Mini review

Non-coding antisense transcripts: fine regulation of gene expression in cancer

Francisco Santos 1, Ana Maria Capela 1, Francisca Mateus 1, Sandrina Nóbrega-Pereira, Bruno Bernardes de Jesus *

Department of Medical Sciences and Institute of Biomedicine – iBiMED, University of Aveiro, 3810-193 Aveiro, Portugal

Abstract

Natural antisense transcripts (NATs) are coding or non-coding RNA sequences transcribed on the opposite direction from the same genomic locus. NATs are widely distributed throughout the human genome and seem to play crucial roles in physiological and pathological processes, through newly described and targeted mechanisms. NATs represent the intricate complexity of the genome organization and constitute another layer of potential targets in disease. Here, we focus on the interesting and unique role of non-coding NATs in cancer, paying particular attention to those acting as miRNA sponges.

© 2022 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Over the years, several studies pointed out to the presence of RNA molecules that were transcribed but did not encode proteins, making up the so-called non-coding genome. In fact, comparative analyses carried out between mammalian genomes and transcriptomes led to the important discovery that around three-quarters of genomic DNA were transcribed [1–3], greatly contrasting with the 2% that is translated into proteins [1]. Extensive annotation has revealed that about 80% of the human genome is functional. Indeed, several biochemical functions have been assigned to the non-coding portion of the genome, including transcription, association of transcription factors, chromatin structure and histone modification. This number overwhelmingly exceeds the percentage of the genome assigned to coding proteins, making it clear that the non-coding portion of the genome is not randomly transcribed and is involved in several important biological processes [4].
Non-coding RNAs can be divided according to their length into small and long non-coding RNAs (lncRNAs). The latter constitute a class of RNA molecules of more than 200 nucleotides, and account for the largest, yet less described, class of non-coding RNAs [5–6]. Many efforts have been applied to distinguish lncRNA from coding RNA transcripts [5,7–9]. In some cases, the absence of a detectable ORF is the only biochemical dissimilarity between lncRNAs and messenger RNA (mRNA).

LncRNAs are often classified according to their location relative to neighboring genes [10]. As such, these transcripts can be divided into sense, antisense, intronic, intergenic and bidirectional according to their positioning to proximal protein coding genes [11]. Here, we will focus only on the particular characteristics of antisense IncRNAs. These transcripts overlap at least one exon of another gene, and originate from the antisense strand [11]. Antisense IncRNAs usually overlap protein-coding genes, but some may overlap other lncRNAs [8,12–13]. Our main goal is to exemplify antisense transcripts in cancer, with particular emphasis on those acting as microRNA (miRNA) sponges and their role in cancer aggressiveness, metabolism, response to chemotherapy and the epithelial-mesenchymal transition (EMT). MiRNA sponges are lncRNAs shown to regulate gene expression through direct binding of specific miRNAs [14–18]. These lncRNAs have been termed competing endogenous RNAs (ceRNAs), and have been implicated in several biological processes and diseases, including cancer [19–22].

2. General overview of antisense lncRNAs

Natural antisense transcripts (NATs) have been considered to be extremely common throughout the mouse and human genomes [23]; in fact, around 70 % of mammalian genes are known to produce NATs [24]. Although these RNA molecules may include coding or non-coding sequences that are complementary to either coding or non-coding transcripts [25–26], here we will specifically describe the biological roles of non-coding NATs (hereafter referred to as ncNATs) complementary to either coding or non-coding sequences.

ncNATs can arise from a variety of promoters, such as independent, bidirectional or cryptic. Bidirectional promoters are shown to generate large numbers of ncRNAs, including in humans. Several factors have been shown to influence promoter bidirectionality, such as chromatin organization and polyadenylation signals that surround the promoter [27–28]. In fact, bidirectional transcription can also be originated from double strand breaks, demonstrating the interactions between transcription and DNA damage [29].

ncNATs can also be generated from cryptic promoters that are positioned either within the transcribed or from the termination regions of the sense gene [30]. Regarding regulation at the RNA level, several ncNATs, in yeast, for example, are considered cryptic unstable transcripts and are targeted for early degradation by nuclear exosomes. Other ncNATs are degraded by cytoplasmic RNA exonuclease [31] or regulated by 5′ decapping activity [32].

Depending on the location of their target genes, ncNATs can be classified as: cis-NATs, if transcribed from the opposite strand at the same genomic locus (and exhibiting perfect sequence complementarity to their targets); or trans-NATs, if transcribed from different genomic loci (and usually displaying imperfect sequence complementarity) [33–34]. Stalled RNA polymerases, R-loops and triple helices have been shown to help retain NATs at their site of transcription, allowing cis-acting NATs to exert their function. On the other hand, the three-dimensional organization of chromatin can lead to the interaction of regions of antisense transcription with other loci, thus mediating trans effects (reviewed in [35]).

cis-NATs are further classified according to how they overlap with their targets, and are categorized as: head-to-head, whose 5′ end overlaps with the 5′ of the target transcript; tail-to-tail, whose 3′ overlaps with the 3′ of the target gene; and internal, which entirely bind to their target. cis-NATs can also be classified as nearby NATs, which are located very close to their targets, but do not, in fact, overlap [24,36] (Fig. 1).

