



Image-Based Network Analysis of DNp73 Expression by Immunohistochemistry in Rectal Cancer Patients

Tuan D. Pham^{1,2*}, Chuanwen Fan^{3,4}, Daniella Pfeifer³, Hong Zhang⁵ and Xiao-Feng Sun^{3*}

¹ Department of Biomedical Engineering, Linkoping University, Linkoping, Sweden, ² The Center for Artificial Intelligence, Prince Mohammad Bin Fahd University, Al Khobar, Saudi Arabia, ³ Department of Oncology, Clinical and Experimental Medicine, Linkoping University, Linkoping, Sweden, ⁴ Institute of Digestive Surgery, West China Hospital, Sichuan University, Chengdu, China, ⁵ Department of Medical Sciences, Orebro University, Orebro, Sweden

OPEN ACCESS

Edited by:

George Bebis, University of Nevada, Reno, United States

Reviewed by:

Neda Slade, Rudjer Boskovic Institute, Croatia Sanjay Ram Kharche, University of Western Ontario, Canada Stella Logotheti, University of Rostock, Germany

*Correspondence:

Tuan D. Pham tuan.pham@liu.se; tpham@pmu.edu.sa Xiao-Feng Sun xiao-feng.sun@liu.se

Specialty section:

This article was submitted to Computational Physiology and Medicine, a section of the journal Frontiers in Physiology

Received: 21 December 2019 Accepted: 09 December 2019 Published: 08 January 2020

Citation:

Pham TD, Fan C, Pfeifer D, Zhang H and Sun X-F (2020) Image-Based Network Analysis of DNp73 Expression by Immunohistochemistry in Rectal Cancer Patients. Front. Physiol. 10:1551. doi: 10.3389/fphys.2019.01551 **Background:** Rectal cancer is a disease characterized with tumor heterogeneity. The combination of surgery, radiotherapy, and chemotherapy can reduce the risk of local recurrence. However, there is a significant difference in the response to radiotherapy among rectal cancer patients even they have the same tumor stage. Despite rapid advances in knowledge of cellular functions affecting radiosensitivity, there is still a lack of predictive factors for local recurrence and normal tissue damage. The tumor protein DNp73 is thought as a biomarker in colorectal cancer, but its clinical significance is still not sufficiently investigated, mainly due to the limitation of human-based pathology analysis. In this study, we investigated the predictive value of DNp73 in patients with rectal adenocarcinoma using image-based network analysis.

Methods: The fuzzy weighted recurrence network of time series was extended to handle multi-channel image data, and applied to the analysis of immunohistochemistry images of DNp73 expression obtained from a cohort of 25 rectal cancer patients who underwent radiotherapy before surgery. Two mathematical weighted network properties, which are the clustering coefficient and characteristic path length, were computed for the image-based networks of the primary tumor (obtained after operation) and biopsy (obtained before operation) of each cancer patient.

Results: The ratios of two weighted recurrence network properties of the primary tumors to biopsies reveal the correlation of DNp73 expression and long survival time, and discover the non-effective radiotherapy to a cohort of rectal cancer patients who had short survival time.

Conclusion: Our work contributes to the elucidation of the predictive value of DNp73 expression in rectal cancer patients who were given preoperative radiotherapy. Mathematical properties of fuzzy weighted recurrence networks of immunohistochemistry images are not only able to show the predictive factor of DNp73 expression in the patients, but also reveal the identification of non-effective application of radiotherapy to those who had poor overall survival outcome.

Keywords: fuzzy weighted recurrence networks, network properties, multi-channel images, DNp73, immunohistochemistry, predictive biomarker, rectal cancer, survival outcome

1

1. INTRODUCTION

Colorectal cancer (CRC) is the third most common cancer in the world (Arnold et al., 2017). There are around 6,500 new diagnosed cases of CRC yearly among the population of Sweden. The causes for CRC are considered to be associated with gene mutations, gene variants, and changed expression of proteins. The combination of surgery and radio-chemotherapy is the most beneficial regimens in current treatment of advanced rectal cancer. Preoperative radiotherapy (RT) is often given to rectal cancer patients as a complement to surgery to improve treatment outcome. However, tumor recurrence plays a major cause of death for progressive rectal patients after surgery. As a result, a significant proportion of patients did not benefit from preoperative RT (Fan et al., 2013).

It remains up to date that the clinical testing of specific mutations in KRAS, BRAF, RAS, and RAF genes along with mismatch repair gene deficiency assists either as prognostic or predictive biomarkers in CRC (Sinicrope et al., 2016; Zarkavelis et al., 2017). Other methods for identifying biomarkers in the treatment of CRC include molecular subtype classification (Cuyle and Prenen, 2017), identifying molecular signatures at protein and RNA levels by microarray analysis (Rahman et al., 2019), protein identification in cell proliferation and new blood vessels (Chatterjee et al., 2019), changes in the amounts of certain proteins (Letellier et al., 2017), and proteomic strategies (Lee et al., 2018). A recent review of methods for discovering prognostic and predictive biomarkers in CRC for personalized therapy can be found in Patel et al. (2019).

The role of predictive biomarkers is known to be essential for the field of radiation oncology (Yaromina et al., 2012). While most efforts aim to improve cancer treatment with respect to physical conditions and technology, such as precision in treatment plans and dose administration (Sonke and Belderbos, 2010), the inclusion of patient-specific biological characteristics into cancer treatment decision would be very useful for personalized treatment. However, such important biological information of individual patients is not well explored. To achieve this purpose, predictive biomarkers are needed to guide radiation oncologists to determine optimal dose prescription, select patient-specific schemes, and treatments for individual cancer patients (Yaromina et al., 2012).

However, it is a big challenge to find predictive biomarkers that can select patients who can benefit from RT, although our and other groups have spent much effort to identify potential predictors for the RT response (Ryan et al., 2016; Ye and Guo, 2019). A previous study of our group suggested that p73 independently predicted poor prognosis in colorectal cancer and p73-negative tumors tended to have a lower local recurrence after RT compared with unirradiated case (Ye and Guo, 2019).

