Gene 549 (2014) 33–40

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Analysis of gene expression for studying tumor progression: the case of glucocorticoid administration

Mario Huerta [a,1], José Fernández-Márquez [b], Jose Luis Cabello [b], Alberto Medrano [b], Enric Querol [a], Juan Cedano [c,*,1]

[a] Institut de Biotecnologia i Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
[b] Escola Tècnica Superior de Ingenieria, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain
[c] Laboratory of Immunology, Regional Norte, Universidad de la Republica, Gral. Rivera 1350, Salto 50.000, Uruguay

ARTICLE INFO

Article history:
Received 11 June 2013
Received in revised form 10 June 2014
Accepted 10 July 2014
Available online 11 July 2014

Keywords:
Gene expression
Glucocorticoid
Cancer
Cell stress
NEDD4L

ABSTRACT

Background: Glucocorticoids are commonly used as adjuvant treatment for side-effects and have anti-proliferative activity in several tumors but, on the other hand, their proliferative effect has been reported in several studies, some of them involving the spread of cancer. We shall attempt to reconcile these incongruities from the genomic and tissue-physiology perspectives with our findings.
Methods: An accurate phenotype analysis of microarray data can help to solve multiple paradoxes derived from tumor-progression models. We have developed a new strategy to facilitate the study of interdependences among the phenotypes defined by the sample clusters obtained by common clustering methods (HC, SOTA, SOM, PAM). These interdependences are obtained by the detection of non-linear expression-relationships where each fluctuation in the relationship implies a phenotype change and each relationship typology implies a specific phenotype interdependence. As a result, multiple phenotypic changes are identified together with the genes involved in the phenotype transitions. In this way, we study the phenotypic changes from microarray data that describe common phenotypes in cancer from different tissues, and we cross our results with biomedical databases to relate the glucocorticoid activity to the phenotypic changes.
Results: 11,244 significant non-linear expression relationships, classified into 11 different typologies, have been detected from the data matrix analyzed. From them, 415 non-linear expression relationships were related to glucocorticoid activity. Studying them, we have found the possible reason for opposite effects of some stressor agents like dexamethasone on tumor progression and it has been confirmed by literature. This hidden reason has resulted in being linked with the type of tumor progression of the tissues. In the first type of tumor progression found, new cells can be stressed during proliferation and stressor agents increase tumor proliferation. In the second type, cell stress and tumor proliferation are antagonists so, therefore, stressor agents stop tumor proliferation in order to stress the cells. The non-linear expression relationships among DUSP6, FERMT2, FKBP5, EGFR, NEDD4L and CITED2 genes are used to synthesize these findings.

1. Introduction

A phenotype is a consequence of the interaction among genetic and environmental factors. There are several thousands of disorders caused by single genes, in these cases the variation in certain traits or cellular functions can be controlled by single-expression variations. But this is not the general rule. Usually, traits or cellular functions are controlled by multiple genes, and a combination of multiple traits or cellular functions leads to the final phenotype. Only by using specific tools to study the dependence among phenotypes is it possible to deal with this complexity. This is the case of our study about the dual effect of glucocorticoids. The dramatic increase of microarray sample series motivates a more subtle analysis of gene dependences throughout these large series in order to infer phenotypes from gene behavior. There are several analytical methods that perform a global clustering of the microarray samples, such as Self-Organizing Maps, or which perform local clusterings by considering only a subset of co-expressed genes, such as biclustering, and so on. All of them build the clusters based on gene-expression levels. The methodology developed improves these clustering methods

Abbreviations: HC, hierarchical clustering; SOTA, self-organizing tree algorithm; SOM, Self-Organizing Map; PAM, partitioning around medoids; PC, principal components; NEDD4L, neural precursor cell expressed, developmentally down-regulated 4-like; DUSP6, dual specificity phosphatase 6; FERMT2, Fermitin family member 2; ENaC, Epithelial Na^+ Channel; CACNB1, voltage-dependent calcium channel; FKBP5, FK506 binding protein 5; EGFR, epidermal growth factor receptor; CITED2, Cbp/p300-interacting transactivator, with Glu/Asp-rich carboxy-terminal domain, 2.

* Corresponding author at: Laboratory of Immunology, Regional Norte, Universidad de la Republica, Gral. Rivera 1350, Salto 50.000, Uruguay.
E-mail address: jcedano@unorte.edu.uy (J. Cedano).

[1] Both authors contributed equally to this work.

http://dx.doi.org/10.1016/j.gene.2014.07.027
0378-1119/

because it enriches the final study of the generated clusters and also directs this study to the researcher's objectives. Initially, our system detects different phenotypic changes involving a limited number of relevant genes which are greatly affected by these phenotypic changes. This detection is not based upon the study of gene-expression levels; instead, it is based upon: a) the full expression-dependence among genes, and b) expression-dependence fluctuations. Finally, the phenotypic changes found by our system are linked with the global clusters obtained by the different classical methods. This fact will facilitate the subsequent study of complex phenotype interdependences and regulation.