ncNATs, just like other lncRNAs, exhibit tissue-specific expression, which suggests that there are many evolutionarily conserved functional roles associated with these transcripts [37]. Regarding the effect in relation to the cognate sequence, their expression...
A variety of techniques can be used to detect novel NATs [86]. Most of the methods include sequencing, being RNA-seq the most basic and versatile application of this principle, followed by microarrays. Depending on the length or other specificities of the target RNAs, enrichment or size selection can be done to optimize the screening. A strand-specific protocol should be preferred for the analysis of the antisense strand and its characteristics. It is also possible to analyze specifically the expression of nascent transcripts, i.e., GRO-seq, CAGE and SAGE are examples of methods that sequence the ends of the transcripts. Techniques such as ChIP-seq or ChIRP analyze the indirect effects of their transcription in chromatin signatures, but this might prove a limitation when NATs overlap coding genes. Polyadenylation is also uncommon in NATs so poly(A)-dependent methods, such as 3P-seq, might not detect all the antisense transcripts present in a sample [87–88]. When NAT structure is already known and annotated, its expression can be further evaluated using RT-qPCR or in situ imaging techniques. There has also been a quick evolution with single-cell adaptations, and it is now already possible to apply several omics at once. Finally, only a handful of the techniques listed have been put to practice in NAT identification, so further advances in this field are to be expected, with the consideration that the best analyses may come from a combined application of several techniques.

### Table 1

| Method                             | Principle                                                                 | Advantages                                                                 | Limitations                                                                          | Ref Original |
|------------------------------------|---------------------------------------------------------------------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------------------|--------------|
| 4Su-seq                            | Metabolic labelling of nascent RNA using 4Su-tagging                      | Changes in RNA kinetics are more visible than with total RNA; more sensitive to transient RNA than RNA-seq | More input required; library preparation may introduce some bias for large transcripts; only a short amount of 3’ regions are labelled during the short 4Su exposure leading to 5’ bias | [89]         |
| CAGE (Cap Analysis of Gene Expression) | Sequencing method that analyzes 5’ end termini of transcripts            | High accuracy and single nucleotide resolution in a high-throughput manner  | Not suited for the characterization of novel NATs                                     | [90]         |
| ChIP-seq (Chromatin immunoprecipitation)  | Detection of chromatin signatures through combination of chromatin immunoprecipitation with RNA-seq, allowing evaluation of expression and primary structure of genes encoding transcripts | High-throughput, sensitivity and specificity, and low cost and input requirement | High complexity; inappropriate to nRNAs overlapping protein-coding genes and short transcripts; detects all RNA polymerase II (Pol II)–RNA complexes, not just those that are actively engaged | [91]         |
| ChIRP-seq (Chromatin Isolation by RNA Purification) | Identification of binding sites and interactions between transcripts and chromatin using tiling oligonucleotides retrieval followed by sequencing | Indicative of putative transcription factor activity for a given ncRNA | Noise caused by precipitation of non-specific DNA fragments from off-target hybridization; prior knowledge of NATs sequence needed | [92]         |
| CLIP (Cross-linking immunoprecipitation) | Detects transcripts that interact with RNA Binding Proteins (RBP); RIP followed by sequencing, with variations on the technique depending on the purpose i.e., HITS-CLIP, PAR-CLIP, iCLIP, CRAC | High resolution; can detect unstable transient interactions as well; covalent binding allows washing and purification | UV cross-linking may result in partial degradation of RNA samples; PCR amplification needed to compensate low efficiency if crosslinking | [93]         |
| GRO-seq (Global Run-On sequencing) | Sequencing method that analyzes nascent transcripts thus differentiating between transcriptionally active and inactive regions | High resolution and specificity | Detects low but significant amount of antisense transcripts; does not necessarily reflect transcription in vivo | [28]         |
| Imaging                            | Several imaging techniques allow for in situ hybridization of transcripts directly on tissues i.e., smFISH, MEFISH, seqFISH, FISSEQ and SPOTS | Highly specific probes that are easy to use and allow for spatial and contextual information | Small probes can lead to false positives, high background, overcrowding, less sensitivity and spatial resolution | SmFISH–[94–96] MEFISH–[97] SeqFISH–[98] FISSEQ–[99–100] SPOTS–[101] |
| Microarrays                        | Detection of transcripts through hybridization with nucleic acids         | Well established, affordable, flexible customization and easy analysis; the use of artificial anti-sense sequence (AFAS) probes identifies unannotated NATs undetectable with cDNA; Affymetrix chips allow a sensitive analysis of most protein coding mRNAs in a more affordable and timely manner | Many non-coding transcripts cannot be detected with standard microarrays due to their design | [102–105]   |
| Tiling microarrays                 | Detection of transcripts through hybridization with nucleic acids         | Identifies more unknown transcripts than RNA-Seq  | High false-positive rate; inadequate for low expression transcripts                  | [106]        |
| Nanostream                         | Directly measures expression of transcripts using probe hybridization and imaging | High automation, specificity, and sensitivity; also detects low expression transcripts without amplification and cDNA production | Requires previous knowledge of the transcript sequence                             | [107]        |
| Nascent-seq                       | Isolation and sequencing of nascent transcripts                          | Changes in RNA kinetics are more visible than with total RNA                  | Many NATs are not polyadenylated and cannot be detected through polyadenylation isolation and sequencing | [108–109]   |
| Northern Blot                      | Probe labelling of RNAs transferred to a membrane that were size separated using gel electrophoresis | Allows estimation of RNA size; cheap and simple                              | No signal amplification; larger probes are more specific but are more prone to create background; labor-intensive; poor reproducibility; sensitivity impacted by RNA degradation; poor quantitative technique | [110]        |
| NET-seq (Native elongating transcript sequencing) | Sequencing of the whole nascent transcriptome that is attached to polymerase | Single-nucleotide resolution of RNA polymerase II (Pol II)–associated transcripts | Only detects still attached to RNA polymerase II (Pol II); cannot distinguish between Pol II accumulation and co-transcriptional cleavage | [48]         |
| PRO-seq (Precision nuclear Run-)   | Analyzes nascent RNAs from their 3’ ends that are                        | High sensitivity and base-pair resolution                                    | Does not distinguish between Pol I, II and III                                       | [111]        |

