Long non-coding RNAs (lncRNAs) are a class of intensely studied, yet enigmatic molecules that make up a substantial portion of the human transcriptome. In this work, we link the origins and functions of some lncRNAs to retroposition, a process resulting in the creation of intronless copies (retrocopies) of the so-called parental genes. We found 35 human retrocopies transcribed in antisense and giving rise to 58 lncRNA transcripts. These lncRNAs share sequence similarity with the corresponding parental genes but in the sense/antisense orientation, meaning they have the potential to interact with each other and to form RNA:RNA duplexes. We took a closer look at these duplexes and found that 10 of the lncRNAs might regulate parental gene expression and processing at the pre-mRNA and mRNA levels. Further analysis of the co-expression and expression correlation provided support for the existence of functional coupling between lncRNAs and their mate parental gene transcripts.

Key words: lncRNAs, long non-coding RNAs, retroposition, retrocopies, antisense transcription, RNA:RNA duplexes

Received: 04 June, 2016; revised: 30 June, 2016; accepted: 19 July, 2016; available on-line: 02 November, 2016

INTRODUCTION

In higher eukaryotes, non-coding RNAs, such as miRNAs (microRNAs) and lncRNAs (long non-coding RNAs), represent considerable portions of the transcriptome, with the latter class represented by 28,031 transcripts in humans (Ensembl 83), compared to 79,930 protein-coding transcripts. Some other sources provide even higher numbers of these RNAs, such as NONCODE (Zhao et al., 2016), with 141,353 lncRNAs. This abundance of lncRNAs sparked interest in deciphering their functions, origins, and evolution. However, the tasks appear to be even more demanding than in the case of protein-coding genes, and as a result, the vast majority of lncRNAs has no biological role assigned. In particular, due to poor evolutionary conservation of their sequences, homology-based functional assignment can be applied to only a small subset of lncRNAs. Additionally, detailed studies of the selected lncRNAs, such as HOTAIR (Tsai et al., 2010), ANRIL (Yap et al., 2010), and ZEB2-NAT (Beltran et al., 2008), indicate a high heterogeneity of their modes of action, making the in silico functional studies quite inaccurate. The accumulated data associate lncRNAs with biological processes such as transcription, splicing, translation, protein localization, cell cycle and apoptosis. They have also been linked to a number of human diseases, including cancers. It is possible that a large portion of lncRNAs has no biological role and represent a mere transcriptional noise or that the act of their transcription itself has a biological meaning, rather than their sequence does (Kornienko et al., 2013). Regarding the modes of action, a number of scenarios has been proposed, with transcriptional regulation being the best studied and being achieved through several mechanisms, such as promoter modifications, creating a permissive chromatin environment or binding transport factors to inhibit the nuclear localization of specific transcription factors (Kugel & Goodrich, 2012). In contrast to transcription-related mechanisms, little is known about the roles lncRNAs play upon base-pairing with fully or partially complementary mate miRNAs. In that scenario, lncRNAs could affect the stability, processing and expression levels of other transcripts (Geisler & Coller, 2013). One possibility is modulating the pre-mRNA splicing by splice site masking and subsequent blocking of the spliceosome assembly, which requires an extensive complementarity with a regulated pre-mRNA molecule. Such complementarity occurs by definition between the natural antisense transcripts (antis-NATs), but interactions in trans are also possible. Several lncRNAs are known to be involved in this type of regulation. For instance, it was shown that NATs influence the splicing patterns of mRNAs at the neuroblastoma MYC, c-ErbAalpha and ZEB2 loci in mammals (Beltran et al., 2008). In the case of neuroblastoma MYC and c-ErbAalpha, this was suggested to be achieved through formation of RNA:RNA duplexes, which then inhibit splicing. At the ZEB2 locus, lncRNA expression inhibits splicing of an intron that contains an internal ribosome entry site (IRES). Translation of ZEB2 relies on this IRES; therefore, expression of the NAT indirectly facilitates expression of ZEB2 protein. In addition to splicing modulation, other regulatory mechanisms triggered by lncRNA:RNA duplexes are possible in humans, and they include adenine to inosine RNA editing at dsRNA regions, mRNA stability control by abrogation of miRNA-induced repression and guiding protein-coding genes to degradation within a Staufen-mediated decay (SMD) pathway (Geisler & Coller, 2013). Recently, the potential...
of lncRNAs to exert regulatory roles through RNA:RNA base-pairings has been assessed for the human transcriptome (Szczesniak and Makalowska, 2016) and for several model plant species (Szczesniak et al., 2016).

