Gene 550 (2014) 81–92

Contents lists available at ScienceDirect

Gene

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

A behavioral study of healthy and cancer genes by modeling electrical network

Tanusree Roy *, Soma Barman

Institute of Radio Physics and Electronics, University of Calcutta, 92, APC Road, Kolkata 700009, India

ARTICLE INFO

Article history:
Received 13 April 2014
Received in revised form 18 June 2014
Accepted 6 August 2014
Available online 10 August 2014

Keywords:
Genomics
Genetic instability
Cancer
Network simulation
System model
LabVIEW

ABSTRACT

In recent years, gene network modeling is gaining popularity in genomics to monitor the activity profile of genes. More specifically, the objective of the network modeling concept is to study the genetic behavior associated with disease. Previous researchers have designed network model at nucleotide level which produces more complexity for designing circuits mostly in case of gene expression studies. Whereas the authors have designed the present network model, based on amino acid level which is simpler as well as more appropriate for prediction of the genetic abnormality. In the present concept, SISO continuous and discrete system models of genes are realized using Foster network. The model is designed based on hydropathy index value of amino acids to study the biological system behavior. The time and phase response in continuous (s) domain and pole-zero distribution in discrete (z) domain are used as measurement metric in the present study. The simulated responses of the system show genetic instability for cancer genes which truly reflects the medical reports. The proposed modeling concept can be used, to accurately identify or separate out the diseased genes from healthy genes.

1. Introduction

Genomics entails the study related to large sets of genes with the goal of understanding the collective gene function rather than the individual gene. The genes are responsible for the creation of amino acids (Anastassiou, 2001; Vaidyanathan, 2004). The twenty amino acids of genes are combined in various ways to create different types of proteins, present in all living organisms; deficiency of which can cause problems ranging from indigestion to depression, stunted growth and physical disorders. Some of the amino acids play significant role in growth of cancer genes. Inhibition of these amino acids may be beneficial for curing cancer patients (https://www.apjohncancerinstitute.org/caat-protocol).

Cancer is a genetic disease (Vogelstein and Kinzler, 2004) and it results from the accumulation of mutations in the genes (Lengauer et al., 1998). Cancer genes produce inordinate amount of lactic acid (Warburg, 1956) and arginine can modulate the growth of breast cancer (Singh et al., 2000). Several researchers have used the statistics of amino acids to study the gene behavior and predict the cancer and healthy genes (Barman et al., 2011a,b; Das and Mitra, 2011; Roy et al., 2013). But still determination of cancer genes accurately is tedious, inefficient and also a significant challenge for oncologists and computational biologists.

Since network modeling is gaining importance in genomic research to learn about the structure and function of DNA, different studies have mostly focused on developing and manipulating algorithms based on molecular computing, molecular algebra, simulation and modeling (Alfinito et al., 2008; Hodzic and Newcomb, 2007; Marshall, 2009, 2010a,b; Sampath, 2006). The genetic networks are very much complex as large amount of uncertainty associated with this model, most of the models proposed are probabilistic. For example, the DNA/RNA strings are modeled using passive electrical components at nucleotide level (Marshall, 2010a) and the ultimate circuitry is incredibly complex when the modeling concept is applied on *Homo sapiens* genes i.e. composed of long nucleotide sequence. The amino acid property based methods have proven more sensitive and responsive than nucleotide based methods (McClellan, 2012); therefore the present network model is much more appropriate as modeling is done on amino acid level. Considering circuit complexity and computational load, the concept is novel in the present era of technology.

In our study, 'genetic instability (Lengauer et al., 1998) and majority of hydrophilic amino acid composition (Stranzl et al., 2012) in cancer associated genes' are the key features to distinguish cancer genes from healthy genes. The main objectives of this paper include:

- Designing electrical circuit model for individual amino acid based on their hydropathy index value (Kyte and Doolittle, 1982) of the side chain from the elementary level using passive electrical components R, L and C.

Abbreviations: DNA, deoxyribonucleic acid; RNA, ribonucleic acid; NCBI, National Center For Biotechnology Information; CGAP, Cancer Genome Anatomy Project; R, resistor; C, capacitor; L, inductor; SISO, single input single output; ATF3, activating transcription factor 3; R_L, load resistance; LabVIEW, Laboratory Virtual Instrument Engineering Workbench; CD, continuous domain; ROC, region of convergence.

* Corresponding author.
E-mail address: tanusree.rinki@gmail.com (T. Roy).

http://dx.doi.org/10.1016/j.gene.2014.08.020
0378-1119/

- Modeling electrical network for genes by cascading individual amino acid model.
- Investigating the network responses for identification of cancer and healthy genes associated with breast, prostate and lung cell in continuous and discrete domain; downloaded from NCBI and CGAP homepages (http://www.ncbi.nlm.nih.gov; http://cgap.nci.nih.gov).

2. Methods

The simulated system model is realized using LabVIEW (version 2012) and MATLAB (version R2009b) to study the genetic behavior, associated with cancerous and healthy Homo *sapiens* cells.

2.1. Model realization of basic amino acid structure

The basic structure of amino acid having three groups i.e. the carboxyl group (COOH), amino group (NH_2) and the variable side chain (r) group; which are attached with a central alpha-carbon (Fig. 1A). The chemical compositions of side chain determine the chemical specificity of an amino acid (Voet et al., 2001).