**Notes:**
- 4Su-seq: 4SU tagging
- CAGE: Cap Analysis of Gene Expression
- ChIP-seq: Chromatin immunoprecipitation
- ChIRP-seq: Chromatin Isolation by RNA Purification
- CLIP: Cross-linking immunoprecipitation
- GRO-seq: Global Run-On sequencing
- Imaging: Several imaging techniques
- Microarrays: DNA microarrays
- Tiling microarrays: In situ hybridization
- Nanostream: RNA microarrays
- Nascent-seq: Native transcriptome sequencing
- Northern Blot: Northern hybridization
- NET-seq: Native elongating transcript sequencing
- PRO-seq: Precision nuclear Run-on sequencing
| Method                                      | Principle                                                                 | Advantages                                                   | Limitations                                                                                     | Ref Original |
|--------------------------------------------|---------------------------------------------------------------------------|--------------------------------------------------------------|-------------------------------------------------------------------------------------------------|--------------|
| On and sequencing assay POINT (Polymerase intact nascent transcript) | Purification and sequencing of nascent transcripts                       | Intact nascent RNA and fast                                 | transcription only detects still attached RNA polymerase II (Pol II); cannot distinguish between Pol II accumulation and co-transcriptional cleavage | [49]         |
| PolyA-seq                                  | Sequencing of poly(A)-associated regions of transcripts                   | High sensitivity; quantitative analysis of novel 3' UTRs without low false positives; better than RNA-seq to detect short transcripts; already strand-specific | Many NATs are not polyadenylated                                                              | [112]        |
| Poly(A)-position profiling by sequencing (3P-Seq) | Sequencing of 3' UTR of polyadenylated transcripts                      | High-throughput                                              | Many NATs are not polyadenylated                                                              | [113]        |
| RT-qPCR                                    | PCR amplification of cDNA                                               | Specific and widely used to evaluate expression              | Requires previous knowledge of the transcript sequence                                         | [114]        |
| RACE (Rapid amplification of cDNA ends)    | Sequencing method following PCR amplification                           | Efficient and precise annotation of single and low abundance transcripts | Less efficiency with larger fragments, in which cases a cDNA library could be done; requires previous knowledge of a short part of the sequence for primer construction | [115]        |
| RIP (RNA immunoprecipitation)              | Determination of transcript targets of a given RNA-binding protein (RBP) in vivo | Gold standard that also allows to detect nascent NATs       | Only captures the relatively stable RNA–protein complexes and may result in under-representation of transient interactions | [116]        |
| RNA CaptureSeq                             | Deep-sequencing method that constructs tiling arrays of regions of interest against which captured cDNAs of transcripts are hybridized | Quantitative analysis with increased sequencing depth, that allows recognition of rare and unannotated NATs | Hybridization might add undesired artifacts                                                   | [117]        |
| RNA-PET (RNA paired-end ditags) or GIS-PET | Sequencing of 5' and 3' end regions of transcripts                      | Validates and maps boundaries of polyadenylated transcripts | Requires large amount of sample; short tags render less mapping specificity; many NATs are not polyadenylated | [118–119]   |
| RNA-seq                                    | Sequencing method that identifies and characterizes transcripts; adaptations can be made to sequence the wanted targets i.e., Targeted RNA-seq; Strand-specific RNA (ssRNA-seq); Short RNA-seq; Single-cell RNA-seq (scRNA-seq) | Can detect novel transcripts, of various sizes, expression levels and kinds with a high resolution and sequencing depth | Long RNAs need to be fragmented and assembled in libraries, which may lead to unprecise genomic boundaries; needs complementary experimentation to exclude transcriptional noise | [119–120]   |
| RT-PCR-seq                                 | Deep-sequencing method combined with highly multiplexed PCR amplification of cDNA of transcripts | Identifies novel exons of known transcripts                  | Primer design limits the number of known or predicted testable junctions to be validated       | [127]        |
| SAGE (Serial Analysis of Gene Expression)  | Sequencing method that analyzes 3' end termini of transcripts (poly(A)-associated regions/sites) | Quantitative and qualitative analysis; can identify novel transcripts | Many NATs are not polyadenylated                                                              | [128]        |
| Single-cell multiomics                     | Simultaneous sequencing of the transcriptome and other omics, such as genomics i.e., DR-seq and G&T-seq; epigenomics i.e., scMeth-seq and scMT-seq; and proteomics i.e., PLAYR and SPARC | Allows a contextual analysis of the cell, useful in cancer studies and diagnosis                  | Expensive; amplification errors; allelic and locus dropout; challenging single-base resolution | [129–134]   |
| TimeLapse-seq                              | Metabolic labelling of nascent RNA using 4sU-tagging followed by sequencing | Results in temporal information; already has internal normalization; low input required; changes in RNA dynamics are more visible than with total RNA | Reads beyond the poly-A termination signal are rare, and with restricted time, only 3' ends are labelled, which enables detection of many NATs that are not polyadenylated with restricted time, only 3' end is labelled and many NATs are not polyadenylated | [135]        |
| TT-seq (Transient transcriptome sequencing) | Metabolic labelling of nascent RNA using 4sU-tagging followed by sequencing | Low input; better than SLAM-seq in analyzing transient RNA species since most ncRNAs are quickly degraded; allows tracking of RNA kinetics; overcomes 4su-seq 5' bias with RNA fragmentation before isolation thus mapping transcribed regions uniformly | Many NATs are not polyadenylated; the cellular 4sU uptake kinetics, transcriptional activity and library sequencing depth may limit the detection of nascent labelled transcripts; an assessment should be made to know toxic concentrations of 4sU for the cell type | [136]        |
| SLAM-seq (Thiol (SH)-Linked Allylation for the Metabolic sequencing of RNA) | Metabolic labelling of nascent RNA using 4sU-tagging followed by sequencing | Allows tracking of RNA and kinetics; rapid, low input and high throughput | Many NATs are not polyadenylated; the cellular 4sU uptake kinetics, transcriptional activity and library sequencing depth may limit the detection of nascent labelled transcripts; an assessment should be made to know toxic concentrations of 4sU for the cell type | [137]        |
may be concordant (associated with an increased expression) [38] or discordant (associated with a decrease in expression levels) [39]. ncNATs may affect negatively or positively the expression of sense transcription, and their expression is closed related with the regulation of their sense gene [40–41]. Indeed, the regulation of promoter specificities is a potential mechanism of ncNATs in cancer. For instance, HNF4A-AS1 selectively activates the HNF4A P1 promoter via HNF1A, which upregulates the expression of tumor suppressor P1-driven isoforms [42]. RNA-seq data has identified several ncNATS whose expression correlates specifically with the activity of one promoter of their sense gene. Silencing of these ncNATs was shown to alter the promoter usage, demonstrating how these transcripts may regulate promoter-specific programs under different contexts [42].