In this work, we scanned the human lncRNAs to identify those transcribed in antisense to retroposition-derived copies (retrocopies) of protein-coding genes. In retroposition, an mRNA molecule is reversely transcribed into cDNA, which occasionally becomes inserted into the genome at a random location (Fig. 1A). The resulting new copy (retrocopy) of the so-called parental gene typically is not functional because it lacks the core promoter and other regulatory sequences that would enable its transcription. In some cases, however, retrocopies use upstream promoters, either new ones (exaptation of cryptic promoter sequences) or the ones from other genes; such new genes, called retrogenes, constitute ca. 7.4% of the human gene set. They might evolve functions other than those of parental genes (neofunctionalization), play the same roles but with different spatiotemporal pattern (subfunctionalization) or replace the parental gene (orphan retrogenes) (Ciomborowska et al., 2013). Finally, some of them are transcribed from the antisense strand, resulting in production of long non-coding RNAs, as shown in this study. These lncRNAs are expected to have functions other than the corresponding retrocopies or parental genes due to the lack of sequence similarity in the sense/sense orientation. A key to understanding their functions might be the observation that as a consequence of their origin, these lncRNAs are fully or partially complementary to their parental genes and are thus able to interact with each other at the RNA level. Keeping this in mind, we performed in silico base-pairing of the antisense lncRNAs with their parental genes and tried to determine whether these results could be linked to the abovementioned functions of RNA:RNA interactions. We found 10 lncRNAs transcribed in antisense to retrocopies and predicted to modulate processing and expression of their parental genes (Suppl. Table 1 at www.actabp.pl). The subsequent analysis of co-expression, expression correlation and sequence conservation led us to the conclusion that retroposition, already known to be one of the most important processes shaping mammalian genomes, might also contribute to the evolution of antisense lncRNAs and be a key to understanding their biological roles.

MATERIALS AND METHODS

Data download. The GENCODE 24 (Derrien et al., 2012) annotation data for human (Homo sapiens), mouse (Mus musculus) and chimp (Pan troglodytes) were downloaded from the Ensembl release 83 (Herrero et al. 2016) using BioMart. To obtain long non-coding RNAs, only sequences classified as 3prime_overlapping_ncrna, antisense, lincRNA, macro_lncRNA, retained_intron, sense_intron, or sense_overlapping were kept. Retrocopy-annotated data were obtained from the Retropredatabase (Kabza et al., 2014). The retrocopies that are known to exist in Ensembl and have assigned Ensembl gene IDs were mapped to Ensembl release 83 using the biomaRt R package, which enabled access to updated, cross-reference information on the genes, including transformation of the genomic coordinates from the human genome version hg19 to hg38. The retrocopy genes that are present only in the retropredatabase were transformed into hg36 coordinates using the LiftOver tool available at the UCSC Genome Browser website (Speir et al., 2016). Retrocopy-
Vol. 63  Retroposition and long non-coding RNAs: possible functional links

- max_target_seqs 1, -outfmt 6, -evalue 1e-5. Additionally, the PFAM profile-HMM database was searched with hmmscan from the HMMER-3 package (http://hmmer.org/) to identify common protein domains. In the third step, the TransDecoder.predict utility was run to obtain only high-confidence proteins based on the BLASTP and hmmscan results, as well as a built-in model for protein classification. Sequences that passed all filtering steps and were not recognized as protein-coding by TransDecoder were classified as long non-coding RNAs.

