Editor-in-Chief
Kuo-Chen Chou

Journal Editorial Board

TABLE OF CONTENTS

Volume 3, Number 1, January 2010

Journal of Biomedical Science and Engineering (JBiSE)

COPYRIGHT

The interaction between the 2009 H1N1 influenza A hemagglutinin and neuraminidase: mutations, co-mutations, and the NA stalk motifs

Wei Hu

Department of Computer Science, Houghton College, Houghton, NY, USA.
Email: wei.hu@houghton.edu

ABSTRACT

As the world is closely watching the current 2009 H1N1 pandemic unfold, there is a great interest and need in understanding its origin, genetic structures, virulence, and pathogenicity. The two surface proteins, hemagglutinin (HA) and neuraminidase (NA), of the influenza virus have been the focus of most flu research due to their crucial biological functions. In our previous study on 2009 H1N1, three aspects of NA were investigated: the mutations and co-mutations, the stalk motifs, and the phylogenetic analysis. In this study, we turned our attention to HA and the interaction between HA and NA. The 118 mutations of 2009 H1N1 HA were found and mapped to the 3D homology model of H1, and the mutations on the five epitope regions on H1 were identified. This information is essential for developing new drugs and vaccine. The distinct response patterns of HA to the changes of NA stalk motifs were discovered, illustrating the functional dependence between HA and NA. With help from our previous results, two co-mutation networks were uncovered, one in HA and one in NA, where each mutation in one network co-mutates with the mutations in the other network across the two proteins HA and NA. These two networks residing in HA and NA separately may provide a functional linkage between the mutations that can impact the drug binding sites in NA and those that can affect the host immune response or vaccine efficacy in HA. Our findings demonstrated the value of conducting timely analysis on the 2009 H1N1 virus and of the integrated approach to studying both surface proteins HA and NA together to reveal their interdependence, which could not be accomplished by studying them individually.

Keywords: Co-Mutations; Entropy; Epitope; H1N1; Hemagglutinin; Influenza; Mutation; Mutual Information; Neuraminidase; Phylogenetic Analysis; Stalk Motif; Swine Flu

1. INTRODUCTION

Influenza viruses caused several pandemics in history such as the Spanish flu (H1N1, 1918), the Asian flu (H2N2, 1958), and the Hong Kong flu (H3N2, 1968), where the H1N1 virus has the longest recorded history of human infection. In March and April 2009, a new A (H1N1) influenza virus first emerged in Mexico and the United States. Antigenically the new virus is similar to North American swine A (H1N1) viruses but distinct from seasonal human A (H1N1). This virus consisting of gene segments in swine or humans has acquired the capacity to spread quickly by human-to-human transmission across the globe and therefore has attracted international attention. On June 11, 2009, the World Health Organization (WHO) declared the H1N1 virus a pandemic.

The influenza A viral genome is composed of 8 genes encoding 11 proteins, including two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA). The main influenza antigens targeted by the human immune system are these two proteins, and influenza A subtypes are classified by the antigenic distinctions of the HA and NA proteins. There are 16 subtypes of HA and 9 subtypes of NA. HA mediates virus binding to sialic acid receptor on a host cell surface to initiate infection, and NA cleaves the binding to promote release of viral progeny. Otherwise, the viral progeny particles will remain aggregated at the cell surface. HA is also the main target of the host immune system. Once in a host cell, the HA protein comes under selective pressure for change to evade the host immune response. HA and NA have been of great interest in flu research due to their pivotal role in viral infection and replication.

The HA is a cylindrically shaped homotrimer molecule composed of three identical HA polypeptides, which are cleaved by protease into two subunits HA1 and HA2 during virus maturation. The globular region of the molecule is based mainly on the HA1 residues, and

the stem contains some residues of HA1 and all of HA2. The enzymatic domain of NA is held away from the virus surface by a long, thin stalk of variable length. The replication efficiency of viruses in eggs and mice correlates the NA stalk length [1]. NA with a short stalk was found to be inefficient in virion progeny since an active site located too close to the viral envelop could not access its substrate correctly [2]. The percentage of viruses with a short NA stalk has increased steadily in recent years [3].

There is a balanced interplay between HA and NA; one serves as receptor binder and one as receptor destroyer, to facilitate efficient virus replication in host cells [4]. Influenza viruses can overcome host restriction and become adapted to a new host by making changes in HA or/and NA [5], i.e., concomitant changes in HA and NA are required for influenza viruses to survive in host cells. The deficiency in NA activity conferred by the shortened protein stalk could be compensated by modulating the receptor binding affinity of HA to restore the functional balance between HA and NA [4]. In [6] a special stalk motif, commonly found in H5N1 in the past, was discovered in the 2009 H1N1 strains for the first time. This finding is significant given the fact that the viruses with this motif tend to have high virulence [3]. One of the goals of this study was to investigate the impact of the NA motifs on HA in the 2009 H1N1 strains.

As the 2009 H1N1 virus continues to transmit effectively from human to human, the occurrence of drug-resistant viruses is expected. One recent study [7] showed that the novel mutations of the 2009 H1N1 virus NA are located at sites that do not interfere with the active site so the currently used three drugs oseltamivir (Tamiflu®), zanamivir (Relenza®), and peramivir remain effective. Another study [6] identified two networks of co-mutations of 2009 H1N1 NA that may affect the active site from a greater distance.

As the principal antigen on the virus surface, HA is the main viral target for the human immune system, which can neutralize the virus through blocking viral binding to the receptors on host cells. An epitope is a region on the surface of an antigen, such as the HA in this study, capable of eliciting an immune response. Antigenic variation of HA is one mechanism employed by the flu virus to escape the response of the host immune system. One report found that one single amino acid substitution in 1918 H1N1 HA changes receptor binding specificity [8]. Influenza HA evolution is typically a combination of functional constraint and positive selection in epitope regions. As such, the identification of epitope regions on HA is important for both drug and vaccine development. Epitope mapping using monoclonal antibodies and the availability of the 3-dimensional structure have identified five antigenic sites in the HA of H3 subtype [9,10]. Corresponding antigenic sites have subsequently been mapped to H1 and H2 subtypes

[11,12]. With new technology as the one employed in [9,10], a recent refinement of the definition of H1 epitopes was conducted in [13]. One of the tasks of our study was to map the sequence mutations of the 2009 H1N1 HA relative to the five epitope regions of H1. We were also interested in finding the co-mutations in HA and co-mutations between HA and NA of 2009 H1N1.