ncNATs are more frequently located in the nucleus, but they can also be found in the cytoplasm [24], and, like other IncRNAs, contain specific motifs that interact with DNA, RNA and proteins [5], meaning that these transcripts can affect gene expression at the transcriptional, post-transcriptional and translational levels. As such, ncNATs can regulate gene expression by DNA methylation (eg. LUC7L), histone modification (eg. XIST/TSIX, HOTAIR), transcriptional interference (eg. GNIT12-AS), regulation of alternative splicing (Zeb2-NAT), and regulation of mRNA stability, either by masking of miRNA binding sites (eg. BACE1-AS) or sponging miRNAs (eg. HOTAIR, NR2FI-AS1) [24,43]. Interestingly, some ncNATs regulate gene expression through DICE-dependent mechanisms (eg. NAT8B531), giving rise to small RNAs with roles in the epigenetic regulation of gene expression. The regulation of mRNA stability through miRNA sponging is of particular interest, as several recently identified ncNATs function as ceRNAs (discussed below). Furthermore, several ncNATs have been detected in both the nucleus and cytoplasm, highlighting their role in the regulation of gene expression through multiple mechanisms.

2.1. Challenges in the detection of ncNATs and how it impacts on their understanding

The overlap of antisense transcripts with their sense counterparts, their relative low expression levels and limited evolutionary conservation, have limited the use of high-throughput approaches to identify these transcripts (reviewed in [35]). In fact, antisense transcription has been carefully exploited at a large scale even before the advent of massive sequencing [44–47]. Further, detecting active transcription using global run-on sequencing (GRO-seq), native elongating transcript sequencing (NET-seq) or, more recently, POINT-technology, have been shown to be essential to identify antisense transcription and, therefore, antisense transcripts [28,48,49]. It is important, however, to discriminate between antisense transcription and antisense transcripts, because the former not always gives rise to functional RNAs, either coding or non-coding. In fact, antisense transcription is known to impact locally on gene expression programs through a variety of mechanisms not always mediated through ncRNAs. The advent of new technologies has been supporting the understanding of transcription complexity (Table 1). Indirect detection of antisense transcripts through chromatin modification states has also been employed, although this approach is rather limited due to the lack of strand specificity of chromatin modification and the generally higher expression levels of sense transcripts [50] (reviewed in [43]). The genomic arrangement of ncNATS has made it difficult to determine their function by loss-of-function studies without affecting the expression of the corresponding sense transcript [35,51]. To bypass this problem, the analysis of genome-wide gene expression levels at several time points and its comparison to the transcriptional response of cells after modulation of components that are involved in gene regulation by antisense transcription (such as histone-modifying enzymes) has proven to be a more suitable approach [52–53].

3. Selected cancer-related ncNATs

Several ncNATs have been associated to cancer, and their roles have been elucidated in several reviews (see references [24,43]). For instance, TALAM1 (MALAT1 antisense RNA) is a ncNAT that has been linked to the expression of IncRNA MALATI (metastasis-associated lung adenocarcinoma transcript 1). Zong and colleagues revealed that TALAM1 interacts with MALATI at its site of transcription, promoting its RNase P-mediated 3’ end cleavage. Interestingly, MALATI positively regulates the transcription and RNA stability of TALAM1. The dynamic between these two IncRNAs has been explored in breast cancer. Both MALAT1 and TALAM1 were shown to be upregulated in MCF7 (luminal A) and MDA-MB-231 (triple negative) cell lines. An upregulation of both MALAT1 and TALAM1 was also observed in MCF 10A cells (derived from normal breast epithelium) after exposure to TGF-β. Silencing TALAM1 negatively impacted the ability of breast cancer cells to migrate in vitro and to develop lung metastasis in immunocompromised mice. Ultimately, this study suggests that MALAT1 and TALAM1 act together to regulate breast cancer aggressiveness and malignancy [54].