One of important reasons is that the TP73 gene expresses isoforms with divergent and/or opposing roles in cancer. These are mainly categorized in two classes, the anti-oncogenic TAp73 isoforms, which contain an intact N-terminal, transactivation domain, and the oncogenic DNp73 isoforms, which lack part or whole of the transactivation domain and act as dominant negative forms of TAp73 proteins (Logotheti et al., 2013). The TAp73 isoforms are generated from an external P1 promoter. The DNp73 proteins are transcribed (a) by the P1 promoter, followed post-transcriptionally by alternative splicing in exons 2 and/or 3 at the 5' end (Stiewe et al., 2002), or (b) by an alternative, internal P2 promoter which generates variants lacking exons 2 and 3, but instead containing an exon 3' that encodes for a unique 13-amino acid domain (Irwin, 2006). Additional complexity is created by alternative splicing in the 3' end, which gives rise to a large number of C-terminal variants of the abovementioned isoforms (Logotheti et al., 2013). Altogether, the TP73 gene expresses at least 35 mRNA variants, which can encode theoretically 29 different p73 protein isoforms (Murray-Zmijewski et al., 2006). Notably, the ratio between TAp73 and DNp73 isoforms has essential effects on the cellular response (Dulloo et al., 2010; Rufini et al., 2011).

The imbalance between TAp73 and DNp73 isoforms may be useful to predict response to chemotherapy and prognosis (Muller et al., 2005; Lucena-Araujo et al., 2015). High DNp73 expression has strong correlation with unfavorable prognosis in several types of cancer patients, and DNp73-positive tumors show a reduced response to chemotherapy and irradiation (Uramoto et al., 2004; Di et al., 2013; Zhu et al., 2015). The upregulation of DNp73 was frequently detected in radioresistant cervical cancers (Liu et al., 2006). Our previous findings indicated that DNp73 is increased in colon cancer cell line that is resistant to γ -irradiation (Pfeifer et al., 2009). Thus, these findings suggested that DNp73 expression may play an important role in the regulation of radiosensitivity. However, the prognostic and preditive role of DNp73 in rectal cancer patients with radiation still remains unclear.

This study aimed to elucidate the role of DNp73 as a predictive biomarker by investigating if DNp73 was related to the survival time of rectal cancer patients who were administered with RT before surgery. To overcome the subjective and time-consuming task of pathologist-based analysis of immunohistochemistry (IHC) images stained for DNp73 expression, we carried out a study by means of a novel image-based recurrence network approach. The motivation for developing this new image-based network analysis was based on the recurrence of image attributes inherently existing in the complex nature of IHC images of rectal cancer tissue arrays.

In fact, network analysis in graph theory has been increasingly recognized as a useful tool for studying cancer. Such studies include the prediction of outcomes of ovarian cancer treatment (Zhang et al., 2013), analysis of breast cancer progression and reversal (Parikh et al., 2014), drug response prediction in cancer cell lines (Zhang et al., 2018), identification of novel cancer gene candidates (Josef Gladitz et al., 2018), tumor biology for precision cancer medicine (Ozturk et al., 2018), and prediction of cancer recurrence (Ruan et al., 2019).

In this present study, we introduce a new method of fuzzy weighted recurrence networks of multi-channel images for computing useful properties of the complex networks of the expression patterns of the DNp73 IHC. The ratios of these network properties discover the predictive value of DNp73 in rectal cancer patients in the Swedish Rectal Cancer Trial. **TABLE 1** Demographic information of the rectal cancer patients who had a median age of 68 years (range: 39–78 years), were followed for a median period of 81 months (range: 0–129 months), and had the median time to disease free of 101 months after surgery (range: 15–288 months).

	Number of patients
Male	16 (64%)
Female	9 (36%)
Shorter survival time (15–75 months)	11 (44%)
Longer survival time (101–288 months)	14 (56%)

2. MATERIALS AND METHODS

2.1. Rectal Cancer Patients

This study included the patients with rectal adenocarcinoma from the Southeast Swedish Health Care region who participated in a clinical trial of preoperative RT for rectal cancer (Swedish Rectal Cancer Trial et al., 1997). Samples of biopsy and primary tumor from the same patients were selected for the analysis. In this Swedish Rectal Cancer Trial study, we collected samples from both pre-radiotherapy and non-radiotherapy rectal cancer patients. The biopsy samples were taken from the rectal cancer before the RT and went through the routine pathological process, and eventually embedded in paraffin blocks. The primary tumor samples were taken from the primary rectal cancer after the RT.

There were 25 patients with RT whose demographic information is given in **Table 1**. This study was carried out in accordance with the recommendations of Good Clinical Practice, the Research Ethics Committee in Linkoping, Sweden with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Research Ethics Committee in Linkoping, Sweden. The clinico-pathologic characteristics of the patients are listed in **Table 2**.

2.2. Immunochemistry and Image Extraction

The five-micrometer paraffin-embedded tissue micro-array (TMA) sections were deparaffinized in xylene and rehydrated with a series of gradient ethanol to water. The sections were heated to boiling point in citrate buffer (pH 6.0) for 30 min to unmasked antigen, followed by a washing in phosphate-buffered saline (PBS). Endogenous peroxidase activity was blocked with 3% H₂O₂ in methanol followed by washing three-times in PBS. The sections were incubated with protein block (Dako, Carpinteria, CA) for 10 min and then incubated with anti-DNp73 antibody (clone 38C674.2, Novus Biologicals, 1:200), which specifically recognized DNp73 isoforms, but not TAp73.

After that, the sections were washed in PBS and then incubated with goat anti-mouse secondary antibody (Dako) at room temperature for 25 min. Next, the sections were subjected to 3,3'-diaminobenzidine tetrahydrochloride for 8 min and then counterstained with hematoxylin. Negative and positive controls were added in each staining run. All slides were scored by two independent investigators. Whole-slide images of entire sections

TABLE 2 | Clinico-pathological characteristics of the rectal cancer patients.

Parameters		Number of cases
Age	<60	6
	>60	19
Gender	Male	9
	Female	16
Growth pattern	Expansion	11
	Infiltration	13
	Null	1
Grade	Well	2
	Moderate	14
	Poor	9
Pathological stages	I	8
	II	6
	III	8
	IV	3

were captured with an Aperio CS2 slide scanner system (Leica Biosystems, Wetzlar, Germany) using a 40x magnification.