2. MATERIALS AND METHODS

2.1. Sequence Data

Published HA sequences of 3936 influenza A virus from 2005 to 2009, H1 sequences of 1900 from 1918 to 2008, and H1 sequences of 508 in 2009 were downloaded from the Influenza Virus Resource (http://www.ncbi/nlm.nih-giv/genomes/FLU/FLU.html) of the National Center for Biotechnology Information (NCBI) on Oct. 13, 2009. We were mainly interested in the sequences in 2009, but also needed the sequences in several years before 2009 to provide comparison in the study. All the sequences used in the study were aligned with MAFFT [14].

2.2. Entropy and Mutual Information

In information theory [15], entropy is a measure of disorder or randomness associated with a random variable. Let x be a discrete random variable that has a set of possible values $\{a_1, a_2, a_3, ...a_n\}$ with probabilities $\{p_1, p_2, p_3, ...p_n\}$ where $P(x = a_i) = p_i$. The entropy H of x is

$$H(x) = -\sum_i p_i \log p_i$$

The mutual information of two random variables is a quantity that measures the mutual dependence of the two variables or the average amount of information that x conveys about y, which can be defined as

$$I(x, y) = H(x) + H(y) - H(x, y)$$

where $H(x)$ is the entropy of x, and $H(x, y)$ is the joint entropy of x and y. $I(x, y) = 0$ if and only if x and y are independent random variables.

In current study, each of the n columns in a multiple sequence alignment of a set of HA sequences of N residues is considered as a discrete random variable x_i (1 ≤i ≤N) that takes on one of the 20 (n=20) amino acid types with some probability. $H(x_i)$ has its minimum value 0 if all the residues at position i are the same, and achieves its maximum if all the 20 amino acid types appear with equal probability at position i, which can be verified by the Lagrange multiplier technique. A position of high entropy means that the amino acids are often varied at this position. While $H(x_i)$ measures the genetic diversity at position i in our current study, $I(x_i, y_j)$ measures the correlation between residue sub-
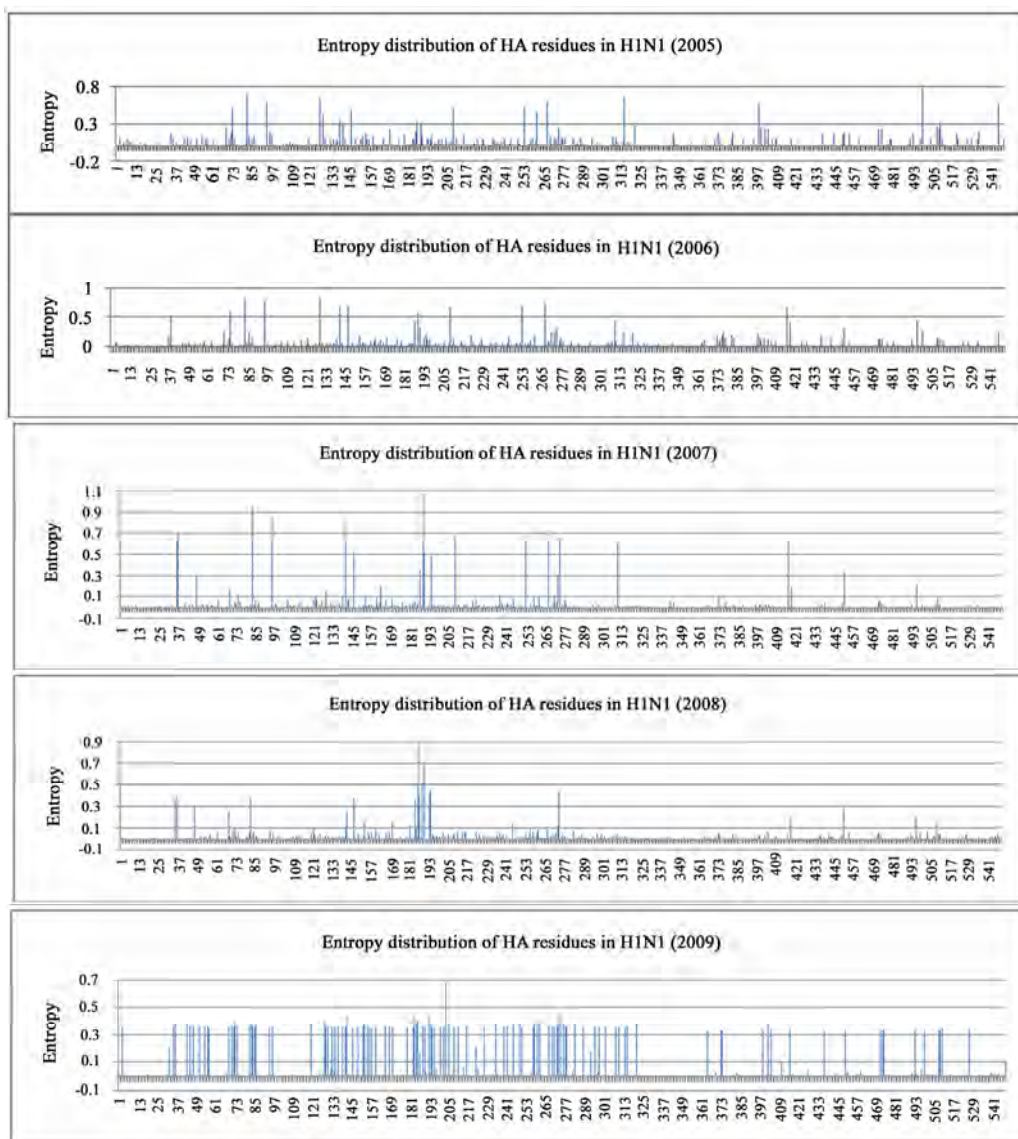
Figure 1. These five plots show the entropy distribution of HA residues in H1N1 from 2005 to 2009.

stitutions at positions i and j. A brief overview of the extensive applications of entropy and mutual information in sequence analysis, in particular the flu virus sequences, can be found in [6].

