

Mammalian microRNAs derived from genomic repeats

Neil R. Smalheiser and Vetle I. Torvik

University of Illinois at Chicago, UIC Psychiatric Institute, MC 912, 1601 W. Taylor Street, Chicago, IL 60612 USA

Corresponding author: Smalheiser, N.R. (smalheiser@psych.uic.edu).

In the accompanying PDF (supplementary file 1), the genomic structure of all human, mouse and rat microRNA precursors identified in this study that contain genomic repeats are shown.

In the accompanying Excel file (supplementary file 2), the numbers of sequences in RefSeq, EST and genome databases that have perfect complementarity (with up to five G:U matches) with microRNAs are shown. The results of all perfect hits occurring within human and mouse RefSeq mRNAs, human and mouse ESTs, and human genomic sequences are shown. If two sequences in the same database shared the hit region plus 25 nucleotides on either side, they were considered redundant and not counted twice. Values shown are # of hits produced by a microRNA / average # of hits made by scrambled versions of that microRNA (see main article for methods). Those entries in RefSeq having 20 times more microRNA hits than scrambled hits are highlighted in yellow. For human ESTs, the data separately show all hits (reading ESTs in both directions) as well as only those hits where, based on EST annotation, the microRNA sequence is complementary to the transcript (i.e. appropriate for a functional microRNA-target interaction). Because many transcripts in the EST databases contain microRNA precursors in sense orientation [1], and will be detected as perfect hits having no G:U matches, we only highlighted such hits in the table where EST annotation indicated the microRNA sequence is complementary to the transcript.

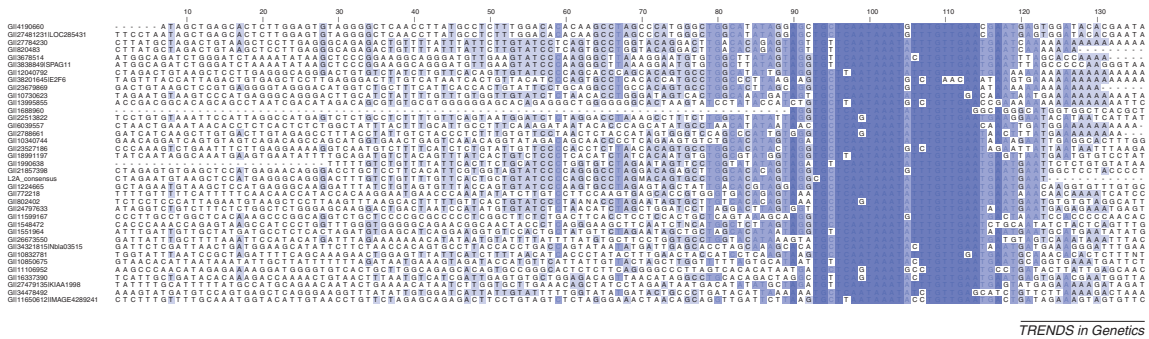


Figure S1. Multiple sequence alignment of mRNAs and ESTs that exhibited perfect complementarity to miR-95. miR-95 hit perfectly (including two-to-four G:U matches) on three mRNAs and 94 ESTs comprising 31 distinct clusters (Unigene clusters or singletons when not belonging to a Unigene cluster). Multiple sequence alignment performed using ClustalW (<http://www.ebi.ac.uk/clustalw>) shows that these putative targets are not related to each other except in the microRNA hit region and in nearby LINE-2 homologous sequences (L2A consensus from Repbase, <http://www.girinst.org>).

Reference

- 1 Smalheiser, N.R. (2003) EST analyses predict the existence of a population of chimeric microRNA precursor-mRNA transcripts expressed in normal human and mouse tissues. *Genome Biol.* 4, 403