Saudi Journal of Biological Sciences (2015) 22, 674–678

King Saud University

Saudi Journal of Biological Sciences

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

Entropic analysis reveals a connection between the recurrence of cancer and chemotherapy

Chih-Yuan Tseng [a,*], Jack Tuszynski [b,c]

[a] MDT Canada Inc., Edmonton, AB, Canada
[b] Department of Oncology, University of Alberta, Edmonton, AB T6G 1Z2, Canada
[c] Department of Physics, University of Alberta, Edmonton, AB, Canada

Received 25 March 2015; accepted 5 May 2015
Available online 14 May 2015

KEYWORDS

Maximum entropy;
β-Tubulin isotype;
Wilms' tumor;
Relapse

Abstract In this study, we proposed an entropic analysis to overcome limitations of conventional statistical methods to analyze clinical data for cancer patients who experienced relapse of tumors following chemotherapy. We have applied this entropic method to reveal potential mechanisms that lead to a relapse of Wilms' tumor in pediatric patients. Results indicate β-tubulin isotype III up-regulation is likely the primary cause of the relapse.

1. Introduction

Relapse of tumors has been a major clinical challenge after patients received treatments either involving surgical removal, radiotherapy or chemotherapy (Boisgerault et al., 2013). Researchers and clinical practitioners have attempted to correlate clinical results and potential risk factors to identify molecular mechanisms of relapse including EGF-like growth factor over-expression (Tagliabue et al., 2003) in breast cancer and up-regulation of several genes (e.g. CLDN16 and TJP3) in ovarian cancer (Laios et al., 2008). Besides biological factors, healthcare providers have reported that a common question from patients is whether stress factors may be playing a role in triggering the relapse of tumors. However, some studies did not find any relationship between stress response and cancer relapse (Todd et al., 2014).

Therefore, many attempts have been subsequently focused on determining biological mechanisms of relapse at a molecular level aimed at eventually developing improved treatments. For example, an integrative model for relapse in ovarian cancer was developed to identify genes that can be related to the relapse and can become the target for better therapy aimed at decreasing drug resistance and optimizing the efficacy of existing drugs (Laios et al., 2008).

Our goal is to develop an information-driven approach in order to interpret clinical data based on the available information relevant to binding properties of drugs and biological targets. Results of this approach may reveal deeper insights into mechanisms of the relapse of cancer after patients receive chemotherapy. We hope this will unveil new directions in the redesign of drugs to overcome cancer recurrence.

The foundation of the proposed approach is based on information theory. Information theory has been shown to be an appropriate and robust method to make inference from insufficient data to resolve problems in complex systems

* Corresponding author.
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.sjbs.2015.05.005
1319-562X

(biological systems and drug discovery are but two examples of numerous applications) with the least bias (Tseng and Tuszynski, 2014). Particularly, the method of maximum entropy (ME) has been proved to be a powerful tool for inductive inference to objectively process information (Giffin et al., 2007; Caticha et al., 2007). Tseng and Tuszynski further reviewed and demonstrated that drug discovery can substantially benefit from the method of ME (Tseng and Tuszynski, 2014). Therefore, based on this method, we proposed an entropic analysis approach. Since the method of ME is a universal inductive inference tool (Giffin et al., 2007; Caticha et al., 2007), the proposed method can be applied to study the relapse of different types of tumors as long as information used is relevant to the problem of interests.

To demonstrate the use of the entropic analysis approach, we consider a case study, which was designed to uncover possible mechanisms of the relapse of Wilms' tumor in pediatric oncology patients. As has been already shown through metastudies, although the treatment protocols developed either by the National Wilms' Tumor Study Group (NWTSG) or the International Society of Pediatric Oncology (SIOP) over the years have an increased patient survival rate, there is still approximately 15% of WT patients who developed relapse of the disease (Geller and Dome, 2010). Therefore, one clinical challenge in treating Wilms' tumor is to investigate mechanisms of the relapse of Wilms' tumor in patients who received tubulin-target based chemotherapy including the drug vincristine. Specifically, the meta-study conducted by Grundy in the Pediatric Oncology Group at the Cross Cancer Institute, Edmonton, was aimed to determine whether variations in the tubulin isotype expression may be responsible for drug resistance and correlate these findings with two stages of relapse (early and late).

