NCBI-Retrieval

genomes of human, mouse and rat

Algorithm

select of non-overlapping, 30-100 nt flanked by 10-30 nt perfect repeats

2.768.940 sequences

RepeatMasker

exclusion of all known repeats except those derived from tRNA

1.033.824 sequences

Local BLAST

against all published tRNA-genes

2.442 sequences

399 tRNA-derived

1899 4.5S$_I$-derived

144 BCI/ID-derived

Supplementary Figure 1: Flow chart of the computational strategy

We wrote a computer algorithm in C programming language to extract from GenBank sequences of 30 to 100 nucleotides in length that are flanked by at least 10 nt-long perfect direct repeats. Note: Because of the search restrictions as shown in the upper right box, the numbers of selectively extracted tailless retropseuodgenes only represent a fraction of the total genomic numbers.