2.3. Mutual Information Evaluation

In order to assess the significance of our mutual information values of residue pairs of HA, it is necessary to show that these values are significantly higher than those based on random sequences. For each residue position of HA, we randomly permuted the amino acids from different sequences at that position and calculated the mutual information of these random sequences. This procedure was repeated 1000 times. The P value was calculated as the percentage of the mutual information values

of the permuted sequences that were higher than those of the sequences of HA.

2.4. Random Forest Clustering

Random Forest, proposed by Leo Breiman in 1999 [16], is an ensemble classifier based on many decision trees. The structure of a single tree could be easily altered by a small perturbation of data. Random Forest overcomes this problem by averaging across different decision trees. For many data sets, Random Forest produces a highly accurate classifier for supervised learning, comparable to Support Vector Machine, the state of the art machine-learning algorithm. It computes proximities between cases and this technique can be extended to unlabeled data, leading to unsupervised clustering.

Table 1. Amino acids on epitopes A, B, C, D, and E of H1 (A/California/04/2009 numbering) from [13].

Epi-tope	Amino acids	Number of amino acids
A	118,120,121,122,126,127,128,129,132,133,134,135,137,139,140,141,142,143,146,147,149,165,252,253	24
B	124,125,152,153,154,155,156,157,160,162,183,184,185,186,187,189,190,191,193,194,195,196	22
C	34,35,36,37,38,40,41,43,44,45,269,270,271,272,273,274,276,277,278,283,288,292,295,297,298,302,303,305,306,307,308,309,310	32
D	170,171,172,173,174,176,179,198,200,202,204,205,206,207,208,209,210,211,212,213,214,215,216,222,223,224,225,226,227,235,237,239,241,243,244,245	48
E	47,48,50,51,53,54,56,57,58,66,68,69,70,71,72,73,74,75,78,79,80,82,83,84,85,86,102,257,258,259,260,261,263,267	34

Table 2. 2009 H1N1 HA mutations on the five epitopes.

Epitope	Mutation Residues	Number of Mutations
A	126, 127, 128, 129,132,134,137,140,141,165, 252	11
B	152, 154,155,156, 183,184,185,189,193,194,195	11
C	35, 44, 269, 271,272,273,276,277,297,307	10
D	95, 167, 169, 204,206,207, 208, 210,215,223,226,244	12
E	50, 53, 56, 68, 70, 71, 72, 73, 82, 83, 84, 85, 257,259,260	15

To view the clusters formed by Random Forest, multidimensional scaling [17] was utilized to project high-dimensional data down into a low-dimensional space while preserving the distances between them. First the proximities between cases i and j form a symmetric and positive definite matrix {prox(i,j)}. Then a second positive definite and symmetric matrix {cv(i,j)} is constructed using the entries of {prox(i,j)}. Random Forest extracts a few largest eigenvalues of the cv matrix and their corresponding eigenvectors. The values of $\sqrt{e(i)}v(i)$ are referred to as the ith scaling coordinate, where $e(i)$ and $v(i)$ are the ith eigenvalue and eigenvector of matrix cv. In this study, the first and second scaling coordinates were utilized to visualize the data.

2.5. Important Sites in HA and NA

The NA active site is a shallow pocket constructed from conserved residues, some of which contact the substrate directly and participate in catalysis, while others provide a structural framework [18]. According to the numbering in [7], these residues of N1 are 118, 119, 151, 152, 156, 179, 180, 223, 225, 228, 247, 277, 278, 293, 295, 368, and 402. The antigenic sites of N1 are residues 83 – 143, 156 – 190, 252 – 303, 330, 332, 340 –345, 368, 370,387 – 395, 431 – 435, 448 – 468.

The HA active site located in a cleft is composed of the residues 91, 150, 152, 180, 187, 191, and 192. The active site cleft of HA is formed by its right edge (131_GVTAA) and left edge (221_RGQAGR) [19]. The human immune system responds primarily to the five epitope regions, A, B, C, D, and E, of the HA protein in H1N1. **Table 1** presents the 160 amino acids on the five eiptope regions of HA in H1N1 as discovered in [13].

3. RESULTS

3.1. Unusually High Entropy Activities of HA in 2009 H1N1

HA is the primary target for neutralizing antibodies, and the gradual accumulation of substitutions at the antibody sites of HA is the main cause for flu virus to resist human immunity. As entropy measures the disorder of amino acid frequency at each residue of HA, we sought to compare the entropy activities of 2009 H1N1 HA with those in the previous years. Due to the rapid spread of the 2009 H1N1 A virus around the world, unusual entropy patterns of its sequences are anticipated. The sequence variation within the 2009 H1N1 strains as reflected by its entropy distribution along with other H1N1 HA sequences from 2005 to 2008 are illustrated in **Figure 1**, where the high entropy activities of 2009 H1N1 HA were observed, especially in the HA1 domain, indicating the 2009 H1N1 strains are under high immune pressure.

3.2. Mutations of HA in 2009 H1N1

To find the sequence variation of HA in 2009 H1N1, three strains (A/California/04/2009(H1N1), A/South Carolina/1/1918(H1N1), and A/Mississippi/UR06-0537/ 2007(H1N1)) were aligned with MAFFT [14] and the resulting multi-sequence alignment was visualized in Jalview [20] (**Figure 1**). There were 118 mutations, 59 of which (50%) were mutations on the five epitopes, implying that HA in particular has a high amino acid substitution rate in its epitope regions. More precisely, 11 mutations were on epitope A, 11 mutations on B, 10 mutations on C, 12 mutations on D, and 15 mutations on E. The detailed distribution of these mutations on the

Table 3. Comparison of amino acid residues of 2009 H1N1 HA near the receptor-binding sites. The numbers in parenthesis indicate the entropy of the amino acids of HA at that position.

