



# Challenging paradigms: long non-coding RNAs in breast ductal carcinoma *in situ* (DCIS)

Mary A. Kosir<sup>1,2\*</sup>, Hui Jia<sup>3</sup>, Donghong Ju<sup>1</sup> and Leonard Lipovich<sup>3</sup>

<sup>1</sup> Department of Surgery, Wayne State University School of Medicine, Detroit, MI, USA

<sup>2</sup> Department of Oncology, Karmanos Cancer Institute, Wayne State University School of Medicine, Detroit, MI, USA

<sup>3</sup> Center for Molecular Medicine and Genetics, and Department of Neurology, Wayne State University School of Medicine, Detroit, MI, USA

\*Correspondence: kosirm@karmanos.org

## Edited by:

Rami I. Aqeilan, Hebrew University-Hadassah Medical School, Israel

## OVERVIEW

The continuing challenges to the treatment of breast ductal carcinoma *in situ* (DCIS) warrant a new avenue of investigation. In parallel, a startling revelation of the first post-genomic decade has been the rising prominence of long non-coding RNAs (lncRNAs) as key gene regulators, including in breast cancer. High-throughput experimental surveys of the human transcriptome show that lncRNAs are abundant. In contrast to microRNAs (miRNAs), lncRNAs possess multiple, diverse, heterogeneous molecular mechanisms of action making them high-level master regulators of gene regulatory networks and pathways in cancer. In particular, lncRNAs directly regulate the activity of transcription proteins and epigenetic modifiers, which, in turn, govern their own downstream sets of target genes. Therefore, provided that such lncRNA regulatory activities are functional in cancer, including DCIS, targeting the lncRNAs or the regulatory networks which they mediate can represent an important new avenue of treatment. Development of lncRNA-targeting therapies may emerge into a key innovation generated in clinical and translational research by the innovative use of the latest human genome annotation datasets combined with high throughput transcriptome analyses. We propose a paradigm shift: direct interrogation of lncRNAs in the molecular characterization of DCIS subtypes, including the idea that lncRNAs may be the trigger for progression from *in situ* to invasive cancer.

## CHALLENGES OF DCIS

The National Institutes of Health recommendations for research on DCIS (Allegra et al., 2010) remain active at this time.

The challenges that make DCIS intractable to treatment have been the heterogeneity of DCIS and its shared expression of molecules with invasive breast cancer. While the phenotype is preinvasive, the proliferative, angiogenic, and even invasive characteristics of DCIS are shared with its invasive counterpart. Kretschmer et al. (2011) have identified overexpressed genes in both DCIS and invasive breast cancer. More recently, non-coding RNA has become a focus of interest as this might organize (or disorganize) the expression of molecules for the DCIS phenotype. Hannofon et al. (2011) evaluated miRNAs expressed in the change from “normal” to preinvasive DCIS. However, there are two reasons for the importance of characterizing and functionally interrogating lncRNAs in DCIS. miRNAs are primarily post-transcriptional suppressors, but lncRNAs allow us to access a diversity of epigenetic, co-transcriptional, and post-transcriptional mechanisms of gene regulation, which are not used by miRNAs. Secondly, lncRNAs themselves represent direct functional targets, while miRNAs are not because each microRNA is a post-transcriptional suppressor of hundreds or thousands of potential distinct gene targets. So miRNAs are not necessarily useful themselves as direct targets for therapies, although they may lead to mRNAs of target genes. lncRNAs are frequently involved in regulatory modalities, such as unique protein–RNA interactions in ribonucleoprotein complexes (RNP), or sense–antisense RNA regulation, which are highly specific and therefore lack counterparts in the microRNA world. Because lncRNAs are high-level regulators of gene expression in cancer, targeting lncRNAs may likely allow therapeutics directed

toward a cause, rather than toward a consequence, of the regulatory networks operant in DCIS (Lipovich et al., 2010). There is current and rapidly growing evidence supporting distinct lncRNA roles in breast cancer, and existing public gene expression datasets, including DCIS microarray data and RNAsequencing (RNAseq), can be interrogated for lncRNA differential expression.