**Expression analysis.** To identify IncRNAs co-expressed with the corresponding parental genes in humans, expression data from ENCODE at the gene level was used (Suppl. Table 2 at www.actabp.pl). In this analysis, only gene pairs with expression values >0.1 TPM in at least one sample were considered to be co-expressed. Expression correlation analysis was performed in R using the same data and requiring that the Spearman's rank correlation coefficient was >0.6 or <-0.6 and the p-value <0.05. Both genes were required to have expression values of 0.2 TPM or higher; otherwise, that particular sample was removed from the correlation testing.

**Identification of IncRNA-RNA interactions and their possible functions.** The IncRNA interactions with their parental genes were predicted using a recently described strategy that is proven to achieve good performance and is able to identify experimentally validated RNA:RNA duplexes (Szczesniak & Małaśkowska, 2016). Briefly, it uses 

**RESULTS AND DISCUSSION**

In this work, we took a closer look at IncRNAs transcribed in antisense to retrocopies, focusing on possible RNA:RNA interactions between them and the corresponding parental genes, both at the mRNA and pre-mRNA levels (Fig. 1B, C). To achieve this, we first...

---

**Figure 1.** (A) Schematic representation of the retroposition process. (B) Mechanism behind the creation of IncRNA base-pairings with parental genes at the mRNA level. Once a retrocopy is transcribed in the antisense orientation, the resulting IncRNAs share sequence similarity with the parental genes in the sense/antisense orientation, meaning they are able to interact and form RNA:RNA duplexes with possible regulatory implications. (C) Evolutionary mechanism that enables the formation of IncRNA interactions with the pre-mRNA of parental genes. A retrocopy is created from one of many splice forms of the parental gene. Its antisense IncRNAs are complementary to the pre-mRNAs of the parental gene. Although retrocopies typically are devoid of introns, some of the retroposition-derived IncRNAs are able to base-pair with intronic parts of the parental gene’s pre-mRNAs and mask the intronic splicing signals. This is possible if the process of retroposition and the formation of IncRNA duplexes engages different splice forms of the parental gene, as shown in the figure.
collected a set of 4,675 human retrocopies from retrogeneDB and 57,145 lncRNAs (25,296 genes) from Ensembl. Using BEDTools *intersect*, we found 58 lncRNAs that were transcribed in antisense to 35 retrocopies. With Ensembl’s annotation data for chimp, we found no antisense lncRNA-retrocopy overlaps. We attributed this to the quality of the available data; therefore, we re-annotated the chimp transcriptome, taking advantage of stranded RNA-Seq data available in the NCBI’s Sequence Read Archive database (Suppl. Table 3 at www.actabp.pl). Altogether, we identified 167,182 transcripts belonging to 104,272 genes, including 36,010 lncRNA transcripts (14-fold more than in Ensembl). With these new data, we discovered 23 antisense lncRNA-retrocopy overlaps. We also identified 6 antisense overlaps in mouse (Ensembl’s annotation data).