Fig. 1 shows the mean expression levels of 13 samples (sample ID and six β-tubulin isotype types are listed in Table 1) from 78 measurements in this case study (Grundy). Fig. 2 shows the

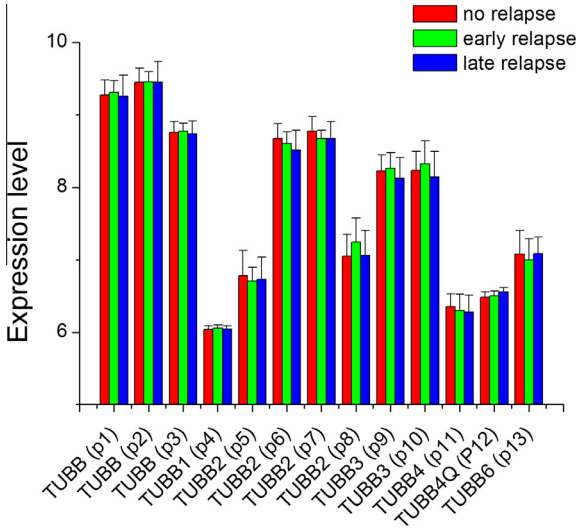

Figure 1 Histogram of mean expression level of β-tubulin isotypes, the gene names are recorded in the x-axis, for the treated patients who are at three stages of relapse of Wilms' tumor. Note that the label inside parentheses is the short annotation for probe sets used in the studies. The original annotations are shown in first column in Table 1.

detailed distribution of all expression levels in each sample. It includes 56 cases where there is no relapse of tumor after tubulin-target chemotherapy treatments, 10 cases are found to have early relapse and 12 cases show late relapse. As shown in Fig. 1, it appears that there are no significant differences in the mean expression values among the different stages of relapse of tumors for each tubulin isotype in each sample. Yet the differences in the expression level distribution for each tubulin isotype in each sample in Fig. 2 seem to disagree with the results from the mean values. Since there are no significant differences in the mean values, we can expect that commonly used statistical hypothesis tests are unlikely to provide meaningful insights. The question that arises is whether we can make any inferences based on the histograms in Fig. 2 to gain insights into the role of β-tubulin isotypes in the relapse of tumors. This question is exactly the problem that the method of ME is designed to answer based on insufficient data.

According to the proposed entropic analysis, we showed that the late relapse is primarily associated with β-tubulin isotype III expression levels. Namely, β-tubulin isotype III over-expression is likely to be the key reason for drug resistance. This result combined with the conclusions by Tseng et al. (2010), which indicate β-tubulin isotype III as the most important molecular target to inhibit microtubule polymerization and eliminate cancer cells, suggests a potential combination chemotherapy, which targets not only β-tubulin isotype III but also its possible mutations.

2. Methods and materials

2.1. Entropic analysis

The proposed approach is based on the method of ME. The foundation of the method of ME hinges on the Bayesian interpretation of probability and the rules of probability theory. The former treats probability as a measure of our state of knowledge about the system of interest rather than a frequency of occurrence of an event. The latter demonstrates that this type of probability can be manipulated by the rules of subtraction, multiplication and addition given by the standard probability theory (Giffin et al., 2007; Caticha et al., 2007). These two tools form the building blocks for inductive inference and its core is the concept of entropy. The most important factor in applying this scheme to solve specific problems is to ask "the right question". Afterward, the principle of ME indicates the most preferable inference one can make to answer that question is the one that maximizes entropy of the problem of interests. Furthermore, the method of ME indicates that given two probability distributions of a system at a specific state i, $P_1(i)$ and $P_2(i)$, the relative entropy of these two probability distributions:

$$S(P_1, P_2) = -\sum_i P_1(i) \log \frac{P_1(i)}{P_2(i)} \leqslant 0, \tag{1}$$

is shown to quantify the difference between these two probability distributions. Furthermore, when $P_2(i)$ is set to a reference distribution (for example, it can be a uniform distribution to represent our complete ignorance of the system of interest), one can utilize this relative entropy to represent a preference scale to rank possible probability distributions of the system (Tseng, 2006; Chen et al., 2007).

Table 1 Results of entropic analysis on investigating the relapse of Wilms' tumor.