## CURRENT STATE OF THE FIELD

Computational integration of the human genome with experimentally determined transcript sequences has revealed that half of human genes encode non-protein-coding RNAs (ncRNAs): miRNAs and the more numerous lncRNA transcripts. Key technological advances have led to large-scale cloning and sequencing of complementary DNAs (cDNAs), which are derived by reverse transcription from mature, full-length, 5'-capped, polyA-tailed RNAs, thereby validating, and inferring new gene structures empirically from cDNA to genome alignments, and to evaluate each experimentally documented transcript for protein-coding potential. This method was used by the Functional Annotation of Mouse (FANTOM) project to show that non-coding transcription is widespread, and that the number of lncRNA transcriptional units in the mammalian genome approximately equals that of protein-coding genes (Carninci et al., 2005; Carninci and Hayashizaki, 2007). Next Generation Sequencing (NGS) technologies, including RNAseq, have further increased the number of known lncRNAs (Cabili et al., 2011).

lncRNAs exert key regulatory functions, and control many normal and disease processes. Knockdown and

overexpression of specific lncRNAs generate reproducible phenotypes, because the lncRNAs regulate protein-coding genes, by multiple mechanisms (Lipovich et al., 2010). Many lncRNA mechanisms directly function in cancer. Specific known lncRNAs promote transformation as oncogene targets; are diagnostic markers and cell cycle regulators in cancer; mediate cell survival; regulate mRNA splicing; and directly regulate epigenetic and transcription factors (Lipovich et al., 2010). Gutschner and Diederichs (2012) have catalogued lncRNAs into cancer-specific activities which are relevant to DCIS such as sustained signaling for proliferation, immortality, angiogenesis, resisting cell death, avoiding suppressors. In more aggressive DCIS, there may also be activation of invasion and ultimately metastasis as DCIS progresses.

With a small number of lncRNAs with known functions, the abundance of lncRNA functions in the cancer literature is noteworthy. P53, perhaps the most famous tumor suppressor, is co-activated by MEG3, an imprinted lncRNA (Zhou et al., 2007), and in turn, regulates an additional lncRNA as a global effector of gene expression changes (Huarte et al., 2010). Another imprinted lncRNA, H19, is a direct target of C-Myc and is functional in cell proliferation (Barsyte-Lovejoy et al., 2006). Sense-antisense mRNA/lncRNA gene pairs affect several tumor suppressors which are cis-regulated by antisense lncRNAs transcribed from the same locus, including p15 and p21 (Lipovich et al., 2010). A prostate-specific lncRNA, PCGEM1, is associated with prostate cancer and regulates apoptosis (Fu et al., 2006), while RNAseq has uncovered a second lncRNA in prostate cancer progression (Prensner et al., 2011). Highly Upregulated in Liver Cancer (HULC), identified as the most-upregulated gene in an HCC microarray study of 7000 genes, is a canonically spliced, polyadenylated lncRNA whose knockdown in a cell line induces CDK8 and a tumor suppressor candidate (Panzitt et al., 2007). The nuclear lncRNAs NEAT1 and NEAT2, resident in and essential for nuclear paraspeckles and speckles respectively, have diverse cancer functions as well. NEAT2 (MALAT-1) is a marker of diverse human carcinomas, and a predictor of

survival and metastasis across multiple cancer types (Lin et al., 2007). Telomerase, a RNP with crucial functions in cancer, which contains a highly conserved lncRNA (Blackburn, 2005), became the subject of a Nobel Prize in 2009. Additional lncRNAs functional in proliferation and with anti-apoptotic roles are still being discovered (Hu et al., 2011), while other lncRNAs which have not yet been functionally validated are emerging as highly significant cancer biomarkers (Gloss et al., 2012).

