'Gene shaving' as a method for identifying distinct sets of genes with similar expression patterns

ABSTRACT

By using the gene shaving method, it is possible to examine gene expression data and identify relevant gene clusters that require further investigation.

INTRODUCTION

The role played by tandem repeats in tandem mapping of genetic maps has made them a crucial aspect of mammalian genetics, and they continue to be utilized for forensic DNA fingerprinting. The classification of tandem repeats is usually based on their location among satellites (encompassing megabases of DNA, associated with heterochromatin), minisatellites (repeat units in the range of 6-100 bp, spanning hundreds of base-pairs) and microsales (ranging from 1-5 dp). The discovery of tandem repeats, such as mini and microsatellites, has been supported by several studies that suggest they can be a crucial source of marker information for identifying pathogenic bacteria, even in the absence of newly evolved monomorphic pathogens. The significance of tandem repeats in the adaptation of pathogens to their host is likely reflected in this. Tandem repeat activity in bacteria appears to vary in two ways: within the regulatory region of a gene, they can cause an on/off switch in gene expression at the transcriptional level. The presence of tandem repetition within coding regions with repeat unit lengths not greater than three can result in a reversible premature end of translation caused by mutations that alter the number of repeats. In such cases, changes in the number of copies alter the gene product itself. The mutation mechanisms of micro and minisatellites have been studied in detail in eukaryotes, specifically human and yeast (refer to). The data obtained thus far indicates that microsatellites undergo mutativity via replication slippage processes, with mutation rates being influenced by the effectiveness of mismatch repair mechanisms and internal heterogeneity within the array that helps stabilize the tandem repeat. Minisatellites undergo mutation in response to a double strand break that is initiated within or near the tandem repeat, but these events can occur naturally in eukaryotes. The minisatellite mutation rate in eukaryotes seems to be unresponsive to mismatch repair efficiency, and internal heterogeneity is compatible with a high mutation. In bacteria, simple sequence contingency loci (repeat units of 1-8 bp) have been identified as such. The altered number of repeats permits the corresponding gene to enter and exit reversible on and off states of expression. The mutation rate of a tetranucleotide tract in Haemophilus influenzae is above 10-4, which helps the pathogen adapt to its hosts as the infection progresses. When faced with extreme circumstances, the microsatellite is not very useful for identifying strains, performing epidemiological and phylogenetic studies. The tandem repeat array is made up of perfect copies of the elementary unit, and different alleles are detected in a single culture. On the other hand, DNA sequencing can be used to verify the phylogenetic identity of minisatellite alleles of the same size, as the units are frequently duplicated but not perfect. The pattern of variants along the array provides an extra layer of allle identification and chromosomal information. Moreover, the use of ordinary horizontal gel electrophoresis can easily type tandem repeats with longer repeat unit length within the size range of a few hundred base-pairs. We will initially explore how a tandem repeats database can be used to sequence bacterial genomes and briefly compare the overall features of tandem

repetitions in various endemic bacteriological species whose sequence has been determined and made available to the public. The demonstration will demonstrate how this device can be used to quickly characterize new and polymorphic markers in two pathogens, Y. pestis and B. anthracis. Pestis, a high-resolution typing tool, is based on RFLP analysis of IS100 locations. However, this technology is more complex than PCR typing, which is why it has been developed. In the case of B. anthracis in the U.S., polymorphisms were initially identified essentially using AFLP (Applied Functional Functional Protein Polymerase) typing. Subsequent analyses revealed that the most informative fragments in AFLP patterns were generated by variation in tandem repeat array length (five minisatellite loci).

CONCLUSION
Specifically, we describe how the mouse RPTP (mouse phage-tumor-like) cDNA is cloned into the human skeletal muscle, the way in which the mice and humans are genetically expressed through alternative splicing of their respective genes, and the presence of an 8 kb 3'UTR in human "RPTTP" (human minus 2), The largest PCRTA gene known to date, spanning over 1 megabase pairs of genomic DNA, with its considerable length, mostly due to expanded introns in this region). Encoding the protein domains in the extracellular segment consist of modules that are flanked by phase 1 introns, while the majority of intracellular segments are phase 0 and relatively small. These data indicate that the ectodomain originated from exon shuffling and duplication and eventually fused with another phosphatase domain at a later time. The MAM domain, which is the region defining type IIB phosphatases, has a genomic structure that is typical of all these domains when located at the N-terminus. Additionally, three exons encode the fourth fibronectin repeat in RPTP, an extra property present in only type I ATPaSeq. At least two spliced exons flank the transmembrane domain, which is the region of greatest variation between the four IIB phosphatases; another resembling an alternatively arranged exonet precedes the catalytic core of the first ATPase. The genomic structure of representative members of the RPTP family (types I-V) shows that the intron/exon organization of both phosphatase domains is highly conserved. Significant variation exists in the length of their 3' UTRs; the longest known record of a regulated transcriptional region at 8 kb is characterized as 3'UTR, or "under regulation" UTM. We have achieved the first-ever characterization of the genomic structure of an RPTP type IIB gene. This knowledge will assist in future research on the regulatory factors that influence tissue specificity of gene expression.