Methods xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Methods

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

Induction of apoptosis through ER stress and TP53 in MCF-7 cells by the nanoparticle $[Gd@C_{82}(OH)_{22}]_n$: A systems biology study

Lin Wang [a], Jie Meng [b,*], Weipeng Cao [b], Qizhai Li [a], Yuqing Qiu [a], Baoyun Sun [c], Lei M. Li [a,*]

[a] NCMIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[b] CAS Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety, National Center for Nanoscience and Technology of China, Beijing 100190, China
[c] CAS Key Laboratory for Biomedical Effects of Nanomaterials and Nanosafety and Key Laboratory for Nuclear Techniques, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Nanoparticle
Microarray normalization
Transcriptional inference
Apoptosis
ER stress
TP53

ABSTRACT

The nanoparticle gadolinium endohedral metallofullerenol $[Gd@C_{82}(OH)_{22}]_n$ is a new candidate for cancer treatment with low toxicity. However, its anti-cancer mechanisms remain mostly unknown. In this study, we took a systems biology view of the gene expression profiles of human breast cancer cells (MCF-7) and human umbilical vein endothelial cells (ECV304) treated with and without $[Gd@C_{82}(OH)_{22}]_n$, respectively, measured by the Agilent Gene Chip G4112F. To properly analyze these data, we modified a suit of statistical methods we developed. For the first time we applied the sub–sub normalization to Agilent two-color microarrays. Instead of a simple linear regression, we proposed to use a one-knot SPLINE model in the sub–sub normalization to account for nonlinear spatial effects. The parameters estimated by least trimmed squares- and S-estimators show similar normalization results. We made several kinds of inferences by integrating the expression profiles with the bioinformatic knowledge in KEGG pathways, Gene Ontology, JASPAR, and TRANSFAC. In the transcriptional inference, we proposed the BASE2.0 method to infer a transcription factor's up-regulation and down-regulation activities separately. Overall, $[Gd@C_{82}(OH)_{22}]_n$ induces more differentiation in MCF-7 cells than in ECV304 cells, particularly in the reduction of protein processing such as protein glucosylation, folding, targeting, exporting, and transporting. Among the KEGG pathways, the ErbB signaling pathway is up-regulated, whereas protein processing in endoplasmic reticulum (ER) is down-regulated. CHOP, a key pro-apoptotic gene downstream of the ER stress pathway, increases to nine folds in MCF-7 cells after treatment. These findings indicate that ER stress may be one important factor that induces apoptosis in MCF-7 cells after $[Gd@C_{82}(OH)_{22}]_n$ treatment. The expression profiles of genes associated with ER stress and apoptosis are statistically consistent with other profiles reported in the literature, such as those of HEK293T and MCF-7 cells induced by the miR-23a∼27a∼24-2 cluster. Furthermore, one of the inferred regulatory mechanisms comprises the apoptosis network centered around TP53, whose effective regulation of apoptosis is somehow reestablished after $[Gd@C_{82}(OH)_{22}]_n$ treatment. These results elucidate the application and development of $[Gd@C_{82}(OH)_{22}]_n$ and other fullerene derivates.

1. Introduction

Cancer is one of the most serious diseases worldwide. Conventional methods such as radiation therapy and chemotherapy are routinely applied in clinical treatment. However, these therapy methods have many adverse side effects. For example, many chemotherapy methods destroy the immune system of patients [1] and cause gastrointestinal side effects such as mouth sores, nausea, loss of appetite, and diarrhea [2]. Therefore, developing a new anti-tumor component with minimal toxicity is highly essential.

Gadolinium endohedral metallofullerenol $[Gd@C_{82}(OH)_{22}]_n$ is a promising anti-cancer nanoparticle that shows low toxicity to normal tissue. $[Gd@C_{82}(OH)_{22}]_n$ exhibits strong inhibition activity on the propagation of implanted hepatoma cells (H22) [3], Lewis lung cancer [4] in mice and f-NP-treated human breast cancer cells (MCF-7) in nude mice [5]. In addition, $[Gd@C_{82}(OH)_{22}]_n$ shows lower toxicity *in vitro* and *in vivo* experiments compared with other anti-tumor drugs, such as cyclophosphamide and paclitaxel [6].

Abbreviations: ER, endoplasmic reticulum; ROS, reactive oxygen species; GO, gene ontology; qRT, quantitative real-time; TF, transcription factor; LTS, least trimmed squares; LMS, least median of squares; MLE, maximum likelihood estimator; HEK293T, human embryonic kidney.

* Corresponding authors. Address: Chinese Academy of Sciences, No. 55, Zhongguancun East Road, Haidian, Beijing 100190, China (L.M. Li).
E-mail addresses: mengj@nanoctr.cn (J. Meng), lilei@amss.ac.cn (L.M. Li).

1046-2023/$ - see front matter © 2014 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.ymeth.2014.01.007

Some studies have been conducted on $[Gd@C_{82}(OH)_{22}]_n$'s anti-tumor and low-toxicity mechanisms, which include the enhancement of cell-mediated immunity [4], inhibition of angiogenesis [6], and regulation of reactive oxygen species (ROS) production [7]. However, an overall and systematic picture of the cellular activity changes caused by $[Gd@C_{82}(OH)_{22}]_n$ treatment is yet to be achieved. In addition, no research on the inhibition of proliferation and the induction of apoptosis induced by $[Gd@C_{82}(OH)_{22}]_n$ treatment has been found in the literature.