There is increasing evidence for lncRNA function in human breast cancer. Specific lncRNAs are predictors of metastasis and survival; directly function in cell proliferation and cell cycle in breast cancer; and directly regulate nuclear hormone receptor function in breast cancer (Gupta et al., 2010; Lipovich et al., 2010; Silva et al., 2011). LSINCT5 is a polyadenylated lncRNA expressed at greater levels in breast and ovarian tumor tissues and cell lines relative to normal tissues. Knockdown of LSINCT5 decreases cellular proliferation (Silva et al., 2011). HOTAIR, an lncRNA whose expression is increased in primary breast tumors and their metastases, induces genome wide retargeting of the Polycomb repressor complex 2 (PRC2), which causes global gene expression changes that facilitate invasiveness and metastasis (Gupta et al., 2010). The lncRNA PINC, expressed in the regressed terminal ductal lobular unit-like structures of the mammary gland, is functional in cell survival and cell cycle progression (Ginger et al., 2006). The bifunctional lncRNA SRA is a co-activator of estrogen receptor alpha (Lipovich et al., 2010). Zfas1, another bifunctional lncRNA, which can serve as a snoRNA host transcript or as a stand alone lncRNA molecule, localizes to the ducts and alveoli of the mammary gland, and has an expression profile independent of the encoded snoRNAs, supporting its distinct role. Zfas1 is downregulated in breast tumors, potentially due to a tumor suppressor function (Askarian-Amiri et al., 2011).

### PERSPECTIVE ON FUTURE RESEARCH FOR DCIS

Despite the wealth of recent evidence for lncRNA function in cancer, including breast cancer, no studies to date have pinpointed lncRNAs as functional

determinants, biomarkers, or treatment targets in DCIS. A strategy to assess lncRNA differential expression in DCIS samples to identify functional lncRNAs includes the use of existing (catalog) microarray designs, custom microarrays, and RNAseq. Although RNAseq is unbiased, while custom microarrays can maximize the coverage of poorly-annotated lncRNAs, most DCIS gene expression profiling to date has used conventional microarray designs from established microarray manufacturers. We have recently performed genome wide identification of lncRNA genes from public ncRNA datasets and full-length cDNA data (Jia et al., 2010). To determine the extent to which existing commercial microarray platforms, such as the still widely used Affymetrix U133A and B chips, can be employed for lncRNA differential expression discovery, we performed genomic positional overlaps between commonly used commercial microarray probesets and our over 6700 lncRNAs. We show that 43% of our lncRNAs have some U133 probe-set representation, and list these lncRNAs in Supplemental Data Set 4 (Jia et al., 2010). In another disease system, we have validated the paradigm of interrogating existing Affymetrix U133 datasets to identify lncRNA differential expression (Michelhaugh et al., 2011). Here, we propose to apply the same paradigm to discover DCIS-specific lncRNA differential expression. We have previously used our annotation approach (which relates public microarray probes to genes which we know to be lncRNAs) to identify putative functional lncRNAs. Data re-mining of existing public cancer array datasets can easily identify lncRNA genes associated with the phenotypes or regulatory systems for DCIS. Depending on the sample properties and the experimental design of the DCIS microarray experiments whose results are in the public domain, specific lncRNAs can be identified directly from the public datasets, by determining which differentially expressed unannotated probesets in the public data reflect lncRNAs. These lncRNAs can be further validated by qRT-PCR to confirm their differential expression in other samples or in cell line models, and they can then be interrogated by reverse genetic

experiments in applicable cell line models (system perturbations: RNAi and overexpression) to elucidate any potential causal relationship to a mechanism or phenotype. Such validation of these lncRNAs as contributors to the pathogenesis of DCIS is a prerequisite for declaring them to be putative drug targets.

We have identified at least four studies which generated Affymetrix U133 microarray public data repositories from clinical DCIS samples (Schuetz et al., 2006; Turashvili et al., 2007; Hawthorn et al., 2010; Gabrovská et al., 2011). These studies included lncRNAs because 43% are represented by unannotated, accession-number-only, and “hypothetical-protein” gene probesets on the microarray. Schuetz et al. (2006) focused on differentially expressed genes marking the transition of stationary epithelial cells to migrating invasive cells, and used patient-matched DCIS/invasive ductal carcinoma (IDC) samples. Gabrovská et al. (2011) aimed to develop grade-specific gene expression signatures from paraffin-embedded IDC tissue samples, including benign tumors, and identified a larger set of unannotated gene signature members which can be mined for lncRNAs. Turashvili et al. (2007) identified genes whose expression differed between IDC and invasive lobular carcinoma, as well as between IDC and normal cells. Hawthorn et al. (2010) simultaneously examined genomic copy number variation and gene expression in IDC, highlighting genes whose transcript levels change as a result of genomic deletions.

