Transcriptomics: Open Access

Santucci-Pereira et al., Transcriptomics 2014, 2:1
DOI: 10.4172/2329-8936.1000104

Short Communication Open Access

Use of Next Generation Sequencing in the Identification of Long Non-Coding RNAs as Potential Players in Breast Cancer Prevention

Julia Santucci-Pereira[1*], Maria Barton[1,2] and Jose Russo[1]

[1]*The Irma H. Russo, MD, Breast Cancer Research Laboratory, Fox Chase Cancer Center- Temple University Health System, Philadelphia, PA, USA*

[2]*Temple University School of Medicine, Philadelphia, PA, USA*

[*]**Corresponding author:** Julia Santucci Pereira, PhD, The Irma H. Russo, MD, Breast Cancer Research Laboratory, Fox Chase Cancer Center- Temple University Health System, 333 Cottman Ave, P2051, Philadelphia, PA, 19111, USA, Tel: (215) 728 5294; Fax: (215) 728 4083; E-mail: julia.pereira@fccc.edu

Received date: August 02, 2014, **Accepted date:** September 02, 2014, **Published date:** September 09, 2014

Introduction

The development of new technologies, such as Next Generation Sequencing (NGS), and methods to improve the capabilities of this technology has been a revolution for the study of genomics and transcriptomics. NGS opens doors for progress in a variety of biological fields, including biomedical research. NGS allows the sequencing of the whole genome and transcriptome in a massive scale, accessible price, and it is not limited to previous knowledge. Genome sequencing has been applied for the development of a variety of research areas such as, characterization of ancient genomes [1], sequencing of different species [2], risk assessment of genetic diseases [3], molecular diagnosis of various diseases including cancer [4,5] among other, leading to a road for personalized medicine. With NGS, detailed analyses of the transcriptome have been made possible. Detailed information about not only messenger RNA, but also ribosomal RNA, transfer RNA, small RNAs are now accessible. Due to the fact that this technology needs no previous knowledge about the systems being studied compared to other high-throughput technologies (e.g. oligo-microarrays), NGS allows for novel transcripts to be discovered. Thus, alternative splicing, novel microRNAs, and non-coding regions which produce long non-coding RNAs (lncRNAs) can now be explored. Non-coding regions of the genome were originally described as junk, or a by-product of sloppy transcriptional machinery. It was not until the ENCODE project that these non-coding regions were shown to be functional parts of the genome. Indeed, they have been shown to have important roles in gene regulation [6].

LncRNAs have been classified as RNAs which are longer than 200 nucleotides (a length which was set arbitrarily to distinguish them from small RNAs) and do not appear to code for a protein (e.g. have no significantly large open reading frame) [6]. Despite making up for the majority of the human genome, only less than 180 lncRNAs have been annotated [7] and the understanding of their biological role is still a matter for active research. LncRNAs have been described as important key players in cell differentiation and cell transformation. However, thus far, not many lncRNAs have been directly linked to breast cancer; H19 and HOTAIR are some of lncRNAs over expressed in breast cancer [8]. The mechanisms of action of lncRNAs vary, they can scaffold proteins complexes needed for transcription, act as decoys that drive away DNA-binding proteins, such as transcription factors, or guide proteins to the genome [6]. These proteins can either work as enhancers or recruit chromatin modification enzymes [6]. Indeed, lncRNAs have been shown to target several chromatin modification complexes inducing either gene silencing or activation of chromatin [6].

An early full term pregnancy confers a protection against breast cancer, and this protection is induced by breast differentiation accompanied by chromatin remodelling [9,10]. In previous studies, we have observed that women who have completed a full term pregnancy have higher levels of expression of genes related to differentiation than nulliparous women [9-12]. Interestingly, we also observed the up-regulation of some lncRNAs [9]. Among the lncRNAs up-regulated in the parous women were nuclear paraspeckle assembly transcript 1 (NEAT1), metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), and X inactive specific transcript (XIST) [9]. Morphologically, we observed differences in the chromatin conformation of the breast epithelial cells of parous and nulliparous women. These were followed by differences in chromatin activation state [9,10].

