A tandem repeats database for bacterial genomes: application to the genotyping of Yersinia pestis and Bacillus anthracis

ABSTRACT

The analysis of the bacterial genome sequences currently available indicates that Bacillus anthracis and Yersinia pestis exhibit an average density of tandem repeat arrays longer than 100 bp (approximately 30 per Mb). A set of over fifteen informative markers, some of which exhibit a high degree of polymorphism, could be quickly produced in both situations by testing varying fractions of these sequences for polyhomogeneity. In one case, the polymorphism information content index is 0.82, and allele lengths between 600 and 1950 bp are extensive.

INTRODUCTION

The absence of dystrophin and its associated proteins can result in the phenotype of muscular dystrophy, as they are believed to play a crucial role in maintaining the integrity of the extracellular matrix and the membrane of muscle cells. The DAPC is composed of various protein complexes that are either directly or indirectly linked to dystrophin. The four transmembrane proteins known as sarcoglycans are organized by a fifth protein called spirochaplasia, which is believed to play cAMP signalling roles at the cell membrane. The dystroglycan complex, which interacts directly with dystrophin in the cytoplasm and laminin on the extracellular matrix, serves as a structural link between the interior and exterior of the cell. A third subcomplex includes dystobruvines and syntrophines, both of which have an unknown function. Recently, the yeast two-hybrid method was used to identify desmuslin (DMN), an -dystrobrevin-interacting protein. Both mRNA and protein are expressed mainly in cardiac and skeletal muscle and contain genes that encode a novel intermediate filament (IF) protein of 1253 amino acids. Electron microscopic analysis indicates that cessnin and desmin can colocalize with each other. During co-immunopreciptation experiments, it was discovered that the desmuslin and -dystrobrevin interaction involves the region of protein encoded by exons 8-16 of etanocellulones (precisely similar to human cDNA) and domains 1A-2A of the demineralin rod domain. Desmuslin is hypothesized to act as a mechanical support for the muscle myofibers by creating an unrecognized interface between the extracellular matrix and the Z-discs through desmin and plectin. Human genetic disorders, such as congenital and adult onset myopathies, have been associated with the involvement of several IF proteins, including duns (desmin), which may also play a role in myopathy. This possibility is supported by the exclusive expression of DMN in skeletal and cardiac muscle. We examined 71 patients with different forms of muscular dystrophy and myopathy for mutations in the DMN gene, finding 9 single-nucleotide polymorphisms (SNPs) that do not alter the protein sequence but 12 that modify the residue they encode. Our research has revealed that no controls are probable origins of the phenotype, but our findings are applicable for disequilibrium studies of this region of chromosome 15q26.3 and for studying mutation analysis and association in other genetic disorders.

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.