**Antisense transcripts of retrocopies are quite poorly conserved**

The evolutionary conservation of human lncRNA transcription in antisense to retrocopies was tested by comparing human cases with the corresponding chimp and mouse homologs for retrocopies and checking whether there is antisense lncRNA transcription, like in human. We found 8 homologs in mouse and 6 in chimp; however, only one, in mouse, had antisense lncRNAs. We assumed this observation could be partially attributed to poor annotation of lncRNAs in chimp and mouse because Ensembl has only 2,586 chimp lncRNAs, as opposed to 28,031 for human. To obtain more reliable cross-species comparison, we performed de novo assembly of the chimp transcriptome using an extensive set of 19 stranded RNA-Seq libraries, followed by lncRNA identification, which resulted in a set of 36,010 lncRNAs. With this new dataset, we found 23 retrocopies with antisense lncRNAs, as opposed to no cases found for the Ensembl data. However, none of them were conserved in humans, showing that antisense transcription of retrocopies is poorly conserved across the analyzed species. This is not surprising because a large fraction of lncRNAs represents species-specific transcripts, and approximately 60–70% are not detectable outside of primates (Neacsulea et al., 2014; Washietl et al., 2014; Derrien et al., 2012). However, the poorly resolved orthology relations for retrocopies and their relatively low conservation across species are also a factor: we were able to find 1-to-1 orthologs in chimp and mouse only for ca. 20% of all human retrocopies, which considerably reduced the chances for finding conserved antisense lncRNA-retrocopy pairs. Considering the facts listed above, we manually checked all previously identified 1-to-1 orthologs for human retrocopies with antisense lncRNAs and found a mouse ortholog of the DNAJB8 retrocopy, which also had lncRNAs transcribed in antisense. We used Clustal Omega (Sievers & Higgins, 2014) to align these antisense transcripts with the corresponding human lncRNA and found that the sequence identity is only 46%. Moreover, both mouse antisense RNAs overlap the translated sequence of DNAJB8, while human antisense transcript overlaps only a 5'UTR region (Fig. 3). These observations led us to the startling conclusion that orthologous retrocopies might possess antisense lncRNAs that originated independently and therefore are not orthologous.

**Selected antisense lncRNAs show correlation of expression with their parental genes**

Considering the lack of conservation of antisense lncRNA-retrocopy pairs, we aimed to provide more support for the supposed functionalities by analyzing the expression values of RNAs that are expected to interact. First, we checked whether they are co-expressed by analyzing human expression data from 153 strand-specific
Retroposition and long non-coding RNAs: possible functional links

RNA-Seq libraries from ENCODE, and we found that 27 of 35 lncRNA genes are co-expressed with their parental genes (Fig. 4, Suppl. Table 4 at www.actabp.pl). Importantly, two RNAs are expected to be co-expressed if they interact in a cell, but co-expression itself does not imply they base-pair, for instance, they could be expressed in different cellular compartments. We therefore hypothesized that functionally coupled lncRNAs and parental genes should, in addition to being co-expressed, show some level of expression correlation. Therefore, we calculated the Spearman Rho correlation coefficient for co-expressed pairs. Requiring the correlation coefficient to be greater than 0.6 or less than -0.6 with a p-value <0.05, we found two pairs with statistically significant positive expression correlation and one pair with negatively correlated expression (Table 1). Using R’s uniReg package, isotonic regression models were built for the AC021224.1-HNRNPA1 and RP11-3P17.5-RPL23A correlated pairs and an antitonic regression model was constructed for the RP11-78A19.3-CHMP1A pair (negatively correlated) (Fig. 5A, B and C, respectively). These results provide indirect evidence for the functionality of these three cases. The remaining lncRNA:RNA pairs may not be functional or alternative scenarios could apply, for example: i) the transcripts are co-expressed in a small subset of samples, producing statistically insignificant results for correlation testing, ii) some of the modes of action, such as splicing modulation or triggering mRNA editing events, do not involve changes in gene expression levels; thus, one does not expect to observe (anti-) correlation of expression, and iii) other factors, such as miRNAs and transcription factors, being involved in the regulatory processes.

Functional insights

Next, we identified possible base-pairings between lncRNAs transcribed in antisense to retrocopies and the parental genes using a previously proposed procedure (Szczesniak & Makałowska, 2016). The subsequent analysis of the RNA:RNA duplexes revealed 10 lncRNAs with potential regulatory roles exerted on their parental genes (Suppl. Table 1 at www.actabp.pl), which included stability control (masking miRNA target sites, guiding to the SMD pathway), pre-mRNA processing (modulating alternative splicing) events and mRNA processing (RNA editing). Three previously described examples with statistically significant correlations of expression were among those pairs with possible base-pairings. Therefore, we focused on them in further analysis. These cases include the following parental genes: hnRNPA1, CHMP1A, and RPL23A.