Altogether, these evidences led us to focus our interest on the epigenetic phenomena triggered by lncRNAs (Figure 1). Therefore, we evaluated the expression levels of lncRNAs in the breast of healthy post-menopausal women comparing parous and nulliparous women using RNA sequencing. We identified 42 lncRNAs differentially expressed between parous and nulliparous women. Of which, 21 were up-regulated and 21 were down-regulated in the parous. An additional eight non-coding regions presented statistically significant correlation in expression with their nearby gene, indicating a possible role of the lncRNAas a cis-regulatory element. The roles of these eight lncRNAs are unknown; however, seven of the nearby genes are linked to cancer or development (Table 1). Neither functional information, nor expression levels in cancer tissues or cell lines of these fifty lncRNAs have been described in the scientific literature. Thus, functional studies of a set of these lncRNAs are currently being performed to determine the role of these lncRNA in the differentiation, chromatin remodeling and protection against breast cancer. In addition, their expression levels in breast cancer tissues are also being evaluated in our laboratory.

Gene	Function
CPEB4	Promotestumorigenesis [14]; candidate biomarker for defining metastatic cancers [15]
MPPED2	Decreased levels are associated with predisposition to Wilms tumors [16]

VGLL4	Plays a role in apoptosis [17]
KLHDC7B	Has been shown to be up-regulated in breast cancer tissues [18]
KRT81	Expressed in breast carcinomas [19]
GPBP1	Plays a role in the development of atherosclerosis [20]
THSD4	Plays a role in connective tissue formation [21]
C4orf36	May have an effect on eating disorders [22]

Table 1: Biological functions of the genes with significant expression correlation with their nearby lncRNA.

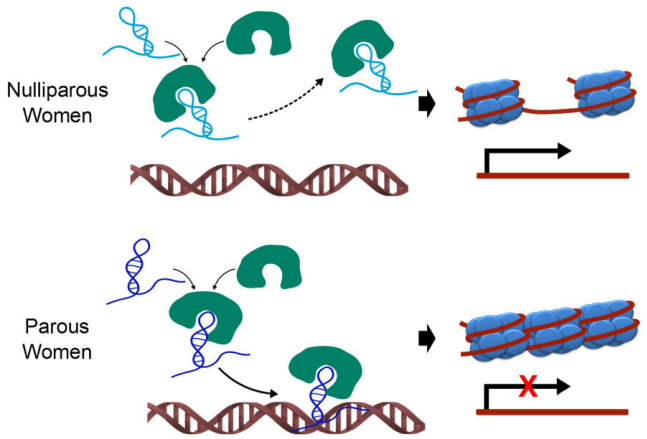

Figure 1: Proposed model of lncRNAs' participation in the chromatin modification observed in breast epithelial cells of parous women. In the nulliparous women, the up-regulated lncRNAs might decoychromatin remodeling factors, allowing the chromatin to stay on its open state (euchromatin). In the parous women, the lncRNAs might be acting as guides or recruiters for chromatin modification complexes. Thus, compacting the chromatin (heterochromatin) and repressing transcription.

In order to select candidates to evaluate the roles of these lncRNAs in the breast, we first analyzed the expression levels of these regions using the genome browser Integrated Genomics Viewer (IGV) [13]. The goal was to define regions which showed higher levels of readings consistently with a defined difference in expression between parous and nulliparous samples. In addition, we also identified potential areas for the development of probes/primers for further validation. Once the sequences of these areas were identified, they were run through NCBI's Basic Local Alignment Search Tool (BLAST) in order to check for sequence specificity and through Custom TaqMan® Assay Design Tool software (Applied Biosystems) to create custom probes for each lncRNA.

The use of next generation sequencing was essential in our project to identify that there are differences in the expression of several lncRNAs comparing the breast of parous versus nulliparous women.In addition, with data generated by this RNA sequencing, we also found that there are significant differences in splicing events between these two groups. These findings will help us to understand the roles of the lncRNAs in gene expression during cellular processes such as differentiation/development, chromatin remodeling and cancer progression. Understanding the link between the lncRNAs with the other genomic, transcriptomic and morphologic changes in the breast cells induced by full term pregnancy will contribute to identify key players in breast cancer prevention.

Acknowledgments

The sample collection was supported by Avon Foundation for Women Breast Cancer Research Program grant 02-2010-117 and the RNA sequencing studies by NIH core grant CA06927 to Fox Chase Cancer Center.

