AntiHunter 2.0: increased speed and sensitivity in searching BLAST output for EST antisense transcripts

Lavorgna G.(1), Triunfo R.(2), Santoni F.(2), Orfanelli U.(1), Noci S.(1), Bulfone A.(1), Zanetti G.(2) , Casari G1)

(1) DIBIT-San Raffaele Scientific Institute, Via Olgettina 58, 20132 Milan, Italy. (2) CRS4-Centro di Ricerca, Sviluppo e Studi Superiori in Sardegna, Parco Scientifico e Tecnologico POLARIS Loc. Pixina Manna Building 1 CP 25 - 09010 Pula

Motivation

An increasing number of eukaryotic and prokaryotic genes are being found to have natural antisense transcripts (NATs). Also, there is a growing evidence to suggest that antisense transcription might have a key role in a range of human diseases. Consequently, there have been several recent attempts to set up computational procedures aimed at identifying novel NATs. Our group has developed the AntiHunter program for the identification of expressed sequence tag (EST) antisense transcripts from BLAST output.

Methods

In order to perform an analysis, program requires as input a genomic sequence plus an associated list of transcript names and coordinates of the genomic region. After masking the repeated regions, program performs a BLASTN search of the input sequence versus the selected EST database, reporting by Email the EST entries that reveal an antisense transcript with respect to the user supplied list. Here, we present the newly developed 2.0 version of the AntiHunter tool. Several improvements have been added to this version of the program in order to increase its capability to detect a larger number of antisense ESTs.

Results

As a result, AntiHunter can now detect, on average, over 45% more antisense ESTs with little or none increase of the percentage of the false positives. Also, we raised the maximum query size to 3 mega bases (previous limit was one mega base). Moreover, we found that a reasonable trade-off between the program search sensitivity and the maximum allowed size of the input query sequence could be obtained by querying the database with the MEGABLAST program, rather than by using the BLAST one. We now offer this new opportunity to the users, i.e. if choosing the MEGABLAST option, the user can input a query sequence up to 30 mega bases long, considerably improving the possibility to analyse longer query regions. The AntiHunter tool is publicly available at http://bioinfo.crs4.it/AH2.0.

Contact email: giovanni.lavorgna@hsr.it
URL: http://bioinfo.crs4.it/AH2.0