

5th Annual Joint Bioinformatics Workshop

July 19, 2005
2229 Seamans Center
University of Iowa

by

Dr. Terry Braun
University of Iowa

Techniques for Pseudogene Identification and Prediction of Gene Conversion

ABSTRACT

Pseudogenes are remnants of gene duplication events – or non-processed and processed pseudogenes. There have been multiple studies to identify pseudogenes in the human genome and pseudogenes have also been identified to be involved in gene conversion events that lead to disease. This study describes methods for predicting pseudogenes and predicting involvement in gene conversion. A genome-wide survey of all human pseudogenes is performed to accurately identify non-processed pseudogenes. A total of 14,476 pseudogenes are identified including 3,426 non-processed pseudogenes. In addition, 1,945 non-processed pseudogenes that are co-localized to the genuine gene are evaluated for their possible role in gene conversion and disease. All eleven known cases of gene conversion (with deleterious effects) involving pseudogenes are successfully identified by these methods. Among the pseudogenes identified is an RP9 pseudogene that carries the Asp170Gly mutation associated with the RP9 form of adRP. The Asp170Gly mutation in RP9 is a potentially novel example of gene conversion between the genuine gene and pseudogene. Ranking of non-processed pseudogenes as candidates for gene conversion is also performed using sequence characteristics of published cases of gene conversion.