Revista Mexicana de Física
ISSN: 0035-001X
rmf@ciencias.unam.mx
Sociedad Mexicana de Física A.C.
México

Hernández-Lemus, E.; Baca López, K.
Bursting and synchronization in gene regulatory dynamics
Revista Mexicana de Física, vol. 58, núm. 1, junio-, 2012, pp. 63-68
Sociedad Mexicana de Física A.C.
Distrito Federal, México

# Bursting and synchronization in gene regulatory dynamics

E. Hernández-Lemus* and K. Baca López
*Departamento de Genómica Computacional, Instituto Nacional de Medicina Genómica*
*Periférico Sur No. 4124, Torre Zafiro 2, Piso 6 Col. Ex Rancho de Anzaldo, Álvaro Obregón 01900, México, D.F., México.*
**Also at Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México*
*and at Facultad de Ciencias, Universidad Autónoma del Estado de México, respectively.*

Recibido el 23 de Marzo de 2010; aceptado el 27 de Abril de 2011

Within the enormous complexity associated with gene regulatory dynamics, recently has been discovered a sudden synchronization phenomenon that often occurs within sets of genes (commonly associated with energy and metabolism). This process has been called transcriptional bursting. To study to what extent bursting and synchronization occur in cancer-associated gene de-regulation, we applied the tools of spectral time series analysis, detrended fluctuations analysis and information theory to categorize a set of 1191 whole genome gene expression (steady-state) profiles of both primary breast cancer and normal breast tissue. Results are discussed towards molecular signatures associated with disease.

*Keywords:* Transcriptional pulses; biological synchronization.

Dentro de la enorme complejidad asociada con la dinámica de la regulación genética, recientemente se ha descubierto un fenómeno de sincronización repentina que frecuentemente ocurre dentro de grupos de genes (comúnmente asociados con la energía y el metabolismo). Este proceso ha sido llamado estallido transcripcional. Con el fin de estudiar hasta qué punto ocurren estos estallidos y la sincronización en la des-regulación asociada al cáncer, aplicamos los métodos del análisis espectral de series de tiempo, el análisis centralizado de fluctuaciones y la teoría de la información a un conjunto de 1191 experimentos de perfilado de expresión de genoma completo (en estado estacionario), tanto de tumores primarios de mama, como de tejido sano. Se discuten algunos resultados relacionados a la presencia de firmas moleculares de la enfermedad.

*Descriptores:* Pulsos transcripcionales; sincronización biológica.

PACS: 87.10.Vg; 87.14.gn; 87.18.Tt; 02.50.Fz

## 1. Introduction

It is known that the process of Gene Expression (GE) which transforms the information encoded in DNA into a *functional* RNA message sometimes occurs in bursts or pulses of activity. GE bursting may result from a series of stochastic biochemical events. It is thus important to consider the role that noise and fluctuations play in the complex series of chemical reactions and physical interactions related to transcriptional processes within the cell's mesoscopic environment; since this is an important source of variation on cell behavior and thus on phenotypic conditions and disease. Experimental data on the fluctuations in genetic activity, both between individual cells and within the same cell over time, have showed GE as a noisy process. Analyses of gene expression at the single-cell level had provided insight into oscillatory or nonlinear behavior in asynchronous cells and has revealed cell-to-cell differences arising due to the stochastic nature of gene expression [1]. Noise in gene expression arises not only from the inherent randomness of biochemical processes just described, but also from the fluctuations in cellular components [2]. The phenomenon of gene expression (also known as mRNA transcription or simply transcription) is a complex one. There is a set of control mechanisms collectively called *transcriptional regulation* that take the duty to control *when* transcription occurs and also *how much* mRNA is created. The transcription of a given gene by means of the RNA polymerase enzyme (RNApol) can be regulated or controlled by at least five different biochemical mechanisms. Different levels of gene regulation are thus strongly coupled. An important case of such cooperativity phenomena is that of transcriptional bursts (TBs). It could be observed that protein production often occurs in pulses, each due to a single promoter or transcription factor binding event [3]. TBs have been observed in a variety of biological settings such as: ulcerative colitis [4], endocrine system anomalies [5], and cellular differentiation in developmental processes [6]. Positive feedback between messenger ribonucleic acid (mRNA) and regulatory-protein production may result in bi-stability and stochastic bursts in gene transcription [7–9]. Some researchers even suggest that TBs are responsible for whole genome expression coordination [10].

