BBA - Reviews on Cancer 1868 (2017) 449–455

Contents lists available at ScienceDirect

BBA - Reviews on Cancer

journal homepage: www.elsevier.com/locate/bbacan

Review

# The role of a new class of long noncoding RNAs transcribed from ultraconserved regions in cancer

Daniela Terracciano[a,1], Sara Terreri[b,1], Filomena de Nigris[c], Valerio Costa[b], George A. Calin[d], Amelia Cimmino[b,*]

[a] Dept. of Translational Medical Sciences, University of Naples "Federico II", Naples, Italy
[b] Institute of Genetics and Biophysics "A. Buzzati Traverso", National Research Council (CNR), Naples, Italy
[c] Department of Biochemistry, Biophysics and General Pathology, University of Campania "Luigi Vanvitelli", Naples, Italy
[d] Departments of Experimental Therapeutics and Leukemia, and the Center for small interfering RNA and non-coding RNAs, University of Texas MD Anderson Cancer Center, Houston, TX, USA

ARTICLE INFO

*Keywords:*
lncRNAs
miRNAs
T-UCRs
Cancer

ABSTRACT

Ultraconserved regions (UCRs) represent a relatively new class of non-coding genomic sequences highly conserved between human, rat and mouse genomes. These regions can reside within exons of protein-coding genes, despite the vast majority of them localizes within introns or intergenic regions. Several studies have undoubtedly demonstrated that most of these regions are actively transcribed in normal cells/tissues, where they contribute to regulate many cellular processes. Interestingly, these non-coding RNAs exhibit aberrant expression levels in human cancer cells and their expression profiles have been used as prognostic factors in human malignancies, as well as to unambiguously distinguish among distinct cancer types. In this review, we first describe their identification, then we provide some updated information about their genomic localization and classification. More importantly, we discuss about the available literature describing an overview of the mechanisms through which some transcribed UCRs (T-UCR) contribute to cancer progression or to the metastatic spread. To date, the interplay between T-UCRs and microRNAs is the most convincing evidence linking T-UCRs and tumorigenesis. The limitations of these studies and the future challenges to be addressed in order to understand the biological role of T-UCRs are also discussed herein. We envision that future efforts are needed to convincingly include this class of ncRNAs in the growing area of cancer therapeutics.

## 1. The ultraconserved regions: definition

Ultraconserved regions (UCRs), otherwise known as ultraconserved elements (UCEs), are highly conserved sequences of human genome that are maximally conserved with orthologous segments in rodents. These regions have been computationally identified in 2004 by *Bejerano* and colleagues as the longest segments of the human genome with the maximal conservation rate (100% identity without indels and excluding rRNAs) in human, mouse and rat genomes. By using this approach, the authors could systematically define genomic segments longer than 200 bp represented in single copy in the haploid genome, widely distributed on all chromosomes except on chromosomes 21 and Y [1]. Interestingly, despite the relatively low percentage of identity between the human and chicken genome, UCRs have extremely high levels of conservation even in this specie. High percentage of identity was also described in dog and fugu genomes, strongly indicating that most of these genomic elements have undergone extreme negative selection in many species [1]. Most of these segments do not overlap with any coding region and the distribution of these ultraconserved genetic elements within the genome is not random. Indeed, they mostly appear in clusters, more often than expected by chance, near coding regions for transcription factors and molecules involved in developmental processes. These observations have suggested the hypothesis that UCRs are candidate regulatory elements with a relevant role in the early stages of vertebrate development, differentiation and coordination between cells. In addition, they are also frequently located at both fragile sites and genomic regions involved in cancer. These functions have been experimentally confirmed for a number of elements [2,3]. Moreover, SNPs falling within these regions are extremely under-represented (by 20-fold) across human populations, as well as in primates [4], suggesting that these UCRs are under high positive selection in these genomes. Surprisingly, the generation of transgenic mice with the

* Corresponding author at: Institute of Genetics and Biophysics "Adriano Buzzati Traverso", CNR, Via Pietro Castellino 111, 80131 Naples, Italy.
*E-mail address:* amelia.cimmino@igb.cnr.it (A. Cimmino).

[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.bbcan.2017.09.001
Received 9 June 2017; Received in revised form 7 September 2017; Accepted 9 September 2017
Available online 13 September 2017
0304-419X/ © 2017 Elsevier B.V. All rights reserved.

deletion of each of the four elements contained in a large cluster of UCE, uc.248 +, uc.329 +, uc.467 +, and uc.482 +, has no apparent effect on the phenotype [5]. Other studies have indicated that some UCRs seem to be "phenotypically redundant" [6,7]. Therefore, it has been hypothesized that their deletion can lead to visible phenotypes only over many generations. These studies have clearly highlighted that the functional conservation cannot be inferred solely from sequence identity and vice versa [8,9].

## 2. UCR classification

The ultraconserved regions of the human genome were first classified according to their genomic localisation in exonic, non-exonic and "possibly" exonic. Later, they were re-annotated according a more recent human genome assembly (i.e. hg18) and then re-classified into 5 subgroups: intergenic, intronic, exonic, partly exonic, and exon containing [8]. In 2007 *Calin* and collaborators demonstrated, by using a custom microarray, that a large fraction of these non-coding UCRs are transcribed [10]. As transcriptome is routinely updated by large-scale studies of international consortia, the genomic annotation of some T-UCRs varied because of the presence of different splice variants for a given gene. Indeed, as different isoforms may be generated from the transcription of a given gene locus, each UCR may be labelled as "exonic" in one transcript or "intronic" in another arising from the same gene [11]. The UCRs are commonly defined as "multiple". Since new updates of the human genome are released, here we provide a schematic re-annotation of all 481 UCRs according to the last available version of the human genome and of GENCODE transcriptome assembly that, at the time of writing, are hg38 and GENCODE v26. Fig. 1 schematically shows the results of this analysis. In brief, 135 and 181 UCRs overlap with exons and introns, respectively, of annotated human transcripts (both protein coding and not). The remaining 165 fall within intergenic regions with variable distance from annotated gene loci, ranging from 330 bp for the uc.335 + and LMO3 gene to ~1.4 Mb for the uc.81 + and ACVR2A gene.

