Integrative Biology-2013
OMICS Group Conferences
Accelerating Scientific Discovery
International Conference on
Integrative Biology Summit
August 5-7, 2013 Embassy Suites Las Vegas, USA

Multi-null hypotheses method as an expert knowledge generator for AI-based detection of selection operating at molecular level

Krzysztof A. Cyran

Silesian University of Technology, USA

Abstract

Detection of signatures of natural selection operating at molecular level is one of the important problems in contemporary evolutionary biology. There has been designed a number of statistical neutrality tests for that purpose, however they often give false results as the actual population dynamic rarely satisfies the classical null hypothesis assumptions such as constancy of the population in time, no sub-population structuring, and no recombination. Therefore, artificial intelligence (AI) based methods can be used to analyze the results of a battery of such neutrality tests applied against classical null. However, in order to apply the AI-based methodology, an expert knowledge is required for the learning phase of the classifier construction. The paper present the multi-null hypotheses method, which by incorporating the actual population growths models, sub-structuring and estimated level of recombination to subsequent nulls, substantially increases the validity of the obtained results in neutrality testing. This increase in accuracy is achieved by eliminating other than selectionbased influence on the tests outcomes, by assuming these other factors in appropriately modified null-hypotheses. This approach requires however large computational effort in order to determine by computer simulations the critical values of the neutrality tests applied against modified nulls (for classical null these critical values are known). Therefore, multi-null hypotheses methodology cannot be considered as a simple alternative to testing against classical null in a large number of genes. However, it can be used as an expert-knowledge generator used for training the AI-classifiers. After they become trained, they can be efficiently used in detection of selection in other genes without need to perform complex computer simulations. The whole methodology is illustrated by search for signatures of balancing selection operating at molecular level in human helicases as examples of genes implicated in human familial cancer.