All sections were reviewed to remove images containing tissue-processing artifacts, including bubbles, section folds and poor staining. A total of 46 whole-slide images from the 25 unique patients were extracted from the TMA slides using ObjectiveViewer (https://www.objectivepathology.com/ objectiveview) with the original resolution.

2.3. Multi-Channel Fuzzy Weighted Recurrence Networks

The term "channel" is a conventional expression used to refer to a certain component of an image. For example, an RGB image has 3 channels that are red (R), green (G) and blue (B) components. A grayscale image has only one channel. Let $\mathbf{I} = [f_{ijk}]$ be a multi-channel image of size $M \times N \times K$, where $i = 1, \ldots, M$, $j = 1, \ldots, N$, and $k = 1, \ldots, K$. Let $m \ge 1$ be an integer, a local image window $\mathbf{W}_{ij}^k \in \mathbf{I}$ of size $(2m+1) \times (2m+1)$ is constructed for each pixel located at ij in each of the k components of the multi-channel image, where ij is the center of the window. This window can be considered as embedding dimensions in two-dimensional space, which considers the local spatial distribution around f_{ij} of the k-th image channel. The Frobenius norm can be used to transform each local window into a scalar measure that has the useful property of invariance under rotations as

$$\|\mathbf{W}_{ij}^{k}\|_{F} = \sqrt{\sum_{i=m}^{i+m} \sum_{j=m}^{j+m} |f_{ijk}|^{2}},$$
(1)

where $(i - m), (j - m) > 0, (i + m) \le M, (j + m) \le N$, and any pixel at the center of the window that requires values from beyond the image boundaries is skipped.

We can then obtain a set of feature vectors \mathbf{y}_{ij} , (i-m), (j-m) > 0, by joining the Frobenius norms computed for each window of

the *k*-th image channel at the same location, for example, a color image of 3 channels:

$$\mathbf{y}_{ij} = \left(\|\mathbf{W}_{ij}^{1}\|_{F}, \|\mathbf{W}_{ij}^{2}\|_{F}, \|\mathbf{W}_{ij}^{3}\|_{F} \right),$$
(2)

where $(i - m), (j - m) > 0, (i + m) \le M, (j + m) \le N$.

Since the Frobenius norm induced feature vector set \mathbf{y}_{ii} can be computed for the multi-channel image I, the multichannel fuzzy weighted recurrence network (MC-FWRN), which is an extension of the FWRN of time series (Pham, 2019), can be constructed as follows. To simplify the notation in subsequent mathematical presentation, \mathbf{y}_{ii} is now denoted as \mathbf{x}_n , $n = 1, \dots, L$, where L is the total number of feature vectors, and some same indices are used but defined differently.

Let $\mathbf{X} = {\mathbf{x}_n}, n = 1, \dots, L, c$ a given number of clusters of the feature space, and a set of *c* fuzzy clusters, $\mathbf{V} = {\mathbf{v}_i : i = 1, ..., c}$. Fuzzy clusters are groups that contain data points, where every data point has a degree of fuzzy membership of belonging to each group. A fuzzy relation $\tilde{\mathbf{R}}$ between \mathbf{v}_i and \mathbf{v}_i , $i, j = 1, \dots, c$, is characterized by a fuzzy membership function $\mu \in [0, 1]$, which expresses the degree of similarity of each pair $(\mathbf{v}_i, \mathbf{v}_i)$ in $\tilde{\mathbf{R}}$. This fuzzy relation has the following three properties (Zadeh, 1971):

- 1. Reflexivity: $\mu(\mathbf{v}_i, \mathbf{v}_i) = 1, \forall \mathbf{v}_i \in \mathbf{V}$.
- 2. Symmetry: $\mu(\mathbf{v}_i, \mathbf{x}_n) = \mu(\mathbf{x}_n, \mathbf{v}_i), \forall \mathbf{x}_n \in \mathbf{X}, \forall \mathbf{v}_i \in \mathbf{V}.$
- 3. Transitivity: $\mu(\mathbf{v}_i, \mathbf{v}_j) = \bigvee_{\mathbf{x}_n} [\mu(\mathbf{v}_i, \mathbf{x}_n) \land \mu(\mathbf{v}_j, \mathbf{x}_n)], \forall \mathbf{x}_n \in$ **X**, \forall **v**_{*i*}, **v**_{*i*} \in **V**, where the symbols \lor and \land stand for max and min, respectively.

The computation of $\mu(\mathbf{v}_i, \mathbf{x}_n)$, $i = 1, \dots, c, n = 1, \dots, L$, which are necessary for the construction of the fuzzy relation \tilde{R} can be carried out by means of the fuzzy c-means (FCM) algorithm (Bezdek, 1981) as follows.

Let μ_{nj} denote a fuzzy membership grade of \mathbf{x}_n , $n = 1, \ldots, L$, which belongs to a cluster j, j = 1, ..., c, whose center is \mathbf{v}_j . This fuzzy membership is calculated by the FCM as

$$\mu_{nj} = \frac{1}{\sum_{i=1}^{c} \left[\frac{d(\mathbf{x}_{n}, \mathbf{v}_{j})}{d(\mathbf{x}_{n}, \mathbf{v}_{i})} \right]^{2/(\alpha - 1)}},$$
(3)

where $1 \leq \alpha < \infty$ is the weighting exponent, and $d(\mathbf{x}_n, \mathbf{v}_i)$ is used as a Euclidean distance between \mathbf{x}_n and \mathbf{v}_i .

Using the fuzzy membership grades, each cluster center \mathbf{v}_i is computed as

$$\mathbf{v}_{j} = \frac{\sum_{n=1}^{L} (\mu_{nj})^{\alpha} \mathbf{x}_{n}}{\sum_{n=1}^{L} (\mu_{nj})^{\alpha}}, \forall j.$$
(4)

The iterative procedure of the FCM is outlined as follows.