Gene name	Sample ID (UG133 Probe set ID)	$S(P_{\text{early}}, P_{\text{non}})$	$P'(X_{\text{early}})\delta X$	$S(P_{\text{late}}, P_{\text{non}})$	$P'(X_{\text{late}})\delta X$
TUBB	209026_x_at	−0.025	0.024	−114.86	1
	211714_x_at	−0.293	0.254	−0.194	0.177
	212320_at	−0.089	0.085	−0.117	0.111
TUBB1	208601_s_at	−0.506	0.397	−0.128	0.12
TUBB2	204141_at	−0.471	0.376	−0.298	0.257
	208977_x_at	−0.197	0.179	−57.454	1
	213726_x_at	−0.419	0.342	−57.677	1
	209372_x_at	−0.258	0.227	−0.457	0.367
TUBB3	202154_x_at	−0.075	0.072	−57.503	1
	213476_x_at	−0.362	0.303	−57.525	1
TUBB4	212664_at	−0.151	0.14	−0.164	0.152
TUBB4Q	211915_s_at	−0.057	0.055	−0.039	0.322
TUBB6	209191_at	−0.221	0.198	−0.289	0.251

Relative entropy $S(P_1, P_2)$ was calculated from Eq. (1).
Probability $P'(X_1)\delta X$ was calculated from Eq. (3).