hnRNPA1

hnRNPA1 belongs to the A/B subfamily of ubiquitously expressed heterogeneous nuclear ribonucleoproteins (hnRNPs), RNA-binding proteins that associate with pre-mRNAs in the nucleus and influence pre-mRNA processing, as well as other aspects of mRNA metabolism and transport (Han et al., 2010). It represents one of the most abundant core proteins of hnRNP complexes and plays a key role in the regulation of alternative splicing (Mayeda et al., 1998). Overexpressed hnRNPA1 effectively downregulates the expression of the transcriptional transacti-

Table 1. Summary of antisense lncRNA and parental gene pairs with statistically significant expression correlation.

<table>
<thead>
<tr>
<th>Ensembl gene IDs</th>
<th>Gene symbols</th>
<th>Correlation measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSG00000262477</td>
<td>DNAJB8-AS1-001 &gt; antisense</td>
<td>Spearman Rho: 0.77, p-value: 0.008, No. of samples: 11</td>
</tr>
<tr>
<td>ENSG000000198242</td>
<td>RPI1-3P17.5-RPL23A</td>
<td>Spearman Rho: 0.7, p-value: 0.0002, No. of samples: 23</td>
</tr>
<tr>
<td>ENSG000000131165</td>
<td>RP11-78A19.3-CHMP1A</td>
<td>Spearman Rho: -0.65, p-value: 0.0008, No. of samples: 23</td>
</tr>
</tbody>
</table>
vator Tat, which in HIV-1 infected cells results in a sharp reduction in the transcription of the viral genome and a 100-fold drop in the production of new HIV-1 virions (Jablonski & Caputi, 2009). As many as 66 retrocopies across the human genome can be found for the hnRNPA1 gene at retrogeneDB (Kabza et al., 2014). One of them, retro_hsap_1933, has an antisense transcript ENST00000573479, also known as AC021224.1-201 or NONHSAT058863.2, that is classified as a long non-coding RNA at NONCODE (Zhao et al., 2016). One of the hnRNPA1 splice variants, ENST00000547276, lacks domains necessary for the major functions of hnRNPA1 (Fig. 6B), i.e., those required for alternative splicing activity, stable binding of RNAs and optimal RNA annealing (Mayeda et al., 1994). This isoform, however, plays regulatory roles in HIV-1 splicing and replication. Our bioinformatics predictions link the generation of this splice form to the absence of IncRNA:RNA base-pairing, which normally would lead to masking of the 5' splice site in the 6th intron and emergence of longer isoforms with extended functionality (Fig. 6). Importantly, the IncRNA and parental gene display a statistically significant correlation of expression, with a Spearman Rho coefficient of 0.77 (Fig. 5A), which supports the idea of their functional coupling.

![Figure 5. (A) Isotonic regression model of co-expression for AC021224.1 and HNRNPA1. (B) Isotonic regression model of co-expression for RP11-3P17.5 and RPL23A. (C) Antitonic regression model of co-expression for RP11-78A19.3 and CHMP1A.](image)

Figure 5. (A) Isotonic regression model of co-expression for AC021224.1 and HNRNPA1. (B) Isotonic regression model of co-expression for RP11-3P17.5 and RPL23A. (C) Antitonic regression model of co-expression for RP11-78A19.3 and CHMP1A.

