The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Molecular cell*. 2013; 52(1):101–112. doi: [10.1016/j.molcel.2013.08.027](https://doi.org/10.1016/j.molcel.2013.08.027) PMID: [24055342](https://pubmed.ncbi.nlm.nih.gov/24055342/)
78. Jeck WR, Sharpless NE. Detecting and characterizing circular RNAs. *Nature biotechnology*. 2014; 32(5):453. doi: [10.1038/nbt.2890](https://doi.org/10.1038/nbt.2890) PMID: [24811520](https://pubmed.ncbi.nlm.nih.gov/24811520/)
79. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*. 2013; 495(7441):333–338. doi: [10.1038/nature11928](https://doi.org/10.1038/nature11928) PMID: [23446348](https://pubmed.ncbi.nlm.nih.gov/23446348/)
80. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, et al. Natural RNA circles function as efficient microRNA sponges. *Nature*. 2013; 495(7441):384–388. doi: [10.1038/nature11993](https://doi.org/10.1038/nature11993) PMID: [23446346](https://pubmed.ncbi.nlm.nih.gov/23446346/)
81. Zlotorynski E. Non-coding RNA: Circular RNAs promote transcription. *Nature Reviews Molecular Cell Biology*. 2015;. doi: [10.1038/nrm3967](https://doi.org/10.1038/nrm3967) PMID: [25714680](https://pubmed.ncbi.nlm.nih.gov/25714680/)
82. Rybak-Wolf A, Stottmeister C, Glažar P, Jens M, Pino N, Giusti S, et al. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. *Molecular cell*. 2015; 58(5):870–885. doi: [10.1016/j.molcel.2015.03.027](https://doi.org/10.1016/j.molcel.2015.03.027) PMID: [25921068](https://pubmed.ncbi.nlm.nih.gov/25921068/)
83. Memczak S, Papavasileiou P, Peters O, Rajewsky N. Identification and characterization of circular RNAs as a new class of putative biomarkers in human blood. *PLoS one*. 2015; 10(10):e0141214. doi: [10.1371/journal.pone.0141214](https://doi.org/10.1371/journal.pone.0141214) PMID: [26485708](https://pubmed.ncbi.nlm.nih.gov/26485708/)
84. Capel B, Swain A, Nicolis S, Hacker A, Walter M, Koopman P, et al. Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell*. 1993; 73(5):1019–1030. doi: [10.1016/0092-8674\(93\)90279-Y](https://doi.org/10.1016/0092-8674(93)90279-Y) PMID: [7684656](https://pubmed.ncbi.nlm.nih.gov/7684656/)
85. Hansen TB, Kjems J, Damgaard CK. Circular RNA and miR-7 in cancer. *Cancer research*. 2013; 73(18):5609–5612. doi: [10.1158/0008-5472.CAN-13-1568](https://doi.org/10.1158/0008-5472.CAN-13-1568) PMID: [24014594](https://pubmed.ncbi.nlm.nih.gov/24014594/)
86. Paci P, Colombo T, Farina L. Computational analysis identifies a sponge interaction network between long non-coding RNAs and messenger RNAs in human breast cancer. *BMC Syst Biol*. 2014; 8:83. doi: [10.1186/1752-0509-8-83](https://doi.org/10.1186/1752-0509-8-83) PMID: [25033876](https://pubmed.ncbi.nlm.nih.gov/25033876/)
87. Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*. 2013; 45(10):1113–20. doi: [10.1038/ng.2764](https://doi.org/10.1038/ng.2764) PMID: [24071849](https://pubmed.ncbi.nlm.nih.gov/24071849/)
88. Tomczak K, Czerwinska P, Wiznerowicz M, et al. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)*. 2015; 19(1A):A68–A77.