## 3. Transcribed-UCRs

These transcripts containing UCRs have been named transcribed UCRs or T-UCRs, as well as ultraconserved genes (uc.) [10]. According to the genomic orientations they can be sense (same orientation) or antisense (opposite orientation). This pioneer paper by *Calin* and colleagues also described an important feature of T-UCRs, i.e. the high tissue-specific expression. Indeed, they described that the highest number of T-UCRs are expressed in B cells, while the lowest levels are described in ovary. Noteworthy, *Calin* and colleagues also found that the expression of this new class of transcribed sequences was altered in several human cancers. In particular, by using a high throughput microarray technology they investigated the expression of T-UCRs in a panel of 173 samples: 133 specimens were from human cancers [e.g., chronic lymphocytic leukaemia (CLL), colorectal (CRC), and hepatocellular carcinomas (HCC)] and 40 from normal tissues. They also reported that T-UCRs have distinct expression signatures in human leukaemia and carcinomas. The above results provided the first link among this class of genomic elements and cancer [11]. Following the initial report that systematically measured T-UCRs' expression in B-cell chronic lymphocytic leukaemia [3], other groups analysed T-UCRs levels in other human malignancies. Interestingly, it has been shown that T-UCRs expression profiles can be successfully used to distinguish among different cancer types, indicating them as new relevant tumour biomarkers [12]. Rapid amplification of cDNA ends (RACE) methodology has been used for some T-UCRs to define the full-length transcripts [10,11,13–15]. Although UCRs' genomic length ranges from 200 to 779 bp, their full-length transcriptional units usually overcome 2 kb [10,11,13–15]. In this review, we summarize the current knowledge about T-UCRs expression in human diseases. Improved understanding of T-UCRs in cancer will ultimately lead to design new strategies for prevention and/or early detection, also providing relevant information to improve the therapeutic approaches currently available in cancer treatment. T-UCRs are reported in literature under different names used by different groups, on the basis of probe spotted on chip microarray. In particular, in this review we will always use the abbreviation uc. (named '+') when the probes were designed according to the sense of the genomic sequence, indeed we will use the name uc. A ("+ A") when the probes are antisense to the genomic DNA sequence. Therefore, each genomic UCR can have two types of transcripts, sense and antisense.

## 4. Associations of T-UCRs with cancer

Multiple recent reports have described that altered expression of T-UCRs is often associated with specific tumour types. These studies have largely highlighted a contribution of T-UCRs in carcinogenesis [16]. Table 1 summarizes the main findings on T-UCRs in different tumours. Among them, *Braconi* and colleagues, by using a genome-wide expression profile analysis, identified 56 T-UCRs aberrantly expressed in malignant HepG2 cells compared with non-malignant human hepatocytes. Lastly, by using in situ hybridization assay they analysed 221 HCC samples in two tissue microarrays, reporting uc.338 + as the most up-regulated UCR in HCC compared to the normal counterpart. They also showed that its knockdown inhibits both anchorage-dependent and -independent cell growth, suggesting that the high levels of this T-UCR measured in tumour samples, affect the malignant phenotype [13]. *Wang* and collaborators have also described the same T-UCR, as the most up regulated in CRC and CRC cell lines. They associated its expression with lymph node metastasis. Small interfering RNA (siRNA)-mediated knockdown of uc.338 + markedly inhibited cell migration and invasion in the microsatellite stable SW480 and in the microsatellite instable HCT116 CRC cell lines [17], supporting a role for this long non-coding RNA in the metastatic process. In addition, the uc.73 + A was described in CRC as one of the most frequently up regulated. *Calin* and co-workers investigated the effects of uc.73 + A down regulation in COLO-320 colorectal cancer cells, using SW620 colon cancer cells—in which uc.73 + A expression levels are similar to those of normal colon cells—as a control [10]. They first described that the two-above-mentioned different colon cancer cell lines, display different uc.73 + A levels. Interestingly, its siRNA-mediated silencing significantly reduced COLO-320 cells' growth, whereas in SW620 cells it did not, as these cells with very low level of this T-UCR were not addicted to its oncogenic effects. Cell cycle analysis revealed an increased fraction of COLO-320 cells in sub-G1 phase, suggesting the induction of apoptosis in silenced cells. Annexin V and caspase 3 assays confirmed that uc.73 + A knockdown was capable to induce apoptosis in cancer cells, suggesting that high expression of this T-UCR associates with colon cancer tumorigenesis [10]. *Mestdagh* and colleagues [8] found that uc.73 + is involved in *TP53* response pathway in neuroblastoma cell line, further supporting its effect on apoptosis. *Sana* and collaborators showed that the expression of both uc.73 + and uc.388 + were down regulated in CRC (54 samples) compared to normal tissue (15 samples), suggesting their potential role as tumour-suppressors [18]. Besides uc.73 + indicated a positive correlation with overall survival, no significant associations of uc.73 + and uc.388 + with clinical stage, grade, and tumour diameter were discovered. These contrasting results with previous study [10] may be due to the different methods used to detect the expression of T-UCRs in CRC samples. In the former, the authors used hybridization microarrays containing probes that detected UCRs transcripts (in sense or antisense orientation), while in the latter study a standard real-time PCR approach was employed, without providing strand specificity to the analysis.

More recently a gene expression profiling for T-UCRs has been performed in human prostate cancer. In particular, analysing 57 tumours and 7 non-cancerous prostate tissues many of the UCRs transcripts were significantly differentially expressed in tumour vs non-