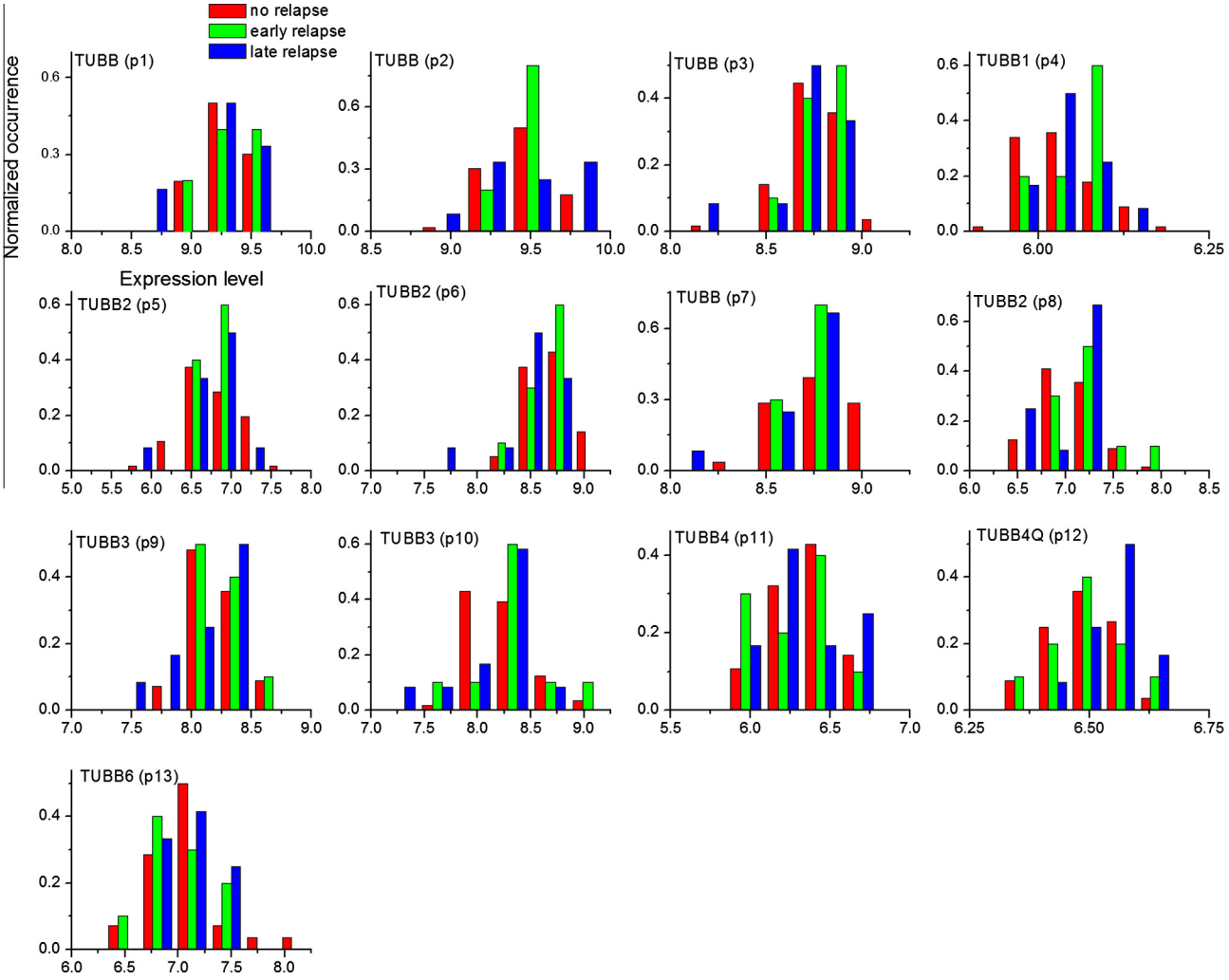

Figure 2 Comparisons of histograms of normalized occurrences of specific expression levels of β-tubulin isotypes from the treated patients at three stages of relapse of Wilms' tumor.

In investigations of possible mechanisms of tumor relapse here, we then consider "*what is the probability that the expression levels of biological targets of chemotherapy in tumor-relapse patients differ from the ones in normal conditions*" as "the right question" to ask. Specifically, we want to focus on β-tubulin based chemotherapy and quantify whether the tubulin isotype expression levels represent mechanisms to cause relapse of tumors. Since there are many factors influencing the measurement of β-tubulin isotypes expression levels (Leandro-García et al., 2010), the concept of probability distribution to comprehend these kinds of uncertainty is an appropriate usage of the ME method. Namely, we are interested in determining the differences between the probability distributions of β-tubulin isotype expressions when cells are subjected to different conditions (for example, patients with no relapse vs early stage of relapse). The degree of these differences then provides essential information to indicate the influence of specific β-tubulin isotype expression levels on potential causes of tumor relapse.

Furthermore, we considered the expression level of a specific β-tubulin isotype to represent the human body at a specific state, which can be the patient who received chemotherapy and had either no, early or late relapse afterward. Next, we simply utilize the normalized histogram of expression levels for each tubulin isotype obtained from patient data at this stage to define the probability of a tubulin isotype having expression level i, $X(i)$, given the patient who has either early or late relapse, $P_1(i)$. The bin size for appropriately sampling expression level data for each sample is chosen to be within the corresponding standard deviation. Next, for the purpose of investigating the mechanisms of relapse in the framework of ME, we considered the probability distribution of a tubulin isotype having expression level i given the patient who has no relapse as the reference distribution $P_2(i)$. The mean reference expression level is given by $X = \sum_i P_2(i)X(i)$. For patients who had relapse, the mean expression level is given by $X' = \sum_i P_1(i)X(i)$ assuming that $X' = X + \delta X$. Therefore, the relative entropy, Eq. (1), of probability distributions for specific tubulin isotype expression levels in patients who had relapse and no relapse allows us to quantify their differences. The larger the difference, the higher the chances that the corresponding tubulin isotype is likely responsible for the different stages of relapse. Note that the relative entropy measure provides a scale to quantify the differences between two distributions. However, the statistical significance of these differences is still unclear.

Therefore, we proposed to apply the Einstein fluctuation theory to resolve this issue (Caticha et al., 2007; Callen et al., 1985). Based on the fluctuation theory, one can show that the given relative entropy of two distributions, Eq. (1), can determine the probability of observing X' to deviate from reference X by δX, which is given by:

$$P(X')\delta X = \exp S(P_1, P_2). \tag{2}$$

On the other hand, within the difference δX, $P(X')\delta X$ represents the probability of $P_1(i)$ being identical to $P_2(i)$. Therefore, we can define the probability of $P_1(i)$ being different from $P_2(i)$ by:

$$P'(X')\delta X = 1 - \exp S(P_1, P_2). \tag{3}$$

Namely, this probability indicates that the chances of either early or late relapse (represented by $P_1(i)$) being different from no relapse ($P_2(i)$) is under consideration of a specific tubulin isotype. This will then be the indicator used in the case study to determine whether differences between tubulin isotype expression levels are critical factors that affect the probability of relapse.