- 1. Given c, α , step t, $t = 0, \dots, T$, initialize matrix $\mathbf{U}^{(t=0)} = [\mu_{nj}]^{(t=0)}$
- 2. Compute $\mathbf{v}_i^{(t)}, j = 1, \dots, c$, using Equation (4).
- 3. Update $\mathbf{U}^{(t+1)}$ using Equation (3). 4. If $\|\mathbf{U}^{(t+1)} \mathbf{U}^{(t)}\| < \epsilon$ or t = T, stop. Otherwise, set $\mathbf{U}^{(t)} = \mathbf{U}^{(t+1)}$ and return to step 2.

The predefined FCM parameters α , T and ϵ usually take the values of 2, 100, and 0.00001, respectively. The number of clusters can be estimated using a cluster validity measure such as the partition entropy, denoted by H, which is defined as (Bezdek, 1981)

$$H = \frac{1}{L} \sum_{j=1}^{c} \sum_{n=1}^{L} \mu_{nj} \log(\mu_{nj}).$$
 (5)

This cluster validity works by computing the partition entropy H for a range of a given number of clusters, $c \ge 2$, and the number of clusters that has the minimum value of H is considered as an optimal *c* for the FCM algorithm.

Finally, an $N \times N$ MC-FWRN can be constructed with the fuzzy relation $\tilde{\mathbf{R}}$ as

$$\mathbf{W} = \tilde{\mathbf{R}} - \mathbf{I},\tag{6}$$

where **W** is an $N \times N$ adjacency matrix of edge weights, and **I** is the $N \times N$ identity matrix. The interested reader is referred to the work described in Pham (2019) to obtain more detailed information about the concept of fuzzy weighted recurrence networks originally developed for time series.

2.4. Network Properties

Two most well-known measures of the statistical characterization of a complex network are the average clustering coefficient and characteristic path length (Watts and Strogatz, 1998; Albert and Barabasi, 2002; Barrat et al., 2004). The clustering coefficient of a node in a network is a numerical indicator of a node that tends to cluster with other neighboring nodes. The average clustering coefficient expresses the average amount of connectivity around individual nodes of a network, whereas the characteristic path length is considered as a measure of the efficiency of transfer of information in a network.

The average clustering coefficient for an unweighted network represented with an $N \times N$ (binary) adjacency matrix $\mathbf{A} = [a_{ii}]$, $i, j = 1, \ldots, N$, is defined as

$$C = \frac{1}{N} \sum_{i=1}^{N} C_i,$$
 (7)

where C_i is the local unweighted clustering coefficient for node *i*, and defined as

$$C_{i} = \frac{\sum_{j,k} a_{ij}a_{jk}a_{ki}}{k_{i}(k_{i}-1)}, \ k_{i} \neq 0, 1,$$
(8)

where k_i is the degree of node *i*, which is the number of links of node *i*.

The average clustering coefficient for a weighted network is defined as

$$CC = \frac{1}{N} \sum_{i=1}^{N} C_i^{w},$$
 (9)

where C_i^w is the local weighted clustering coefficient for node *i*, and defined as (Fagiolo, 2007)

$$C_i^w = \frac{\sum_{j,k} [w_{ij} w_{ik} w_{jk}]^{1/3}}{k_i (k_i - 1)}, k_i \neq 0, 1,$$
(10)

where $w_{ij}, w_{ik}, w_{ik} \in W$.

In general, the clustering coefficient of a node is the ratio of existing links connecting a node's neighbors to each other to the maximum possible number of such links. The clustering coefficient for the entire network is the average of the clustering coefficients of all the nodes.

The characteristic path length of a network is defined as the average of all shortest path lengths:

$$CP = \frac{1}{N(N-1)} \sum_{i \neq j, i, j=1}^{N} d_{ij},$$
 (11)

where d_{ij} is the length of the shortest path between nodes *i* and *j*. The Dijkstra's algorithm (Newman, 2010) was used for computing the shortest weighted path in this study. The characteristic path length is calculated by finding the shortest path between all pairs of nodes, adding them up, and then dividing by the total number of pairs. This operation shows on average the number of steps it takes to get from one node of the network to another.

2.5. Algorithm for Computing Network Properties From MC-FWRN

- 1. Given a multi-channel image I, window parameter *m*, number of clusters *c*, and FCM parameters.
- 2. Using Equation (1) to compute the Frobenius norm for each window $(2m+1) \times (2m+1)$ of each image channel, and using Equation (2) to form a matrix of vectors of length 3 with the number of pixels that can be used to construct the windows.
- 3. Compute the fuzzy weighted adjacency matrix **W** using Equation (6) via the FCM.
- 4. Using W to calculate the clustering coefficient with Equation (9), and the characteristic path length with Equation (11).

3. RESULTS AND DISCUSSION

Table 3 shows the screening results of the 25 rectal cancer patients. Patient numbers 1-11 are those who had shorter survival time, and patient numbers 12–25 are those who had longer survival time. The evaluation of the IHC-stained color intensity of the whole slide of a tissue core with brown antibody stain and blue counter-stain were assessed as being positive and negative, respectively. The positive stain is subjectively classified as weak = 1 (light brown), moderate = 2 (moderate brown), and strong = 3 (dark brown), whereas the negative stain = 0 (blue). Figure 1 shows representative IHC staining for DNp73 expression on the biopsy and primary tumor tissue images obtained from a rectal cancer patient survived 40 months after radiotherapy, and biopsy and primary tumor tissue images obtained from a rectal cancer patient who survived 255 months after radiotherapy at the censoring date.

Patient #	Disease-free time	Recurrence status	Survival time	IHC score	
				Primary tumor	Biopsy
1	0	Yes	15	1	1
2	6	Yes	19	1	3
3	20	Yes	25	2	1
4	13	Yes	40	3	3
5	37	Yes	60	3	3
6	44	Yes	62	2	3
7	0	Yes	15	3	2
8	63	Yes	75	3	3
9	26	Yes	27	3	3
10	12	No	26	3	3
11	34	Yes	43	2	2
12	100	Yes	101	1	1
13	0	Yes	180	1	3
14	114	Yes	114	2	2
15	122	Yes	255	3	2
16	167	Yes	167	2	2
17	81	No	101	3	2
18	129	Yes	129	2	3
19	129	Yes	288	2	3
20	126	Yes	126	3	2
21	186	Yes	238	3	3
22	122	Yes	122	2	2
23	168	Yes	288	2	3
24	151	Yes	151	3	3
25	168	Yes	168	2	2

Time is in months. For IHC score, 0 = negative, 1 = weak, 2 = moderate, and 3 = strong.