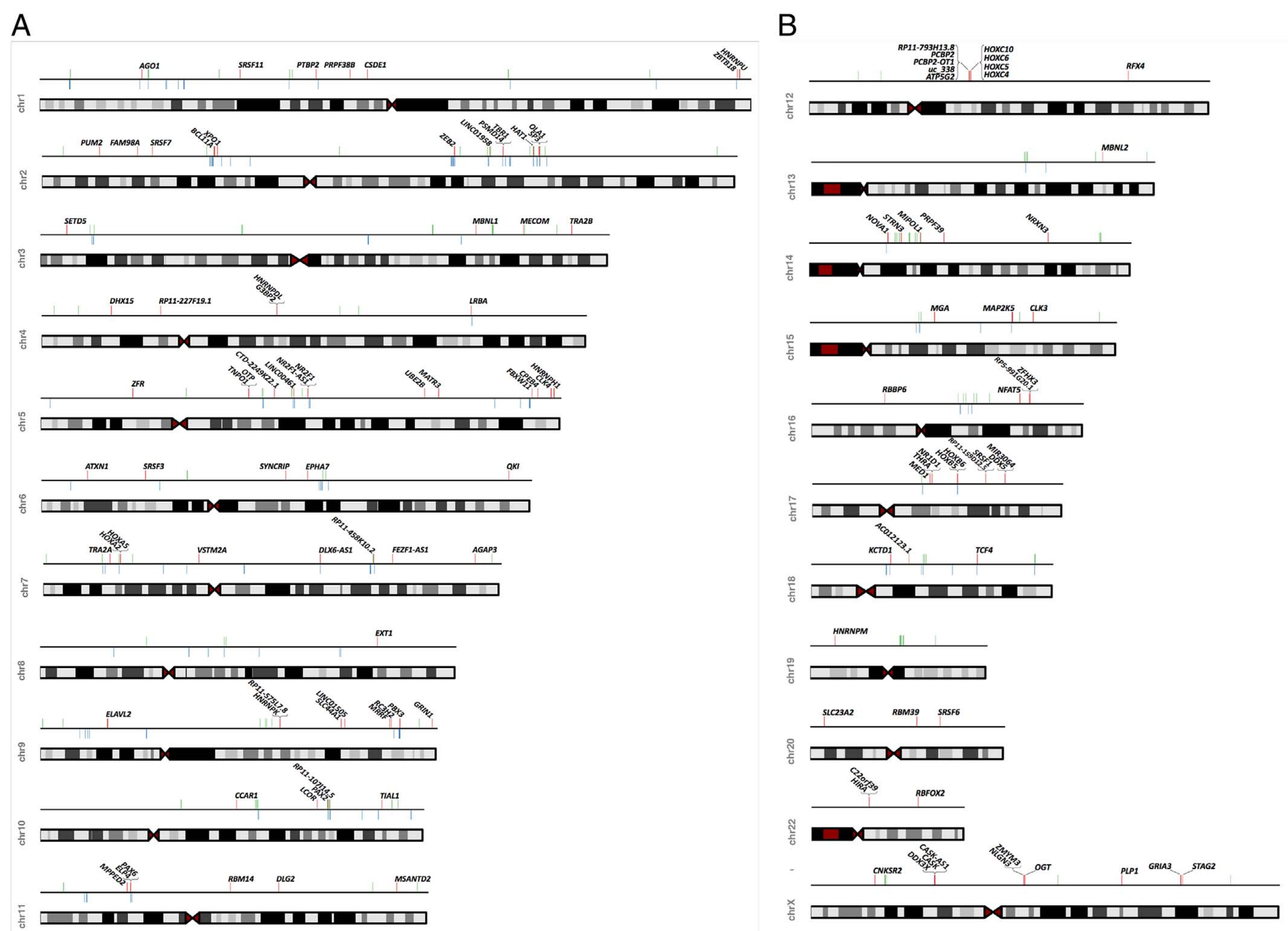


**Fig. 1.** A–B. Schematic representation of the chromosomal localization of the 481-ultraconserved regions (UCRs) on ideograms of the 24 human chromosomes. UCRs genomic coordinates were first mapped on the latest release of the human genome (hg38) and then intersected with human transcriptome (GENCODE v26 release) using BedTools. Data were plotted along chromosomal ideograms using IdeoViz package in R language. Exon-overlapping UCRs are reported as red marks extending above the chromosome with the corresponding official gene symbol; intronic elements are reported as blue marks below the chromosome; intergenic UCRs as green marks above the chromosomes. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tumour tissues: the uc.106 + was found up-regulated in tumours, but tended to be down-regulated in tumours with high Gleason grade and in patients with extra prostatic extension of the disease (EPE) vs patients without EPE [19]. In addition, uc.106 + knockdown in the prostate cancer cell line LNCaP leads to the deregulation of genes involved in proliferation and migration, indicating that uc.106 + may influence cancer progression [19].

In 2010, *Mestdagh* and colleagues [8] demonstrated that T-UCRs expression is correlated with bad prognosis and MYCN amplification in neuroblastoma. In particular, uc.347 +, uc.350 +, uc.279 +, uc.460 +, uc.379 +, uc.446 +, and uc.364 + are over-expressed in MYCN-amplified (MNA) tumours compared to MYCN-non-amplified (MNN) ones, moreover the amplification of MYCN is associated with poor clinical outcome in patients [8]. Accordingly, up-regulated T-UCRs were found to be associated with several cancer-related pathways, such as those one involved in DNA damage, cell cycle regulation, differentiation and immune response.

The expression of T-UCRs is altered even in several hematological malignancies. In particular, a signature of five T-UCRs, uc.269 + A, uc.160 +, uc.215 +, uc.346 + A, and uc.348 +, was able to specifically distinguish between two main subgroups of CLL patients with different prognosis [10], previously classified according to the expression of 70-kDa zeta-associated protein (ZAP-70), one of the several prognostic markers for patients with CLL [20]. A T-UCR signature has also been associated with histological lesions involved in the morphogenesis of Barrett's adenocarcinoma by *Fassan* and collaborators through T-UCRs microarray. Progression from normal to Barrett's oesophagus (BE) and further to adenocarcinoma has been associated with the over-expression of uc.58 + A, uc.202 + A, uc.207 + A, and uc.223 + A and the down-regulation of uc.214 + [21]. A 9 T-UCRs signature, described in Table I, allowed researchers to distinguish BE from squamous epithelium (Sq). This finding is particularly relevant since very little is known about the molecular alterations occurring during the transition from benign squamous oesophageal epithelium to metaplastic Barrett's mucosa and then to neoplastic transformation. In 2013, *Watters* and co-workers identified an expressional alteration in specific T-UCRs in response to all-trans retinoic acid-induced (ATRA) differentiation in neuroblastoma cells [22]. In particular, a gene expression microarray analysis following knockdown of uc.300A +, revealed a number of genes whose expression was altered by changing the lncRNA levels and that might play a role in the increased proliferation and invasion of neuroblastoma cells prior to all-trans-retinoic acid (ATRA)-treatment. The authors demonstrated that the function of uc.300A + is connected with proliferation, invasion and the inhibition of differentiation of neuroblastoma cell lines prior to ATRA treatment [22]. Using a custom microarray, which was used previously to examine T-UCR expression profiles in human leukaemia, colon and hepatocellular cancers [3,10,23], *Olivieri* and colleagues recently identified the differential

