Methods 83 (2015) 36–43

Contents lists available at ScienceDirect

Methods

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

Cerebrospinal fluid protein dynamic driver network: At the crossroads of brain tumorigenesis

Zhou Tan [a,b,1], Rui Liu [b,c,1], Le Zheng [b,d,1], Shiying Hao [b,1], Changlin Fu [b,e], Zhen Li [b], Xiaohong Deng [b], Taichang Jang [b], Milton Merchant [b], John C. Whitin [b], Minyi Guo [e], Harvey J. Cohen [b,2], Lawrence Recht [b,2], Xuefeng B. Ling [b,2,*]

[a] Hangzhou Normal University, Zhejiang 311121, China
[b] Stanford University, Stanford, CA 94305, USA
[c] South China University of Technology, Guangzhou 510640, China
[d] Tsinghua University, Beijing 100084, China
[e] Shanghai Jiao Tong University, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 2 May 2015
Accepted 5 May 2015
Available online 14 May 2015

Keywords:
Transition state
Dynamical driver network (DDN)
Critical transition
Tumorigenesis progressing
Network entropy

ABSTRACT

To get a better understanding of the ongoing *in situ* environmental changes preceding the brain tumorigenesis, we assessed cerebrospinal fluid (CSF) proteome profile changes in a glioma rat model in which brain tumor invariably developed after a single *in utero* exposure to the neurocarcinogen ethylnitrosourea (ENU). Computationally, the CSF proteome profile dynamics during the tumorigenesis can be modeled as non-smooth or even abrupt state changes. Such brain tumor environment transition analysis, correlating the CSF composition changes with the development of early cellular hyperplasia, can reveal the pathogenesis process at network level during a time before the image detection of the tumors. In our controlled rat model study, matched ENU- and saline-exposed rats' CSF proteomics changes were quantified at approximately 30, 60, 90, 120, 150 days of age (P30, P60, P90, P120, P150). We applied our transition-based network entropy (TNE) method to compute the CSF proteome changes in the ENU rat model and test the hypothesis of the critical transition state prior to impending hyperplasia. Our analysis identified a dynamic driver network (DDN) of CSF proteins related with the emerging tumorigenesis progressing from the non-hyperplasia state. The DDN associated leading network CSF proteins can allow the early detection of such dynamics before the catastrophic shift to the clear clinical landmarks in gliomas. Future characterization of the critical transition state (P60) during the brain tumor progression may reveal the underlying pathophysiology to device novel therapeutics preventing tumor formation. More detailed method and information are accessible through our website at http://translationalmedicine.stanford.edu.

1. Introduction

The influence of the local environment in cancer development, clearly established in several systemic neoplasms including colon, breast and prostate cancers [1–3], remains unexplored in gliomas. An ideal approach to study the early cancer development preceding the clinical landmark of brain tumor is to analyze abnormalities in distinct time-series prior to the detection of the apparent malignancy. However, brain tumor develops with abnormal cells forming inside the brain, which significantly limits the study of its origin due to the limited access to the tissue.

Approximately 10–30% of all cerebrospinal fluid (CSF) is extrachoroidal in origin and is represented by bulk flow of the interstitial fluid from brain parenchyma into the ventricles and subarachnoid space [4–6]. With this readily accessible sample source, we previously profiled CSF proteome to survey brain environment alterations prior to impending hyperplasia by surface-enhanced laser desorption/ionization TOF mass spectrometry (SELDI-TOF-MS). SELDI-TOF-MS has been used successfully to identify biomarkers in blood from various malignancies using comparative proteomic strategies [6–8].

While there have been several clinical studies that attempted to identify biomarkers of brain tumor using comparative proteomic techniques [9–11], failure in controlling variables such as age,

* Corresponding author.
E-mail address: bxling@stanford.edu (X.B. Ling).
[1] Authors contributed equally as first authors.
[2] Authors contributed equally as last authors.

http://dx.doi.org/10.1016/j.ymeth.2015.05.004
1046-2023/

space occupying volume and tissue permeability prevented these studies from recognizing whether a changed protein expression pattern accurately represented an effect of the neoplastic process. To control these variables, we assessed changes in CSF proteome at days P30, P60, P90, P120 and P150 in a rat model, of which gliomas invariably developed after a single *in utero* exposure to the neurocarcinogen ethylnitrosourea (ENU).