To capture the local information of the DNp73 expression over the whole IHC-stained slides, images of biopsy and primary tumor of each of the 25 rectal cancer patients were divided into subimages of 150×150 pixels. The subimages that contain either the background or a large portion of the background were excluded in the analysis. To construct the FWRNs of the IHC-stained subimages, we selected the FWRN parameters m=3 to establish a reasonable local window size of 7×7 , c = 20 that was approximately based on the partition entropy, and the FCM parameters $\alpha = 2$, T = 100, and $\epsilon = 0.00001$, which are widely adopted for the FCM analysis. The clustering coefficient and characteristic path length were calculated for each subimage of each patient, and the total average values of the clustering coefficients and characteristic path lengths of all subimages represent the reported values.

Figures 2, **3** show the clustering coefficients and characteristic path lengths of the FWRNs of the biopsy and primary tumor images obtained from the 25 rectal cancer patients, respectively. The scatter plot of the survival time against the ratios of the clustering coefficients of the primary tumors to those of the biopsies, and the ratios of the characteristic path lengths of the primary tumors to those of the biopsies are shown in **Figure 4**.



Based on the visualization of the scatter plot, we discovered the predictive value of DNp73 in the rectal cancer patients in terms of the clustering-coefficient and characteristic-pathlength ratios, which are shown in **Figure 5**. The probability (p) for the predicted survival time based on the clusteringcoefficient ratio was computed as the number of patients who lived between 101 and 288 months divided by the total number of patients whose clustering-coefficient ratios are within the ratio range (p = 11/15 = 0.7333). The probability (p) for the predicted survival time based on the characteristic-path-length ratio was computed as the number of patients who lived between 126 and 288 months divided by the total number of patients whose characteristic-path-length ratios are within the ratio range (p = 7/9 = 0.7778).

Both intensity and percentage of the IHC staining have to be considered when we score the slides. We have been working with such a classic scoring system for many years. We have realized that even two experienced pathologists score the slides, there is still difficulty to make clear decisions for about 10% of the cases. In this study, a new image-based network analysis was developed to analyze the immunostaining array slides and to extract patterns of the IHC staining, including both intensity and percentage in the whole arrays. We further analyzed the associations of the immunostaining patterns with our clinical data to provide more precise information for rectal cancer.

The mean values of both clustering coefficients and characteristic path lengths of the rectal cancer patients of shorter survival are lower than those of longer survival. There is no correlation between the ratios of the clustering coefficients of the tumor to those of the biopsy and the survival time (correlation coefficient R = 0.0120, *p*-value = 9.7656e-04) among the shorter-surviving rectal cancer patients whose maximum survival time was about over 6 years (75 months). This can be observed from **Figure 4**. There is also no correlation between





the ratios of the characteristic path lengths of the tumor to those of the biopsy and the survival time (R = -0.0780, pvalue = 9.7656e-04) among the shorter-surviving patients. This can also be observed from **Figure 4**. There is an indication of correlation between the ratios of the clustering coefficients of the tumor to those of the biopsy and the survival time (R = 0.4924, p-value = 1.2207e-04) among the longer-surviving rectal patients whose maximum survival time was 24 years (288 months). This can be observed from **Figure 4**. There is also evidence of correlation between the ratios of the characteristic path lengths of the tumor to those of the biopsy and the survival time (R = 0.4778, p-value = 1.2207e-04) among the longer-surviving rectal cancer patients. This can also be observed from **Figure 4**.

Figure 4 shows similar plots of the ratios of the two networkproperty parameters of the tumor to biopsy against the survival time, suggesting the consistency of the results. It is reported



(A) ratios of clustering coefficients of primary tumor [CC(tumor)] to those of biopsy [CC(biopsy)], and (B) ratios of characteristic path lengths of primary tumor [CP(tumor)] to those of biopsy [CP(biopsy)]. Symbols "O" and "□" indicate patient groups with shorter and longer times of survival, respectively.

that rectal patients who survive at least 5 years (60 months) are likely to die from causes that are common in the general population (London, 2017). This finding highlights the predictive value of DNp73 revealed by the image-based FWRN analysis among the cohort of rectal cancer patients whose survival time was between 8.4 years (101 months) and 24 years (288 months) correlated with the clustering-coefficient ratios, and 10.5 years (126 months) and 24 years (288 months) correlated with the characteristic-path-length ratios.

The lack of correlation of the ratios of the MC-FWRN parameters and the (shorter) survival time may suggest an implication of poor responses or non-effective treatment of the RT provided to the rectal cancer patients. Meanwhile, those patients who have positive correlation between the ratios of the FWRN parameters and the (longer) survival time were very likely to have a good or better response to the RT. General findings are that higher values of the ratios of the MC-FWRN parameters indicate longer survival time. The longest survival time is found with the values of the MC-FWRN parameters of the 25 rectal cancer patients and their survival months, we can predict



deemed to have longer survival time as being between 101 and 288 months, in terms of ratios of clustering coefficient (**A**), and characteristic path length (**B**) of image-based FWRNs. CC(PT) and CC(B) denote clustering coefficients of primary tumor and biopsy, respectively; and CP(PT) and CP(B) denote characteristic path lengths of primary tumor and biopsy, respectively.

the survival time between 101 months (8.42 years) and 288 months (24 years) with a probability of 73% for those patients whose clustering-coefficient ratio is within the range between 0.97 and 1.05. Similarly, the survival time between 126 months (10.5 years) and 288 months (24 years) with a probability of 78% for those patients whose characteristic-path-length ratio is within the range between 0.99 and 1.10.

The application of a novel image-based network analysis presented in this study was able to discover the predictive factor of DNp73 biomarker in rectal cancer patients having preoperative RT. Predictive biomarkers provide useful information on the probability of obtaining a response to treatment (Walther et al., 2009) and support the process of therapeutic decision for personalized cancer treatment (Voon and Kong, 2011). Such a discovery of DNp73 expression as a predictive biomarker in rectal cancer patients is expected to provide early assessment of the patient outcome, clinical value in the diagnostics of the disease, identification of targeted postoperative therapy.