Residues/Amino Acids						
Right edge	131	132	133	134	135	
A/California/04/2009	G (0.0)	V (0.06)	T (0.36)	A (0.04)	A (0.36)	
A/Mississippi/UR06-0537/2007	G	V	S	A	S	
A/South Carolina/1/1918	G	V	T	A	A	
Left edge	221	222	223	224	225	226
A/California/04/2009	R (0.0)	D (0.22)	Q (0.05)	E (0.01)	G (0.0)	R (0.0)
A/Mississippi/UR06-0537/2007	R	D	Q	E	G	R
A/South Carolina/1/1918	R	D	Q	A	G	R
Receptor binding	91	150	152	180	187	191	192
A/California/04/2009	Y (0.0)	W (0.0)	V (0.37)	H (0.0)	D (0.18)	L (0.03)	Y (0.0)
A/Mississippi/UR06-0537/2007	Y	W	T	H	D	L	Y
A/South Carolina/1/1918	Y	W	T	H	D	L	Y
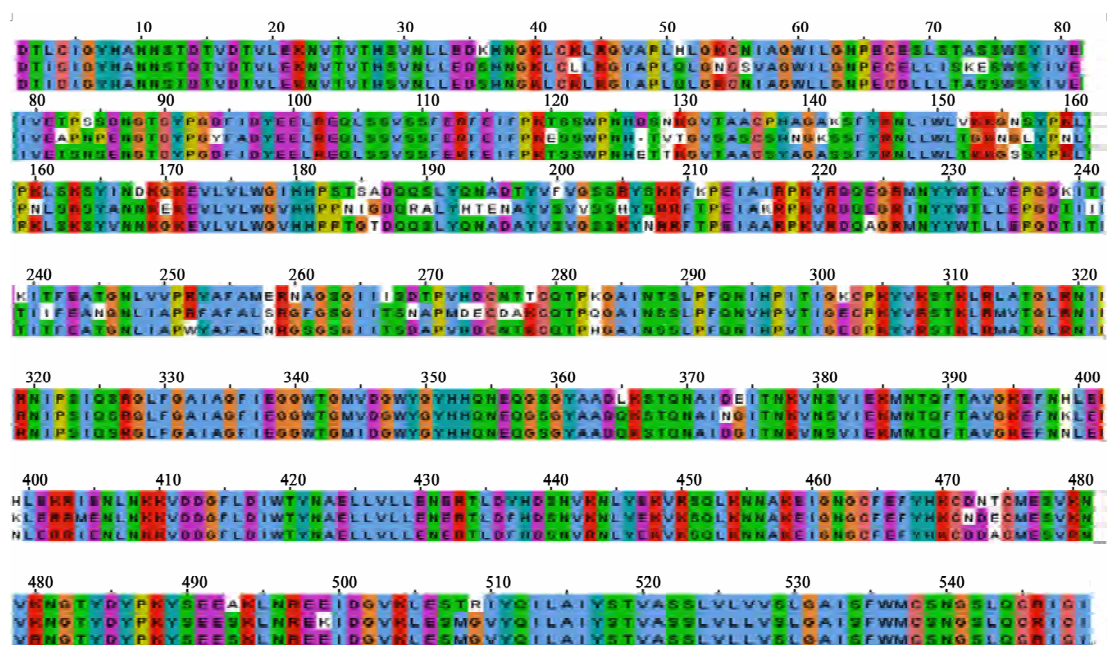
Figure 2. Three sequence alignment using MAFFT and visualized by Jalview: the top sequence is A/California/04/2009(H1N1), the middle one is A/Mississippi/UR06-0537/2007(H1N1), and the bottom one is A/South Carolina/1/1918(H1N1).

five epitopes is in **Table 2**. Epitopes A and B are the dominant ones as they have the highest mutation rate among the five, suggesting that A and B are under the most pressure from the immune system. The information about the dominant epitopes can be used to calculate the Pepitope, a specific measure of antigenic distance between two strains of influenza to estimate the vaccine efficacy [21]. We first displayed the five epitope regions of 2009 H1N1 HA in the homology 3D model built in [13] in **Figure 3** and then mapped these 118 mutations of 2009 H1N1 HA to this 3D model (**Figure 4**).

To learn the sequence variation at or around the active site of 2009 H1N1 HA, we built **Table 3** to show that the amino acids at these sites were highly conserved, which

was in agreement with the previous findings in [22]. There were three residues, 133 and 135 on the left edge and 152 on the active site, that had high entropy and amino acid substitution. In [19] residue 152 was found to allow the substitution for a hydrophobic residue based on an investigation of 191 different sequences of 15 subtypes of HAs, a discovery illustrated in our analysis as well (**Table 3**).

