Genome Medicine

METHOD Open Access

Comprehensive analysis of the genome transcriptome and proteome landscapes of three tumor cell lines

Pelin Akan*, Andrey Alexeyenko, Paul Igor Costea, Lilia Hedberg, Beata Werne Solnestam, Sverker Lundin, Jimmie Hällman, Emma Lundberg, Mathias Uhlén and Joakim Lundeberg

Abstract

We here present a comparative genome, transcriptome and functional network analysis of three human cancer cell lines (A431, U251MG and U2OS), and investigate their relation to protein expression. Gene copy numbers significantly influenced corresponding transcript levels; their effect on protein levels was less pronounced. We focused on genes with altered mRNA and/or protein levels to identify those active in tumor maintenance. We provide comprehensive information for the three genomes and demonstrate the advantage of integrative analysis for identifying tumor-related genes amidst numerous background mutations by relating genomic variation to expression/protein abundance data and use gene networks to reveal implicated pathways.

Background

Human cancer cell lines have been an invaluable and practical resource for cancer research. The availability of genomic, transcriptomic and proteomic data on these lines is expected to further increase their utility. To this end, we conducted whole-genome and transcriptome sequencing on three tumor cell lines (A431, U251MG and U2OS) for which there is a large body of proteomics data [1]. The choice of these lines was also motivated by their origin from different lineages (tumor cell lines from mesenchymal, epithelial and glial tumors) and abundance of literature.

A431 is used as a model cell line for epidermoid carcinoma and there are currently 3,359 publications describing studies using this cell line. It was established from an epidermoid carcinoma in the vulva of an 85-year-old patient [2]. This cell line expresses high levels of epidermal growth factor receptor (EGFR) and is often used to investigate cell proliferation and apoptosis. U251MG is a commonly used glioblastoma cell line (over 1,200 published articles) established from a male's brain tissue [3]. U2OS is an osteosarcoma cell line derived from a 15-year-old female [4]. Osteosarcoma tumors arise from cells of mesenchymal origin that differentiate to osteoblasts. It is the most common form of bone cancer, responsible for 2.4% of all malignancies in pediatric patients, and its triggers are currently not known [5]. U2OS is a common choice for osteosarcoma research: 35% of the articles associated with the osteosarcoma Medical Subject Headings (MeSH) term in the PubMed database have used this cell line.

Using modern technologies, we subjected these three cell lines to genome and RNA sequencing in order to identify genomic alterations and expression of messenger and microRNAs. A review by Ideker and Sharan summarized studies that demonstrate how genes with a role in cancer tend to cluster together on well-connected sub-networks of protein-protein interactions [6]. We also earlier demonstrated that somatic mutations in a glioblastoma cancer genome produced a pathway-like pattern of enriched connectivity in the gene interaction network. Hence, in this work we analyzed functional relations between all detected somatic mutations, structural variations (altered copy number) and allelic imbalances of expression via network enrichment analysis (NEA) [7,8]. A biological pathway could be seen as an area of densely connected genes in a functional gene network. The idea of NEA when applied to cancer-related genes is that multiple key mutations (which are believed to be

* Correspondence: pelin.akan@scilifelab.se
KTH - Royal Institute of Technology, Science for Life Laboratory, School of Biotechnology, SE-171 65 Solna, Sweden