## 2. Bursting and synchronization analyses

### 2.1. Spectral analysis

To understand the non-linear behavior of gene regulatory interactions it will be useful to consider periodic or quasi-periodic expression levels of groups of genes. One important class of analysis that could be made is by means of *Power Spectral Density* (PSD) calculations. PSD gives us a two-fold understanding of transcriptional bursts and gene synchronization. In the one hand, PSD describes how the *energy* of a signal or a time series is distributed with frequency. This

could give us clues as to what is the dynamic behavior of the local chemical potentials [11] that are, in turn, related to intracellular biochemical activity patterns. In the other hand, PSD describes the evolution of the variance, giving us thus insights on the correlation structure of the associated dynamic processes. For quasi-stationary signals which are evolving in an *unpredictable* (*i.e.* complex) way, we require a frequency up-dating (FU) PSD analysis. The PSD of stochastic quasi-stationary processes can be estimated by using short time Fourier series [12], linear autoregressive models [13] and window spectral algorithm [14]. An appropriate FU algorithm based on a *Nyquist cut-off* is already implemented in the [R]-package *spectrum* [15].

Let us consider $\Gamma$ as a series containing the successive expression levels of a gene, then the corresponding PSD, $I(\omega)$ is given by:

$$I(\omega) = \frac{1}{N}\left|\sum_{t}^{N}\Gamma(t)\exp(-i\,\omega t)\right|^2 ; \quad \omega \in [0,\pi] \tag{1}$$

The presence of periodic behaviors could be detected by considering the following *linear* model for $\Gamma$:

$$\Gamma(t) = \beta\ \cos(\omega t + \phi) + \epsilon_i \tag{2}$$

Here $\beta$ is a positive constant (amplitude), $\omega \in [0,\pi]$, $\phi$ is a uniformly distributed *phase shift* ($\phi \in (-\pi,\pi]$) and $\{\epsilon_i\}$ is a sequence of uncorrelated random variables with mean 0 and variance $\sigma^2$ independent of $\phi$ (*i.e.* a Gaussian noise). Under this model periodic behavior could be traced-off by means of looking at *significant* peaks in the PSD, either within an $\omega$-continuous process or more commonly with $\omega$ taking discrete values

$$\frac{2\pi k}{N};\ \ k = 0,1,2\ldots,\left[\frac{N}{2}\right],$$

each of these discrete values is known as a Fourier frequency. If a given time series $\Gamma$ has hidden periodic components, say with a given frequency $\omega^\star$, then the PSD will show a peak at $\omega^\star$. If, in the other hand $\Gamma$ is a random, aperiodic signal, then the $I(\omega)$ vs $\omega$ plot will be a straight line. This corresponds to $\beta = 0$ in the linear model, Eq. 2. It is important to notice that only positive values of $\Gamma(t)$ are physically sound, for that reason sinusoidal models like the one in Eq. 2 are only approximations to the actual expression levels. There are two ways to deal with this limitation in practice, either you round off negative values to zero or add a positive *background constant* to all the values in the series. We took here the second option since it preserves the relative importance of all the peaks. In the PSD analysis of stochastic variables, one is to test the null hypothesis $\beta = 0$ versus the data in order to explicitly state what a *significant* peak is. There are many ways to do this, but almost all of them deal with asymptotic considerations. Nevertheless, an early result from Fisher applies also to finite time series, the so called g-statistic [16]:

$$g = \frac{\max_k I(\omega_k)}{\sum_{k=1}^{N/2} I(\omega_k)} \tag{3}$$

Values of $g$ larger than *expected* lead to the rejection of the null hypothesis of a purely random process. The exact distribution of $g$ is given by:

$$P(g > x) = \alpha(1-x)^{\alpha-1} - \frac{\alpha(\alpha-1)}{2}(1-2x)^{\alpha-1} + \cdots + (-1)^r\frac{\alpha!}{r!(\alpha-r)!}(1-rx)^{\alpha-1} \tag{4}$$

Here $\alpha = N/2$ and $r$ is the largest integer less than $1/x$. So if the observed value of $g$ is $g^\star$ then there is a p-value $P(g > g^\star)$ to evaluate the null hypothesis.