**Table 1**
Main findings on T-UCRs in different tumours. a

| T-UCR name | Genomic localization | Cancer type | T-UCR expression | Biological or molecular function in cancer | Reference |
|---|---|---|---|---|---|
| uc.8+ | Chr.1p36.22 | BlCa | Up regulated | KD inhibits cell migration, invasion and proliferation in J82 cell line, promoting bladder cancer tumorigenesis | 11 |
| uc.73+ | Chr.2q22.3 | CRC | Down regulated | Associated with poor survival of CRC patients | 18 |
| uc.73+ | Chr.2q22.3 | NB | Up regulated | TP53 response pathway: apoptotic effect | 8 |
| uc.73+A | Chr.2q22.3 | CRC | Up regulated | KD in COLO-320 cell line, increases sub-G1 phase cells, indicating a role in apoptosis | 10 |
| uc.106+ | Chr.2q31.1 | PCa | Up regulated | KD leads to a down regulation of genes involved in proliferation and migration | 19 |
| uc.300+A | Chr.10q24.31 | NB | Up regulated | Involved in the proliferation, invasion and inhibition of differentiation of neuroblastoma cell lines prior to ATRA treatment. | 22 |
| uc.338+ | Chr.12q13.13 | HCC | Up regulated | KD reduces anchorage dependent and independent growth oh HCC cells | 13 |
| uc.338+ | Chr.12q13.13 | CRC | Up regulated | KD inhibits cell migration and invasion in SW480 and HCT116 cell lines, promoting invasion and metastasis | 17 |
| uc.388+ | Chr.15q21.3 | CRC | Down regulated | Associated with distal location of CRC | 18 |
| uc.460+ | Chr.Xp22.11 | NB | Up regulated | Activated by amplified MYCN in tumours | 8 |
| uc.269+A<br>uc.160+<br>uc.215+<br>uc.346+A<br>uc.348+ | Chr.9q33.3<br>Chr.5q14.1<br>Chr.7p14.1<br>Chr.12q23.3<br>Chr.13q21.33 | CLL | Up regulated | CLL prognosis | 20 |
| uc.58+A<br>uc.202+A<br>uc.207+A<br>uc.223+A | Chr.2p16.1<br>Chr.6q16.3<br>Chr.7p15.3<br>Chr.7q31.1 | Barrett's adenocarcin oma | Up regulated | Progression from normal to Barrett's oesophagus | 21 |
| uc.214+ | Chr.7p14.3 | | Down regulated | KD leads to Barrett's oesophagus carcinogenesis | |
| uc.161+A<br>uc.165+A<br>uc.327+A | Chr.5q14.1<br>Chr.5q14.3<br>Chr.11p13 | | Down regulated | Progression from squamous oesophagus to Barrett's oesophagus carcinogenesis | |
| uc.153+A<br>uc.158+A<br>uc.206+A<br>uc.274+A<br>uc.472+A<br>uc.473+A | Chr.5q13.2<br>Chr.5q14.1<br>Chr.7p15.3<br>Chr.9q33.3<br>Chr.Xp11.4<br>Chr.Xq13.1 | | Up regulated | | |
| uc.347+<br>uc.350+<br>uc.279+<br>uc.460+<br>uc.379+<br>uc.446+<br>uc.364+ | Chr.13q21.33<br>Chr.13q21.33<br>Chr.9q33.3<br>Chr.Xp22.11<br>Chr.14q32.2<br>Chr.19q12<br>Chr.14q12 | NB | Up regulated | Involved in in MYCN-amplified (MNA) tumours | 20 |

ATRA – all trans-retinoic acid; KD – knock down; BlCa – bladder cancer; CRC – colorectal carcinoma; NB – neuroblastoma; PCa – prostate cancer; HCC – hepatocellular carcinoma; CLL – chronic lymphocytic leukaemia.
[a] A + means the transcription is in antisense with the genomic orientation (telomere to centromer for the short arm of chromosomes and centromere to telomere, for the long arm of chromosomes).

expression of T-UCRs in malignant and non-malignant urothelial tissues [11]. In particular, they investigate the functional characterization of uc.8 + as the most up regulated T-UCR in bladder cancer tissues. In vitro experiments showed significant decreased capacities for bladder cancer cell invasion, migration, and proliferation, following uc.8 + knockdown.

## 5. Genetic variants in ultraconserved regions associate with cancer

*Bejerano* and co-workers reported that only six single-nucleotide polymorphisms (SNPs) in a total of 481 UCRs were validated in NCBI's SNP database (dbSNP) of human genome [1]. These six SNPs are located in the region of uc.53 +, uc.140 +, uc.252 +, uc.295 +, uc.353 + and uc.374 +, respectively. The evolutionary analysis by *Katzman* and collaborators further demonstrated that ultraconserved regions are under a stronger selection compared to protein-coding genes, suggesting that these DNA portions are selectively constrained and contain functional variation [24]. In addition, *Drake* and colleagues measured in these positions the allele frequency distributions in the human population describing that these are rarer than new alleles in non-conserved regions [12]. Furthermore, *Silla* and colleagues identified that prevalent variants in UCR sequences are less likely to overlap transcription factor binding site, meaning that allele frequencies of variants identify selective pressure on transcription factor binding [25].

In 2010, *Wojcik* and collaborators demonstrated that germline mutations, as well as single nucleotide polymorphisms (SNPs) in UCR, occur more frequently in patients with colon cancer and chronic leukaemia than in the general population [26]. Thus, these SNPs could be excellent predictors of individual risk of cancer, treatment efficacy, and prognosis. Several studies have investigated the association between genetic polymorphisms within UCRs and cancer risk. In particular *Yang* and colleagues [27], by studying the potential impact of the six under-represented SNPs identified by *Bejerano* and collaborators [1] within UCRs, have reported that two of them (rs9572903 and rs2056116) are associated with familial breast cancer (BC) risk by comparing 1214 German patients negative for *BRCA1* and *BRCA2* genes' mutations and 2084 German female blood donors. Conversely, *Shen* and collaborators