2.2. Case study: Wilms' tumor

This case study was designed following our earlier work (Tseng et al., 2010; Huzil et al., 2007) on drug resistance using β-tubulin, a major component of microtubules, as a model. According to this earlier work (Tseng et al., 2010; Huzil et al., 2007), binding sites for taxol-based and vinca alkaloids in β-tubulin have been identified and are considered as hot spots for mutations that hamper therapeutic effects of those drugs. Moreover, each tubulin isotype has a different affinity for various chemotherapy agents and this can be used as a strategy to mitigate the effects of chemotherapy drugs by tumor cells (Huzil et al., 2006). Therefore, the goal of this study is to quantify tubulin expression levels and potentially screen for mutations in the β-tubulin gene, specifically in those locations which correspond to the binding sites for the chemotherapy drugs acting on tubulin in the cells of Wilms' tumors. The present study aimed to answer the following two questions. First, what is the level of β tubulin isotype expression in patients whose disease recurs and second, what is the probability that a difference in the expression levels of tubulin isotypes exist between those who relapse and those who do not. A future study should also investigate similar questions directed at specific somatic single point mutations in tubulin genes.

Table 1 lists 14 UG133 probe sets from Affymetrix Inc. used to explore the interactions of six β-tubulin isotypes and vinca alkaloids. The first column records the name of β-tubulin isotype gene (HUGO ID) and the second column denotes probe set ID. Sequences bearing the predicted mutations (S172A, P173A, D197N, K350N) will be exclusively targeted for sequencing.

Note that the data received (Grundy) refer to the expression levels of β-tubulin rather than the mutation rate. The latter should be investigated by an additional study. Also note that for some genes such as TUBB, there are three samples while some others have only one sample. This is due to a varying level of sample quality.

3. Results and discussion

We conclude the following based on the data analysis performed: (1) β-tubulin isotype III (gene name is TUBB3) expression plays the major role and (2) β-tubulin isotype II (TUBB2) expression as the secondary factor in causing late relapse of tumor. Fig. 2 plots the normalized histogram of expression levels of each tubulin isotype in each sample at three stages of relapse. By defining the probability distribution for specific tubulin isotype expression levels in each sample at early and late stages of relapse and no relapse based on these histograms, the relative entropy of the probability distributions for either early or late relapse and no relapse, the results of Eq. (1) are calculated and shown in the third and fifth columns in Table 1, respectively. Thereafter, one can calculate the probabilities of either early $P'(X_{early})\delta X$ or late $P'(X_{late})\delta X$ relapse being different from no relapse under consideration of a specific tubulin isotype, Eq. (3), as shown in the fourth

and sixth columns, respectively, in Table 1. These two probability distributions are shown to have large fluctuations except the one for TUBB3 in the late relapse case, whose expression level distributions in both samples are at 100% probability different from the distributions of no relapse. Furthermore, the inconsistency between the $P'(X_{early})\delta X$ and $P'(X_{late})\delta X$ results suggests that the corresponding tubulin isotypes do not have strong correlations with the cause and the stage of relapse. TUBB3 among all isotypes likely plays a major role in triggering late relapse of tumor. Furthermore, since there are two out of four samples (208977_x_at and 213726_x_at) in TUBB2, we also have $P'(X_{late})\delta X = 1$, and we may infer that TUBB2 may play a secondary role in late relapse.

As mentioned previously, because $P'(X_{early})\delta X$ shows large fluctuations for all tubulin isotypes, it is difficult to conclude that the tubulin isotype expression level distributions for samples with early relapse are different from the one with no relapse. Therefore, this indicates that all β-tubulin isotypes mutations are unlikely to correlate with causes of the early stage of relapse.

4. Summary

In this paper, we have proposed an approach based on entropic inductive inference to interpret clinical results for pediatric oncology patients with Wilms' tumor who received chemotherapy utilizing information regarding the molecular target (tubulin isotypes) expression levels relevant for the chemotherapy drugs used. This approach reveals probabilities of biological targets responsible for the relapse of cancers after patients received chemotherapy. We have applied the maximum entropy method to investigate the relapse of Wilms' tumor as only one example. Our studies indicate that β-tubulin isotype III plays a relatively more important role than other isotypes and shows resistance to vinca-based chemotherapy in late stages of treatments and it is correlated with the causes for tumor recurrence. It further suggests that an effective treatment plan to prevent the recurrence should include either modified vinca-based drugs or alternatives that can bind to tubulin isotype III with a high affinity. A method of optimizing drug selection for specific tubulin expression profiles in both tumor cells and healthy tissues has been recently published (Ravanbakhsh et al., 2013) and it can be used in future work on rational drug design for Wilms' tumor and other neoplastic diseases with tubulin as a therapeutic target.