### 2.2. Detrended fluctuations analysis

It is known that a bounded time series could be mapped to a self-similar process by integration. Peng, *et al.* [17] introduced a modified root mean square analysis of the underlying random walk that has been termed detrended fluctuation analysis (DFA). DFA detects self-similar patterns even if they are embedded in an apparent non-stationary frame, and also avoids the spurious detection of artificial self-similarity due to trending of the probability distribution function. To do so, DFA first integrates the time series as follows. Let $\Gamma(i)$ be a dynamic-series (with discrete steps *i*), then the k-integrated value $\Gamma(k)$ is given by:

$$\Gamma(k) = \sum_{i=1}^{k}\left(\Gamma(i) - \hat{\Gamma}\right) \tag{5}$$

here $\hat{\Gamma}$ is the average value of $\Gamma(i)$ in the considered interval. $\Gamma(k)$ gives a mapping from a time series to a self similar process. To measure the characteristic scale for the integrated time series, time is divided into isometrical bins of length $n$. For every bin it is performed a least squares linear fitting of the data (called the *local trend*). The calculated coordinate of the straight line is denoted by $\Gamma_n(k)$. To *detrend* $\Gamma(k)$ we substract the linear local trend $\Gamma_n(k)$. For a given bin size $n$, the characteristic length-scale for the fluctuations in the integrated and detrended series is:

$$\mathcal{F}(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\Gamma(k) - \Gamma_n(k)\right)^2} \tag{6}$$

In the DFA algorithm, the value of $\mathcal{F}(n)$ is calculated to get a relationship between $\mathcal{F}(n)$ and the bin size $n$. $\mathcal{F}(n)$ will usually increase with $n$. A linear relationship on a log-log plot (*i.e.* a power law) implies self-similarity of the related fluctuations. The slope of this log-log plot determines the scaling exponent $\alpha$.

Thus,

$$\mathcal{F}(n) \sim n^\alpha \tag{7}$$

DFA has revealed the extent of long-range correlations in apparently irregular time series [17, 18]. A value of $\alpha$ greater

than 0.5 indicates the presence of persistent long-range power law correlations. The case $\alpha$ = 1.0 has raised a lot of of interest both from physicists and biologists for many years, it corresponds to $1/f$ noise [19]. $0 < \alpha < 0.5$ denotes the presence of power-law anti-correlations, such that large values are more likely to be followed by small values and vice versa [17].

### 2.3. Information based similarity

The analysis of complex time series, is usually intended to discover hidden information within the repetitive appearance of certain patterns embedded in the given signals. To detect and quantify such underlying structures several methods have been developed. One approach is based on the quantification of the so called Information-Based Similarity Index (IBS) [20], and has proved to be a very powerful tool in the comparison of the dynamics of highly nonlinear processes [21, 22].

For $\Gamma(i) = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_N\}$, we map each pair of successive points into one of the following binary states $I_n$, if $(\Gamma_{n+1} - \Gamma_n) < 0$ then $I_n = 0$; in the other case $((\Gamma_{n+1} - \Gamma_n) \geq 0)$ $I_n = 1$. This turns the $N$ step real-valued time series $\Gamma(i)$ into an $N-1$ step binary-valued series $I(i)$. It is possible then, to define a binary sequence of length $m$ (called an $m$-bit word). Each m-bit word, $w_k$ represents a unique pattern of fluctuations. For every unitary time-shift $\lambda$, the algorithm makes a different collection $W_\lambda$ of m-bit words over the whole time series, $W_\lambda = \{w_1, w_2, \ldots, w_n\}_\lambda$. It is expected that the frequency of occurrence of these m-bit words will reflect somehow the underlying dynamics of the original (real-valued) time series.

