Putative Estrongen-Responsive Genes database (PERG)

[1]Cordero F, [1]Lazzarato F, [2]Debortoli M, [3]Weisz A, [3]Cicatiello G, [3]Scafoglio C, [3]Basile W and [1]Calogero RA

[(1)]Dip. Scienze Cliniche e Biologiche, c/o Az. Ospedaliera S. Luigi, Regione Gonzole 10, Orbassano (TO).
raffaele.calogero@unito.it
[(2)]Istituto per la Ricerca e Cura del Cancro (IRCC), Candiolo (TO)
michele.debortoli@ircc.it
[(3)]Dipartimento di Patologia generale, Seconda Università degli Studi di Napoli.
alessandro.weisz@unina2.it

Keywords. Genomics, Estrogen, Microarrays, Breast cancer

Introduction

Estrogens are known to regulate the proliferation of breast cancer cells and to alter their cytoarchitectural and phenotypic properties, but the gene networks and pathways by which estrogenic hormones regulate these events are only partially understood.
As starting point to obtain a genome-wide picture of the genes modulated by estrogens we have built a database of the genes having in their putative promoter region Estrogen-responsive Element (ERE).

Results

ZR-75 and MCF7 up-modulated genes, derived by transcriptional profiling experiments [1, 2] were scanned for the presence of ERE sequences (GGTCANNNTGACC, a maximum of two mismatches was permitted [3]) the region -2000/+500 with respect to the 1st transcribed nucleotide. The scanning was performed by PATSEARCH [4] and 82 ERE elements were identified in 49 genes.
The genome-wide scanning for putative estrogen responsive genes was done using a modified version (Lazzarato, unpublished) of PATSER program [5] together with the ERE weighted alignment matrix (ERE-m) generated using the 82 ERE elements previously identified (Fig. 1).

Fig. 1: SLOGOS representation of the EREs used to generate the ERE weighted alignment matrix (ERE-m)

The statistical threshold (10^{-8}) used to identified the presence of EREs was defined selecting the p-value by which all the 49 genes used to generate the ERE-m could be identified in a genome-wide exploration as ERE containing genes.
The human genome-wide exploration yielded a total of 7054 putative ERE-containing genes. Using the GeneOntology (www.geneontology.org) and the EASE tool [6] we could not find any specific enrichment for Biological process subclasses. However, 589/7054 genes were annotated as belonging to "DNA binding" class (GO Molecular function).
To have a more robust description of ERE-containing genes we repeated the genome-wide exploration on two human orthologs (mouse and rat) [7]. This analysis yielded a total of 1631 genes preserving the presence of at least an ERE in the rat or mouse orthologs and 397 preserving at least one ERE in both the orthologs. Since the definition of ERE-containing genes via ERE preservation in orthologous organism might be too stringent, we have also annotated the probability of the ERE containing genes to be markers associated to ER-responsive breast cancers. The dataset of ERE-

containing genes base on orthologous (1631/7054) was integrated with those genes, out of the 7054 human ERE-containing genes, found differentially expressed between ER+ and ER- tumors. This analysis was done using the data sets of Sotiriou [8] and van't Veer [9]. Furthermore, we profiled the expression of the ERE-containing genes using the human and mouse tissues atlas [10] in order to define which are the tissues in which ERE-containing genes are found expressed in physiological conditions.
All the described information are now annotated as part of our RRE database [11]. The access to RRE is possible trough a Spitfire server (www6.unito.it:8443) that has the double advantage of securing sensitive data and avoiding unauthorized access by defining access policies for authorized users (X509 certificate).

Acknowledgements

F. Lazzarato is the recipient of a fellowship supported by a PRIN 2002053274 grant. This work is supported by FIRB RBAU01JTHS and RBNE0157EH.

References

[1] L Cicatiello, C Scafoglio, L Altucci, M Cancemi, G Natoli, A Facchiano, G Iazzetti, R Calogero, N Biglia, M De Bortoli, C Sfiligoi, P Sismondi, F Bresciani and A Weisz J. A Genomic View of Estrogen Actions in Human Breast Cancer Cells by Expression Profiling of the Hormone Responsive Transcriptome *Mol. Endocrinol*. In press.

[2] J Frasor , JM Danes, B Komm, KC Chang, CR Lyttle and BS Katzenellenbogen. Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. *Endocrinology*, 144:4562-4574, 2003.

[3] MD Driscoll, G. Sathya, M Muyan, CM Klinge, R Hilf, RA Bambara Sequence requirements for estrogen receptor binding to estrogen response elements *J Biol Chem.,* 273:29321-29330, 1998.

[4] G Grillo, F Licciulli, S Liuni, E Sbisa and G Pesole PatSearch: A program for the detection of patterns and structural motifs in nucleotide sequences. *Nucleic Acids Res*., 31:3608-3612, 2003.

[5] GZ Hertz and GD Stormo Identifying DNA and protein patterns with statistically significant alignments of multiple sequences *Bioinformatics*., 15:563-577, 1999.

[6] DA Hosack, G Jr Dennis, BT Sherma, HC Lane and RA Lempicki dentifying biological themes within lists of genes with EASE. *Genome Biol.*, 4:R70, 2003.

[7]Lenhard B, Sandelin A, Mendoza L, Engstrom P, Jareborg N, Wasserman WW Identification of conserved regulatory elements by comparative genome analysis. *J Biol*, 2:13, 2003.

[8] C Sotiriou, SY Neo, LM McShane, EL Korn, PM Long, A Jazaeri, P Martiat, SB Fox, AL Harris and Liu Breast cancer classification and prognosis based on gene expression profiles from a population-based study *Proc Natl Acad Sci U S A*., 100:10393-10398, 2003.

[9] LJ van 't Veer, H Dai, MJ van de Vijver, YD He, AA Hart, M Mao, HL Peterse, K van der Kooy, MJ Marton, AT Witteveen, GJ Schreiber, RM Kerkhoven, C Roberts, PS Linsley, R Bernards, SH Friend Gene expression profiling predicts clinical outcome of breast cancer *Nature*, 31, 415:530-536, . 2002.

[10] AI Su, MP Cooke, KA Ching, Y Hakak, JR Walker, T Wiltshire, AP Orth, RG Vega, LM Sapinoso, A Moqrich, A Patapoutian, GM Hampton, PG Schultz, LB Hogenesch. Large-scale analysis of the human and mouse transcriptomes. *Proc Natl Acad Sci U S A.*, 99:4465-4470, 2002

[11] Lazzarato F, Franceschinis G, Botta M and Calogero RA RRE: a tool for the extraction of non-coding regions surrounding annotated genes from genomic datasets Bioinformatics, accepted for publication.