**Table 2**
SNPs contained in T-UCRs in different tumours.

| T-UCR name | Genomic Localization | Single Nucleotide Polymorphism (SNP) | Cancer type | Reference |
|---|---|---|---|---|
| uc.53+<br>uc.140+<br>uc.252+<br>uc.295+<br>uc.353+<br>uc.374+ | Chr.2p16.1<br>Chr.4p15.33<br>Chr.9p22.2<br>Chr.10q24.31<br>Chr.13q21.33<br>Chr.14q13.3 | rs1861100<br>rs2056116<br>rs1538101<br>rs7092999<br>rs9572903<br>rs7143938 | None | 1 |
| uc.140+<br>uc.353+ | Chr.4p15.33<br>Chr.13q21.33 | rs2056116<br>rs9572903 | BC in German Population | 27 |
| uc.140+<br>uc.353+ | Chr.4p15.33<br>Chr.13q21.33 | rs2056116<br>rs9572903 | BC in Chinese Population | 28 |
| uc.140+<br>uc.353+ | Chr.4p15.33<br>Chr.13q21.33 | rs2056116<br>rs9572903 | BC in Italian Population | 29 |
| uc.368+ | Chr.14q13.1 | rs8004379 | PCa | 28, 30 |

[28] failed to support any significant association of the six identified SNPs including rs9572903 and rs2056116 with BC risk in Chinese population. *Catucci* and collaborators [29] did not find any significant association between these two SNPs and familial BC risk in Italian population without mutations in BRCA1/2 genes. As the authors suggest in their manuscript, these conflicting results rely on differences in allele frequency of different populations. However, in another study carried out on patients with locally advanced colorectal adenocarcinoma treated with adjuvant fluoropyrimidine-based chemotherapy, the authors selected 48 potentially functional SNPs within UCEs and systematically evaluated their individual and joint associations with clinical outcomes [29]. Only for stage II patients with at least one variant allele of rs7849, consistent association with increased recurrence risk was observed in the training validation, and meta-analysis set [29].

Recently, another study reported that the SNPs rs8004379 contained in the uc.368 + was significantly associated with biochemical recurrence in two independent cohorts of localized prostate cancer patients (PCa) [28]. A consistent association of the same SNP with a decreased risk for prostate cancer-specific mortality was also observed. *Bao* and collaborators systematically evaluated 14 common SNPs within UCRs in three cohorts of PCa patients, to test the hypothesis that these UCR SNPs might influence clinical outcomes. Multivariate analysis adjusting for known clinical parameters found an association between rs8004379 and recurrence in localized disease. Remarkably, a consistent association of rs8004379 with a decreased risk for prostate cancer-specific mortality was also observed in the advanced PCa patient group. Their finding suggests that SNPs within UCRs may be valuable prognostic biomarkers for assessing PCa treatment response and survival [30]. In Table 2 we summarize the main findings on SNPs located in several T-UCRs in different tumours. As increasing number of studies on T-UCRs are going to be published, it is reasonable to speculate that even more details about the specific roles of T-UCR in cancer will be demonstrated soon.

## 6. Regulation of T-UCRs by direct interaction with microRNAs

To date, experimental evidence supports that the aberrant regulation of T-UCRs expression in cancer is driven by two main mechanisms: the altered interaction with regulatory non-coding RNA, such as miRNAs, and the hypermethylation of CpG island located in T-UCR promoters.

In 2007 *Calin* and colleagues gave the first evidence that T-UCRs are functional miRNA targets [11]. They demonstrated that some T-UCRs are directly regulated by highly expressed microRNAs via base pairing in CLL cells. In particular, uc.160 +, uc.346 + A and uc.348 + showed significant antisense complementarity with miR-24-1 miR-155, miR-146, miR-29b respectively. A negative expression correlation has been reported for each T-UCR with the corresponding microRNA levels in CLL samples. Furthermore, uc.348 +, uc.346 + A and uc.160 + were cloned into luciferase reporter vectors to confirm their interaction with miR-155, miR-24-1 and miR-29-b in vitro [11]. Each T-UCR::microRNA interaction was extensively validated by measuring the expression levels of uc.160 + and uc.346 + A after the overexpression of miR-155 in leukaemia cells. In addition, uc.160 + and uc.346 + A were down regulated after miR-155 transfection in human megakaryoblastic leukaemia cell line MEG-01 [11]. Similarly, *Scaruffi* and colleagues found a negative correlation between expression values of nine specific T-UCRs (uc.209 +, uc.271 +, uc.312 +, uc.330 +, uc.371 +, uc.411 +, uc.421 +, uc.435 +, uc.452 +) and five miRNAs, which allow distinguishing between long- and short-surviving high-risk neuroblastoma (NB) patients [31]. In 2016, *Goto* and colleagues found an overexpression of uc.416 + A in gastric cancer (GC) biopsies and identified a miR-153 binding site in the uc.416 + A RNA transcript. Modulation of miR-153 changed uc.416 + A expression in GC cells line MKN-74, suggesting a direct regulation of uc.416 + A expression by miR-153 [32]. In all these studies, the authors focused on the high sequence complementarity between the two non-coding RNA species, indicated by the huge number of miRNA::T-UCR interacting pairs. Most importantly, these findings suggest that both RNA species are involved in, and cooperate to, initiating or sustaining human tumorigenesis.

## 7. Regulation of microRNAs processing by direct interaction with T-UCRs

Sequestering of miRNA by endogenous RNAs (both coding and noncoding) has been only recently demonstrated in plants [33] and mammals [5,34]. Interestingly, the first report describing a direct regulation of miRNAs by an endogenous ultraconserved RNA was in HCT-116 human colorectal cancer cells by *Liz* and co-workers in 2014 [9]. Authors reported that the uc.283 + A directly binds the pri-miR-195. The complementarity involved 11 nucleotides situated immediately upstream of the Drosha cropping site within the miR-195 primary transcript. As validated in vivo and in vitro, the overexpression of uc.283 + A in HCT-116 colorectal cancer cells impair the proper processing of miR-195. *Carotenuto* and co-workers provided the first evidence that uc.158 + A is dependent by the Wnt pathway activation in liver cancer and drives its growth [15]. In particular, they demonstrated that inhibition of uc.158 + A reduced cell growth, and spheroid-based cell migration, and increased apoptosis in human malignant hepatocytes cells line HepG2 and SW1 [15]. miR-193b was predicted to have binding sites within the uc.158 + A conserved sequence. Experimental modulation of uc.158 + A affected miR-193b expression in human malignant hepatocytes supporting its role as a competing endogenous long non-coding RNAs. In addition, co-transfection of uc.158 + A inhibitor and anti-miR-193b rescued the effect of uc.158 + A inhibition on cell viability. More recently, it was demonstrated that the uc.8 +, up regulated in bladder cancer, is able to interact with miR-596, and that it is involved in the promotion and development of BlCa. By using in vitro