89. Lemay G, Jolicoeur P. Rearrangement of a DNA sequence homologous to a cell-virus junction fragment in several Moloney murine leukemia virus-induced rat thymomas. *Proceedings of the National Academy of Sciences*. 1984; 81(1):38–42. doi: [10.1073/pnas.81.1.38](https://doi.org/10.1073/pnas.81.1.38)
90. Graham M, Adams JM, Cory S. Murine T lymphomas with retroviral inserts in the chromosomal 15 locus for plasmacytoma variant translocations. *Nature*. 1984; 314(6013):740–743. doi: [10.1038/314740a0](https://doi.org/10.1038/314740a0)
91. Villeneuve L, Rassart E, Jolicoeur P, Graham M, Adams J. Proviral integration site Mis-1 in rat thymomas corresponds to the pvt-1 translocation breakpoint in murine plasmacytomas. *Molecular and cellular biology*. 1986; 6(5):1834–1837. doi: [10.1128/MCB.6.5.1834](https://doi.org/10.1128/MCB.6.5.1834) PMID: [3785181](https://pubmed.ncbi.nlm.nih.gov/3785181/)
92. Graham M, Adams JM. Chromosome 8 breakpoint far 3' of the c-myc oncogene in a Burkitt's lymphoma 2; 8 variant translocation is equivalent to the murine pvt-1 locus. *The EMBO journal*. 1986; 5(11):2845. PMID: [3024964](https://pubmed.ncbi.nlm.nih.gov/3024964/)
93. Huppi K, Siwarski D. Chimeric transcripts with an open reading frame are generated as a result of translocation to the Pvt-1 region in mouse B-cell tumors. *International journal of cancer*. 1994; 59(6):848–851. doi: [10.1002/ijc.2910590623](https://doi.org/10.1002/ijc.2910590623) PMID: [7989128](https://pubmed.ncbi.nlm.nih.gov/7989128/)
94. Hodgson G, Hager JH, Volik S, Hariono S, Wernick M, Moore D, et al. Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. *Nature genetics*. 2001; 29(4):459–464. doi: [10.1038/ng771](https://doi.org/10.1038/ng771) PMID: [11694878](https://pubmed.ncbi.nlm.nih.gov/11694878/)
95. Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbelt JO, et al. What is a gene, post-ENCODE? History and updated definition. *Genome research*. 2007; 17(6):669–681. doi: [10.1101/gr.6339607](https://doi.org/10.1101/gr.6339607) PMID: [17567988](https://pubmed.ncbi.nlm.nih.gov/17567988/)
96. Colombo T, Farina L, Macino G, Paci P. PVT1: a rising star among oncogenic long noncoding RNAs. *Biomed Res Int*. 2015; 2015:304208. doi: [10.1155/2015/304208](https://doi.org/10.1155/2015/304208) PMID: [25883951](https://pubmed.ncbi.nlm.nih.gov/25883951/)

97. Huppi K, Siwarski D, Skurla R, Klinman D, Mushinski J. Pvt-1 transcripts are found in normal tissues and are altered by reciprocal (6; 15) translocations in mouse plasmacytomas. *Proceedings of the National Academy of Sciences*. 1990; 87(18):6964–6968. doi: [10.1073/pnas.87.18.6964](https://doi.org/10.1073/pnas.87.18.6964) PMID: [2402486](https://pubmed.ncbi.nlm.nih.gov/2402486/)
98. Guan Y, Kuo WL, Stilwell JL, Takano H, Lapuk AV, Fridlyand J, et al. Amplification of PVT1 contributes to the pathophysiology of ovarian and breast cancer. *Clinical Cancer Research*. 2007; 13(19):5745–5755. doi: [10.1158/1078-0432.CCR-06-2882](https://doi.org/10.1158/1078-0432.CCR-06-2882) PMID: [17908964](https://pubmed.ncbi.nlm.nih.gov/17908964/)
99. Tseng YY, Moriarity BS, Gong W, Akiyama R, Tiwari A, Kawakami H, et al. PVT1 dependence in cancer with MYC copy-number increase. *Nature*. 2014;. doi: [10.1038/nature13311](https://doi.org/10.1038/nature13311) PMID: [25043044](https://pubmed.ncbi.nlm.nih.gov/25043044/)
100. Brooksbank C, Bergman MT, Apweiler R, Birney E, Thornton J. The european bioinformatics institute's data resources 2014. *Nucleic acids research*. 2014; 42(D1):D18–D25. doi: [10.1093/nar/gkt1206](https://doi.org/10.1093/nar/gkt1206) PMID: [24271396](https://pubmed.ncbi.nlm.nih.gov/24271396/)