To apply this concept to symbolic sequences, one should consider the frequency of every m-bit word and then sort them in descending order by frequency of occurrence, in this way we are able to write down a probability distribution function in the *rank-frequency* representation (RF-PDF). This RF-PDF represents the statistical hierarchy of symbolic words of the original time series [20]. Two given symbolic sequences are taken as *statistically equivalent* if they give rise to similar probability distribution functions. Yang and coworkers [20] defined a measure of similarity (akin to statistical equivalence) between two time series by plotting the rank number of every m-bit word in the first time series with the rank for the same m-bit word in the second time series. Obviously if the two RF-PDFs are statistically equivalent, then the scattered points will lie *almost surely* in the diagonal line. The average deviation of these points from the diagonal is then a measure of the distance (or dissimilarity) between these two time series. For finite time series the m-bit words are not equally likely to appear, then the method introduces the likelihood of each word by defining a weighted distance $\Delta_m$ between two given symbolic sequences $\sigma_1$ and $\sigma_2$ as follows:

$$\Delta_m(\sigma_1, \sigma_2) = \frac{1}{2^m - 1} \sum_{k=1}^{2^m} |R_1(w_k) - R_2(w_k)| F(w_k) \quad (8)$$

$F(w_k)$ is the normalized likelihood of the m-bit word k, weighted by its given Shannon entropy, *i.e.*:

$$F(w_k) = \frac{1}{Z}[-p_1(w_k)\log p_1(w_k) - p_2(w_k)\log p_2(w_k)] \quad (9)$$

in this case, $p_i(w_k)$ and $R_i(w_k)$ represent the probability and rank of a given word $w_k$ in the i-th series. The normalization factor in Eq. 9 is the total Shannon's entropy of the ensemble and is calculated as

$$Z = \sum_k [-p_1(w_k)\log p_1(w_k) - p_2(w_k)\log p_2(w_k)] .$$

$\Delta_m(\sigma_1, \sigma_2)$ is called the Information Based Similarity Index (IBS) between series $\sigma_1$, and $\sigma_2$. One notices that $\Delta_m(\sigma_1, \sigma_2) \in [0, 1]$; $\forall \sigma_1, \sigma_2$; $\forall m$. In fact one is able to consider $\Delta_m(\sigma_1, \sigma_2)$ as a probability measure. In the situation in which $\lim \Delta_m(\sigma_1, \sigma_2) \to 1$ the series are absolutely dissimilar, whereas in the opposite case given by $\lim \Delta_m(\sigma_1, \sigma_2) \to 0$ the two series become identical (in the statistical sense).

## 3. Results and discussion

The above mentioned techniques were applied to a set of publicly available whole genome (steady state) gene expression microarray experiments. In the case of human mRNA samples taken directly from organ tissue (by a biopsy) and not from cultured cell-lines, it is extremely difficult to design time-course experiments. As a proxy to time-courses we propose the following alternative: after quality control preprocessing, background correction and normalization of the microarrays [26], the samples were prioritized (ordered) according to their *BNIP3* (Affymetrix-probe ID 201848_s_at) expression level, in such a way that sample 1 has the smallest value of gene expression for *BNIP3* and sample 1191 the largest one. *BNIP3* is a well known marker of progression and malignancy in primary breast cancer that correlates both with lab tests and clinical trials [27]. A tumor progression marker is a molecule whose concentration steadily increases as the cancer cells are growing and invading healthy tissue. Thus within clinical standards, a higher concentration of such a molecule is then considered to be equivalent to a later, more advanced degree of cancer. Given these facts, our ansatz is that a higher level of BNIP3 implies a later stage of cancer progression, hence a later point in the corresponding dynamic evolution. By ordering the independent, steady-state samples in this way, to constitute a proxy to a time series, we intend to mimic the dynamics of cancer progression.

### 3.1. Spectral analysis of TBs

As it could be already seen in the example given by Figs. 1 and 2, the dynamic behavior of the expression levels of genes

