Cancer and antisense transcription: a bioinformatic strategy for the identification of putative antisense-regulated tumor suppressor genes

Cazzola M.(1), Cremona M.(2), Monti L.(1), Vignati F.(1), Lavorgna G.(3), Taramelli R.(1), Acquati F.(1)and Guffanti A.(4)

(1) Dipartimento di Biotecnologie e Scienze Molecolari, Via JH Dunant 3 - 21100 Varese Italy (2) Laboratory of proteomics, Department of experimental oncology and laboratory, National Cancer Institute, Milan Italy (3) DIBIT-Ospedale San Raffaele Via Olgettina, 58 20132 Milano Italy (4) IFOM - Fondazione Istituto FIRC di Oncologia Molecolare Via Adamello, 16b - 20139 Milano Italy

Motivation

Computational analysis of genomic sequence databases has recently revealed a striking abundance of Natural Antisense Transcripts within the mouse and human genomes. Since antisense transcription is increasingly recognized as a molecular mechanism involved in the regulation of gene expression, a significant proportion of human disease genes could potentially display antisense-mediated abnormal patterns of gene expression. Cancer is a pathological phenotype that could represent a potential target for such gene regulation mechanism, given the high number of genes governing cancer-related cellular functions such as proliferation, differentiation and apoptosis. Preliminary experimental evidence has been reported recently for the occurrence of natural antisense transcript for several genes whose function has been implicated in cancer pathogenesis. Therefore, a targeted in silico survey of antisense transcription, coupled with a detailed inspection of annotated gene features, could represent a novel tool for the identification of candidate cancer-related genes. We have performed an in silico search for "sense-antisense gene clusters" within two regions from human chromosome 6 (6q21 and 6q27) that have long been reported to carry cancer-associated deletions and rearrangements, but for which no tumor suppressor genes has been unambiguously identified. Experimental validation of each sense-antisense cluster detected in this study, followed by definition of bona fide tumor suppressor candidates based on the available annotation features, confirmed the feasibility of this approach to better define candidate cancer-associated genes.

Methods

the procedure that has been carried out in this work is based on a recently described antisense detection algorithm known as AntiHunter, whose task is to scan the publicly available databases to find ESTs that display antisense orientation with respect to a gene structure present in a given genomic sequence. Data retrieval was performed by collecting all available sequence and annotation informations for the above mentioned genomic regions from the EnsEMBL databases by means of the EnsMart annotation system and 'ad hoc' created scripts. Further informations were derived from LocusLink and other sources by using the GeneWebEx software. The retrieved information was subsequently organized in a dedicated database, in order to facilitate successive analyses. One hundred-one genes from the two genomic regions (both RefSeq genes and novel transcripts) were then inserted into the database and subsequently analyzed with the AntiHunter algorithm in order to detect putative antisense EST sequences. A graphical interface was added using the Generic Genome Browser database and interactive web pages system.

Results

the AntiHunter output consisted in 5 genes within 6q27 and 26 within 6q21 for which antisense EST sequences were reliably detected in silico. EST alignment on the human genome was verified for each of these genes by means of a BLAT search. We also created a graphical representation of the EST alignments to the genome sequence on the UCSC Genome Browser. The CAP assembler and ORF finder tools were used to produce contigs of the antisense EST sequences and to detect putative coding potential, respectively. We next tried to further support the biological significance of the identified sense-antisense gene clusters by searching for mouse orthologs displaying the same sense-antisense transcription mechanisms, identifying twenty-one sense-antisense gene pairs. All these informations were organized in a dedicated database with a web interface. To validate experimentally the results obtained by our search, strand-specific RT-PCR analysis was carried out for 18 randomly chosen sense-antisense gene pairs. Using human brain total RNA as a substrate, 13 gene pairs (72%) clearly showed expression of both sense and antisense transcripts. Since just one tissue was used as a source of RNA for the RT-PCR assays, we cannot rule out that the five negative gene pairs could also be subjected to bidirectional transcription, too. Taken together, these results suggest that our strategy provides a reliable system for detecting sense-antisense gene pairs within a genomic region of interest. Finally, we were left with the need to define candidate cancer-related genes on the basis of the available annotation for the genes displaying antisense transcription. A detailed inspection of all 21 evolutionary conserved sense-antisense gene pairs within 6q21 and 6q27 allowed us to define nine genes controlling cellular or molecular functions that are compatible with a putative role in cancer development or progression. Interestingly, seven among these gene pairs were scored positive for experimental validation by strand-specific RT-PCR as described above. These seven gene pairs are actually under investigation in our laboratory to further evaluate their potential role in human cancer pathogenesis.

Contact email: francesco.acquati@uninsubria.it
URL: http://www.ifom-ieo-campus.it/RESEARCH/Bioinformatics/index.html