p53FamTaG : a database resource of human p53, p63 and p73 direct target genes combining in silico prediction and microarray data

Sbisà E (1), Catalano D (1), Gisel A (1), Grillo G (1), Licciulli F (1), Turi A (1), Liuni S (1), Pesole G (1,2), De Grassi A (2), Caratozzolo MF (2), D'Erchia AM (1,2), Navarro B (1), Tullo A (1), Saccone C (1,2)

(1) Istituto di Tecnologie Biomediche- Sede di Bari, CNR, , Via Amendola, 122/D 70126 Bari, Italy
(2) Dipartimento di Biochimica e Biologia Molecolare, "Ernesto Quagliariello", Università degli Studi di Bari, Via Orabona, 4, 70126 Bari, Italy

Motivation

The p53gene family is composed of three genes, p53, p63 and p73, with polyhedrical functions in pivotal cellular processes such as DNA synthesis and repair, growth arrest, apoptosis, genome stability, angiogenesis, development and differentiation. p53, p63 and p73 encode sequence-specific nuclear transcription factors which recognise the same responsive element (RE), but with a degree of specificity for the target genes that is quantitatively distinct. The three genes are differentially regulated and carry out specialized, non-overlapping functions. Their inactivation or aberrant expression may determine tumour progression or developmental disease. The discovery of several protein isoforms with antagonistic roles which are produced by the expression of different promoters and alternative splicing, widened the complexity of the scenario of the transcriptional network of the p53 family members. Therefore, the identification of the genes transactivated by p53 family members is crucial to understand the specific role for each gene in cell cycle regulation. To identify new direct target genes, we combined a genome-wide computational search of p53 family REs and microarray analysis. The huge amount of biological results produced raised a critical need for bioinformatic instruments able to manage and integrate the data and facilitate their retrieval and analysis. We have developed the p53FamTAG database (p53 FAMily TArget Genes), which contains p53 family direct target genes selected in the human genome searching for the presence of the REs and the expression profile of the target genes obtained by microarray experiments.

Methods

The genome-wide computational analysis was performed by using PatSearch, a pattern matching program implemented in DNAfan tool (DNA Feature Analyzer) developed in our Lab. These data were integrated with the microarray results produced in our Lab from the overexpression of different isoforms of p53, p63 and p73 stably transfected in isogenic cell lines, allowing to study in a comparable way the transcriptional activity of all the proteins in the same cellular background. p53FamTAG is a relational database, designed in a modular way so that, new data coming from different public resources and experimental analyses can be integrated and updated independently when needed. The p53FamTAG was implemented using MySQL as DBMS and the query/retrieval system was built using PHP Seagull Framework.

Results

p53FamTAG represents a unique integrated resource of human direct p53 family target genes, linked to other public databases (HUGO, EnsEmbl, RefSeq), and provides the user with an efficient query/retrieval system which allows the export of the RE sequences and results of our microarray experiments. p53FamTAG also contains 83 experimentally verified p53 family target genes, and 341 p53REs recently identified by ChIP-PET strategy linked to the relevant UCSC genome tracks and to PubMed entries. The database was developed for supporting and integrating high-throughput in-silico and experimental analyses and represents an important reference source of knowledge for research groups involved in the field of oncogenesis, apoptosis and cell cycle regulation.

Contact email: elisabetta.sbisa@ba.itb.cnr.it