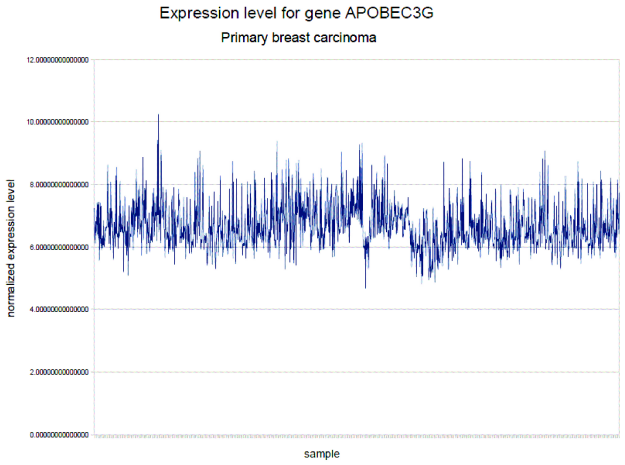


FIGURE 1. Expression level time series for gene *APOBEC3G*

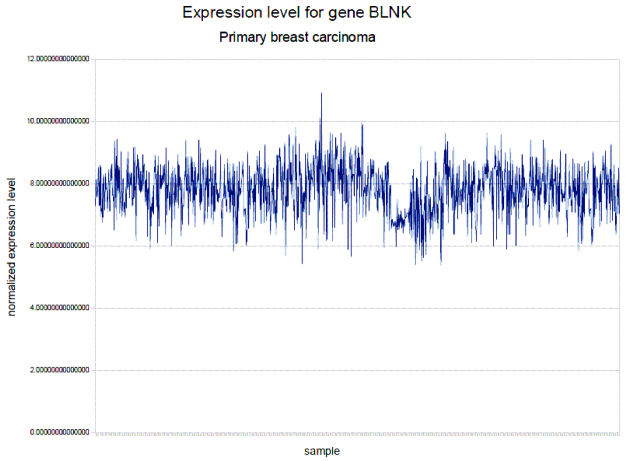


FIGURE 2. Expression level time series for gene *BLNK*

is quite complex. In order to somehow detect the presence of patterns of periodic or quasi-periodic behavior (more appropriately of *hints* of it) we calculate the PSD for the associated series, since even in the power spectra there are many signals we choose to follow only the most significant ones. The presence of statistically significant enrichment of Fourier modes (quasi-periodicities) is shown in Table I (last column). Of the 30 genes considered just 14 of them showed quasi-periodic behavior. In some cases the Fourier modes of sets of genes are correlated. For example in the case of gene *APOBEC3G* the most significant peaks are showed on Fig. 3 (left panel). By studying the frequency distribution of such significant peaks we found that in some cycles there are a considerable overlap with the PSD of gene *BLNK* (Fig. 3 right panel).

If we analyze the IBS correlation measure between these two genes we have two independent methods pointing out to synchronization phenomena between these two genes. Other interesting cases of synchronization are present between *APOBEC3G* and *GALC* and also between *GALC* and *MNDA*. If we consider the quasi-periodic coupling between these genes, we could anticipate a transcriptional motif that

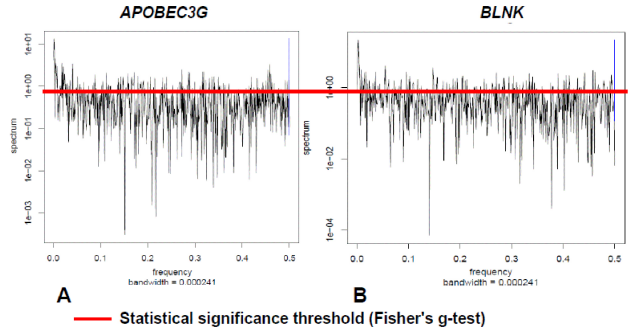


FIGURE 3. Power spectral density (PSD) for the FU-Fourier modes of *APOBEC3G* and *BLNK*. Significant peaks are above a g-test threshold indicated by the red-line (p-value = 0.05 in Eq. 4). We notice that a number of significant frequency peaks are shared by both genes, in the corresponding times these two genes are *synchronized* in their expression

could serve as a basis for the propagation of TF activity from *APOBEC3G* and *MNDA* (for an independent confirmation of this motif from the standpoint of IBS calculations see Fig. 5).