101. Huppi K, Volfovsky N, Runfola T, Jones TL, Mackiewicz M, Martin SE, et al. The identification of microRNAs in a genomically unstable region of human chromosome 8q24. *Molecular Cancer Research*. 2008; 6(2):212–221. doi: [10.1158/1541-7786.MCR-07-0105](https://doi.org/10.1158/1541-7786.MCR-07-0105) PMID: [18314482](https://pubmed.ncbi.nlm.nih.gov/18314482/)
102. Meyer KB, Maia AT, O'Reilly M, Ghousaini M, Prathalingam R, Porter-Gill P, et al. A functional variant at a prostate cancer predisposition locus at 8q24 is associated with PVT1 expression. *PLoS Genet*. 2011; 7(7):e1002165. doi: [10.1371/journal.pgen.1002165](https://doi.org/10.1371/journal.pgen.1002165) PMID: [21814516](https://pubmed.ncbi.nlm.nih.gov/21814516/)
103. Chapman MH, Tidswell R, Dooley JS, Sandanayake NS, Cerec V, Deheragoda M, et al. Whole genome RNA expression profiling of endoscopic biliary brushings provides data suitable for biomarker discovery in cholangiocarcinoma. *Journal of hepatology*. 2012; 56(4):877–885. doi: [10.1016/j.jhep.2011.10.022](https://doi.org/10.1016/j.jhep.2011.10.022) PMID: [22173169](https://pubmed.ncbi.nlm.nih.gov/22173169/)
104. Wang F, Yuan JH, Wang SB, Yang F, Yuan SX, Ye C, et al. Oncofetal long noncoding RNA PVT1 promotes proliferation and stem cell-like property of hepatocellular carcinoma cells by stabilizing NOP2. *Hepatology*. 2014; 60(4):1278–1290. doi: [10.1002/hep.27239](https://doi.org/10.1002/hep.27239) PMID: [25043274](https://pubmed.ncbi.nlm.nih.gov/25043274/)
105. Zhuang C, Li J, Liu Y, Chen M, Yuan J, Fu X, et al. Tetracycline-inducible shRNA targeting long non-coding RNA PVT1 inhibits cell growth and induces apoptosis in bladder cancer cells. *Oncotarget*. 2015; 6(38):41194–41203. doi: [10.18632/oncotarget.5880](https://doi.org/10.18632/oncotarget.5880) PMID: [26517688](https://pubmed.ncbi.nlm.nih.gov/26517688/)
106. Zhou Q, Chen J, Feng J, Wang J. Long noncoding RNA PVT1 modulates thyroid cancer cell proliferation by recruiting EZH2 and regulating thyroid-stimulating hormone receptor (TSHR). *Tumor Biology*. 2016; 37(3):3105–3113. doi: [10.1007/s13277-015-4149-9](https://doi.org/10.1007/s13277-015-4149-9) PMID: [26427660](https://pubmed.ncbi.nlm.nih.gov/26427660/)
107. Cui D, Yu CH, Liu M, Xia QQ, Zhang YF, Jiang WL. Long non-coding RNA PVT1 as a novel biomarker for diagnosis and prognosis of non-small cell lung cancer. *Tumor Biology*. 2016; 37(3):4127–4134. doi: [10.1007/s13277-015-4261-x](https://doi.org/10.1007/s13277-015-4261-x) PMID: [26490983](https://pubmed.ncbi.nlm.nih.gov/26490983/)
108. Haverty PM, Hon LS, Kaminker JS, Chant J, Zhang Z. High-resolution analysis of copy number alterations and associated expression changes in ovarian tumors. *BMC medical genomics*. 2009; 2(1):21. doi: [10.1186/1755-8794-2-21](https://doi.org/10.1186/1755-8794-2-21) PMID: [19419571](https://pubmed.ncbi.nlm.nih.gov/19419571/)
109. Barsotti AM, Beckerman R, Laptenko O, Huppi K, Caplen NJ, Prives C. p53-Dependent induction of PVT1 and miR-1204. *Journal of Biological Chemistry*. 2012; 287(4):2509–2519. doi: [10.1074/jbc.M111.322875](https://doi.org/10.1074/jbc.M111.322875) PMID: [22110125](https://pubmed.ncbi.nlm.nih.gov/22110125/)
110. Hamilton MJ, Young MD, Sauer S, Martinez E. The interplay of long non-coding RNAs and MYC in cancer. *AIMS biophysics*. 2015; 2(4):794. doi: [10.3934/biophy.2015.4.794](https://doi.org/10.3934/biophy.2015.4.794) PMID: [27077133](https://pubmed.ncbi.nlm.nih.gov/27077133/)
111. Sarver AL, Murray CD, Temiz NA, Tseng YY, Bagchi A. MYC and PVT1 synergize to regulate RSP01 levels in breast cancer. *Cell Cycle*. 2016; 15(7):881–885. doi: [10.1080/15384101.2016.1149660](https://doi.org/10.1080/15384101.2016.1149660) PMID: [26889781](https://pubmed.ncbi.nlm.nih.gov/26889781/)
