Genome Medicine

METHOD Open Access

Improving the prediction of the functional impact of cancer mutations by baseline tolerance transformation

Abel Gonzalez-Perez[1*], Jordi Deu-Pons[1] and Nuria Lopez-Bigas[1,2*]

Please see related research highlight: http://genomemedicine.com/content/4/11/88

Abstract

High-throughput prioritization of cancer-causing mutations (drivers) is a key challenge of cancer genome projects, due to the number of somatic variants detected in tumors. One important step in this task is to assess the functional impact of tumor somatic mutations. A number of computational methods have been employed for that purpose, although most were originally developed to distinguish disease-related nonsynonymous single nucleotide variants (nsSNVs) from polymorphisms. Our new method, transformed Functional Impact score for Cancer (transFIC), improves the assessment of the functional impact of tumor nsSNVs by taking into account the baseline tolerance of genes to functional variants.

Background

With the advent of high-throughput sequencing, our ability to identify single nucleotide variants (SNVs) in the genome or exome of individuals has far exceeded our capacity to experimentally validate their impact on disease phenotypes. Therefore, computational methods that predict the impact of non-synonymous SNVs (nsSNVs) on protein function have become very important and of wide interest. Bioinformatics methods have been developed and tested over the past decade that distinguish disease-related nsSNVs from neutral polymorphisms [1-11]. A different, although related, problem is assessing the relevance of nonsynonymous somatic variants in cancer emergence. In principle, functional somatic mutations can only be causative of cancer if they affect cancer driver genes, which upon mutation confer a distinct selective advantage or a newly acquired capability to the cell [12,13].

The need for computational methods to predict the functional impact of cancer-causing somatic variants contrasts with the low number of methods that have been designed or tested specifically for this purpose [4,14]. One likely explanation is the absence of curated sets of true driver and passenger cancer mutations. Many recently published cancer resequencing projects use methods like SIFT [15,16], and PolyPhen2 (PPH2) [17-19] to predict the functional impact of cancer somatic mutations, although these methods were not developed or tested for this purpose and the quality of their performance in this context is not clear.

Existing methods provide a predictive functional impact score (FIS) for each mutation [3]. The FIS calculated for nsSNVs relies mainly on the conservation of single residues across multiple sequence alignments. In other words, these methods employ evolutionary information to assess the likely impact of an amino acid change on the structure or function of the altered protein. Nevertheless, the ultimate effect of this amino acid change on the functioning of a cell depends on other factors as well, such as the particular role played by the altered protein in the cellular machinery. The criticality of that role will determine the protein's tolerance to amino acid changes. Our view is that a score purporting to assess the likelihood of individual mutations to provide a somatic cell with an acquired advantage - and possibly give origin to a tumoral clone - must take this feature into consideration.

* Correspondence: abel.gonzalez@upf.edu; nuria.lopez@upf.edu
[1]Research Programme on Biomedical Informatics - GRIB. Universitat Pompeu Fabra - UPF, Hospital del Mar Medical Research Institute - IMIM. Parc de Recerca Biomèdica de Barcelona (PRBB). Dr. Aiguader, 88, E-08003 Barcelona, Spain
Full list of author information is available at the end of the article