The generalized electrical circuit model of an amino acid consists of two distinct circuits i.e. the backbone circuit and the side chain (r) group circuit (Fig. 1B). The circuits for the backbone structure comprised of carboxyl (COOH) and the amino group (NH_2) are common for all amino acids. The circuit which is connected in parallel with the backbone structure represents the side chain. The impedance of the backbone circuit model is Z_b and the side chain (r) impedance is Z_{SC}.

In the present network model, the backbone structure is represented by resistor R (7 Ω) as the structure comprised of total seven atoms i.e. four atoms from carboxyl group (COOH) and three atoms from amino group (NH_2). The amino acid proline has a slightly different structure among twenty amino acids as the "r" group is bent and attaches itself with nitrogen in one of the hydrogen atom place, so for proline, the backbone structure is represented through resistor R of 6 Ω as amino group (NH) has two atoms. The amino acid structure parameters are used for modeling (Table 1).

Depending on the polarity of the side chain, amino acids vary in their hydrophilic or hydrophobic character. In the present case, the amino acids are classified into two groups i.e. hydrophilic and hydrophobic based on hydropathy index value (Kyte and Doolittle, 1982) of the side chain (Table 2). The hydrophobic amino acids (Ala, Cys, Leu, Met, Ile, Phe and Val) having positive hydropathy index values are designed by inductors L (mH) as the inductor has leading or positive phase and the hydrophilic amino acids (Gln, Asp, Glu, Gly, His, Arg, Lys, Asn, Pro, Ser, Thr, Trp and Tyr) having negative hydropathy index values are modeled through capacitors C (μF) as the capacitor has lagging or negative phase. The values of inductors and capacitors are selected based on the hydropathy index value of individual amino acid (Fig. 2). So the reduced equivalent impedance for the amino acid circuit is given by,

Table 1
Amino acid structure circuit modeling parameters.

Amino acid structure		Circuit element
Backbone structure	COOH	R
	NH_2	R
Side chain structure	Hydrophobic side chain	L
	Hydrophilic side chain	C

COOH = carboxyl group. NH_2 = amino group. R = resistor. L = inductor. C = capacitor.

$$Z_{eq} = \frac{Z_b \cdot Z_{SC}}{Z_b + Z_{SC}}. \tag{1}$$

The equivalent impedance Z_{eq} of twenty amino acids is different and computed using Eq. (1).

2.2. Amino acid chain modeling

The n numbers of amino acids are joined together in series to form 1st Foster network (Daryanani, 1968) and protein structure is formed as shown in Fig. 3. The protein structure is converted into an equivalent single electrical circuit, where the impedances are replaced by equivalent impedance (Z^n_{eq}).

The electrical circuits of n number amino acid strings (Fig. 4A) are lumped into its equivalent network model (Fig. 4B), where the circuit's arms Z^n_b and Z^n_{SC} present the nth amino acid in a string and the previous $n-1$ amino acids are lumped into equivalent impedance Z^{n-1}_{eq}.

For amino acid chain of length two, the electrical circuit is obtained by attaching second amino acid model to the equivalent of the first amino acid model. The process of cascading circuit elements is carried out to obtain the generalized expressions of electrical impedance of amino acid chain of any length. For length of two to arbitrary length (n) of amino acids, the equivalent impedance is given by Eq. (2),

$$\begin{aligned} Z^2_{eq} &= Z^1_{eq} + \frac{Z^2_B \cdot Z^2_{SC}}{Z^2_B + Z^2_{SC}} \\ &\vdots \\ Z^n_{eq} &= Z^{n-1}_{eq} + \frac{Z^n_B \cdot Z^n_{SC}}{Z^n_B + Z^n_{SC}} \end{aligned} \tag{2}$$

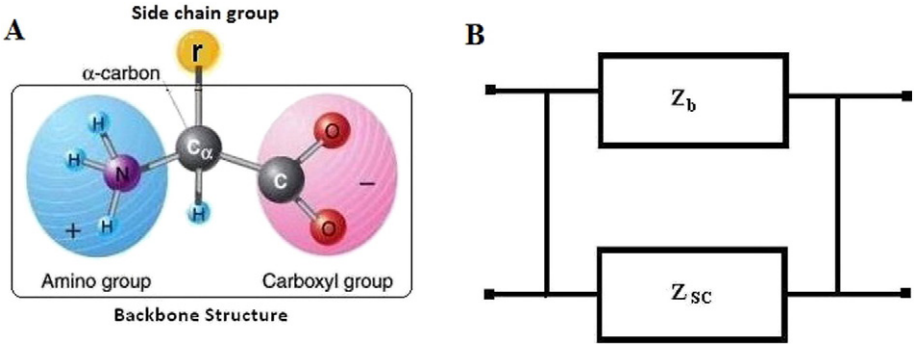

Fig. 1. Amino acid structure. A. Basic structure of amino acid composed of amino group, carboxyl group and side chain group. The backbone structure i.e. amino and carboxyl group is same for all twenty amino acids and the side chain is variable. B. Generalized electrical circuit model for amino acid structure. The impedances of the backbone structure and side chain are Z_b and Z_{sc}.