A recent study [23] indicated that substitutions F71S, T128S, E302K, M314L in HA1 of 2009 H1N1 are essential for the interaction between swine and humans; and residues 94, 196 and 274 are predicted to be "hot spots" for mutations that may increase infectivity of the virus (**Figure 2**). It also found that the highly conserved
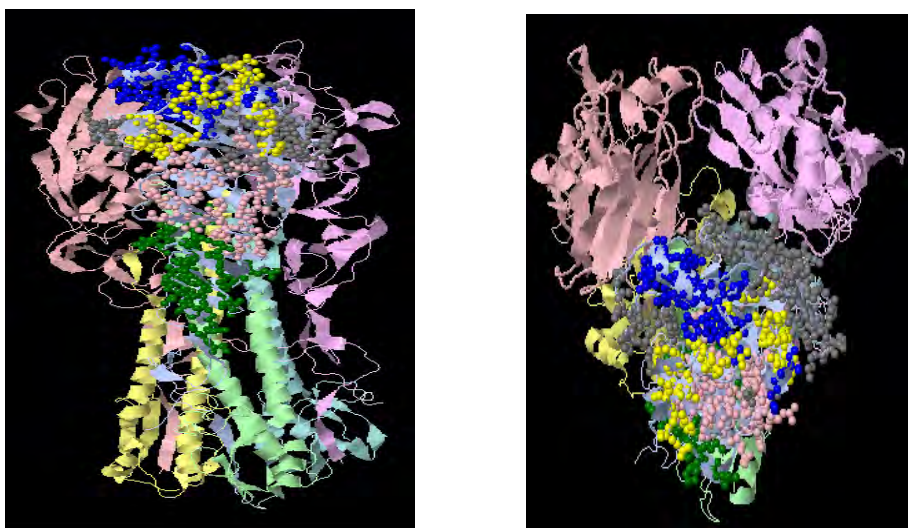
Figure 3. Left plot and right plot display two views of the five epitope regions of H1: A is in yellow, B in blue, C in green, D in grey, and E in pink. The five regions are all shown on one HA monomer within the HA trimer structure (PDB code: 1RU7).
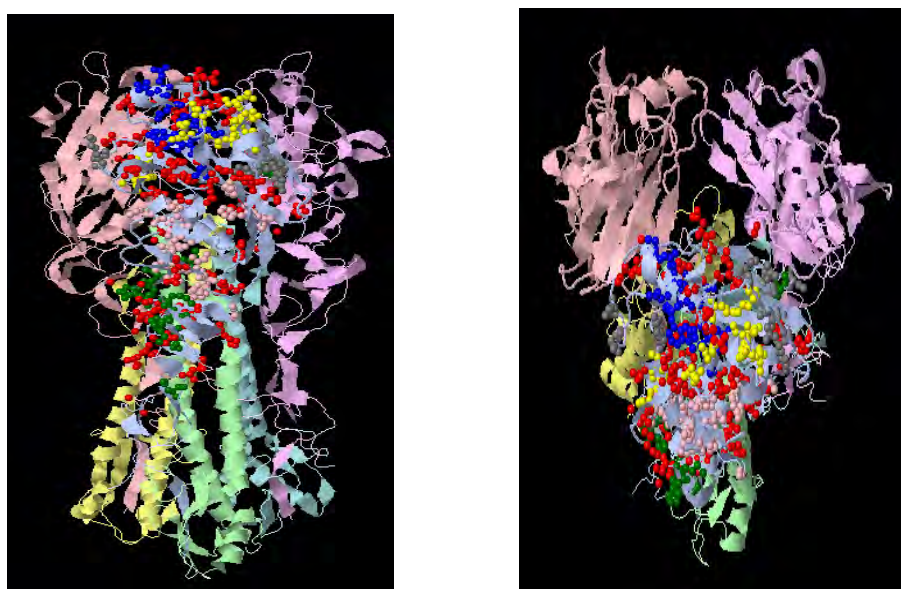
Figure 4. Left and right plots display two views of the 118 mutations of HA in 2009 H1N1: all the mutations are in red except those that are on the five epitope regions are in their own color as in **Figure 3**.

region 286–326 of HA1 is a strong determinant for receptor specificity. As 2009 H1N1 transmits from human to human, additional HA1 sequence variation will likely occur to favor the human interaction.

One of the advantages of using entropy over multi-sequence alignment is that entropy is able to measure sequence variation before a mutation actually occurs. Having knowledge of the potential mutation sites may help us take actions preventatively. Residues 35, 203,310,321, and 416 were not mutations yet, but they had high entropy. Furthermore, residues 35 and 310 were

on epitope C, demonstrating their importance. Among the top 103 high entropy sites in 2009 H1N1 HA in **Figure 5**, there were 91 sites in HA1 domain, 82 of which (90%) were on the five epitope regions, illustrating the high mutational propensity to escape human immune response in these regions.

The five epitopes in the HA1 domain were covered with high entropy sites and there were more of these sites in the HA1 domain than in the HA2 domain (**Figure 5**). In general, the HA1 polypeptide containing the receptor-binding sites and major epitopes is the anti-
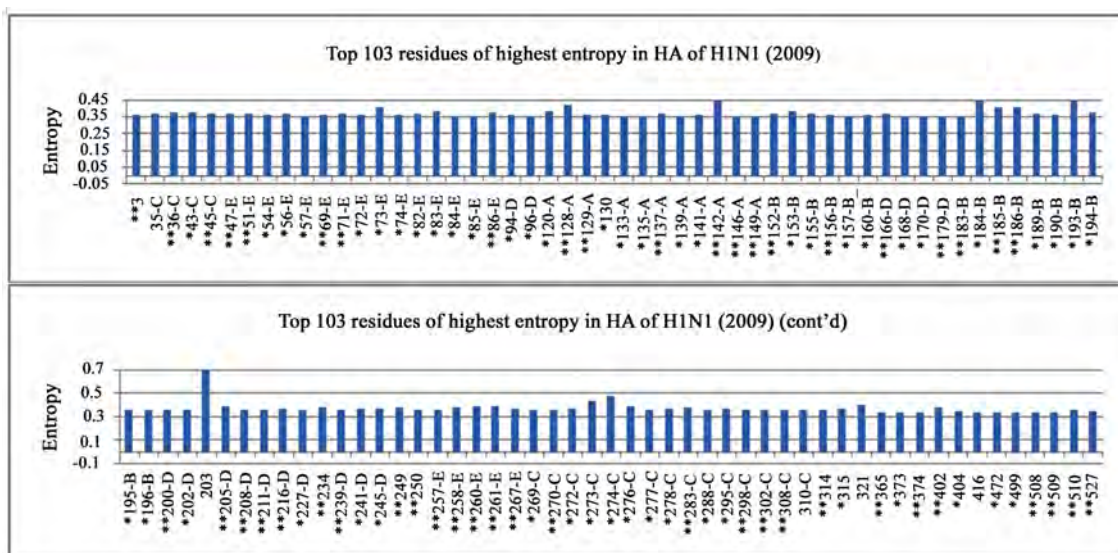
Figure 5. Two plots show the top 103 residues of highest entropy in HA of 2009 H1N1. Residues that had one different amino acid than the two reference strains in **Figure 2** were marked with one asterisk, and those that had two different amino acids were marked with two asterisks. A corresponding letter of A, B, C, D, and E was appended to a residue if it was on one of the five epitope regions.

genically variable region of HA, whereas HA2 is a relatively conserved part of HA. This difference is due to their functions in the HA molecule. HA1 is responsible for the immune response of the host, while HA2 anchors the whole structure in the virus membrane.

