Journal of
Data Mining in Genomics & Proteomics

Editorial   **Open Access**

In the Race towards a Better Diagnosis and Prognostication of Cancer Patients Long Non-Coding Intergenic RNA's (lincrna's) have found their Place

Jian-Hua Mao* and Heinz-Ulli G Weier

Life Sciences Division, E.O. Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

Editorial

Cancer research in various organ sites revealed that most genetic alterations found in human solid tumors did not affect random loci, but specific oncogenic or tumor-suppressor genes or regulatory elements. While many such 'key mutations' were found within the protein-coding sequences of roughly 130 hundred genes [1], it appears as a paradox that next-generation sequencing (NGS) was able to demonstrate that many cancer patients did not harbor mutations in these genes [2]. We interpret this observation as a strong indicator that genomic alterations outside of the protein-coding sequences drive the onset and progression of human tumors [3,4]. It was reported that variants in dinucleotide repeats may alter the physical structure and coding sequences thereby leading to gene amplification and aberrant expression [4,5]. We and other investigators furthermore postulate that mutations in non-coding DNA, many of which lead to transcripts and might represent important regulatory elements [6], while their aberrant expression and function are key events along a path that drives normal epithelial cells and breast tissue to hyperplasia, pre-neoplastic lesions and eventually malignant tumors.

The significance of developing new biomarkers for cancer diagnosis and prognostication is emphasized by breast cancer occurrence and outcome: With more than 200,000 new cases diagnosed every year, breast cancer remains the most frequent cancer in women [7]. Although breast cancer-related deaths have decreased over the last decade, this dreaded disease is the second most common cause or cancer-related deaths in women. Breast cancer has an unpredictable course and even after removal of the primary tumor, the risk of metastasis continues for 20 years or more [8]. Many patients would greatly benefit from detecting the disease in its earliest stages, when chances for a complete cure are still high.

Breast cancer is a complex disease and the cancer cells often show alteration in pathways ranging from signal transduction to DNA repair, drug response and apoptosis to survival in nutrient or oxygen deficient environments [9-14]. Our understanding of the disease seems very limited considering co-founding factors such as ethnicity in the age at onset or diagnosis of breast cancer [4,15]. Breast cancers progress through accumulation of genomic aberrations that enable development of cancer pathophysiological changes such as unlimited growth and metastasis. Accumulated evidence has demonstrated that breast cancer is a complex and intrinsically heterogeneous disease in which patients may exhibit similar symptoms, and appear to have the same disease, for entirely different genetic reasons. Most published studies of breast cancer tumorigenesis have focused on the role of protein-coding genes during the onset and progression of the disease [16-19]. We support plans to investigate the role of large intergenic non-coding RNAs (lincRNAs; also known as 'long non-coding RNAs' or 'lncRNAs') [6,20] and to assess lincRNAs as potential biomarkers for the early detection of breast as well as cancers. Recent discoveries showed thousands of DNA sequences in the human genome potentially coding for lincRNAs with individual sizes ranging from a few hundred to more than a hundred kb

[6]. The expression of lincRNAs is strikingly tissue-specific and they are typically co-expressed with neighboring genes. It is well documented that lincRNAs play key roles in diverse biological processes such as gene dosage compensation, imprinting, chromatin remodeling, mRNA splicing and tumor metastasis [1,21,22]. For example, overexpression of the lincRNA HOTAIR predicts tumor recurrence in hepatocellular cancer, and has been shown to remodel the chromatin state to promote cancer metastasis in breast cancer [23,24].

Based on recent studies published with our Co-Investigators [21,22], we favor to apply NGS RNAseq to determine the full spectrum of expressed DNA sequences [6]. RNAseq, also called 'Whole Transcriptome Shotgun Sequencing (WTSS)', refers to the use of high-throughput sequencing technologies to sequence cDNA in order to get information about a sample's RNA content. This technology is often used in combination with targeted exome sequencing to identify abnormal RNA transcripts, point mutations, genomic variants and copy number changes in a broad range of tumor types including circulating tumor cells [25,26]. Deep RNA and DNA sequencing will provide hints to alterations that parallel tumor progression, and which will can be validated by use of several technologies including qPCR and FISH.

In summary, breast cancer is worldwide the most frequently occurring cancer in women. Most breast cancer patients succumb to their disease as a result of tumor metastasis. It is therefore important to elucidate the factors which effect breast cancer progression, therapy resistance and metastasis. Accumulated evidence could demonstrate that breast cancer is a very complex, heterogeneous disease, involving various cancer cell-specific changes such as unlimited growth in even nutrient limited environments and metastasis to distant organs [27]. Recent advances in technology have made it possible to deconvolute the heterogeneity and complexity of somatic breast cancer genetics. Using RNA microarray-derived multigene expression pattern, breast cancer has been classified into five molecular subtypes (normal breast like, luminal A, luminal B, ERBB2, and basal-like), which are associated with different levels of aggressive growth and poor prognosis [28]. Based on these findings, some improvements have been made in diagnosis and treatment of breast cancer. However, for most patients, the prognosis

*Corresponding author: Jian-Hua Mao, Life Sciences Division, E.O. Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA, E-mail: JHMao@lbl.gov.

and disease-free survival times have not changed dramatically [7]. Thus, mechanistic and functional studies of breast cancer-promoting molecular changes and the development of novel biomarkers for diagnosis and as therapeutic targets are urgently required. The large intergenic non-coding RNA's (lincRNA's) and 'transcripts of uncertain coding potential (TUCP's)' have recently received much attention due to their potential contribution to disease etiology including cancer and have become targets for biomarker development [6,21,22,29].