FCG5-AS1 (FGD5 antisense RNA 1) was shown to be overexpressed in pancreatic cancer and bind to miR-577, which targets β-catenin and LRP6 (low density lipoprotein receptor related protein 6) [55]. Downregulation of FGC5-AS1 led to increased levels of miR-577, resulting in decreased levels of β-catenin and LRP6. This decrease was associated with a reduction in the Wnt signaling pathway and inhibition of cell proliferation, migration and invasion [55]. CERS6-AS1 (ceramide synthesis 6 antisense RNA 1) was also identified as being overexpressed in both pancreatic tissue samples and cell lines [56]. According to study conducted by Gao and colleagues, CERS6-AS1 acts as a ceRNA by interacting with miR-195-5p, resulting in an increase of WIP1. Silencing this antisense transcript resulted in an increase of miR-195-5p and a decrease of WIP2, which was accompanied by a decrease of cell proliferation, evidenced by a reduction in EdU positive cells, and an increase of apoptotic cells. Remarkably, overexpression of WIP2 reversed the effects of CERS6-AS1 silencing. Also in pancreatic cancer, Zhang and colleagues observed increased expression levels of SLC04A1-AS1 (solute carrier organic anion transporter family member 4A1 antisense RNA 1) in cancer samples. This IncRNA also targeted a miRNA, miR-4673, to derepress KIF21B. Comparable to other antisense transcripts identified in pancreatic cancer, silencing SLC04A1-AS1 led to a decrease in cell viability and migration, and an increase of apoptosis [57].

LncRNA AFAP1-AS1 (actin filament-associated protein 1 antisense RNA 1) was discovered to play an important role in retinoblastoma, which was shown to be overexpressed in tumor samples and cell lines. AFAP1-AS1 acts as a ceRNA to suppress miR-545-3p, which in turn derepresses its target GNB1. By silencing AFAP1-AS1, miR-545-3p is able to inhibit GNB1, and this was associated with decreased cell proliferation and migration [58].

EMT plays a pivotal role in cancer progression [59]. Recently, Bozgeyik and colleagues identified VIM-AS1 (vimentin antisense RNA 1), which is transcribed opposite of VIMENTIN (VIM), a well-known EMT marker, and both are upregulated in oral cancer. Accordingly, the expression of E-cadherin exhibited an opposite trend, being downregulated in tumor samples. The higher expression of VIM-AS1 was also associated with an advanced clinical stage and the presence of lymph node metastasis [60]. Other crucial EMT-related ncNAT is ZEB2-NAT (ZEB2 natural antisense transcript), which was shown to be important for the maintenance of 5’-UTR
ZEB2 intron through direct overlapping with the 5′ splice site in the intron [61]. ZEB2 is a transcriptional repressor of E-cadherin and a major activator of EMT, and expression of ZEB2-NAT prevents splicing of the first intron of ZEB2, increasing the levels of ZEB2 protein. Thus, ZEB2-NAT expression favors EMT, and is therefore associated with cancer progression and cellular reprogramming [38,62]. This ncNAT has been shown to play an important role in the conservation of stemness properties in quiescent cancer stem cells, supporting the existence of a cellular population with chemoresistance potential [63]. In fact, several ncNATs have been shown to be equally important in cancer and stem cells and/or to modulate response of cancer cells to chemotherapeutic agents. One example is a ncNAT occurring in the SOX9 locus (SOX9-NAT). SOX9-NAT expression was significantly lower in cancer tissues or human embryonic stem cells, compared with their matched normal tissues, suggesting that slight modifications in SOX9-NAT may result in remarkable changes in SOX9 expression, making this ncNAT a potential therapeutic target in regenerative medicine and cancer treatment [64]. Another example, TRPM2-AS (transcription receptor potential cation channel subfamily M member 2 antisense RNA), was shown to confer paclitaxel resistance in prostate cancer. Shi and colleagues showed that both prostate cancer tissues and cell lines exhibited higher levels of TRPM2-AS, compared to normal tissues and cells [65]. Moreover, paclitaxel-resistant prostate cancer cell lines demonstrated even higher expression levels of TRPM2-AS. Mechanistically, this transcript was shown to bind to miR-497-5p, whose expression was decreased in prostate cancer tissues, when compared to normal tissues. Furthermore, it was revealed that miR-497-5p negatively regulates FOXK1, whose expression levels are higher in prostate cancer tissues and cell lines. The authors revealed that silencing TRPM2-AS led to increased and decreased levels of miR-497-5p and FOXK1, respectively, resulting in decreased proliferation, migration and invasion and increased apoptosis, and even suppressed tumor growth in vivo [65].

Another example of how ncNATs are associated with chemoresistance was brought to light by Ling and colleagues [66]. They identified FOXD3-AS1 (foxbos fine box D3 antisense RNA 1) as an overexpressed antisense transcript in glioblastoma associated with a worse prognosis. Furthermore, temozolomide-resistant glioblastoma cell lines evidenced higher FOXD3-AS1 expression levels compared to those that were sensitive [66]. In fact, overexpression of FOXD3-AS1 increased tolerance of temozolomide in sensitive cells. Additionally, this lncRNA was revealed to interact with miR-128-3p, which in turn was shown to negatively regulate cell cycle protein WEE1. The authors demonstrated that silencing FOXD3-AS1 resulted in decreased and increased expression levels of miR-128-3p and WEE1, respectively. Furthermore, reduced expression of FOXD3-AS1 sensitized resistant glioblastoma cells to temozolomide, inhibiting cell growth and inducing an increase of cleaved caspase-3, resulting in apoptosis [66].