3.3. Co-mutations of HA in 2009 H1N1

Inter-residue interactions in proteins are commonly reflected by mutations at one site that compensate for mutations at another site. Simultaneous mutations at antigenic sites collectively enhance antigenic drift in addition to the single mutations. To find the co-mutation pairs, we calculated the mutual information of each possible residue pairs from 548 residues of HA in 2009 H1N1. The 40 top pairs (0.026%) in HA were selected out of 149878 pairs, and all of them had a P value of zero. Four networks of co-mutations were identified. The first one was composed of residues 269, 276, and 309, which were all on epitope C. The second one had residues 34, 167, 195, and 268. The third one had residues 129, 210, and 238. All these three networks had one interesting feature, namely, that each one residue in the network co-mutated with all the others in the network. The fourth one had residues 297, 56, 178, 303, and 509, where residue 297 co-mutated with all the others and residues 56, 178, 303, and 509 co-mutated with each other. In the above four co-mutation networks, there were several antigenic sites: 34, 56, 129, 167, 195, 210, 269, 276, 297, 303, and 309, indicating a selective advantage for novel amino acid sequences among the antigenic regions. As in the single mutation case, most co-mutation pairs in **Table 4** were in the HA1 domain.

3.4. Interaction between HA and NA in 2009 H1N1

HA and NA depend on each other for efficient virus exit from and entry into cells, since there must be a balance between HA activity (binding to sialic acid) and NA activity (removing sialic acid). Such balance could be impaired under various circumstances such as transmission to a new host, reassortment, or therapeutic intervention. A previous report [24] found that a non-optimal combination of HA and NA in a reassortant may be overcome by specific mutations in HA.

In [6], we categorized the NA stalk motifs in H1N1 and H5N1 in 2007, 2008, and 2009. To continue our studies on NA stalk motifs, we aimed to investigate the impact of the length of the NA stalk motifs on HA. To this end, the pairs of HA and NA sequences from the same patient in 2009 were collected and each pair of HA and NA sequences were concatenated to form a single sequence of length 1017, where the HA sequences had a length of 548 and the NA sequences had a length of 469. There were 144 such sequences assembled and then divided into three categories. The first category (n=88) had full-length stalk motifs, the second (n=39) had partially deleted stalk motifs, and the third (n=17) had deleted stalk motifs. It turned out that the HA sequences in the first category had high entropy in both the HA1 and HA2 domains. The HA sequences in the second category had high entropy only in the HA1 domain, and the HA sequences in the third category had high entropy only in the HA2 domain (**Figure 6**). These three distinct entropy responses from HA to the changes of NA motifs pro-

Table 4. Top 40 pairs of co-mutations in HA of 2009 H1N1. All have a P value of zero. A corresponding letter of A, B, C, D, and E was appended to a residue if it was on one of the five epitope regions.

(34-C,167-D)	(34-C,195-B)	(34-C,268)	(44-C,449)	(46,248)	(56-E,297-C)	(68-E,156-B)	(71-E,132-A)
(71-E,287)	(81,277-C)	(84-E,182)	(95-D,165-A)	(127-A,189-B)	(129-A,210-D)	(129-A,238)	(132-A,287)
(134-A,492)	(145,207-D)	(145,307-C)	(155-B,259-E)	(167-D,195-B)	(167-D,268)	(167-D,471)	(178,294)
(178,297-C)	(178,509)	(195-B,268)	(195-B,372)	(195-B,471)	(167-D,471)	(207-D,307-C)	(207-D,401)
(210-D,238)	(268,372)	(269-C,276-C)	(269-C,309-C)	(276-C,309-C)	(294,297-C)	(297-C,509)	(307-C,401)

Table 5. Top 53 pairs of co-mutations in between HA and NA of 2009 H1N1. All have a P value of zero.

(46,13)	(248,13)	(320,13)	(226,15)	(269,16)	(276,16)	(309,16)	(226,19)	(55,21)	(403,23)
(95,34)	(165,34)	(55,42)	(373,47)	(240,48)	(55,59)	(314,75)	(44,173)	(155,189)	(259,189)
(44,220)	(314,232)	(271,234)	(46,241)	(248,241)	(44,257)	(56,263)	(178,263)	(297,263)	(509,263)
(151,264)	(46,257)	(256,269)	(56,288)	(178,288)	(297,288)	(509,288)	(507,289)	(56,321)	(178,321)
(297,321)	(509,321)	(154,336)	(275,341)	(472,341)	(154,382)	(55,385)	(509,389)	(209,427)	(207,432)
(307,432)	(401,432)	(249,453)							

vided another support of the notion that HA and NA need to maintain a functional balance.

In addition to the study of the impact of the NA stalk motifs on HA, we also discovered co-mutation pairs, one mutation in HA and one in NA, that correlated each other across the two proteins. We calculated the mutual information of each possible residue pairs from 1017 residues of HA and NA (as a single sequence) in 2009 H1N1. The top 57 pairs (0.011%) were selected out of 516636 pairs in the sequences of HA and NA, and they all had a P value of zero. Discarding the pairs in HA or NA, and retaining only those with one residue in HA and one residue in NA resulted in 53 pairs (**Table 5**). Among the pairs found, there were two co-mutation networks, one in HA consisting of residues 56, 178, 297, and 509 and one in NA consisting of residues 263, 288, and 321, where each residue in the network co-mutated with all residues in the other network.