Regarding the MC-FWRN introduced in this study, this new method appears to be the first of its kind mathematically formulated to capture the recurrence features of multi-channel data inherently existing in complex histology images in a way that is both effective and easily implemented for practical use. Complex networks consist of certain attributes that can be computed to analyze the properties and characteristics of the networks. Mathematical properties of these networks are utilized to define network models and to elucidate how certain models different to each other. The proposed MC-FWRN allows the calculation of the clustering coefficients and characteristic path lengths of DNp73 expression in the primary tumors and biopsies. These values can used to predict the survival time of a cohort of rectal cancer patients who were deemed to be positively influenced by preoperative RT.

The fuzzy weighted recurrence network analysis proposed herein is not supposed to be the study of the complexity of DNp73-controlled networks, but the derivation of structural properties of DNp73 expression from complex microscopy images that can be difficult to understand by pathologists. The results suggest that there are relationships between the graph properties of fuzzy weighted recurrence networks and the color distribution of the stained images. Hence, the network analysis vields new quantitative characteristics of the complexity of the IHC detection of the protein in tissue sections. From a molecular biology perspective, the average clustering coefficient and characteristic path length of the image-based fuzzy weighted recurrence network provide a mathematical measure of the heterogeneity of DNp73 in IHC staining, in correlation with clinicopathological characteristics. This heterogeneity may reflect diverse cell populations in expressing different levels of DNp73.

In this study, we have shown significant results concerning the DNp73 protein expression in predicting the outcome for the rectal cancer patients with the proposed mathematical approach. A limitation of this study is a relatively small number of the rectal cancer patients selected in the analysis. Therefore, future studies with more subgroups of rectal patients will be considered. It should be pointed out that although the total samples of the RT clinical trail from the Southeast Swedish Health Care region included 216 cases, only 102 cases randomly received preoperative RT. Given the aim of this study, only the paired samples of biopsy and primary tumors that are from the same patients were selected for the analysis. Many of the biopsy samples are too tiny to be used for IHC staining, constituting to the limitation of the sample size carried out in this pilot study, which still can provide some representative indication due to the paired samples from the same patient and all the samples derived from the random clinical trial.

Furthermore, results from rectal cancer patients with and without preoperative RT will be obtained and compared. Images of biopsies, primary cancers and metastatic cancers should be further investigated. Eventually, we will analyze the associations of the reactions from tumor invasive margins and stroma with the patients' prognosis.

Another limitation in this study is that the TAp73 expression was not performed in the present 25 pairs of rectal cancer samples. It is known that TAp73 acts as a tumor suppressor, while DNp73 exerts as an oncogene that is opposite to TAp73 (Amelio et al., 2014; Stantic et al., 2015). Therefore, it is necessary to expand the sample size and simultaneously evaluate TAp73 and DNp73 in the future, based on the methodology we have developed in the current study. DNp73 links to the ability to act as dominant-negative of the TAp73 isoforms and p53. This negative regulation by DNp73 forms an autoregulatory feedback loop, since both TAp73 and p53 can induce expression of DNp73 isoforms by direct binding to the P2 promoter (Irwin, 2006; Rufini et al., 2011; Di et al., 2013). A newest evidence showed that DNp73 isoforms has higher applicant potential in colorectal cancer patients than the canonical p73 protein (Garranzo-Asensio et al., 2019). Thus, it is reasonable to focus on DNp73. In addition, in the present paper, we mainly focused on the automated quantification of IHC expression using an image-based complex network model. The expression of TAp73 and the relationship between TAp73 and DNp73 will be investigated in our future study.

The highlights of the technical development and findings addressed in this paper are summarized as follows. First, the proposed MC-FWRN analysis of DNp73 expression by IHC in rectal cancer is the first of its kind. Second, a new mathematical analysis of IHC-stained biopsy and tumor images reveals the predictive power of DNp73 in rectal cancer patients who received RT. Third, a new method of multi-channel fuzzy weighted recurrence networks is developed for extracting two useful complex network properties of IHC images that can be used as prognostic indicators of rectal cancer. Fourth, the proposed approach for quantifying the expression of IHC is not limited to the study of DNp73, but can also be generally applied to discovering image patterns of other tumor proteins. Fifth, the proposed approach can be utilized as a computerized tool for extracting features from whole slide images in digital pathology.

4. CONCLUSION

The findings presented herein show the useful application of complex network analysis of images for studying the predictive factor of DNp73 biomarker expression in rectal cancer patients. The use of DNp73 biomarker can give insight into preoperative

REFERENCES

- Albert, R., and Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.* 74, 47–97. doi: 10.1103/RevModPhys.74.47
- Amelio, I., Antonov, A. A., Catani, M. V., Massoud, R., Bernassola, F., Knight, R. A., et al. (2014). TAp73 promotes anabolism. *Oncotarget* 5, 12820–12934. doi: 10.18632/oncotarget.2667

RT that has been considered as an important companion in the treatment of rectal cancer. A larger sample size when being available in future clinical trial will further confirm the current findings. Moreover, the proposed approach is not only found useful to rectal cancer but also can be adopted for the analysis of other biomarkers as well as other types of cancer, where human-based pathology practice is of limited capacity. In fact, there are many reports on the computerized image analysis of H&E (Haemotoxylin and Eosin) staining, much less effort has been made to apply computational methods for the automated analysis of IHC staining. The MC-FWRN presented in this paper can be generally applied for studying the expression of other potential biomarkers.

Although there are many studies reported about the association between DNp73 protein biomarker expression and malignant potential, the function of DNp73 still remains unclear. Our work contributes to the elucidation of the predictive value of DNp73 expression in rectal cancer patients who were given preoperative RT. We developed an original method for constructing weighted recurrence networks of multi-channel images. These networks allow the extraction of useful network properties from complex IHC images. The clustering coefficients and characteristic path lengths of the MC-FWRNs are not only able to show the predictive factor of DNp73 expression in the patients, but also reveal the identification of non-effective application of RT to those who had poor overall survival outcome.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Good Clinical Practice, the Research Ethics Committee in Linkoping, Sweden with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the the Research Ethics Committee in Linkoping, Sweden.