Given that the rat gliomas are not generally detectable pathologically until approximately 90 days of age (P90), we hypothesized that brain tumor progression can be modeled into three states: (1) a pre-hyperplasia state with high resilience and robustness to perturbations; (2) a critical state defined as the prelude to catastrophic shift into the hyperplasia state, occurring before the imminent phase transition point is reached, and with low resilience and robustness due to its dynamical structure; (3) a hyperplasia state representing a seriously deteriorated stage possibly with high resilience and robustness, when the system usually finds it difficult to recover or return to the normal state even after intervention. This hypothesis was supported by the observations that there was usually sudden health catastrophic shift during the gradual progression of many chronic diseases [12–17]. The drastic or a qualitative transition in the focal system or network, from a normal state to a disease state, corresponds to a so-called bifurcation point in dynamical systems theory [18,19]. Various critical transition phenomena have been reported in climate and ecosystems [20]. When the system is near the critical point, there exists a dominant group which we defined as dynamic driver network (DDN) of features satisfying the following three conditions: (1) the correlation between any pair of members in DDN becomes very strong; (2) the correlation between one member of DDN and any other molecule of non-DDN becomes very weak; (3) any member of DDN becomes highly fluctuating during transition [21–23]. We previously employed transition-based network entropy (TNE) to effectively identify the DDN as well as the transition state [21]. The TNE was actually an improved Shannon entropy [24] that was conditional on the previous state of a local dynamical network in a Markov process, which was also the entropy rate of the state change in a feature space network, where each node represented a feature and each edge represented a regulatory relation between two features, with the assumption that a Markov process governed the dynamics of each node. Given a high dimensional feature network, we found that the TNE was drastically increasing when the system approached the transition state, whereas there were no significant TNE fluctuations at either normal or disease states.

In this study, we set to assess the CSF proteome profile dynamics and test our hypothesis of non-smooth or even abrupt state changes during the glioma tumorigenesis. Such brain tumor environment transition analysis, correlating the CSF composition changes with the development of early cellular hyperplasia, can reveal the pathogenesis process at network level during a time before the imaging detection of the tumors.

2. Materials and methods

In this section, we describe the experimental procedure and the theoretical basis, i.e., the TNE score, and the mathematical basis of DDN method (Fig. 1); some details are given in Supplementary information.

2.1. Data acquirement and ethics

Case (ENU) and control rat handling was in accordance with guidelines for animal safety and welfare. Rat CSF proteomics experiment was approved by the Stanford IUCAC (Protocol #11936).

2.2. ENU administration, rat CSF collection, histological analysis, and CSF proteomics

ENU rat glioma model, ENU administration, rat CSF collection and subsequent histological analysis were as previously described [6]. CSF proteomics profiling and subsequent data analysis were as previously described (Table 1) [6,25,26].

2.3. Markov process of the network evolution

The dynamics for the progression of complex diseases are very complicated either before or after sudden deterioration, and therefore the state equations are generally constructed in a high-dimensional space with a large number of variables and parameters. However, when the system is driven by a group of parameters to approach to a critical point, theoretically the system can be expressed in a very simple form, generally by one- or two-variable dynamical equations in an abstract phase space around a codimension-one bifurcation point. This is generally guaranteed by both the bifurcation theory and center manifold theory [23]. Based on this special feature during this special phase, we derived the dynamical characteristics of the network at this stage to detect the critical transition.

Specifically, we first defined the network state (or original variables) and transition state of a dynamical network in a Markov process. For an n-node network, let

$$Z(t) = (z_1(t), \ldots, z_n(t))$$

represent the network state at the sampling point t, where $z_i(t)$ denotes the expression value of node (i.e., feature i). Then, $x_i(t) \in \{0, 1\}$ is defined to measure whether or not node i has a large change at t, that is, if $|z_i(t) - z_i(t-1)|$ is sufficiently large ($\geqslant d_i$), $x_i(t) = 1$, otherwise $x_i(t) = 0$, where d_i is a positive constant threshold or the threshold. Thus, $X(t) = (x_1(t), \ldots, x_n(t))$ represents the transition state for the network at t.

Next, a local network structure centered on each node is defined, which is the basis to construct the conditional network entropy. Assume that node i has m linked first-order neighbor nodes $i_1, i_2, \ldots, i_m$, which composes a local network centered on node i with local transition state $X^i(t) = (x_i(t), x_{i_1}(t), \ldots, x_{i_m}(t))$ at t. Clearly, from the current state $X^i(t)$ at time t, there are totally 2^{m+1} possible state transitions (or possible transition states), which is denoted as $\{A_u\}_{u=1,2,\ldots,2^{m+1}}$ for this local network at the time point $t+1$ (see Fig. 2A). To simplify notation, $X^i(t)$ is denoted as $X(t)$, and transition state is denoted as state.

From the network structure, the Markov matrix $P = (p_{u,v})$ can be derived, where $p_{u,v}(t)$ describes the transition rate from state u to state v with

$$p_{u,v}(t) = \Pr(X(t+1) = A_v | X(t) = A_u), \tag{1}$$

where $u, v \in \{1, 2, \ldots, 2^{m+1}\}$ and $\sum_v p_{u,v}(t) = 1$. Thus, we have the following the stochastic Markov process for $X(t)$

$$\{X(t+i)\}_{i=0,1,\ldots} = \{X(t), X(t+1), \ldots, X(t+i), \ldots\} \tag{2}$$

with $X(t+i) = A_u, u \in \{1, 2, \ldots, 2^{m+1}\}$.

2.4. Theoretical derivation near the critical point

Consider the following discrete-time dynamical system representing dynamical evolution of a network

$$Z(t+1) = f(Z(t); P), \tag{3}$$

where $Z(t) = (z_1(t), \ldots, z_n(t))$ is an n-dimensional state vector or variable at time instant k representing feature values,