Caspase cleavage sites in the human proteome: CaspDB, a database of predicted substrates

Piotr Cieplak, Ph.D.
Research Associate Professor, Sanford Burnham Medical Research Institute, 10901 North Torrey Pines Road, La Jolla, California 92037, USA

Piotr Cieplak: pcieplak@sanfordburnham.org

Abstract

Caspases are enzymes belonging to a conserved family of cysteine-dependent aspartic-specific proteases that are involved in vital cellular processes and play a prominent role in apoptosis and inflammation. Determining all relevant protein substrates of caspases remains a challenging task. Over 1500 caspase substrates have been discovered in the human proteome according to published data and new substrates are discovered on a daily basis. To aid the discovery process we developed a caspase cleavage prediction method using the recently published curated MerCASBA database of experimentally determined caspase substrates and a Random Forest classification method. On both internal and external test sets, the ranking of predicted cleavage positions is superior to all previously developed prediction methods. The *in silico* predicted caspase cleavage positions in human proteins are available from a relational database: CaspDB. Our database provides information about potential cleavage sites in a verified set of all human proteins collected in Uniprot and their orthologs, allowing for tracing of cleavage motif conservation. It also provides information about the positions of disease-annotated single nucleotide polymorphisms, and posttranslational modifications that may modulate the caspase cleaving efficiency. The CaspDB database is publicly available at http://caspdb.sanfordburnham.org for all users and no login or registering is required.

Dear Editor

Recently we published a paper about the predicted caspase cleavage sites in human genome that are collected in CaspDB database, which is available at http://caspdb.sanfordburnham.org. The paper has been published in recent edition of PloS One journal: S.Kumar, B.van Raam, G.S.Salvesen, P.Cieplak: "**Caspase cleavage sites in the human proteome: CaspDB, a database of predicted substrates**", PLoS One. 2014 Oct 17;9(10):e110539, doi:10.1371/journal.pone.0110539. We believe that this database offers a robust prediction of caspases cleavage sites positions, the comparison of cleavage positions in orthologous proteins in several organisms, as well as information about known SNPs and posttranslational modifications that could interfere with caspase cleavage events. The basic theoretical method for cleavage recognition is based on MerCasbah - the published database of known more than 500 caspase cleavage events and machine learning approach. We think that our database offers a unique combination of information and could be potentially widely used by all the readers of the Apoptosis journal. Below we include the Abstract of our paper.