validation, uc.8 + was shown to be a natural decoy for miR-596 [10]. In addition, we found that the polycomb protein Yin Yang 1 (YY1) mediates the binding between miR-596 and T-UCR 8 + [35]. Thus, uc.8 + up regulation results in increased expression of MMP9 that increases the invasive potential of BlCa cells. Moreover, uc.8 + silencing increased miR-596 expression, which in turn reduced total uc.283 + [34]. These results show that the perturbation of one element in this network (specifically the T-UCR) is able to modify also the expression of other interactors (such as miRNAs). Thus, this potentially represents a new regulatory layer in which this new class of lncRNA modulate miRNAs levels. This new function opens up new scenarios for the understanding of how the interaction among different non-coding RNA molecules can regulate gene expression in physiologic processes and as its alteration is linked to pathophysiology [35].

## 8. T-UCR and epigenetic modification

As occurs with protein coding genes and miRNAs, hypermethylation of CpGs also regulate the expression of lncRNA and in particular T-UCRs. The presence of a canonical CpG island within a 2000-bp distance upstream the sense transcript of uc.160 +, uc.283 + A, and uc.346 + was found by *Lujambio* and colleagues [14]. In particular, a hypermethylation of a CpG island and the following transcriptional silencing was demonstrated in HCT-116 colorectal cancer cells and in other cell lines, such as breast (MCF-7, MDA-MB-231, CAMA-1), lung (H552, H441, H358, EBC1), lymphoma (Ramos, Raji, Namalwa), and leukaemia cells (HL-60, Jurkat, KG-1a). In addition, also in different human primary malignancies the authors observed the hypermethylation of CpGs nearby uc.160 +, uc.283 + A and uc.346 + in primary colon, breast and lung tumours, in melanomas, leukaemia and lymphomas [36]. This reinforced the link between CpG islands' hypermethylation and T-UCRs silencing.

*Hudson* and co-workers have recently reported that the treatment of LNCaP (prostate cancer cell line) with 5-AzaC—a chemical analogue of cytidine capable to inhibit DNA methyltransferases and to induce DNA hypomethylation—increases the expression of six T-UCRs, uc.308 + A, uc.434 + A, uc.241 + A, uc.283 + A, uc.285 +, uc.85 + [19]. Among them, the uc.283 + A has been previously reported to undergo transcriptional silencing through CpG hypermethylation [37]. These results also indicated that these T-UCRs are epigenetically silenced also in prostate cancer. *Mestdagh* and colleagues [8] evaluated the distance distribution of trimethylation of lysine 4 of histone H3 (H3K4me3)—a marker of transcription initiation—of intergenic and intragenic T-UCRs in four different cell lines. This study has reported that intergenic and intragenic T-UCRs significantly associate with active H3K4me3, but they noted a different distribution compared to protein-coding genes. Conversely, they observed H3K4me3 marks' distribution similar to miRNAs, suggesting common features in the transcriptional organization for these two classes of ncRNAs [8].

In 2016, *Goto* and colleagues, found down regulation of uc.158 + A expression in GC as well as PCa. The expression was consistently restored after 5-Aza-dC (5-aza-2′-deoxycytidine) treatment in four cell lines. Bisulfite genomic sequencing revealed a cancer-specific CpG hypermethylation upstream of uc.158 + A in LNCaP and DU145 cells as well as both in PCa and GC biopsies [31].

Taken together, a large body of evidence indicates that, similarly to miRNAs, also T-UCRs are susceptible to DNA methylation-associated inactivation in transformed cells, supporting the model in which genetic and epigenetic alterations in both coding and noncoding sequences cooperate in human tumorigenesis. Furthermore, given the recent discoveries about the effects of distant regulatory regions—such as enhancers [38] and super-enhancers [39]—on the activation of long non-coding RNAs in tumours, it is reasonable to speculate that also T-UCRs may undergo a similar epigenetic regulatory mechanism. Further studies in this direction are needed in the next future.

## 9. Perspectives and concluding remarks

Since their discovery, establishing the functional role of the ultraconserved transcribed elements of the human genome has been challenging. Most of our knowledge about UCEs mainly derives from computational analysis and thus from informatics predictions. To date, active transcription arising from most of these regions has not been experimentally confirmed nor it has been linked to specific biological processes. In many cases, the full-length sequence of these transcripts has not been determined yet. In addition, it has not been fully established the origin of these patterns with a high conservation across the species. Despite several studies have clearly linked the expression of T-UCRs to human diseases, and particularly to cancer, it has not been well established if their alteration has a causal or secondary role in cancer onset/progression. Therefore, a wide and robust body of experimental evidences is needed to ascertain their possible pathogenic contribution in cancer. However, a systematic characterization of T-UCRs in each tissue and their peculiar expression in specific cancer could be informative also to integrate—and to make more specific and sensitive—currently available cancer panels, widely used for prognostic and diagnostic purposes.

Indeed, it is now widely assumed that the expression of well-studies classes of non-coding RNAs (miRNAs and lncRNAs, including also the primary transcripts of microRNAs, that are less studied these days) [40–41] associates with patients' prognosis in many cancer types. Cancer panels including even T-UCRs signatures can be used in proof-of-principle to predict prognosis or other clinical outcomes (such as the response to therapy). We therefore envisage that T-UCRs-based prognostic/diagnostic kits will be released on the market, with the specific aim to help clinicians in a better, and more accurate, characterization of the histological subtype of cancer. However, their potential therapeutic use is still far, due to a substantial lack about their biological role in many contexts [42]. Since the knowledge of T-UCRs is still in its infancy—especially if compared to other ncRNAs classes—further studies are needed to convincingly include these ncRNAs in the growing area of cancer therapeutics.

## Funding

A.C. was supported by the Italian Association for Cancer Research (MFAG number 11510).

## Transparency document

The http://dx.doi.org/10.1016/j.bbcan.2017.09.001 associated with this article can be found, in online version.