![Figure 6. (A) Exon-intron structures of the two main splice forms of HNRNPA1 and a short isoform (ENST00000547276); the predicted base-pairing region between the transcripts and IncRNA (AC021224.1-201) overlaps the 5' splice site in the 6th intron and is marked with a green line. (B) Schematic representation of the hnRNPA1 protein domain structure, corresponding to these three HNRNPA1 transcripts (from Jean-Philippe et al., 2014, modified). The product of ENST00000547276 lacks an RGG-Box and a portion of the prion-like domain.](image)

Figure 6. (A) Exon-intron structures of the two main splice forms of HNRNPA1 and a short isoform (ENST00000547276); the predicted base-pairing region between the transcripts and IncRNA (AC021224.1-201) overlaps the 5' splice site in the 6th intron and is marked with a green line. (B) Schematic representation of the hnRNPA1 protein domain structure, corresponding to these three HNRNPA1 transcripts (from Jean-Philippe et al., 2014, modified). The product of ENST00000547276 lacks an RGG-Box and a portion of the prion-like domain.
CHMP1A

CHMP1A has two retrocopies in humans (Kabza et al., 2014). One is retro_hsap_75 with antisense transcript ENST00000586474 that is also known as RP11-78A19.3-001, and is classified as a long non-coding RNA at NONCODE. CHMP1A encodes a member of the CHMP/Chmp family of proteins, which are involved in multivesicular body sorting of proteins to the interiors of lysosomes (Howard et al., 2001). Overexpression of CHMP1A in cultured cells leads to gene silencing due to interaction with BMI1 transcriptional repressor and the effect on the chromatin structure (Stauffer et al., 2001). Recent studies link CHMP1A to tumor development because the gene is differentially expressed in diverse tumor types (Li et al., 2008; You et al., 2012). For instance, shRNA knockdown of CHMP1A expression in HEK 293T cells results in increased anchorage-independent growth in vitro and tumor formation in vivo (Li et al., 2008); on the other hand, overexpression of CHMP1A inhibits the growth of pancreatic cancer cells in vitro (Li et al., 2008). Moreover, CHMP1A overexpression suppresses the proliferation of renal carcinoma cells in vitro and leads to suppressed tumor growth of rat renal carcinoma cells in vivo, while inhibition of CHMP1A expression has no effect on tumor cell growth (You et al., 2012).

dsRNA formed by the CHMP1A gene and IncRNA may have some functions related to CHMP1A activity in tumors. Our expression analysis shows that the antisense RNA (RP11-78A19.3-001) is co-expressed with the parental gene (CHMP1A) in 24 of 300 samples, mostly in tumor samples, such as HT1080, A172, SK-MEL-5, and K562. However, with the current knowledge, it is impossible to speculate the relevance of the strong negative correlation between the two genes (Spearman Rho of -0.65 and p-value 0.0006; Fig. 7), especially because we were unable to link their RNA:RNA base-pairing to any of the mechanisms considered in this study.

RPL23A

RPL23A encodes a ribosomal protein that is part of the 60S subunit. This gene contains antisense transcripts that mediate downregulation of RPL23A expression in IFN-b-treated cells; they were identified de novo in tumor cells and were confirmed by northern blot and RT-PCR assays (Jiang et al., 1997). RPL23A has as many as 68 retrocopies across the human genome (Kabza et al., 2014), and three of them, namely, retro_hsap_1775, retro_hsap_2021, and retro_hsap_2874, have antisense IncRNAs that are co-expressed with RPL23A (Table 2). We found that these three IncRNA transcripts show elevated expression in two cell lines: K562 (derived from erythroleukemia cells) and GM12878 (from normal lymphoblastoid cells). AC016629.3, an IncRNA, shows the highest expression in K562 samples derived from a female patient with chronic myelogenous leukemia (ca. 27 TPM), in contrast to GM12878 and other non-cancer cell lines, where its expression is <2 TPM. Our analysis of IncRNA:RNA duplexes shows that AC016629.3
might mask miRNA target sites in seven splice forms of RPL23A. Another lncRNA, RP11-3P17.5, is co-expressed with RPL23A in 21 samples, more than half of which are cancer-related. The analysis of expression values showed strong correlation between the two genes, with a Spearman Rho of 0.70 and p-value of 0.0002. An isotonic regression model was built and is presented in Fig. 5B.