AUTHOR CONTRIBUTIONS

TP, HZ, and X-FS conceived the project. TP developed the fuzzy weighted recurrence networks of multi-channel images. TP, CF, HZ, and X-FS analyzed the results. TP, CF, and X-FS wrote the paper. TP carried out the computer implementation of the computational methods. DP contributed to the Lab work and data preparation. All authors edited and approved the manuscript.

- Arnold, M., Sierra, M. S., Laversanne, M., Soerjomataram, I., Jemal A., and Bray, F. (2017). Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 66, 683–691. doi: 10.1136/gutjnl-2015-310912
- Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. U.S.A.* 101, 3747–3752. doi: 10.1073/pnas.0400087101

- Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. New York, NY: Plenum Press.
- Chatterjee, S. B., Hou, J., Bandaru, V. V. R., Pezhouh, M. K., Syed Rifat Mannan, A. A., and Sharma, R. (2019). Lactosylceramide synthase β -1,4-GalT-V: a novel target for the diagnosis and therapy of human colorectal cancer. *Biochem. Biophys. Res. Commun.* 508:380. doi: 10.1016/j.bbrc.2018. 11.149
- Cuyle, P. J., and Prenen, H. (2017). Current and future biomarkers in the treatment of colorectal cancer. *Acta Clin. Bel.* 72, 103–115. doi: 10.1080/17843286.2016.1262996
- Di, C., Yang, L., Zhang, H., Ma, X., Zhang, X., and Sun, C. (2013). Mechanisms, function and clinical applications of DNp73. *Cell Cycle* 12, 1861–1867. doi: 10.4161/cc.24967
- Dulloo, I., Gopalan, G., Melino, G., and Sabapathy, K. (2010). The antiapoptotic DeltaNp73 is degraded in a c-Jun-dependent manner upon genotoxic stress through the antizyme-mediated pathway. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4902–4907. doi: 10.1073/pnas.0906782107
- Fagiolo, G. (2007). Clustering in complex directed networks. Phys. Rev. E 76:026107. doi: 10.1103/PhysRevE.76.026107
- Fan, C. W., Chen, T., Shang, Y. N., Gu, Y. Z., Zhang, S. L., Lu, R., et al. (2013). Cancer-initiating cells derived from human rectal adenocarcinoma tissues carry mesenchymal phenotypes and resist drug therapies. *Cell Death Dis.* 4:e828. doi: 10.1038/cddis.2013.337
- Garranzo-Asensio, M., Guzmán-Aránguez, A., Povés, C., Fernández-Aceñero, M. J., Montero-Calle, A., Ceron, M. Á., et al. (2019). The specific seroreactivity to ∆Np73 isoforms shows higher diagnostic ability in colorectal cancer patients than the canonical p73 protein. *Sci. Rep.* 9:13547. doi: 10.1038/s41598-019-49960-x
- Irwin, M. S. (2006). ΔNp73: misunderstood protein? *Cancer Biol. Ther.* 5, 804–807. doi: 10.4161/cbt.5.7.3023
- Josef Gladitz, J., Klink, B., and Seifert, M. (2018). Network-based analysis of oligodendrogliomas predicts novel cancer gene candidates within the region of the 1p/19q co-deletion. Acta Neuropathol. Commun. 6:49. doi: 10.1186/s40478-018-0544-y
- Lee, P. Y., Chin, S. F., Rahman, T. Y. L., and Jamal, R. (2018). Probing the colorectal cancer proteome for biomarkers: current status and perspectives. J. Proteom. 118, 93–105. doi: 10.1016/j.jprot.2018.06.014
- Letellier, E., Schmitz, M., Ginolhac, A., Rodriguez, F., Ullmann, P., Qureshi-Baig, K., et al. (2017). Loss of Myosin Vb in colorectal cancer is a strong prognostic factor for disease recurrence. *Brit. J. Cancer* 117, 1689–1701. doi: 10.1038/bjc.2017.352
- Liu, S. S., Chan, K. Y. K., Cheung, A. N. Y., Liao, X. Y., Leung, T. W., and Ngan, H. Y. S. (2006). Expression of ΔNp73 and TAp73a independently associated with radiosensitivities and prognoses in cervical squamous cell carcinoma. *Clin. Cancer Res.* 12, 3922–3927. doi: 10.1158/1078-0432.CCR-05-2573
- Logotheti, S., Pavlopoulou, A., Galtsidis, S., Vojtesek, B., and Zoumpourlis, V. (2013). Functions, divergence and clinical value of TAp73 isoforms in cancer. *Cancer Metastasis Rev.* 32, 511–534. doi: 10.1007/s10555-013-9424-x
- London, S. (2017, May 25). Common causes of death predominate among longterm colorectal cancer survivors. *The ASCO Post*.
- Lucena-Araujo, A. R., Kim, H. T., Thome, C., Jacomo, R. H., Melo, R. A., and Bittencourt, R. (2015). High △Np73/TAp73 ratio is associated with poor prognosis in acute promyelocytic leukemia. *Blood* 126, 2302–2306. doi: 10.1182/blood-2015-01-623330
- Muller, M., Schilling, T., Sayan, A. E., Kairat, A., Lorenz, K., and Schulze-Bergkamen, H. (2005). TAp73/Delta Np73 influences apoptotic response, chemosensitivity and prognosis in hepatocellular carcinoma. *Cell Death Differ*. 12, 1564–1577. doi: 10.1038/sj.cdd.4401774
- Murray-Zmijewski, F., Lane, D. P., and Bourdon, J. C. (2006). p53/p63/p73 isoforms: an orchestra of isoforms to harmonise cell differentiation and response to stress. *Cell Death Differ*. 13, 962–972. doi: 10.1038/sj.cdd.44 01914
- Newman, M. E. J. (2010). *Networks: An Introduction*. Oxford: Oxford University Press.

