

Research Article
Volume 8 Issue 2 - August 2017
DOI: 10.19080/CTBEB.2017.08.555731

**Curr Trends Biomedical Eng & Biosci** 

Copyright © All rights are reserved by Rafiul Amin Laskar

# Protein Sequence Textuary of Biomedical Images Based on Protein Modelization of MRI



### Mohammad Reza Dawoudi<sup>1\*</sup> and Attila Gyenesei<sup>2</sup>

<sup>1</sup>Salo Nursing Campus, Turku University of Applied Sciences, Finland

<sup>2</sup>The Finnish Microarray and Sequencing Centre, Finland

Submission: August 17, 2017; Published: August 29, 2017

\*Corresponding author: Rafiul Amin Laskar, Mutation Breeding Lab., Department of Botany, Aligarh Muslim University, Aligarh, 202002, UP, India, Email: rafihkd@gmail.com

### Abstract

Serving an image as a text is one of the items on the data-base. A variety of methods could be employed. In this paper we have proposed a novel method for converting medical images into textuary formats. Through this method it is possi-ble to create protein sequences of the current biomedical medical images. This is achieved using the protein modelization of digital medical images. Protein modelization of a digital image is an arrangement for allowing numbers of pix-els values to be presented as sequences of characters. In this system every pixel values of color scale is procedure into four protein characters. One advantage of protein sequence textuary compared to other image serving methods is the browsing, retrieving and sequence blasting from database.

# Introduction

This author kit is designed to assist you in preparing your submission. It is an exact representation of the format expected by the editor. A template in Microsoft Word is available on the Medinfo 2013 World Wide Web server at http://www.medinfo2013.dk. This example provides all the information you require to comply with the Medinfo 2013 submission standards for Scientific demonstrations. The official language of the Medinfo 2013 Congress in Co-penhagen, Denmark is English. Values should be in Interna-tional System units.

To ensure that the greatest opportunity for participation in Medinfo 2013 is made available to researchers and institu-tions, authors may appear in multiple submissions but may have only ONE submission per category considered where they are the primary or first author [1-6].

## **Methods and Materials**

Protein Modelization of a digital image is an arrangement for allowing numbers of pixels values and pixels maps to be pre-sented as sequence of protein characters. The possible pixel value is arranged from 0 to 255 and the permutations with repetitions of four letters are 256, therefore every pixel values could be procedure into four letters. The first step is to carry out the direct coding of the digital image, with the four letters A, G, C, and T. Alphabet {A, C, G, T} is chosen, because it is the same as in real DNA sequences. The second stage is translation DNA sequence to protein sequence. This is per-formed with translated tools which allow the translation of DNA sequences to protein sequences. ExPASy [7] and Trans-late Nucleic Acid

Sequence Tool [8] are examples of free online translation tools. Last section will execute with multi-ple sequence alignment. The results will be obtained by the statistical significance of match's ratio. The similar image could be finding by corresponding degree of similarity (match's ratio) with image list.

## **Experiment Results**

# Normal Brain N8450T2 N8850T2 N9150T2 N9350T2 Neoplastic Disease (Sarcoma) S14502R S1550T2 S1750T2 S1650T2 Image

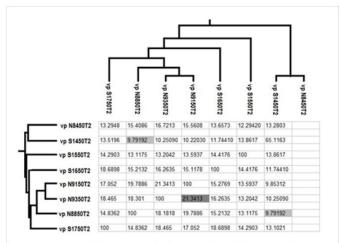
In this section we present the experiment result of protein modelization using Translate tool and Multiple Sequence

Alignment by CLUSTALW. The experiment performs protein modelization of the eight real lateral ventricles (brain) images (Image). In the first stage the pixels values of each images converted into the virtual DNA sequences. These conversions should be done based on Table 1. The second stage is performed with DNA translation in to protein sequences using translated tools. The statistical signif-icance of matches is calculated by sequence alignments. There are various algorithms and tools could be implemented. The DNA and protein sequences of images are shown in Ap-pendix I and II.