Recent research links functions of antisense lncRNAs to their mate retrocopies and parental genes

A growing body of evidence shows that retrocopies play significant biological roles and are also key players in genome evolution (Szczéśniak et al., 2012; Ciomborowska et al., 2013; Navarro & Galante 2015). A number of them constitute long non-coding RNAs (less than 3% of RetrGeneDB retrocopies have protein_coding status in Ensembl and half of them possess premature stop codons and/or frame-shifts compared with the coding sequences of their parental genes), making retroposition a significant source of lncRNAs. Recent studies revealed that retrocopies often express antisense RNAs (asRNAs), which are active regulators of their sense counterparts through transcriptional and post-transcriptional mechanisms. They were shown to participate in controlling the promoters and transcription of the retrocopies (Morris et al., 2008), while suppression of these asRNAs results in transcriptional activation of the retrocopies (reviewed in Weinberg & Morris, 2013). Finally, lncRNAs transcribed in antisense to the retrocopies might act in trans and contribute to regulation of the parental genes. For example, PTEN, a tumor-suppressor gene, is under control of its retrocopy, PTENpogl (Johnson et al., 2013). PTENpogl has two antisense RNAs, α and β, which regulate PTEN transcription and the stability of its transcripts. The α isomorph functions in trans and epigenetically modulates PTEN by recruiting a DNA methyltransferase, while the β isoform interacts with PTENpogl through RNA:RNA base-pairing, which affects the stability of sense PTENpogl and thus enables its sponge activity.

A large-scale analysis of antisense lncRNAs in a recent study (Milligan et al., 2016) found 2277 loci containing exon-to-exon overlaps between long non-coding RNAs and pseudogenes. This dataset included retrocopies and other pseudogenes, such as processed and unprocessed pseudogenes from Ensembl. Further analysis of the full-length cDNAs and ESTs that supported 313 pseudogene-lncRNA overlaps indicated that this phenomenon is prevalent. The use of EST/cDNAs as transcriptional evidence represents a conservative approach, and many more cases likely exist. The subsequent comparison of the parental genes of the pseudogenes to all human genes showed enrichment of several ontology categories; however, no insight into the biology behind these findings was provided. In particular, the biological processes and mechanisms that could possibly underlay the hypothesized lncRNA-pseudogene associations were not investigated. To the best of our knowledge, we performed this type of assessment for the very first time. Our findings suggest that retroposition-derived, antisense lncRNAs might affect the expression and processing of parental genes in a number of ways, which is supported by the in silico base-pairing of the RNA molecules, followed by computational function assignment, co-expression data and, occasionally, correlation of expression and evolutionary conservation.

FINAL REMARKS

In this work, we analyzed the potential roles of retroposition-derived lncRNAs in regulating the expression and processing of the corresponding parental genes. We assumed directionality of the regulatory effect, although the reverse scenario is possible, with lncRNAs being affected by the transcripts of parental genes. Additionally, the antisense lncRNAs could regulate paralogs of the parental genes or any other gene with sufficient sequence similarity, which was not considered in this study. In particular, in cis effects are possible between lncRNAs and retrocopies expressed from the opposite strand because antisense transcripts are expected to base-pair easily due to their 100% sequence similarity (in sense/antisense orientation), and they occupy the same genomic locus, which further facilitates contact between the RNA molecules. Importantly, scenarios other than RNA:RNA interactions are possible, including transcription-dependent and transcription-independent mechanisms leading to chromatin remodeling, which has already been the subject of a number of studies (reviewed in Geisler & Coller, 2013; Milligan & Lipovich 2015).

Acknowledgements

This work was supported by the National Science Centre (grant No. 2014/15/D/NZ2/00525 to M.W.S.; grant No. 2013/11/B/NZ2/02598 to I.M.), the Foundation for Polish Science (START Scholarship grant editions 2014/2015 and 2015/2016 to M.W.S.), and the KNOW Poznan RNA Centre (grant No. 01/KNOW2/2014).

REFERENCES