## References

[1] G. Bejerano, M. Pheasant, I. Makunin, S. Stephen, W.J. Kent, J.S. Mattick, D. Haussler, Ultraconserved elements in the human genome, Science 304 (2004) 1321–1325.
[2] A. Subramanian, P. Tamayo, V.K. Mootha, S. Mukherjee, B.L. Ebert, M.A. Gillette, A. Paulovich, S.L. Pomeroy, T.R. Golub, E.S. Lander, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc. Natl. Acad. Sci. U. S. A. 102 (43) (2005) 15545–15550.
[3] G.A. Calin, M. Ferracin, A. Cimmino, G. Di Leva, M. Shimizu, S.E. Wojcik, M.V. Iorio, R. Visone, N.I. Sever, M. Fabbri, et al., A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia, N. Engl. J. Med. 353 (17) (2005) 1793–1801.
[4] A. Derti, F.P. Roth, G.M. Church, C.T. Wu, Mammalian ultraconserved elements are strongly depleted among segmental duplications and copy number variants, Nat. Genet. 38 (2006) 1216–1220.
[5] M. Cesana, D. Cacchiarelli, I. Legnini, T. Santini, O. Sthandier, M. Chinappi, A. Tramontano, I. Bozzoni, A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA, Cell 147 (2011) 358–369.
[6] F.A. Karreth, Y. Tay, D. Perna, U. Ala, S.M. Tan, A.G. Rust, G. DeNicola, K.A. Webster, D. Weiss, P.A. Perez-Mancera, et al., In vivo identification of tumor-suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma, Cell 147 (2011) 382–395.
[7] P. Sumazin, X. Yang, H.S. Chiu, W.J. Chung, A. Iyer, D. Llobet-Navas,