Table 1: Pixel values inside "[]" are replaced by the string of letters.

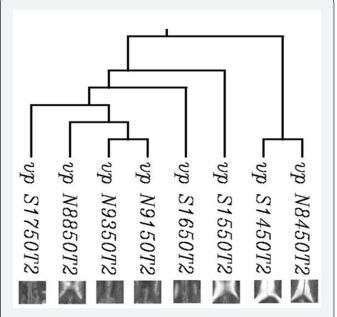
[0]AAAA	[32]AGAA	[64]CAAA	[96]CGAA	[128]GAAA	[160]GGAA	[192]TAAA	[224]TGAA	
[1]AAAC	[33]AGAC	[65]CAAC	[97]CGAC	[129]GAAC	[161]GGAC	[193]TAAC	[225]TGAC	
[2]AAAG	[34]AGAG	[66]CAAG	[98]CGAG	[130]GAAG	[162]GGAG	[194]TAAG	[226]TGAG	
[3]AAAT	[35]AGAT	[67]CAAT	[99]CGAT	[131]GAAT	[163]GGAT	[195]TAAT	[227]TGAT	
[4]AACA	[36]AGCA	[68]CACA	[100]CGCA	[132]GACA	[164]GGCA	[196]TACA	[228]TGCA	
[5]AACC	[37]AGCC	[69]CACC	[101]CGCC	[133]GACC	[165]GGCC	[197]TACC	[229]TGCC	
[6]AACG	[38]AGCG	[70]CACG	[102]CGCG	[134]GACG	[166]GGCG	[198]TACG	[230]TGCG	
[7]AACT	[39]AGCT	[71]CACT	(103)CGCT	[135]GACT	[167]GGCT	[199]TACT	[231]TGCT	
[8]AAGA	[40]AGGA	[72]CAGA	[104]CGGA	[136]GAGA	[168]GGGA	[200]TAGA	[232]TGGA	
[9]AAGC	[41]AGGC	[73]CAGC	[105]CGGC	[137]GAGC	[169]GGGC	[201]TAGC	[233]TGGC	
[10]AAGG	[42]AGGG	[74]CAGG	[106]CGGG	[138]GAGG	[170]0000	[202]TAGG	[234]TGGG	
[11]AAGT	[43]AGGT	[75]CAGT	[107]CGGT	[139]GAGT	[171]GGGT	[203]TAGT	[235]TGGT	
[12]AATA	[44]AGTA	[76]CATA	[108]CGTA	[140]GATA	[172]GGTA	[204]TATA	[236]TGTA	
[13]AATC	[45]AGTC	[77]CATC	[109]CGTC	[141]GATC	[173]GGTC	[205]TATC	[237]TGTC	
[14]AATG	[46]AGTG	[78]CATG	(110)CGTG	[142]GATG	[174]GGTG	[206]TATG	[238]TGTG	
[15]AATT	[47]AGTT	[79]CATT	(111)CGTT	[143]GATT	[175]GGTT	[207]TATT	[239]TGTT	
[16]ACAA	[48]ATAA	[80]CCAA	[112]CTAA	[144]GCAA	[176]GTAA	[208]TCAA	[240]TTAA	
[17]ACAC	[49]ATAC	[81]CCAC	[113]CTAC	[145]GCAC	(177)GTAC	[209]TCAC	[241]TTAC	
[18]ACAG	[50]ATAG	[82]CCAG	[114]CTAG	[146]GCAG	[178]GTAG	[210]TCAG	[242]TTAG	
[19]ACAT	[51]ATAT	[83]CCAT	[115]CTAT	[147]GCAT	[179]GTAT	[211]TCAT	[243]TTAT	
[20]ACCA	[52]ATCA	[84]CCCA	[116]CTCA	[148]GCCA	[180]GTCA	[212]TCCA	[244]TTCA	
[21]ACCC	[53]ATCC	[85]CCCC	[117]CTCC	[149]GCCC	[181]GTCC	[213]TCCC	[245]TTCC	
[22]ACCG	[54]ATCG	[86]CCCG	[118]CTCG	[150]GCCG	[182]GTCG	[214]TCCG	[246]TTCG	
[23]ACCT	[55]ATCT	[87]CCCT	[119]CTCT	[151]GCCT	[183]GTCT	[215]TCCT	[247]TTCT	
[24]ACGA	[56]ATGA	[88]CCGA	[120]CTGA	[152]GCGA	[184]GTGA	[216]TCGA	[248]TTGA	
[25]ACGC	[57]ATGC	[89]CCGC	[121]CTGC	[153]GCGC	[185]GTGC	[217]TCGC	[249]TTGC	
[26]ACGG	[58]ATGG	[90]CCGG	[122]CTGG	[154]GCGG	[186]GTGG	[218]TCGG	[250]TTGG	
[27]ACGT	[59]ATGT	[91]CCGT	[123]CTGT	[155]GCGT	[187]GTGT	[219]TCGT	[251]TTGT	
[28]ACTA	[60]ATTA	[92]CCTA	[124]CTTA	[156]GCTA	[188]GTTA	[220]TCTA	[252]TTTA	
[29]ACTC	[61]ATTC	[93]CCTC	[125]CTTC	[157]GCTC	[189]GTTC	[221]TCTC	[253]TTTC	
[30]ACTG	[62]ATTG	[94]CCTG	[126]CTTG	[158]GCTG	[190]GTTG	[222]TCTG	[254]TTTG	
[31]ACTT	(63)ATTT	[95]CCTT	[127]CTTT	[159]GCTT	[191]GTTT	[223]TCTT	[255]TTTT	