112. Posa I, Carvalho S, Tavares J, Grosso AR. A pan-cancer analysis of MYC-PVT1 reveals CNV-unmediated deregulation and poor prognosis in renal carcinoma. *Oncotarget*. 2016;. doi: [10.18632/oncotarget.9487](https://doi.org/10.18632/oncotarget.9487) PMID: [27366943](https://pubmed.ncbi.nlm.nih.gov/27366943/)
113. Lewis BP, Shih Ih, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell*. 2003; 115(7):787–798. doi: [10.1016/S0092-8674\(03\)01018-3](https://doi.org/10.1016/S0092-8674(03)01018-3) PMID: [14697198](https://pubmed.ncbi.nlm.nih.gov/14697198/)
114. Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, et al. Clustering and conservation patterns of human microRNAs. *Nucleic acids research*. 2005; 33(8):2697–2706. doi: [10.1093/nar/gki567](https://doi.org/10.1093/nar/gki567) PMID: [15891114](https://pubmed.ncbi.nlm.nih.gov/15891114/)
115. Humphries B, Yang C. The microRNA-200 family: small molecules with novel roles in cancer development, progression and therapy. *Oncotarget*. 2015; 6(9):6472. doi: [10.18632/oncotarget.3052](https://doi.org/10.18632/oncotarget.3052) PMID: [25762624](https://pubmed.ncbi.nlm.nih.gov/25762624/)
116. Korpál M, Lee ES, Hu G, Kang Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2.

Journal of Biological Chemistry. 2008; 283(22):14910–14914. doi: [10.1074/jbc.C800074200](https://doi.org/10.1074/jbc.C800074200) PMID: [18411277](https://pubmed.ncbi.nlm.nih.gov/18411277/)

117. Mongroo PS, Rustgi AK. The role of the miR-200 family in epithelial-mesenchymal transition. *Cancer biology & therapy*. 2010; 10(3):219–222. doi: [10.4161/cbt.10.3.12548](https://doi.org/10.4161/cbt.10.3.12548)
118. Cong N, Du P, Zhang A, Shen F, Su J, Pu P, et al. Downregulated microRNA-200a promotes EMT and tumor growth through the wnt/ β -catenin pathway by targeting the E-cadherin repressors ZEB1/ZEB2 in gastric adenocarcinoma. *Oncology reports*. 2013; 29(4):1579–1587. PMID: [23381389](https://pubmed.ncbi.nlm.nih.gov/23381389/)
119. Park SM, Gaur AB, Lengyel E, Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes & development*. 2008; 22(7):894–907. doi: [10.1101/gad.1640608](https://doi.org/10.1101/gad.1640608)
120. NCBI;. ftp://ftp.ncbi.nih.gov/gene/DATA/GENE_INFO/Mammalia/Homo_sapiens.gene_info.gz.
121. Grillo G, Turi A, Licciulli F, Mignone F, Liuni S, Banfi S, et al. UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research*. 2010; 38(suppl 1):D75–D80. doi: [10.1093/nar/gkp902](https://doi.org/10.1093/nar/gkp902) PMID: [19880380](https://pubmed.ncbi.nlm.nih.gov/19880380/)
122. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*. 2005; 120(1):15–20. doi: [10.1016/j.cell.2004.12.035](https://doi.org/10.1016/j.cell.2004.12.035) PMID: [15652477](https://pubmed.ncbi.nlm.nih.gov/15652477/)
123. mirBase—release18;. www.mirBase.org.
124. Ensembl;. <http://www.ensembl.org/>.
125. BioConductor;. <http://www.bioconductor.org>.
126. NCBI—Entrez gene;. <http://www.ncbi.nlm.nih.gov/gene>.
127. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, Van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature biotechnology*. 2010; 28(5):511–515. doi: [10.1038/nbt.1621](https://doi.org/10.1038/nbt.1621) PMID: [20436464](https://pubmed.ncbi.nlm.nih.gov/20436464/)
128. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*. 2011; 29(7):644–652. doi: [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883) PMID: [21572440](https://pubmed.ncbi.nlm.nih.gov/21572440/)
129. Jolliffe I. *Principal component analysis*. Wiley Online Library; 2002.
130. Jackson JE. *A user's guide to principal components*. vol. 587. John Wiley & Sons; 2005.