### 3.2. Detrended fluctuation analysis of TBs

As it could be seen in Table I, gene expression dynamics are positively correlated ($\alpha > 0.5$), some genes however, are strongly correlated (*LAMP3*, *ELMO1* and *TNFRSF13B* for instance) whereas others are just slightly correlated (being the most noticeable examples *POU2AF1*, *IL10RA* and *PSCDBP*, which are almost undistinguishable from white noise (specially *POU2AF1*)). Closer inspection of the $\alpha$-values reveal an interesting trend, since one could notice that genes with transcription factor (TF) activity present, in general, lower values of $\alpha$ (hence lower *correlation lengths*) than non-TF genes. This may point to TF genes as being more prune to show stochastic events (*e.g.* TBs), this in turn, could be related with TFs having lower chemical potentials of transcription [11]. This is the case with the aforementioned transcriptional motif formed between *APOBEC3G* and *MNDA* as transcription factors for *GALC*. A word of caution must be stressed however, since here we studied the evolution of just 30 genes (within a whole genome environment) there is no conclusive evidence supported, for example, via statistical significance analysis. More studies are to be focused on this tendency before any *hard* conclusion could be stated.

### 3.3. Information Based Similarities of TBs

We build several sets of non-linear (IBS) correlation networks, from the series associated with the group of 30 genes under consideration. In each of these networks we set a value of IBS cut-off and also a bin size (*i.e.* a size for the maximum length of the m-bit words considered in the analysis). We could see in Fig. 4 how the number of nonlinear interactions (network links) between the genes (nodes in the network) varies as a function of both $\Delta_{i,j}$ the IBS-value and $m$ the bin-size between 30 trivial interactions (every gene is

TABLE I. Scaling exponent $\alpha$ for the linear Detrended Fluctuations Analysis (DFA) and PSD quasi-periodic modes of expression levels for the set of 30 genes. Asterisks in the last column denote the presence of statistically significant (Eq. 4) Fourier modes in Eq. 1: * $\rightarrow$ less than 20 enriched Fourier modes, ** $\rightarrow$ more than 20 enriched Fourier modes, *** $\rightarrow$ seemingly chaotic behavior (*i.e.* all computed Fourier modes are enriched)

| Gene name | DFA $\alpha$ | DFA $R^2$ | PSD enrichment |
|---|---|---|---|
| TNFRSF1B | 0.83 | 0.98 | * |
| SELL | 0.73 | 0.96 | ** |
| ELMO1 | 1.02 | 0.98 | |
| APOBEC3G | 0.74 | 0.98 | * |
| MNDA | 0.75 | 0.98 | * |
| BLNK | 0.76 | 0.97 | * |
| GLDC | 0.88 | 0.98 | * |
| IL2RB | 0.64 | 0.98 | * |
| CD69 | 0.79 | 0.99 | ** |
| SP140 | 0.64 | 0.99 | |
| IL10RA | 0.58 | 0.99 | |
| LAMP3 | 1.17 | 0.98 | |
| IL12A | 1.04 | 0.98 | |
| CD38 | 0.67 | 0.99 | * |
| IGHM | 0.64 | 0.98 | * |
| PSCDBP | 0.59 | 0.97 | ** |
| POU2AF1 | 0.54 | 0.99 | *** |
| GALC | 0.83 | 0.97 | * |
| SMAD3 | 0.69 | 0.99 | |
| NCKAP1L | 0.71 | 0.98 | |
| TNFRSF14 | 0.9 | 0.98 | |
| PTPN22 | 0.83 | 0.98 | |
| SKAP1 | 0.6 | 0.99 | |
| SPARC | 0.99 | 0.98 | * |
| TNFRSF13B | 1.03 | 0.98 | |
| SYK | 0.88 | 0.98 | |
| SLC2A5 | 0.85 | 0.97 | |
| PLEK | 0.61 | 0.98 | |
| RHOH | 0.73 | 0.98 | |
| MEF2C | 0.96 | 0.99 | |

dynamically correlated with itself) to 900 (all genes are connected). An intersection set of the most significant interactions could be used as a model to built a *synchronization network* of genes that share significant patterns of variation. These genes are said to be *synchronized*. In Fig. 5 we present a transcriptional regulation network of synchronized genes, it is noteworthy the possible role of TFs (genes shown as green nodes) as *initiators* of synchronicity (they are said to be *upstream of the transcriptional cascade*). Thus, if TFs present abnormal stochastic bursts, these would cause de-regulation