In osteosarcoma, ANRIL (antisense non-coding RNA in the INK4 locus) was identified as a potential biomarker for chemosensitivity and clinical outcome. In a study carried out by Lee et al., ANRIL was shown to be upregulated in osteosarcoma cell lines, which was correlated with resistance to anti-cancer drugs cisplatin and doxorubicin [67]. By silencing the expression of ANRIL, cells became more sensitive to these treatments and less proliferative. On the other hand, overexpression of ANRIL resulted in an opposite outcome: increased cell proliferation was observed, and cells developed further resistance to cisplatin and doxorubicin [67]; ncNATs have also been shown to identically play fundamental roles through the regulation of cancer metabolism. Li and colleagues identified IncRNA OIPS-AS1 in cervical cancer, whose expression was shown to be higher in cancer tissues and cell lines, in comparison to their normal counterparts [68]. Furthermore, patients were stratified, according to OIPS-AS1 expression, into low and high expression levels, and the latter was associated with larger tumor size, lymph node metastasis and poor 5-year overall survival. Interestingly, under hypoxic conditions, this transcript became upregulated, along with HIF-1α, GLUT1 and LDHA. OIPS-AS1 was revealed to bind to miR-124-5p, leading to de-repression of IDH2 [68].

3.1. Therapeutic and prognostic value of ncNATs in cancer?

Although ncNATs have been widely associated to cancer as prognostic and diagnostic markers [43,69–73] they are still far from reaching a therapeutic significance. Several clinical trials are in progress to assess their potential role as cancer biomarkers, in particular when explored in circulating particles, such as exosomes (see, for example, [43]). Examples of ncNATs in clinical trials are IncRNA HOTAIR (coexisting in the HOX gene) in thyroid cancers and MFI2-AS1 (melanotransferrin antisense RNA) in localized clear-cell cancers of the kidney. MFI2-AS1 expression was associated with the recurrence and survival of patients with clear-cell kidney carcinoma [74]. Interestingly, MFI2-AS1 may act as a sponge of miR-574-5p, a miRNA with potential roles in cancer metastasis [75–76].

Therapeutics that modulate RNAs in general and ncNATs in particular may be envisioned to play fundamental roles in disease. The development of antisense oligonucleotides (ASOs), duplex RNA technologies (RNAi) or genome editing (e.g. CRISPR), has aided the translational approach of ncNATs to the clinics, in particular when loss of function is projected [77–79]. One example is the application of NATs therapeutics in the Angelman syndrome. Here, targeting NATs may be used to derepress the normally-repressed paternal copy of the UBE3A gene in patients experiencing absent expression of the UBE3A gene-copy of the maternal allele [79–80].

Additionally, using the knowledge acquired from how NATs may interfere with the expression of associated genes, new therapies may evolve. Recently, Taekyu Ha and colleagues explored a gene-targeting method to deplete ephrinB2 using an inducible lentiviral vector. EphrinB2 promotes colorectal cancer and predicts poor patient survival. They demonstrated that integration and expression of the lentiviral construct in the host DNA may drive divergent transcription. Antisense transcription was associated with cell death through activation of a stress response providing evidence that divergent gene transcription from lentiviral vector integration may have an impact on the regulation of gene expression [81].

4. Conclusion

ncNATs are a class of IncRNAs that overlap either protein- or non-coding sequences. Although many of these transcripts can be regarded as junk RNA, several others present a conservation and cellular specificity, guiding them to important roles during normal tissue functions and diseases, including cancer. In particular, ncNATs have been shown to affect either their neighboring genes, and sequences on different genomic loci, thus adding additional layers to the already complex process of gene regulation.

Many recently identified ncNATs have been shown to derepress miRNA targets by functioning as ceRNAs, thus contributing to the regulation of miRNA stability. Interestingly, some of these transcripts also have nuclear functions, making it clear that they can regulate gene expression through several mechanisms.

The advances of gene editing and targeted techniques show great potential for the tissue-specific determination of the role of ncNATs in several biological processes, allowing the characterization of these transcripts as potential therapeutic targets in disease.
Funding
This work was supported by LISBOA-01-0145-FEDER-028534, projeto cofinanciado pelo FEDER através POR Lisboa 2020 - Programa Operacional Regional de Lisboa, do PORTUGAL 2020 e pela Fundação para a Ciência e a Tecnologia (FCT) e Fundação para a Ciência e Tecnologia (FCT) (ERA-CVD 2018 / 3599-PPCDT - ERA-CVD/0001/2018 - INNOVATION). FS has an individual scholarship from FCT (SRH/BD/146204/2019). S. N.-P. has a 2020.00355.CEE-CIND assistant research contract and AM-C and FM have a Verão com Ciência 2022 research grant.

Declaration of Competing Interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments
We thank other members of the iBiMED, Bruno B. de Jesus and Sandrina N-Pereira labs for insightful discussions and advice.