## CONCLUSIONS

There is great potential in studying non-coding RNA in DCIS. In a paradigm-shifting approach, lncRNAs that are currently contained in published data sets, in particular the unannotated lncRNAs represented by Affymetrix probesets, are available to investigate DCIS. To the extent that the corresponding clinical datasets are represented by public microarray repositories, these lncRNAs can be utilized to further characterize the heterogeneous nature of DCIS, and to elucidate the mechanisms that support preinvasive DCIS and the progression to invasive breast cancer. The pace of future functional DCIS lncRNA discovery can

be substantially accelerated if unbiased whole-transcriptome differential expression profiling via RNAseq, further aided by including the diverse public lncRNA datasets such as those from ENCODE and the Broad Institute, becomes a reality in the analysis of DCIS clinical samples.

## ACKNOWLEDGMENTS

Supported by the Karmanos Cancer Institute/American Cancer Society IRG (Leonard Lipovich) and by the Fund for Medical Education and Research (FMRE) (Mary A. Kosir).

## REFERENCES

- Allegra, C. J., Aberle, D. R., Ganschow, P., Hahn, S. M., Lee, C. N., Millon-Underwood, S., et al. (2010). NIH state-of-the-science conference statement: diagnosis and management of ductal carcinoma *in situ* (DCIS). *J. Natl. Cancer Inst.* 102, 161–169.
- Askarian-Amiri, M. E., Crawford, J., French, J. D., Smart, C. E., Smith, M. A., Clark, M. B., et al. (2011). SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. *RNA* 17, 878–891.
- Barsyte-Lovejoy, D., Lau, S. K., Boutros, P. C., Khosravi, F., Jurisica, I., Andrulis, I. L., et al. (2006). The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res.* 66, 5330–5337.
- Blackburn, E. H. (2005). Telomeres and telomerase: their mechanisms of action and the effects of altering their functions. *FEBS Lett.* 579, 859.
- Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927.
- Carninci, P., and Hayashizaki, Y. (2007). Noncoding RNA transcription beyond annotated genes. *Curr. Opin. Genet. Dev.* 17, 139–144.
- Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., and Maeda, N. (2005). The transcriptional landscape of the mammalian genome. *Science* 309, 1559–1563.
- Fu, X., Ravindranath, L., Tran, N., Petrovics, G., and Srivastava, S. (2006). Regulation of apoptosis by a prostate-specific, and prostate cancer-associated noncoding gene, PCGEM1. *DNA Cell Biol.* 25, 135–141.
- Gabrovská, P. N., Smith, R. A., Tiang, T., Weinstein, S. R., Haupt, L. M., and Griffiths, L. R. (2011). Development of an eight gene expression profile implicating human breast tumours of all grade. *Mol. Biol. Res.* 39, 3879–3892.
- Ginger, M. R., Shore, A. N., Contreras, A., Rijnkels, M., Miller, J., Gonzalez-Rimbau, M. E., et al. (2006). A noncoding RNA is a potential marker of cell fate during mammary gland development. *Proc. Natl. Acad. Sci. U.S.A.* 103, 5781–5786.
- Gloss, B. S., Patterson, K. I., Barton, C. A., Gonzalez, M., Scurry, J. P., Hacker, N. F., et al. (2012). Integrative genome-wide expression and promoter DNA methylation profiling identifies a potential novel panel of ovarian cancer epigenetic biomarkers. *Cancer Lett.* 318, 76–85.
- Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464, 1071–1076.
- Gutschner, T., and Diederichs, S. (2012). The hallmarks of cancer. A long non-coding RNA point of view. *RNA Biol.* 9, 703–719.
- Hannofon, B. N., Sebastiani, P., de las Morena, A., Lu, J., and Rosenberg, C. L. (2011). Expression of microRNA and their gene targets are dysregulated in preinvasive breast cancer. *Breast Cancer Res.* 13, R24.
- Hawthorn, L., Luce, J., Stein, L., and Rothschild, J. (2010). Integration of transcript expression, copy number and LOH analysis of infiltrating ductal carcinoma of the breast. *BMC Cancer* 10:460. doi: 10.1186/1471-2407-10-460
- Hu, W., Yuan, B., Flygare, J., and Lodish, H. F. (2011). Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev.* 25, 2573–2578.
- Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M. J., Kenzelmann-Broz, D., et al. (2010). A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* 142, 409–419.
- Jia, H., Osak, M., Bogu, G. K., Stanton, L. W., Johnson, R., and Lipovich, L. (2010). Genome-wide computational identification and manual annotation of human long noncoding RNA genes. *RNA* 16, 1478–1487.
- Kretschmer, C., Sterner-Kock, A., Siedentopf, F., Schoenegg, W., Schlag, P. M., and Kemmner, W. (2011). Identification of early molecular markers for breast cancer. *Mol. Cancer* 10, 15.
- Lin, R., Maeda, S., Liu, C., Karin, M., and Edgington, T. S. (2007). A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene* 26, 851–858.
- Lipovich, L., Johnson, R., and Lin, C. Y. (2010). MacroRNA underdogs in a microRNA world: evolutionary, regulatory, and biomedical significance of mammalian long non-protein-coding RNA. *Biochim. Biophys. Acta* 1799, 597–615.
- Michelhaugh, S. K., Lipovich, L., Blythe, J., Jia, H., Kapatos, G., and Bannon, M. J. (2011). Mining Affymetrix microarray data for long non-coding RNAs: altered expression in the nucleus accumbens of heroin abusers. *J. Neurochem.* 116, 459–466.
- Panzitt, K., Tschernatsch, M. M., Guelly, C., Moustafa, T., Stradner, M., Strohmaier, H. M., et al. (2007). Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as a noncoding RNA. *Gastroenterology* 132, 330–342.
- Prensner, J. R., Iyer, M. K., Balbin, O. A., Dhanasekaran, S. M., Cao, Q., Brenner, J. C., et al. (2011). Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lncRNA implicated in disease progression. *Nat. Biotechnol.* 29, 742–749.
- Schuetz, C. S., Bonin, M., Clare, S. E., Nieselt, K., Sotlar, K., Walter, M., et al. (2006). Progression-specific genes identified by expression