Our methodology to detect phenotypic changes is based upon the following two principles: First, since the different phenotypes are performed by different sets of co-expressed genes not linearly correlated among each other, the interdependence among these phenotypes cannot be described by a linear-expression relationship. Second, each fluctuation in a correlated expression relationship implies a phenotype change. Whereas the classical global and local methods are focused on the detection of co-expressed genes, the proposed strategy here is focused on the detection of non-linear expression relationships between sets of co-expressed genes, because these relationships describe the interdependence among the phenotypes.

Accordingly, we have incorporated a tool for the detection and classification of non-linear expression relationships in our web-application server for microarray analysis. On the other hand, we have incorporated the most common "global" clustering methods applied to sample series. Their output data are represented together with the detected non-linear expression relationships with the aim to link both (Fig. 2). Furthermore, the system allows for crossing the results with GO, KEGG, PubMed, and other Biomedical DB, to direct the final study to genes related to the researcher's objectives. As a result, an accurate dissection of phenotype interdependences can be performed, and this dissection can be linked to biomedical information and current bibliography to focus the data analysis on the researcher's area of interest. The conclusions that can be extracted from each work will depend in a great extent on the gene-expression matrix used as a base for these analyses and on how well this matrix represents the phenotypes that the researchers wish to study.

The paradoxes are common in cancer research; often, the same gene is reported as a marker of tumor progression as well as a marker of tumor suppression. The research topic of the present work is the convenience of glucocorticoid administration in cancer therapy, and it remains as an open question due to these paradoxes. Stressor agents like dexamethasone have shown positive and negative effects on tumor treatment. On the one hand, glucocorticoids are commonly used as adjuvant treatment for side-effects and have anti-proliferative activity in several tumors while, on the other hand, the proliferative effect of synthetic steroids with predominantly glucocorticoid activity has been reported in several studies, some of them involving the spread of cancer.

We have reconciled these apparent incongruities with our findings, facilitated by the mentioned tools. In the present work we present both the procedure and the practice trying to solve these paradoxes. It is commonly accepted that the immunosuppressive activity of glucocorticoids and other stressors can increase cancer incidence, but their causal relationship seems to be more complex than this. We have found that the pro- and anti-tumoral effects of stressor agents like glucocorticoids depend on the tumor-proliferation type of the tissue. In the first type of tumor proliferation we have found that the new cells can be immediately stressed after proliferation without stopping tumor proliferation. In these tissues, stressor agents would increase tumor proliferation progressively. In the second type we have found that cell stress (and function) is incompatible with tumor proliferation, and during tumorigenesis the new cells remain unstressed and undifferentiated. In these tissues, stressor agents would stop tumor proliferation to then stress the cells.

The proliferation type of each tumor tissue is closely related to the physiology of the healthy tissue. In tissues where the proliferation and function are performed in different areas of the tissue (like colon), the new cells are immediately incorporated from proliferation areas to functional areas, in which cells can be stressed. In these tissues, stressor agents increase tumor proliferation progressively. In tissues that cannot be functional during tissue remodeling, the new cells remain non-functional until the remodeling of the tissue area is finished, and it can then be stressed again (like glia or bone). In these tissues, stressor agents stop tumor proliferation in order to stress the cells.

Thus, the reason for the existence of two types of tumor proliferation seems to be that the type of proliferation pattern is an important property of the healthy tissues that tend to remain, at least in part, when these tissues become tumorous.

2. Results

2.1. Expression-relationship typology reveals the role of genes in phenotypic changes

Whereas an expression-relationship fluctuation points to a phenotype change, the relationship typology indicates the kind of interdependence among the phenotypes. Our tools classify gene relationships by typology, because each different typology indicates a different kind of interdependence. These tools have been used to study the role of some genes related to the glucocorticoid signal in tumors, as this glucocorticoid signal does not always promote the same phenotypic changes.

We have performed the analysis of expression data corresponding to the genes from 60 different tumoral cell-lines in response to 118 drugs. The final data analyzed are a robust gene-expression matrix. Our objective in analyzing these data is to study the phenotypes in common among different types of tissue. These data show the phenotypes where each drug acts, ignoring whether these phenotypes belong to one tissue or to another (two samples of the same tissue can be in two different phenotypes, and different tissue samples can be in the same phenotype). For this purpose, the correlation between the level of expression of the genes and the drug's effect has been initially sought, gathering samples from all tumorous tissues. The samples of different tumor cell-lines become the data-cloud, with their corresponding level of expression and drug effectiveness. From these data we obtain the phenotypes, defined by the expression of the genes, in which the drugs act. Thus, in the final matrix we see that each drug will act in a concrete phenotype, which is common among different types of tissue.

The procedure followed trying to solve the paradoxes about the double effect of glucocorticoids in cancer progression is shown in Fig. 1. This includes sample clustering, detection of non-linear expression relationships, classification by type of curve, access to remote databases, and analysis of dual behaviors. Throughout this paper we will outline the different phases of the pipeline and the results we obtain.

In this process, for the 1416 genes of the final matrix, 11,244 significant non-linear expression relationships classified into 11 different typologies have been detected. From them, 122 genes and 415 non-linear expression relationships were related to glucocorticoid activity. Below, some of the main relationship typologies are described together with the expression relationships that can help us to describe our findings.

2.2. Different expression-relationship typologies and our gene relationships of study

$y = \ln(x)$ Expression-relationship typology followed by trigger and enhancer genes. One of the genes must be over-expressed to make the other gene's over-expression possible.