ABSTRACT

METHOD TO GENERATE OR DETERMINE NUCLEIC ACID TAGS CORRESPONDING TO THE TERMINAL ENDS OF DNA MOLECULES USING SEQUENCES ANALYSIS OF GENE EXPRESSION (TERMINAL SAGE)

We describe a method of generating a database comprising a plurality of records, each record comprising an indication whether a gene is expressed by a particular cell, the method comprising steps of: (a) providing a first nucleic acid sequence; (b) linking the first nucleic acid sequence to an linker sequence to form a linked nucleic acid, in which the linker sequence comprises: (i) a first recognition site for a first nucleic acid cleavage enzyme that allows cleavage of the first nucleic acid sequence at a site distant from the first recognition site, and (ii) a second recognition site for a second nucleic acid cleavage enzyme that allows nucleic acid cleavage at a site distant from the second recognition site, said cleavage site located at a position within or about the first recognition site; in which the linked nucleic acid has the structure: 5' – second recognition site – first recognition site – first nucleic acid – 3' (c) cleaving the linked nucleic acid with the first nucleic acid cleavage enzyme to provide (i) a linked tag comprising the linker sequence linked to a nucleotide sequence tag representative of the first nucleic acid sequence and comprising a terminal portion thereof: and (ii) a second nucleic acid sequence comprising a remainder portion of the first nucleic acid; (d) detecting the presence, sequence or identity of the linked tag or the nucleotide sequence tag to provide an indication of an instance of gene expression; and storing an indication of expression of the gene as a record in the database.

Figure 3