P. Rajbhandari, M. Bansal, P. Guarnieri, J. Silva, et al., An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma, Cell 147 (2011) 370–381.
[8] P. Mestdagh, E. Fredlund, F. Pattyn, A. Rihani, T. Van Maerken, J. Vermeulen, C. Kumps, B. Menten, K. De Preter, A. Schramm, et al., An integrative genomics screen uncovers ncRNA T-UCR functions in neuroblastoma tumours, Oncogene 29 (2010) 3583–3592.
[9] J. Liz, A. Portela, M. Soler, A. Gómez, H. Ling, G. Michlewski, G.A. Calin, S. Guil, M. Esteller, Regulation of pri-miRNA processing by a long noncoding RNA transcribed from an ultraconserved region, Mol. Cell 55 (2014) 138–147.
[10] G.A. Calin, C. Liu, M. Ferracin, T. Hyslop, R. Spizzo, C. Sevignani, M. Fabbri, A. Cimmino, E.J. Lee, S.E. Wojcik, M. Shimizu, E. Tili, S. Rossi, et al., Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas, Cancer Cell 12 (2007) 215–229.
[11] M. Olivieri, M. Ferro, S. Terreri, M. Durso, A. Romanelli, C. Avitabile, O. De Cobelli, A. Messere, D. Bruzzese, I. Vannini, et al., Long non-coding RNA containing ultraconserved genomic region 8 promotes bladder cancer tumorigenesis, Oncotarget 7 (15) (2016) 20636–20654.
[12] J.A. Drake, C. Bird, J. Nemesh, D.J. Thomas, C. Newton-Cheh, A. Reymond, L. Excoffier, H. Attar, S.E. Antonarakis, E.T. Dermitzakis, J.N. Hirschhorn, Conserved noncoding sequences are selectively constrained and not mutation cold spots, Nat. Genet. 38 (2) (2006) 223–227.
[13] C. Braconi, N. Valeri, T. Kogure, P. Gasparini, N. Huang, G.J. Nuovo, L. Terracciano, C.M. Croce, T. Patel, Expression and functional role of a transcribed noncoding RNA with an ultraconserved element in hepatocellular carcinoma, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 786–791.
[14] A. Lujambio, M. Esteller, CpG island hypermethylation of tumor suppressor microRNAs in human cancer, Cell Cycle 6 (2007) 1455–1459.
[15] P. Carotenuto, M. Fassan, R. Pandolfo, A. Lampis, C. Vicentini, L. Cascione, V. Paulus-Hock, L. Boulter, R. Guest, L. Quagliata, Wnt signalling modulates transcribed-ultraconserved regions in hepatobiliary cancers, Gut (2016), http://dx.doi.org/10.1136/gutjnl-2016-312278.
[16] M. Esteller, Non-coding RNAs in human disease, Nat. Rev. Genet. 12 (12) (2011) 861–874.
[17] C. Wang, Z. Wang, J. Zhou, S. Liu, C. Wu, C. Huang, Y. Ding, TUC.338 promotes invasion and metastasis in colorectal cancer, Int. J. Cancer 140 (6) (2017) 1457–1464.
[18] J. Sana, S. Hankeova, M. Svoboda, I. Kiss, R. Vyzula, O. Slaby, Expression levels of transcribed ultraconserved regions uc.73 and uc.388 are altered in colorectal cancer, Oncology 82 (2012) 114–118.
[19] R.S. Hudson, M. Yi, N. Volfovsky, R.L. Prueitt, D. Esposito, S. Volinia, C.G. Liu, A.J. Schetter, K. Van Roosbroeck, R.M. Stephens, G.A. Calin, C.M. Croce, S. Ambs, Transcription signatures encoded by ultraconserved genomic regions in human prostate cancer, Mol. Cancer 12 (2013) 13.
[20] G.A. Calin, C.M. Croce, Chronic lymphocytic leukemia: interplay between non-coding RNAs and protein-coding genes, Blood 114 (23) (2009) 4761–4770.
[21] M. Fassan, L. Dall'Olmo, M. Galasso, C. Braconi, M. Pizzi, S. Realdon, S. Volinia, N. Valeri, P. Gasparini, R. Baffa, et al., Transcribed ultraconserved noncoding RNAs (T-UCR) are involved in Barrett's esophagus carcinogenesis, Oncotarget 5 (16) (2014) 7162–7171.
[22] K.M. Watters, K. Bryan, N.H. Foley, M. Meehan, R.L. Stallings, Expressional alterations in functional ultra-conserved non-coding rnas in response to all-trans retinoic acid—induced differentiation in neuroblastoma cells, BMC Cancer 13 (2013) 184.
[23] M. Fabbri, R. Garzon, M. Andreeff, H.M. Kantarjian, G. Garcia-Manero, G.A. Calin, MicroRNAs and non-coding RNAs in hematological malignancies: molecular, clinical and therapeutic implications, Leukemia 22 (2008) 1095–1105.
[24] S. Katzman, A.D. Kern, G. Bejerano, G. Fewell, L. Fulton, R.K. Wilson, S.R. Salama, D. Haussler, Human genome ultraconserved elements are ultraselected, Science 317 (5840) (2007) 915.
[25] T. Silla, K. Kepp, E.S. Tai, L. Goh, S. Davila, Ivkovic T. Catela, G.A. Calin, P.M. Voorhoeve, Allele frequencies of variants in ultra conserved elements identify selective pressure on transcription factor binding, PLoS One 9 (11) (2014) e110692.
[26] S.E. Wojcik, S. Rossi, M. Shimizu, M.S. Nicoloso, A. Cimmino, H. Alder, V. Herlea, L.Z. Rassenti, K.R. Rai, T.J. Kipps, et al., Non-codingRNA sequence variations in human chronic lymphocytic leukemia and colorectal cancer, Carcinogenesis 31 (2) (2010) 208–215.
[27] R. Yang, B. Frank, K. Hemminki, C.R. Bartram, B. Wappenschmidt, C. Sutter, M. Kiechle, P. Bugert, R.K. Schmutzler, N. Arnold, SNPs in ultraconserved elements and familial breast cancer risk, Carcinogenesis 29 (2) (2008) 351–355.
[28] H. Shen, C. Lu, Y. Jiang, J. Tang, W. Chen, H. Zhang, Q. Zhang, J. Wang, J. Liang, Z. Hu, H. Shen, Genetic variants in ultraconserved elements and risk of breast cancer in Chinese population, Breast Cancer Res. Treat. 128 (3) (2011) 855–861.
[29] I. Catucci, P. Verderio, S. Pizzamiglio, S. Manoukian, B. Peissel, M. Barile, L. Tizzoni, L. Bernard, F. Ravagnani, L. Galastri, et al., SNPs in ultraconserved elements and familial breast cancer risk, Carcinogenesis 30 (3) (2009) 544–545.
[30] B.Y. Bao, V.C. Lin, C.C. Yu, H.L. Yin, T.Y. Chang, T.L. Lu, H.Z. Lee, J.B. Pao, C.Y. Huang, S.P. Huang, Genetic variants in ultraconserved regions associate with prostate cancer recurrence and survival, Sci. Rep. 6 (2016) 22124.
[31] P. Scaruffi, S. Stigliani, S. Moretti, S. Coco, C. De Vecchi, F. Valdora, A. Garaventa, S. Bonassi, G.P. Tonini, Transcribed-ultra conserved region expression is associated with outcome in high-risk neuroblastoma, BMC Cancer 9 (2009) 441.
[32] K. Goto, S. Ishikawa, R. Honma, K. Tanimoto, N. Sakamoto, K. Sentani, N. Oue, J. Teishima, A. Matsubara, W. Yasui, The transcribed-ultraconserved regions in prostate and gastric cancer: DNA hypermethylation and microRNA-associated regulation, Oncogene 35 (2016) 3598–3606.
[33] J.M. Franco-Zorrilla, A. Valli, M. Todesco, I. Mateos, M.I. Puga, I. Rubio-Somoza, A. Leyva, D. Weigel, J.A. Garcia, J. Paz-Ares, Target mimicry provides a new mechanism for regulation of microRNA activity, Nat. Genet. 39 (2007) 1033–1037.
[34] L. Poliseno, L. Salmena, J. Zhang, B. Carver, W.J. Haveman, P.P. Pandolfi, A coding-independent function of gene and pseudogene mRNAs regulates tumour biology, Nature 465 (2010) 1033–1038.
[35] S. Terreri, M. Durso, V. Colonna, A. Romanelli, D. Terracciano, M. Ferro, S. Perdonà, L. Castaldo, F. Febbraio, F. de Nigris, et al., New cross-talk layer between Ultraconserved non-coding RNAs, micriRNAs and Polycomb Protein YY1 in bladder cancer, Genes 7 (12) (2016).
[36] A. Lujambio, M. Esteller, How epigenetics can explain human metastasis: a new role for microRNAs, Cell Cycle 8 (2009) 377–382.
[37] A. Lujambio, A. Portela, J. Liz, S.A. Melo, S. Rossi, R. Spizzo, C.M. Croce, G.A. Calin, M. Esteller, CpG island hypermethylation-associated silencing of non-coding RNAs transcribed from ultraconserved regions in human cancer, Oncogene 29 (2010) 6390–6401.
[38] E. Vidal, S. Sayols, S. Moran, A. Guillaumet-Adkins, M.P. Schroeder, R. Royo, M. Orozco, M. Gut, I. Gut, et al., A DNA methylation map of human cancer at single base-pair resolution, Oncogene (2017), http://dx.doi.org/10.1038/onc.2017.176.
[39] H. Heyn, E. Vidal, H.J. Ferreira, M. Vizoso, S. Sayols, A. Gomez, S. Moran, R. Boque-Sastre, S. Guil, et al., Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer, Genome Biol. (2016), http://dx.doi.org/10.1186/s13059-016-0879-2.
[40] F. Bullrich, H. Fuji, G.A. Calin, H. Mabuchi, M. Negrini, Y. Pekarsky, L. Rassenti, H. Alder, J.C. Reed, M.J. Keating, et al., Characterization of the 13q14 tumor suppressor locus in CLL: identification of ALT1, an alternative splice variant of the LEU2 gene, Cancer Res. 61 (18) (2001) 6640–6648.
[41] R.S. Redis, A.M. Sieuwerts, M.P. Look, O. Tudoran, C. Ivan, R. Spizzo, X. Zhang, V. de Weerd, M. Shimizu, H. Ling, et al., CCAT2, a novel long non-coding RNA in breast cancer: expression study and clinical correlations, Oncotarget 4 (10) (2013) 1748–1762.
[42] H. Ling, L. Girnita, O. Buda, G.A. Calin, Non-coding RNAs: the cancere genome dark matter the matters!, Clin. Chem. Lab. Med. 55 (5) (2017) 705–714.