- Ozturk, K., Dow, M., Carlin, D. E., Bejar, R., and Carter, H. (2018). The emerging potential for network analysis to inform precision cancer medicine. *J. Mol. Biol.* 430, 2875–2899. doi: 10.1016/j.jmb.2018.06.016
- Parikh, A. P., Curtis, R. E., Kuhn, I., Becker-Weimann, S., Bissell, M., Xing, E. P., et al. (2014). Network analysis of breast cancer progression and reversal using a tree-evolving network algorithm. *PLoS Comput. Biol.* 10:e1003713. doi: 10.1371/journal.pcbi.1003713
- Patel, J. N., Fong, M. K., and Jagosky, M. (2019). Colorectal cancer biomarkers in the era of personalized medicine. J. Pers. Med. 9:E3. doi: 10.3390/jpm9010003
- Pfeifer, D., Wallin, A., Holmlund, B., and Sun, X. F. (2009). Protein expression following gamma-irradiation relevant to growth arrest and apoptosis in colon cancer cells. J. Cancer Res. Clin. Oncol. 135, 1583–1592. doi: 10.1007/s00432-009-0606-4
- Pham, T. D. (2019). Fuzzy weighted recurrence networks of time series. *Physica A* 513, 409–417. doi: 10.1016/j.physa.2018.09.035
- Rahman, M. R., Islam, T., Gov, E., Turanli, B., Gulfidan, G., Shahjaman, M., et al. (2019). Identification of prognostic biomarker signatures and candidate drugs in colorectal cancer: insights from aystems biology analysis. *Medicina* 55:20. doi: 10.3390/medicina55010020
- Ruan, J., Jahid, M. J., Gu, F., Lei, C., Huang, Y. W., Hsu, Y. T., et al. (2019). A novel algorithm for network-based prediction of cancer recurrence. *Genomics* 111, 17–23. doi: 10.1016/j.ygeno.2016.07.005
- Rufini, A., Agostini, M., Grespi, F., Tomasini, R., Sayan, B. S., Niklison-Chirou, M. V., et al. (2011). p73 in cancer. *Genes Cancer* 2, 491–502. doi: 10.1177/1947601911408890
- Ryan, J. E., Warrier, S. K., Lynch, A. C., Ramsay, R. G., Phillips, W. A., and Heriot, A. G. (2016). Predicting pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer: a systematic review. *Colorectal Dis.* 18, 234–246. doi: 10.1111/codi.13207
- Sinicrope, F. A., Okamoto, K., Kasi, P. M., and Kawakami, H. (2016). Molecular biomarkers in the personalized treatment of colorectal cancer. *Clin. Gastroenterol. Hepatol.* 14, 651–658. doi: 10.1016/j.cgh.2016.02.008
- Sonke, J. J., and Belderbos, J. (2010). Adaptive radiotherapy for lung cancer. Semin. Radiat. Oncol. 20:94e106. doi: 10.1016/j.semradonc.2009.11.003
- Stantic, M., Sakil, H. A., Zirath, H., Fang, T., Sanz, G., Fernandez-Woodbridge, A., et al. (2015) TAp73 suppresses tumor angiogenesis through repression of proangiogenic cytokines and HIF-1α activity. *Proc. Natl. Acad. Sci. U.S.A.* 112, 220–225. doi: 10.1073/pnas.1421697112
- Stiewe, T., Zimmermann S., Frilling A., Esche, H., and Putzer, B. M. (2002). Transactivation-deficient DeltaTA-p73 acts as an oncogene. *Cancer Res.* 62, 3598–3602.
- Swedish Rectal Cancer Trial, Cedermark, B., Dahlberg, M., Glimelius, B., Påhlman, L., Rutqvist, L. E., et al. (1997). Improved survival with preoperative radiotherapy in resectable rectal cancer. *N. Engl. J. Med.* 8, 980–987. doi: 10.1056/NEJM199704033361402
- Uramoto, H., Sugio, K., Oyama, T., Nakata, S., Ono, K., Morita, M., et al. (2004). Expression of △Np73 predicts poor prognosis in lung cancer. *Clin Cancer Res.* 10, 6905–6911. doi: 10.1158/1078-0432.CCR-04-0290
- Voon, P. J., and Kong, H. L. (2011). Tumour genetics and genomics to personalise cancer treatment. Ann. Acad. Med. Singapore 40, 362–368.
- Walther, A., Johnstone, E., Swanton, C., Midgley, R., Tomlinson, I., and Kerr, D. (2009). Genetic prognostic and predictive markers in colorectal cancer. *Nat. Rev. Cancer* 9, 489–499. doi: 10.1038/nrc2645
- Watts, D. J., and Strogatz, S. (1998). Collective dynamics of "small-world" networks. *Nature* 393, 440–442. doi: 10.1038/30918
- Yaromina, A., Krause, M., and Baumann, M. (2012). Individualization of cancer treatment from radiotherapy perspective. *Mol. Oncol.* 6, 211–221. doi: 10.1016/j.molonc.2012.01.007
- Ye, H., and Guo, X. (2019). TP73 is a credible biomarker for predicting clinical progression and prognosis in cervical cancer patients. *Biosci. Rep.* 39:BSR20190095. doi: 10.1042/BSR20190095
- Zadeh, L. A. (1971). Similarity relations and fuzzy orderings. Informat. Sci. 3, 177–200. doi: 10.1016/S0020-0255(71)80005-1
- Zarkavelis, G., Boussios, S., Papadaki, A., Katsanos, K. H., Christodoulouc, D. K., and Pentheroudakis, G. (2017). Current and future biomarkers in colorectal cancer. *Ann. Gastroenterol.* 30, 613–621. doi: 10.20524/aog.2017.0191

- Zhang, F., Wang, M., Xi, J., Yang, J., and Li, A. (2018). A novel heterogeneous network-based method for drug response prediction in cancer cell lines. *Sci. Rep.* 8:3355. doi: 10.1038/s41598-018-21622-4
- Zhang, W., Ota, T., Shridhar, V., Chien, J., Wu, B., and Kuang, R. (2013). Network-based survival analysis reveals subnetwork signatures for predicting outcomes of ovarian cancer treatment. *PLoS Comput. Biol.* 9:e1002975. doi: 10.1371/journal.pcbi.1002975
- Zhu, W., Pan, X., Yang, Z., Xing, P., Zhang, Y., Li, F., et al. (2015). Expression and prognostic significance of TAp73 and ΔNp73 in FIGO stage I-II cervical squamous cell carcinoma. *Oncol. Lett.* 9, 2090–2094. doi: 10.3892/ol.20 15.3052

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Pham, Fan, Pfeifer, Zhang and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.