Construction of a Full-Lenght cDNA Bank from Human Cell Lineages

Carreira, A.C.O., Cruz, L.O., Colin, C., Sodré, F.M.C., Coelho, T.M., Sogayar, M.C.

Chemistry Institute, University of São Paulo, Biochemistry Dept., São Paulo, Brazil

A full-length cDNA sequence, including a transcription initiation site and a polyadenylation site, is the gold standard for transcript definition. The Transcript Finishing Initiative (TFI) (Genome Research, 2004, 14:1413-1423) introduced a strategy to characterize new human transcripts and splicing isoforms expressed at low levels and in restricted sets of tissues. Cell lines, obtained from the American Type Culture Collection, were used in order to generate a cDNA panel representing distinct tissues. We set out to construct a human full-length cDNA bank using these human cell lines and other lineages, in a total of 20 samples. In parallel, we generated a cDNA pool from 20 different human tissues (Clontech). RNA was prepared using the cesium chloride cushion technique (Chirgwin et al., 1979) or commercial RNA extraction kits. cDNA was synthesized using 2 µg total RNA and reverse-transcribed by using oligo (dT)12-18 and *Improm* II Reverse Transcriptase. The quality of each cDNA was evaluated by PCR amplification of the sequence located at the 5´ end of the *NOTCH2* gene (a long transcript), *GAPDH* (an abundant transcript) and *p53* (a low expression transcript). A total of 40 cDNA samples, derived from human cell lines and tissues were analyzed and validated. This cDNA bank allowed us to successfully amplify full-length cDNAs encoding different proteins (cytokines, growth factors, etc.), particularly proteins with biopharmaceutical potential. Upon confirmation by DNA sequencing, these sequences were cloned into vectors for different expression systems (bacteria, insect cells/BEVS and mammalian cells). Our full-length cDNA bank represents a validated source for amplification of human genes and their isoforms.

Supported by: FAPESP, CNPq, FINEP, Ludwig Inst.