Analysis of p63 isoform-driven gene expression: a cDNA array/bioinformatics integrated approach - (session: Structural Genomics)

S. Saviozzi, M. Lo Iacono, F. Lanzarato, G. Franceschini, G. La Mantia, V. Calabrò and R.A. Calogero

Università di Torino

Two homologs of p53 have been identified: p73 and p63 (Kagdah et al. 1997; Yang et al. 1998). The hallmark features of the p53 protein - an acid amino-terminal transactivation domain (TA), a core domain for DNA-binding and a carboxy-terminal oligomerization domain - are shared by both p63 and p73. P73 and p63 are also characterized by the presence at C-terminus of a sterile alpha-motif (SAM)-like sequence. The p63 gene encodes at least six polypeptides by way of two different promoters/ATG (TA and deltaN isoforms) and three alternatively spliced C-terminal regions.
P63-controlled gene expression profiling was investigated hybridising total RNA extracted form p53 -/- cell line (SAOS-2) transiently transfected with the six p63 isoforms on cDNA arrays (8395 cDNAs, Incyte LifeGrid). We identified 384 differentially expressed genes passing two rounds of statistical validations (Tusher et al. 2001; Baldi and Long 2001). We observed that TAp63beta prevalently induces up-regulation, TAp63gamma mainly induces down-regulation and deltaN isoforms, although without transactivation domain, are able to induce a transcriptional response. Transcriptional profiles of genes controlled by the six isoforms were grouped using k-way and adaptive quality-based clustering approaches. Both methods show that three classes are the optimal solution for our data set and the three clusters were defined as UP, i.e. genes up-modulated by all isoforms, DOWN, i.e. genes down-modulated by all isoforms and MIX, genes up-modulated only by TAp63beta, dNp63beta and dNp63gamma isoforms.
We used these p63-controlled genes as starting material to evaluate the presence of common regulative elements in promoters of co-regulated genes. We extracted from NCBI human genomic data 2Kb upstream to the annotated 1st transcribed nucleotide of the p63-controlled genes and we mapped the presence of know human transcriptional elements, described in TRANSFAC professional 6.1, using PATSER program (Hertz and Stormo 1999). Then, we applied a computational data mining technique used in market basket analysis: the Agrawal association rules induction algorithm (Agrawal 1993), which is a powerful method to find regularities in a set of documents/transactions (e.g. a commercial association rule like "If a customer buys wine and bread, he often buys cheese, too." can be rewritten in a gene oriented way: "If an upstream region contains at least one AP2 and one SP1, it often contains HRE, too).
The association rules induction algorithm was used to find specific rules (group of transcriptional elements) associated to genes modulated by p63 isoforms also contain p53 responsive elements (p53RE). Subsequently we have identified the rules containing p53RE and we have evaluated their frequencies in the transcriptional profile clusters (UP, DOWN e MIX). By this approach we have identify respectively 12, 19 and 20 rules which are statistically specific ($p<0.001$) for DOWN, MIX and UP clusters. We are actually investigated the presence of this rules in hortologous genes and their relative distance in p63-controlled genes.