References

1. Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, et al. (2006) Analysis of one million base pairs of Neanderthal DNA. Nature 444: 330-336.
2. Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Molecular Ecology Resources 8: 3-17.
3. Gonzalez-Garay ML, McGuire AL, Pereira S, Caskey CT (2013) Personalized genomic disease risk of volunteers. Proc Natl Acad Sci U S A 110: 16957-16962.
4. Normanno N, Rachiglio AM, Roma C, Fenizia F, Esposito C, et al. (2013) Molecular diagnostics and personalized medicine in oncology: challenges and opportunities. J Cell Biochem 114: 514-524.
5. Zhang J, Ding L, Holmfeldt L, Wu G, Heatley SL, et al. (2012) The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature 481: 157-163.
6. Rinn JL, Chang HY (2012) Genome regulation by long noncoding RNAs. Annu Rev Biochem 81: 145-166.
7. Kurokawa R (2013) Transcriptomics Inspires Long Noncoding RNA. Transcriptomics 1: e102.
8. Vikram R, Ramachandran R, Abdul KS (2014) Functional significance of long non-coding RNAs in breast cancer. Breast Cancer 21: 515-521.
9. Russo J, Santucci-Pereira J, de Cicco RL, Sheriff F, Russo PA, et al. (2012) Pregnancy-induced chromatin remodeling in the breast of postmenopausal women. Int J Cancer 131: 1059-1070.
10. Russo J, Santucci-Pereira J, Russo IH (2014) The genomic signature of breast cancer prevention. Genes (Basel) 5: 65-83.
11. Belitskaya-Levy I, Zeleniuch-Jacquotte A, Russo J, Russo IH, Bordas P, et al. (2011) Characterization of a genomic signature of pregnancy identified in the breast. Cancer Prev Res (Phila) 4: 1457-1464.
12. Peri S, de Cicco RL, Santucci-Pereira J, Slifker M, Ross EA, et al. (2012) Defining the genomic signature of the parous breast. BMC medical genomics 5: 46.
13. Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in bioinformatics 14: 178-192.

14. Ortiz-Zapater E, Pineda D, Martínez-Bosch N, Fernández-Miranda G, Iglesias M, et al. (2012) Key contribution of CPEB4-mediated translational control to cancer progression. Nat Med 18: 83-90.
15. Xu H, Liu B (2013) CPEB4 is a candidate biomarker for defining metastatic cancers and directing personalized therapies. Med Hypotheses 81: 875-877.
16. Liguori L, Andolfo I, de Antonellis P, Aglio V, di Dato V, et al. (2012) The metallophosphodiesterase Mpped2 impairs tumorigenesis in neuroblastoma. Cell Cycle 11: 569-581.
17. Jin HS, Park HS, Shin JH, Kim DH, Jun SH, et al. (2011) A novel inhibitor of apoptosis protein (IAP)-interacting protein, Vestigial-like (Vgl)-4, counteracts apoptosis-inhibitory function of IAPs by nuclear sequestration. Biochem Biophys Res Commun 412: 454-459.
18. Kim TW, Kim YJ, Lee HJ, Min SY, Kang HS, et al. (2010) Hs.137007 is a novel epigenetic marker hypermethylated and up-regulated in breast cancer. Int J Oncol 36: 1105-1111.
19. Nishikawa J, Kiss C, Imai S, Takada K, Okita K, et al. (2003) Upregulation of the truncated basic hair keratin 1(hHb1-DeltaN) in carcinoma cells by Epstein-Barr virus (EBV). Int J Cancer 107: 597-602.
20. Bijnens AP, Gils A, Jutten B, Faber BC, Heeneman S, et al. (2003) Vasculin, a novel vascular protein differentially expressed in human atherogenesis. Blood 102: 2803-2810.
21. Saito M, Kurokawa M, Oda M, Oshima M, Tsutsui K, et al. (2011) ADAMTSL6beta protein rescues fibrillin-1 microfibril disorder in a Marfan syndrome mouse model through the promotion of fibrillin-1 assembly. J Biol Chem 286: 38602-38613.
22. Wade TD, Gordon S, Medland S, Bulik CM, Heath AC, et al. (2013) Genetic variants associated with disordered eating. Int J Eat Disord 46: 594-608.