References
[1] Djebali S, Davis CA, Merkel A, et al. Landscape of transcription in human cells. Nature 2012;489(7414):101–8. https://doi.org/10.1038/nature11233.
[2] Carinci P, Kasukawa T, Katayama S, et al. The transcriptional landscape of the mammalian genome. Science (80-) 2005;309(5740):1559–63. 309/5740/1559.
[3] Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. Cell 2012;149(3):515–24. https://doi.org/10.1016/j.cell.2012.05.014.
[4] Zhao S, Zhang X, Chen S, Zhang S. Natural antisense transcripts in the biological hallmarks of cancer: Powerful regulators hidden in the dark. J Exp Clin Cancer Res 2020;39(1):1–18. https://doi.org/10.1186/s13046-020-01518-0.
[5] 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. https://doi.org/10.1038/359091a4.
[6] Osilok F, Kapranov P, Foissac S, et al. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. Cell 2010;143(6):1018–29. https://doi.org/10.1016/j.cell.2010.11.020.
[7] Najafi S, Tan NC, Rae P, et al. Gene regulation by antisense transcription: A focus on neurological and cancer diseases. Biomed Pharmacother 2022;145:. https://doi.org/10.1016/j.biopha.2021.112625.
[8] Balbin OA, Malik R, Dhanasekaran SM, et al. The landscape of antisense gene expression in human cancers. Genome Res 2015;25(7):1068–79. https://doi.org/10.1101/gr.177342.115.
[9] Ligté G, Poulet C, Bours V, Josse C, Jerusalem G. Natural Antisense Transcripts: Molecular Mechanisms and Implications in Breast Cancers. Int J Mol Sci 2018;19(1):123. https://doi.org/10.3390/ijms19010123.
[10] Seila AC, Calabrese JM, Levine SS, et al. Divergent transcription from active promoters. Science (80-) 2008;322(5909):1849–51. https://doi.org/10.1126/science.1159163.
[11] Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Res 2012;22(9):1775–89. https://doi.org/10.1101/gr.132159.111.
[12] Lee JT, Davidow LS, Warshawsky D. Tsix, a gene antisense to Xist at the X-chromosome inactivation centre. Nat Genet 2000;26(2):168–76. https://doi.org/10.1038/74047.
[13] Zong X, Nakagawa S, Freier SM, et al. Natural antisense RNA promotes 3′ end processing and maturation of MALAT1 IncRNA. Nucleic Acids Res 2016;44(6):2989–98. https://doi.org/10.1093/nar/gkw047.
[14] Ebert MS, Neillson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods 2007;4(9):721–6. https://doi.org/10.1038/nmeth1079.
[15] Ebert MS, Sharp PA. Roles for MicroRNAs in conferring robustness to biological processes. Cell 2012;149(3):515–24. https://doi.org/10.1016/j.cell.2012.05.014.
[16] Verboekom J, Taconet L, Gouache J, et al. Non-coding natural antisense transcripts: analysis and application. J Biotechnol 2012;30(5):453–9. https://doi.org/10.3389/fmicb.2017.02292.
[17] De Jesus BB, Marinho SP, Barros S, et al. Silencing of the IncRNA ZEB2-NAT facilitates reprogramming of aged fibroblasts and safeguards stem cell pluripotency. Nat Commun 2018;9(1):1–11. https://doi.org/10.1038/s41467-017-01921-6.
[18] Modarresi F, Faghihi MA, Lopez-Toledano MA, et al. Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. Nat Biotechnol 2010;28(10):983–8. https://doi.org/10.1038/nbt.1958.
[19] Sin S, Zhang L, Luo W, Zhang X. Characteristics of antisense transcript promoters and the regulation of their activity. Int J Mol Sci 2015;17(1):9. https://doi.org/10.3390/ijms16010009.
[20] Oszilok F, Góhali MA, Lopez-Toledano MA, et al. Inhibition of natural antisense transcripts in vivo results in gene-specific transcriptional upregulation. Nat Biotechnol 2012;30(5):453–9. https://doi.org/10.1038/nbt.1958.
[21] 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. https://doi.org/10.1038/359091a4.
[22] Zong X, Nakagawa S, Freier SM, et al. Natural antisense RNA promotes 3′ end processing and maturation of MALAT1 IncRNA. Nucleic Acids Res 2016;44(6):2989–98. https://doi.org/10.1093/nar/gkw047.
[23] Ebert MS, Neillson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods 2007;4(9):721–6. https://doi.org/10.1038/nmeth1079.
[24] Ebert MS, Sharp PA. Roles for MicroRNAs in conferring robustness to biological processes. Cell 2012;149(3):515–24. https://doi.org/10.1016/j.cell.2012.05.014.
[25] Ebert MS, Sharp PA. Emerging roles for natural microRNA sponges.Curr Biol 2010;20(19):https://doi.org/10.1016/j.cub.2010.08.052.
[26] Kartha RV, Sabal P, Dhanasekaran SM, et al. The landscape of antisense gene expression in human cancers. Genome Res 2015;25(7):1068–79. https://doi.org/10.1101/gr.177342.115.
[27] Salmena L, Polesino L, Tay Y, Kats L, Pandolfi PP. A cerRNA hypothesis: the rosetta stone of a hidden RNA language? Cell 2011;146(3):353–8. https://doi.org/10.1016/j.cell.2011.06.049.
[28] Qi J, Waterfall JJ, Lis JT. Nascent RNA Sequencing reveals widespread transcription initiation from cryptic sites. Science (80-) 2005;309(5740):1559–63. 309/5740/1559.
[29] Vítor AC, Sridharan SC, Sabino JC, et al. Single-molecule imaging of transcription at damaged chromatin. Sci Adv 2019;5(1):https://doi.org/10.1126/sciadv.aau1249.2022/05/16.
Koyosawa H, Yamanaka I, Osato N, et al. Antisense transcripts with FAN10MT2 clone set and their implications for gene regulation. Genome Res 2003;13(8):1324–34. https://doi.org/10.1101/gr.933303.

Kadowaki S, Honma Y, Kucukarslan T, et al. Antisense transcription in the mammalian transcriptome. Science 2005;309(5740):1564–6. https://doi.org/10.1126/science.1117209.

Churchill GA, Rutledge J, Naslund JS. Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature 2011;469(7330):368–73. https://doi.org/10.1038/nature09857.

Souza-Luís R, Dujardin G, Zukher I, et al. POINT technology illuminates the processing of poly(A)-associated intact nascent transcripts. Mol Cell 2021;81(9):1935. https://doi.org/10.1016/j.molcel.2021.02.034.