Acknowledgements

References

1. Niland CN, Merry CR, Khalil AM (2012) Emerging Roles for Long Non-Coding RNAs in Cancer and Neurological Disorders. Front Genet 3: 25.

2. Hudson TJ (2011) Cancer genome variation in children, adolescents, and young adults. Cancer 117: 2262-2267.

3. Brandt B, Hermann S, Straif K, Tidow N, Buerger H, et.al. (2004) Modification of breast cancer risk in young women by a polymorphic sequence in the egfr gene. Cancer Res 64: 7-12.

4. Buerger H, Packeisen J, Boecker A, Tidow N, Kersting C, et.al. (2004) Allelic length of a CA dinucleotide repeat in the egfr gene correlates with the frequency of amplifications of this sequence--first results of an inter-ethnic breast cancer study. The Journal of Pathology 203: 545-550.

5. Welnicka-Jaskiewicz M, Zaczek A, Brandt B, Bielawski KP, Jaskiewicz J, et.al. (2006) (CA)n microsatellite polymorphism of ERBB-1 in breast cancer. Eur J Cancer 42: 1698-1701.

6. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, et.al. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25: 1915-1927.

7. Siegel R, Naishadham D, Jemal A (2012) Cancer statistics, 2012. CA Cancer J Clin 62: 10-29.

8. Keys HM, Bakemeier RF, Savlov ED (1987) Breast Cancer, In Clinical Oncology, Rubin P (Ed), American Cancer Society: Rochester, New York.

9. Wotiz HH, Chattoraj SC, Kudisch M, Muller RE (1978) Impeding estrogens and the etiology of breast cancer. Cancer Res 38: 4012-4020.

10. Dickson RB, Lippman ME, Slamon D (1990) UCLA colloquium. New insights into breast cancer: the molecular biochemical and cellular biology of breast cancer. Cancer Res 50: 4446-4447.

11. Balz LM, Bartkowiak K, Andreas A, Pantel K, Niggemann B, et.al. (2012) The interplay of HER2/HER3/PI3K and EGFR/HER2/PLC-gamma1 signalling in breast cancer cell migration and dissemination. J Pathol 227: 234-244.

12. Helms MW, Kemming D, Contag CH, Pospisil H, Bartkowiak K, et.al. (2009) TOB1 is regulated by EGF-dependent HER2 and EGFR signaling, is highly phosphorylated, and indicates poor prognosis in node-negative breast cancer. Cancer Res 69: 5049-5056.

13. Agelopoulos K, Kersting C, Korsching E, Schmidt H, Kuijper A, et.al. (2007) Egfr amplification specific gene expression in phyllodes tumours of the breast. Cell Oncol 29: 443-451.

14. Agelopoulos K, Buerger H, Brandt B (2008) Allelic imbalances of the egfr gene as key events in breast cancer progression--the concept of committed progenitor cells. Curr Cancer Drug Targets 8: 431-445.

15. Amend K, Hicks D, Ambrosone CB (2006) Breast cancer in African-American women: differences in tumor biology from European-American women. Cancer Res 66: 8327-8330.

16. Stemke-Hale K, Gonzalez-Angulo AM, Lluch A, Neve RM, Kuo WL, et.al. (2008) An integrative genomic and proteomic analysis of PIK3CA, PTEN, and AKT mutations in breast cancer. Cancer Res 68: 6084-6091.

17. Bose R, Kavuri SM, Searleman AC, Shen W, Shen D, et al. (2013) Activating HER2 Mutations in HER2 Gene Amplification Negative Breast Cancer. Cancer Discov 3: 2224-2237.

18. Weitzel JN, Clague J, Martir-Negron A, Ogaz R, Herzog J, et.al. (2013) Prevalence and Type of BRCA Mutations in Hispanics Undergoing Genetic Cancer Risk Assessment in the Southwestern United States: A Report from the Clinical Cancer Genetics Community Research Network. J Clin Oncol 31: 210-216.

19. Shannon KM, Chittenden A (2012) Genetic testing by cancer site: breast. Cancer J 18: 310-319.

20. Guttman M, Amit I, Garber M, French C, Lin MF, et.al. (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458: 223-227.

21. Ren S, Peng Z, Mao JH, Yu Y, Yin C, et.al. (2012) RNA-seq analysis of prostate cancer in the Chinese population identifies recurrent gene fusions, cancer-associated long noncoding RNAs and aberrant alternative splicings. Cell Res 22: 806-821.

22. Cui Z, Ren S, Lu J, Wang F, Xu W, et.al. (2012) The prostate cancer-up-regulated long noncoding RNA PlncRNA-1 modulates apoptosis and proliferation through reciprocal regulation of androgen receptor. Urol Oncol 31: 1117-23.

23. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, et.al. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 464: 1071-1076.

24. Yang Z, Zhou L, Wu LM, Lai MC, Xie HY, et.al. (2011) Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. Ann Surg Oncol 18: 1243-1250.

25. Joosse SA, Gorges TM, Pantel K (2015) Biology, detection, and clinical implications of circulating tumor cells. EMBO Mol Med 7: 1-11.

26. Joosse SA, Muller V, Steinbach B, Pantel K, Schwarzenbach H (2014) Circulating cell-free cancer-testis MAGE-A RNA, BORIS RNA, let-7b and miR-202 in the blood of patients with breast cancer and benign breast diseases. Br J Cancer 111: 909-917.

27. Rice J (2012) Metastasis: The rude awakening. Nature 485: S55-57.

28. Almendro V, Fuster G (2011) Heterogeneity of breast cancer: etiology and clinical relevance. Clinical & translational oncology 13: 767-773.

29. Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, et.al. (2011) Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. Science 331: 593-596.