Acknowledgments

The authors are grateful for valuable discussions with Dr. P. Grundy and providing the patient data to the authors. This research was partially sponsored by the Allard Foundation.

References

Boisgerault, N., Kottke, T., Pulido, J., et al, 2013. Functional cloning of recurrence-specific antigens identifies molecular targets to treat tumor relapse. Mol. Ther. 21 (8), 1507–1516.

Callen, H.B., 1985. Thermodynamics and An Introduction to Thermostatistics, second ed. John Wiley & Sons, Chapter 19.

Caticha, A., 2007. Information and entropy. In: Knuth, K.H., Caticha, A., Center, J.L., Giffin, A., Rodríguez, C.C. (Eds.), . In: Bayesian Inference and Maximum Entropy Methods in Science and Engineering. AIP, Melville, NY, USA, pp. 11–22. AIP Conf. Proc. 954.

Chen, C.-C., Tseng, C.-Y., Dong, J.-J., 2007. New entropy-based method for variables selection and its application to the debris-flow hazard assessment. Eng. Geol. 94, 19–24.

Geller, J.I., Dome, J.S., 2010. Molecular targeted therapy for Wilms' tumor. In: Houghton, P.J., Arceci, R.J. (Eds.), . In: Molecularly Targeted Therapy for Childhood Cancer. Springer, pp. 401–424.

Giffin, A., Caticha, A., 2007. Updating probabilities with data and moments. In: Knuth, K.H., Caticha, A., Center, J.L., Giffin, A., Rodríguez, C.C. (Eds.), . In: Bayesian Inference and Maximum Entropy Methods in Science and Engineering. AIP, Melville, NY, USA, pp. 74–84. AIP Conf. Proc. 954.

P. Grundy, Personal communication.

Huzil, J.T., Luduena, R.F., Tuszynski, J.A., 2006. Modeling human beta tubulin isoforms and their implications for drug binding. Nanotechnology 17 (4), S90–S100.

Huzil, J.T., Chen, K., Kurgan, L., Tuszynski, J.A., 2007. The roles of β tubulin mutations and isotype expression in acquired drug resistance. Cancer Inf. 3, 159–181.

Laios, A., O'Toole, S.A., Flavin, R., 2008. An integrative model for recurrence in ovarian cancer. Mol. Cancer 7, 8.

Leandro-García, L.J., Leskelä, S., Landa, I., Montero-Conde, C., López-Jiménez, E., Letón, R., Cascón, A., Robledo, M., Rodríguez-Antona, C., 2010. Tumoral and tissue-specific expression of the major human beta-tubulin isotypes. Cytoskeleton (Hoboken) 67, 214–223.

Ravanbakhsh, S., Gajewski, M., Greiner, R., Tuszynski, J.A., 2013. Determination of the optimal tubulin isotype target as a method for the development of individualized cancer chemotherapy. Theor. Biol. Med. Model. 10, 29.

Tagliabue, E., Agresti, R., Luisa, M., et al, 2003. Role of HER2 in wound-induced breast carcinoma proliferation. Lancet 362, 527–533.

Todd, B.L., Moskowitz, M.C., Ottati, A., et al, 2014. Stressors, stress response, and cancer recurrence: a systematic review. Cancer Nurs. 37 (2), 114–125.

Tseng, C.-Y., 2006. Entropic criterion for model selection. Physica A 370, 530–538.

Tseng, C.-Y., Tuszynski, J.A., 2014. Maximum entropy in drug discovery. Entropy 16, 3754–3768.

Tseng, C.-Y., Mane, J.Y., Winter, P., et al, 2010. Quantitative analysis of the effect of tubulin isotype expression on sensitivity of cancer cell lines to a set of novel colchicine derivatives. Mol. Cancer 9, 131.