NA residues 263 and 321 were a part of a co-mutation network in NA consisting of residues 149, 263, 321, and 389 discovered in [6], which had a property that each of the residues in this network co-mutated with all the other three. NA residues 263 and 288 were also antigenic sites in NA. NA residue 288 was located in a cluster of mutation sites consisting of residues 285, 286, 287, 288, and 289 in NA. The fact that NA residue 288 was part of a NA network of residues 263, 288, and 321 that co-mutated with a HA network of residues 56, 178, 297, and 509 suggested there might be a link between the mutation cluster in NA near residue 288 and the co-mutation network of residues 149, 263, 321, and 389 found in [6]. This conclusion could not be inferred if we were only using the information from NA along to study NA.

3.5. Phylogenetic Analysis of HA in 2009 H1N1

The HA and NA genes of 2009 H1N1 are in the classical

swine lineage and the Eurasian swine genetic lineage respectively [25]. In [6], it was shown that the eight representatives of the novel NA sequences in 2009 H1N1 were diverse enough to cover the major branches of the phylogenetic tree of past NA strains. With more HA sequences available to date than in May or June 2009, we constructed the phylogenetic tree of the representative HA sequences from 1918 to 2009 with the neighbor-joining method using MEGA 4 software [26]. Seven HA sequences in 2009 were selected using cd-hit with identity removed at 98.5% from all the HA sequences in 2009. HA sequences of 29 were selected using cd-hit with identity removed at 98% from all the HA sequences in the years prior to 2009. In contrast to the diversity of NA sequences in 2009 H1N1 [6], the seven representative HA sequences were mainly clustered together in the phylogenetic tree in **Figure 7**, which implied that the HA sequences had remained the same diversity as they were on 16 May 2009 when a similar phylogenetic tree was constructed from a collection of HA sequences in [25].

As demonstrated in the study in [6], Random Forest-based clustering can reveal some subtle features of sequence clusters that a phylogenetic tree built with the neighbor-joining method cannot. We attempted to employ the Random Forest-based clustering technique to cluster the same sequences used in **Figure 7** and the results are in **Figure 8**. Due to the space limitation, we used a number from 1 to 36 to represent a sequence in **Figure 8**, where the same number is attached to the start of the corresponding sequence name in **Figure 7**. The numbers 1 to 7 were assigned to the seven representative HA sequences in 2009 H1N1. In **Figure 8**, the sequences numbered 1, 2, 4, 5, 6, and 7 were close in the second scaling coordinate, and those numbered 1, 4, 5, and 7 were also close in the first scaling coordinate. This detailed clustering information about the seven sequences
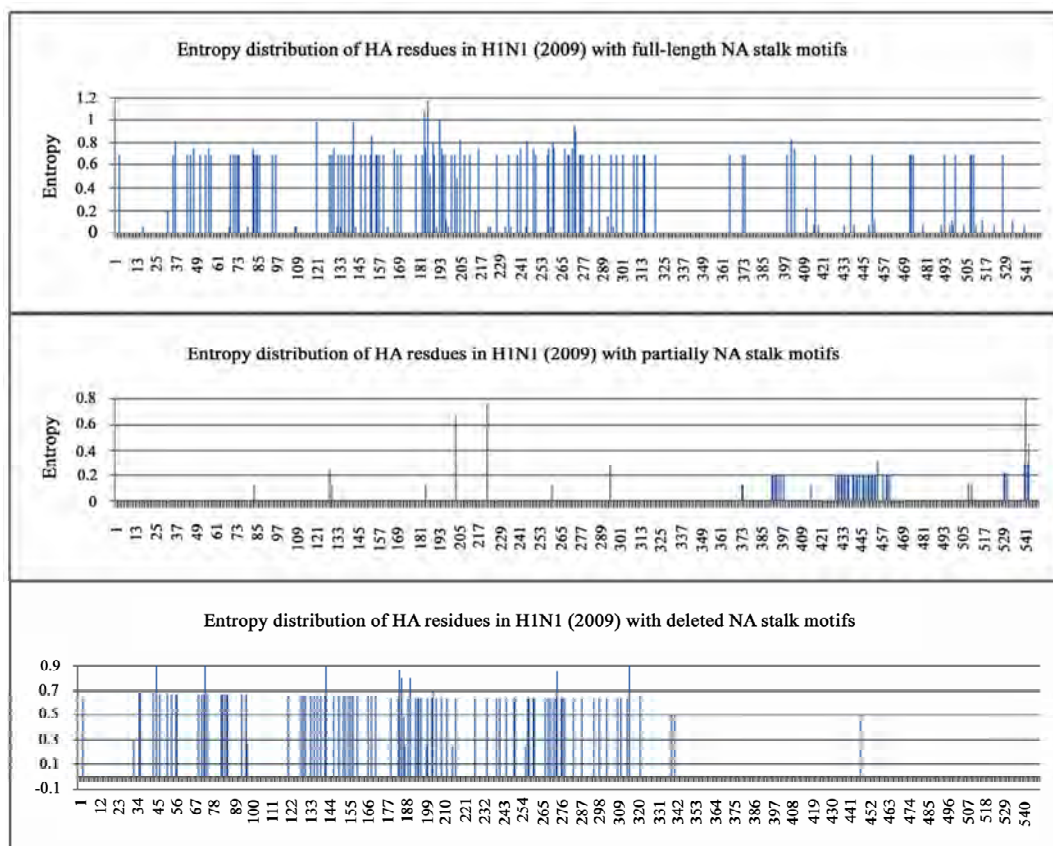
Figure 6. Three plots show the three distinct responses from the HA of 2009 H1N1 to the changes of NA stalk motifs.

provides another view of the current diversity of the HA sequences in 2009 H1N1 relative to the past HA sequences besides the view from the phylogenetic tree in **Figure 7.**

4. DISCUSSION

In this study, we focused on the single mutations and co-mutations in HA of 2009 H1N1. There was extensive research on mutations in H3N2. Studies on H3N2 found that changes at HA residues 183, 186 and 226 could influence HA receptor-binding affinity [27], and that residues 131, 222, 225 and 226 are vital for efficient replication [28]. In [29] 209 complete genomes of the human influenza A virus from 1998 to 2004 were sequenced, and mutations and co-mutations were identified in all the genes in H3N2. Nucleotide co-occurrence networks were constructed in [30] using genome sequences of 1032 H3N2 isolates from 1968 to 2006, and another recent study found co-mutated positions in HA of H3N2 for predicting the antigenic variants using entropy and mutual information [31].