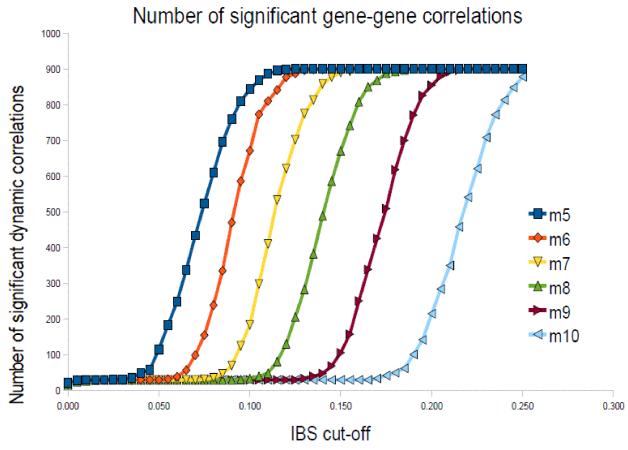


FIGURE 4. Significant gene-gene dynamic correlations versus bin size $m$ (for $m = 5, 6, \ldots, 10$) and IBS-measure value.

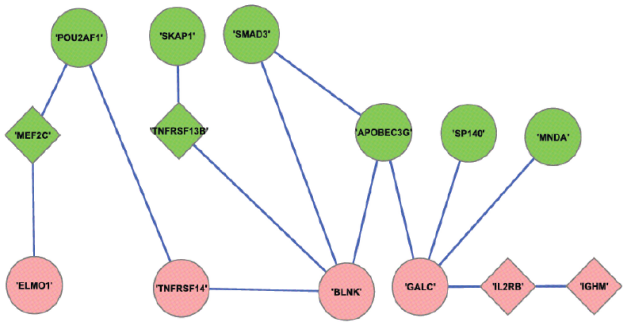


FIGURE 5. Synchronization network (green nodes are transcription factors, pink nodes are target genes). The *degree of synchronicity* is based on an IBS-measure.

phenomena in the entire network of synchronized target genes, hence abnormal systemic (*i.e.* genome-wide) transcriptional behavior. Some instances of this abnormal behavior are related to diseased states.

## 4. Conclusions

By means of non-linear time-series techniques, such as the quantification of the Power Spectral Density profiles (PSD), as well as statistically significant frequency peaks in the related spectra; Detrended Fluctuations Analysis (DFA) of the extent of long-range correlations and Information Based Similarity (IBS) inference of *synchronization networks* we have studied the complex phenomena of anomalous *bursting* (sudden episodes of increased expression levels for certain genes) and *synchronization* (the presence of time-correlations and simultaneous rising of expression within sets of genes) with relation to breast cancer phenomenology. Transcriptional network analysis have showed that, instead of being independent, different levels of gene regulation are strongly coupled. This cooperativity-complexity relationship take special importance with regards to pathologies strongly related to gene de-regulation such as cancer. An important case of such cooperativity phenomena is that of transcriptional bursts (TBs)

which are experimentally verifiable outcomes of this phenomena. It has been stated that positive feedback between mRNA and regulatory-protein production may result in bistability and seemingly stochastic bursts in gene transcription. In fact, when explicit mRNA diffusion is considered, the bursts may be much more irregular; the periods between bursts may be much shorter; and, depending on the circumstances, the burst window may be reduced or extended. All these complex physicochemical phenomena could be related with anomalies in their corresponding dynamics and thus can be observed in a time series of measurements of gene activity.

For example, PSD points out to the presence of quasi-periodic behavior embedded in the apparent random patterns of gene expression levels, some of the related Fourier modes are even shared between groups of genes, this have brought some evidence of gene synchronization. IBS analyses also aport information pointing towards synchronized regulation, even at the network level. In the other hand, DFA apparently shows that transcription factors (TFs) show shorter correlation lengths and are thus, more prune to de-correlate in time and present stochastic phenomena. This could be behind the biological origin of anomalous transcriptional bursts, because stochastic pulses in the levels of TFs (that are up-stream in the transcriptional cascade) would cause their target genes (TGs) to also became disturbed. Nevertheless due to the comparatively longer range of correlation of TGs the system could reach homeostasis in many instances.