To further elucidate the mechanisms of anti-tumor efficacy and low toxicity, we adopted a cell line model and took a systems biology approach in this study. That is, we measured the gene expression profiles of the human breast cancer cells MCF-7 and human umbilical vein endothelial cells ECV304 treated with and without $[Gd@C_{82}(OH)_{22}]_n$, respectively, using the Agilent Gene Chip G4112F. Moreover, we report some experimental results on the cell number reduction from the $[Gd@C_{82}(OH)_{22}]_n$ treatment. After the microarray data preprocessing that aimed to reduce the experimental bias and variation, we integrated the expression profiles with relevant biological information to draw inferences about the changes of cellular activities caused by $[Gd@C_{82}(OH)_{22}]_n$ treatment and about the associated regulatory mechanisms. The biological informatics database included KEGG, Gene Ontology (GO), JASPAR, and TRANSFAC. The expressions of some important genes identified by the study were validated by quantitative real-time (qRT)-PCR.

2. Results

2.1. Preprocessing of microarray data

The experimental raw data were cleaned up by the sub–sub normalization method [8], in which a one-knot SPLINE model was applied to account for possible nonlinear spatial effects. The S- and LTS-estimators were used to estimate the parameters in the SPLINE model and they gave similar normalization results. The effectiveness of a statistical adjustment is usually accessed by two measures: bias and variance. How should we define bias and variance in normalization? In our perspective, normalization is a blind inversion problem [9]. One requirement of the inversion is that the expression differentiations of "un-differentially expressed genes" should be around zero, which can serve as the basis for bias correction and dispersion. But generally we do not directly know which genes are "un-differentially expressed" and this is why we consider normalization as blind inversion. In many microarray experiments, as in our case, the "un-differentially expressed genes" between the two samples account for a large portion of all the genes. We use robust estimator with high break-down value to capture the "un-differentially expressed genes". Control information, if available, can help evaluate bias and variance directly. In this microarray experiment, some spike-in probe sets of given concentrations are provided as negative control, although they are not perfect due to uncontrollable experiment factors.

Insightful graphical displays of data help translate the concepts into common sense. When we compare two expression profiles, their joint distribution can be visualized by the M–A plot [10]. At the top of Fig. S1, the two M–A plots show the distributions of treated vs. untreated samples, respectively, for the MCF-7 cells and ECV304 cells. More points are observed below the horizontal lines $M = 0$, indicating downward biases. This finding is supported by the values of the spike-in probe sets represented in pink [11]. We also plot in Fig. 1 the kernel density estimates of the differentiation values, namely, the M-values in the M–A plots. Both modes of the distributions of the two experimental samples bias from zero. In contrast, after sub–sub normalization, both modes shift close to zero, suggesting that the biases have been corrected to a great extent. This result is corroborated by their M–A plots, in which the spike-in probe sets are shifted toward the horizontal lines $M = 0$ after normalization. We could obtain the same conclusion from the biases and the standard deviations of the expression differentiation values for each chip (Table S1). The absolute values of biases of the whole probe set expression differentiations (both MCF-7 cells and ECV304 cells) in the sub–sub normalization are smaller than those in the original data, respectively. The standard deviations of the spike-in probe set expression differentiations (both MCF-7 cells and ECV304 cells) after normalization are smaller than those in the original data, respectively. In Fig. 1, the peaks around the modes represent the differentiation values of the "un-differentially expressed genes" due to biological or experimental noise. The peaks become higher and narrower after normalization, suggesting that the normalization reduces variations.

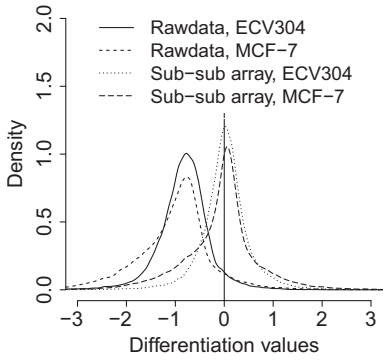

Fig. 1. The kernel density curves of the differentiation values. This figure shows the comparison of the raw differentiation values and the differentiation values after sub–sub normalization. The vertical line is a reference.

These results show that the sub–sub normalization method is effective in reducing the bias and noise of the raw data; thus, it enhances the signal to noise ratio and provides a reliable basis for further analysis. We have compared this normalization results with the normalization results from the Agilent standard approach using the above criteria and the results of our normalization are more reasonable.

In Fig. S1 (A) and (C), when we take the green line as a reference, we can see downward tails at the right ends of the M–A plots. This result indicates that the raw data have some nonlinear effects. The simple SPLINE form was used to adjust for the possible nonlinear effects. After normalization with one-knot SPLINE model, we can take the black line as a reference in Fig. S1 (B) and (D). These two figures show that the downward tails of the M–A plots are mended and the nonlinear effects are adjusted.

$[Gd@C_{82}(OH)_{22}]_n$ induces more expression differentiations in cancer cells. In Fig. S1, figures (B) and (D) show the comparison of the M–A plot for the MCF-7 cells and that for the ECV304 cells according to the gene expression differentiation values. The result indicates that the impact of $[Gd@C_{82}(OH)_{22}]_n$ on the tumor cells is greater than that on the normal cells. In Table S1, the standard deviation of all the expression differentiation values, namely, the M-values, is 0.780 in the MCF-7 cells, which is larger than that in the ECV304 cells (0.508). However, for the spike-in probe sets, the standard deviation is 0.179 in the MCF-7 cells, which is a little larger than that in the ECV304 cells (0.174). These results indicate that the difference of the standard deviations in the MCF-7 cells and in the ECV304 cells for all the probe set expression differentiation values is not due to the normalization but ascribed to the $[Gd@C_{82}(OH)_{22}]_n$ treatment. We next check the density curves of the M-values in Fig. 1. The mass around the near-zero mode is larger in the ECV304 cells' density curve, whereas the two tails are