**Table 2:** Image comparison based on sequence alignment. This scores has to indicate the degree of similarity.



# Multiple sequence alignment by CLUSTALW

Multiple sequence alignment tools determine the ratio of sim-ilarity between protein sequences. The ratio of similarity be-tween eight protein sequences has computed using CLUSTALW. The ratios of similarity between sequences (statistical significance of matches) are shown in Table 2. The ratio of distance between eight protein sequences has computed using CLUSTALW, and the below graph has been obtained (Figure 1).



**Figure 1:** Image clustering with the Multiple Sequence Alignment by CLUSTALW, based on protein Modelization of Images.

### Results

The sequence alignments are executed to detect percentage of similarity ratio in virtual protein sequences. Based the degree similarity between virtual protein sequences we can categories the images. In every category a collection of objects are similar and according to a given score we can identify the group of objects which are "close" to this score.

### Acknowledgement

This article was carried out at The Finnish Microarray and Se-quencing Centre (FMSC) in Turku, Finland at the year 2009 and 2010. Special thanks go to head of the centre Dr. Attila Gyenesei for providing me the possibility to do my research on this significant subject.

### References

- 1. Barnard K, Shirahatti N. A method for comparing content based image retrieval methods.
- Marqes O, Mayron L. System and methods of image retrieval. Patent application. USPC Class: 382305.
- http://www.searchenginejournal.com/a-look-into-reverse-imagesearch-tools/14666/trackback
- 4. http://www.searchenginejournal.com/a-look-into-reverse-image-search-tools/14666/trackback/
- 5. http://au.expasy.org/tools/dna.html
- 6. http://align.genome.jp/
- 7. http://au.expasy.org/tools/dna.html
- 8. http://biotools.umassmed.edu/cgi-bin/biobin/transeq

# **Current Trends in Biomedical Engineering & Biosciences**



This work is licensed under Creative Commons Attribution 4.0 License

DOI: 10.19080/CTBEB.2017.08.555731

# Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- · Global attainment for your research
- Manuscript accessibility in different formats (Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission https://juniperpublishers.com/online-submission.php