This study has focused on a set of 30 well-characterized genes (mainly related to transcription factor activity and metabolism) in a curated set [26] of 1191 whole genome gene expression microarray experiments. Some trends have become apparent, nevertheless more comprehensive studies are needed before any definite conclusion could be drawn.


1. D. Longo, and J. Hasty, *Molecular Systems Biology* **2** (2006) 64. doi:10.1038/msb4100110.
2. A. Becskei, B.B. Kaufmann, and A. van Oudenaarden, *Nature Genetics* **37** (2005) 937-944.
3. P.J. Ingram, M.P.H. Stumpf, and J. Stark, *PLoS Comput Biol* **10** (2008) e1000192.
4. X. Jin, A.P. Malykhina, F. Lupu, and H.I. Akbarali, *Am J Physiol Gastrointest Liver Physiol* **287** (2004) G274-G285.
5. S.L. Shorte, *et al.*, *Endocrinology* **143** (2002) 1126-1133.
6. Y Niwa, Y. Masamizu, T. Liu, R. Nakayama, and C. Deng, R. Kageyama, *Developmental Cell* **13** (2007) 298-304.
7. A. Raj and A. van Oudenaarden, *Annual Review of Biophysics* **38** (2009) 255-270.
8. T. Tripathi *et al.*, *Euro. Phys. Lett.* **84** (2008) 68004.
9. V. Zhdanov, *JETP Letters* **85** (2007) 302-305.
10. L. Cai, C.K. Dalal and M.B. Elowitz, *Nature* **455** (2008) 485-490.
11. E. Hernández-Lemus, *Journal of Non-equilibrium Thermodynamics* **34** (2009) 4 [In press]
12. E. Jacobsen and R. Lyons, *Signal Processing IEEE* **20** (2003) 74-80.
13. L. Marple, *IEEE Transactions on Acoustics, Speech and Signal Processing* ASSP-31, **4** (1983) 56-65.
14. K.M. Aamir and M.A. Maud, *World Academy of Science, Engineering and Technology* **2** (2005) 92-95.
15. For the spectrum calculation [R] package please refer to: http://rss.acs.unt.edu/ under /Rdoc/library/stats/html/spectrum.html For the theory behind see, for example, B. Kedem, and K. Fokianos, *Regression Models for Time Series Analysis* (John Wiley and Sons, N. York, 2002)
16. R.A. Fisher, *Proc. Roy. Soc. A* **125** (1929) 54-59.
17. C-K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, and A.L. Goldberger, *Phys Rev E* **49** (1994) 1685-1689.
18. C-K Peng, S. Havlin, H.E. Stanley, and A.L. Goldberger, *Chaos* **5** (1995) 82-87.
19. W. Li and D. Holste, *Physical Review E* **71** (2005) 0419410-19.
20. A.C. Yang, S.S. Hseu, H.W. Yien, A.L. Goldberger and C-K. Peng, *Phys Rev Lett* **90** (2003) 108103.
21. E. Hernández-Lemus, D. Velázquez-Fernández, J.K. Estrada-Gil, I. Silva-Zolezzi, M.F. Herrera-Hernández, and G. Jiménez-Sánchez, *Physica A* **388** (2009) 5057-5069.
22. E. Hernández-Lemus, J.K. Estrada-Gil, I. Silva-Zolezzi, J.C. Fernández-López, A. Hidalgo-Miranda, and G. Jiménez-Sánchez, *American Institute of Physics Conf. Proc. Volume 978 (Biological Physics)*, **1** (2008) 34-56.
23. D. Ruelle, *Chaotic Evolution and Strange Attractors* (Accademia Nationale dei Lincei, Cambridge University Press, Cambridge 1989).
24. P. Grassberger, I. Procaccia, *Phys. Rev. Lett.* **50** (1983) 346-349, see also P. Grassberger, and I. Procaccia, *Phys. Rev. A* **28** (1983) 2591.
25. www.physionet.org/physiotools/
26. K. Baca-López, E. Hernández-Lemus, and M. Mayorga, *Rev. Mex. Fis.* (2009) [In press]
27. E.Y. Tan, L. Campo, C. Han, H. Turley, F. Pezella, K.C. Gatter, A.L. Harris, and S.B. Fox, *Clin. Cancer Res.* **13** (2007) 467-474.