Guttman M, Amit I, Garber M, et al. An antisense transcript mediates posttranscriptional response to DNA damage in breast cancer. BMC Cancer 2019;19(1):1–11. https://doi.org/10.1186/s12885-019-5662-3.

Zhang W-T, Zhang J-J, Shao Q, et al. FGD3-AS1 is an oncogenic IncRNA in pancreatic cancer and regulates the Wnt/-catenin signaling pathway via miR-545-5p. Oncol Rep 2019;41(7):2471–8. https://doi.org/10.3892/or.2019.9739.

Gao KF, Zhao YF, Liao WJ, Xu GL, Zhang JD. CERS6-AS1 promotes cell proliferation and represses cell apoptosis in pancreatic cancer via miR-195-5p/ WIF2 axis. Kaohsiung J Med Sci 2022;38(1):1–12. https://doi.org/10.1053/j.kjms.2020.12.009.

Kim T, Xu Z, Clauder-Münster S, Steinmetz LM. The powerful world of antisense oligonucleotides: From bench to bedside. Wiley Interdiscip Rev RNA 2016;7:11. https://doi.org/10.1002/wrna.1461.

Robertson G, Hirst M, Bainbridge M, et al. Genome-wide profiles of STAT1 phosphorylation reveal a role in gene expression regulation. Science (80-) 2003;302(5648):1212–5. https://doi.org/10.1126/science.1090923.

Shademan M, Salanghuch AN, Zare K, et al. Expression profile analysis of two antisense IncRNAs to improve prognosis prediction of colorectal adenocarcinoma. Cancer Cell Int 2019;19(1):1–12. https://doi.org/10.1186/s12935-019-0635-1.

Wenisch S, Elguindi S, Cabeza JH, et al. Transcriptome-wide analysis of natural antisense transcripts shows their potential role in breast cancer. Sci Reports 2020;10(1):1–16. https://doi.org/10.1038/s41598-020-65489-z.

Ros G, Pego-Soru S, De Angélica P, et al. HMGAA2 antisense long non-coding RNAs as new players in the regulation of HMGAA2 expression and pancreatic cancer promotion. Front Oncol 2020;10:9. https://doi.org/10.3389/fonc.2020.00152.

Faghihi MA, Wahlestedt C. RNA interference is not involved in natural antisense mediated regulation of gene expression in mammals. Genome Biol 2006;7(5):R38. https://doi.org/10.1186/gb-2006-7-5-r38.

Chen CB, Berx G. Regulatory networks defining EMT during cancer initiation and progression. Oncogene 2017;36(34):4822–34. https://doi.org/10.1038/onc.2017.105.

Wang Y, Li H, Zhang D, et al. LncRNA CASC20 promotes chemoresistance in colorectal cancer. J Exp Clin Cancer Res 2020;39(1):2. https://doi.org/10.1186/s13046-019-1505-4.

Curcio CA, Ruskin MA, Capela AM, et al. Computational and Structural Biotechnology Journal 2020;18(1):253–65. https://doi.org/10.1515/med-2022-0418.

Shi T, Li R, Duan P, et al. TRPM2-AS promotes paclitaxel resistance in prostate cancer. J Cell Physiol 2017;232(5):3233–47. https://doi.org/10.1002/jcp.25723.

Li L, Ma Y, Maerkeya K, Reyanguly D, Han L. LncRNA OIP5-AS1 regulates the kinetics. Cell 2012;150(6):1158–69.https://doi.org/10.1016/j.cell.2012.08.016.

Levin JZ, Yassour M, Adiconis X, et al. Comprehensive comparative analysis of transcriptome-wide IncRNA expression in human diploid fibroblasts. Cell Syst 2015;5(1):22–33. https://doi.org/10.1016/j.cels.2015.12.006.

Kung JTY, Colognori D, Lee JT. Long noncoding RNAs: Past, present, and future. Mol Sci 2021;22(20).https://doi.org/10.3390/ijms222011168.

Liang D, Tan P, Pei G, et al. Non-coding RNA MF2-AS1 promotes colorectal cancer cell proliferation, migration and invasion through miR-574-5p/MYCBP axis. Cell Prolif 2019;52(4).https://doi.org/10.1111/cpr.12632.

Kim J, Hu C, Moufawad El Achkir CA, et al. Patient-customized oligonucleotide therapy for a rare genetic disease. Nat Med 2019;25(6):1644–52. https://doi.org/10.1038/s41591-019-0612-8.

Betram M, Grujovic D, Peck G, et al. A pre-existing population of ZEB2+ quiescent cells with stemness and mesenchymal features dictate chimeroma in colorectal cancer. J Exp Clin Cancer Res 2020;39(1):2. https://doi.org/10.1186/s13046-020-1330-7.

Satija R, Dhopte N, Grujovic D, et al. Point technology illuminates the powerful world of antisense oligonucleotides: From bench to bedside. Wiley Interdiscip Rev RNA 2016;7(2):110. https://doi.org/10.1002/wrna.12522.

Shi T, Li R, Duan P, et al. TRPM2-AS promotes paclitaxel resistance in prostate cancer. J Cell Physiol 2020;235(9):914–24. https://doi.org/10.1002/jcp.28377.

Tan P, Pei G, et al. Non-coding RNA MF2-AS1 promotes colorectal cancer cell proliferation, migration and invasion through miR-574-5p/MYCBP axis. Cell Prolif 2019;52(4).https://doi.org/10.1111/cpr.12632.

Kim J, Hu C, Moufawad El Achkir CA, et al. Patient-customized oligonucleotide therapy for a rare genetic disease. Nat Med 2019;25(6):1644–52. https://doi.org/10.1038/s41591-019-0612-8.