- profiling of matched ductal carcinomas *in situ* and invasive breast tumors, combining laser capture microdissection and oligonucleotide microarray analysis. *Cancer Res.* 66, 5278–5286.
- Silva, J. M., Boczek, N. J., Berres, M. W., Ma, X., and Smith, D. I. (2011). LSINCT5 is over expressed in breast and ovarian cancer and affects cellular proliferation. *RNA Biol.* 8, 496–505.
- Turashvili, G., Bouchal, J., Baumforth, K., Wei, W., Dziechciarkova, M., Ehrmann, J., et al. (2007). Novel markers for differentiation of lobular and ductal invasive breast carcinomas by laser microdissection and microarray analysis. *BMC Cancer* 7:55. doi: 10.1186/1471-2407-7-55
- Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D. L., Gejman, R., et al. (2007). Activation of p53 by MEG3 non-coding RNA. *J. Biol. Chem.* 282, 24731–24742.
- Received: 19 February 2013; accepted: 17 March 2013; published online: 09 April 2013.*
- Citation: Kosir MA, Jia H, Ju D and Lipovich L (2013) Challenging paradigms: long non-coding RNAs in breast ductal carcinoma in situ (DCIS). Front. Genet. 4:50. doi: 10.3389/fgene.2013.00050*
- This article was submitted to Frontiers in Non-Coding RNA, a specialty of Frontiers in Genetics.*
- Copyright © 2013 Kosir, Jia, Ju and Lipovich. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.*