All the flu drugs currently on the market are NA inhibitors; as a result, emergence of resistance mutations in NA could decrease drug effectiveness. In light of the

steady increasing of NA-inhibitor resistant flu strains each year, a new study was conducted in [32] to design a flu drug that can target both HA and NA. For this type of drug, the mutations in HA as well as those in NA will have a direct impact on its outcome. When NA activity is decreased due to NA-inhibitor drug selection, HA mutations to lower the HA receptor-binding affinity are frequently observed. Conversely viruses with reduced HA binding efficiency require less NA activity [5]. Identifying the co-mutations in HA and NA can benefit the design and administration of this type of new drugs.

In our study, HA1 and HA2 displayed different entropy distributions. In general, HA1 had higher entropy than HA2, implying that HA1 is the main responder to the host immunity. This fact should not diminish the value of HA2 being a potential target for vaccine design. The antibodies recognizing HA1 can neutralize virus infectivity, but do not cross-react to the HAs of other subtypes of influenza. One report [33] found that antibodies induced by HA2 are cross-reactive among different subtypes and may moderate virus infection. This result supported the notion that identifying the mutations in HA2 are as important as identifying those in HA1, even though HA1 was the focus of vaccine design in the past.
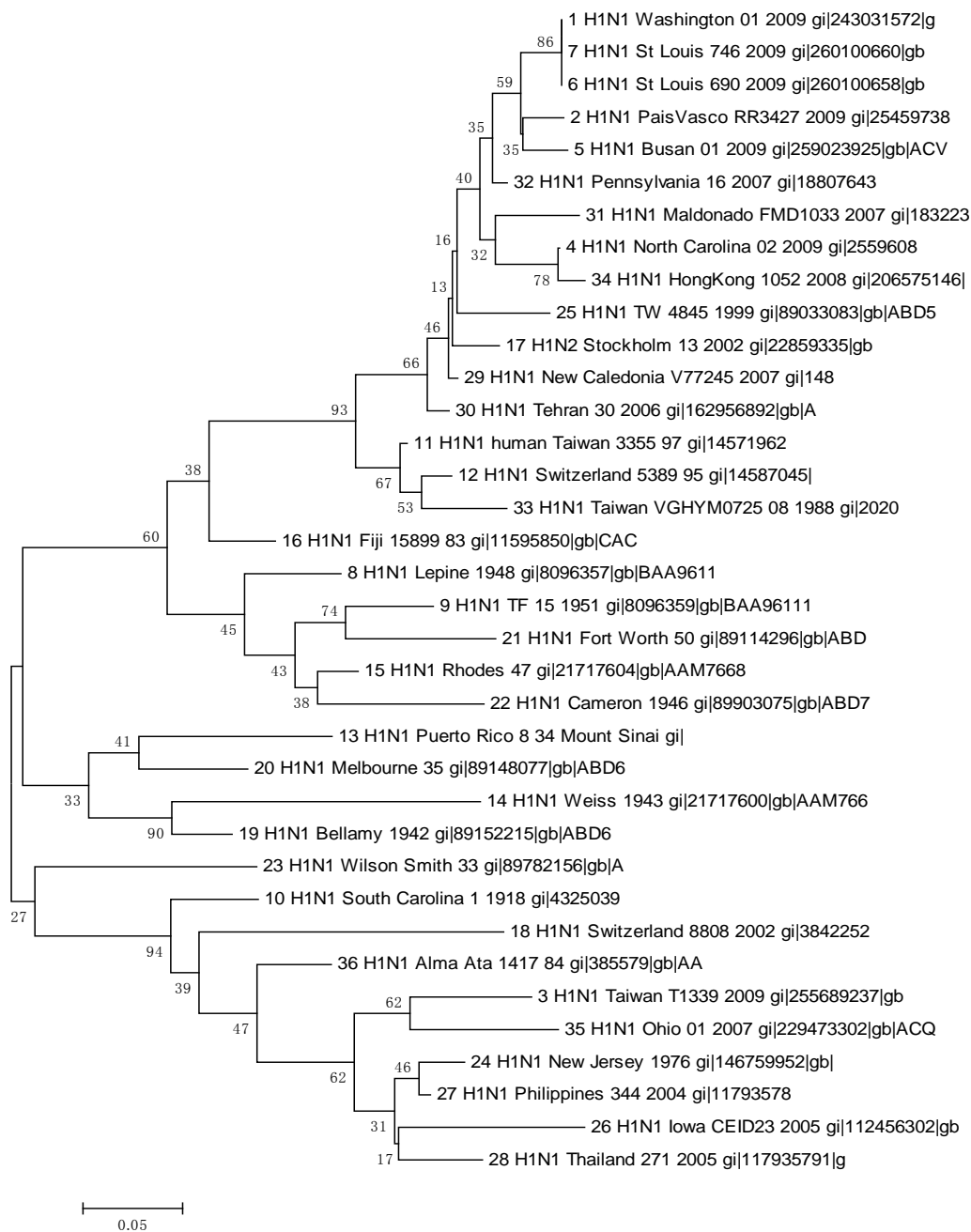
Figure 7. Phylogenetic tree of the HA protein sequences of the H1 subtype family.

5. CONCLUSIONS

There is a great interest in gaining more understanding of the 2009 H1N1 virus given the urgency of the current 2009 flu pandemic. In our previous study on 2009 H1N1 [6], three aspects of NA were investigated: the mutations and co-mutations, the stalk motifs, and the phylogenetic analysis. In this study, we focused on HA and the interaction between HA and NA. The 118 mutations of 2009 H1N1 HA were uncovered and mapped to the 3D ho-

mology model of H1, and the mutations on the five epitope regions on H1 were identified. This information is essential for the development of new drugs and vaccine. With entropy and mutual information analysis, we were able to locate several antigenic sites in HA that could potentially become mutatienal sites. In addition to the identification of single mutations and co-mutations in 2009 H1N1 HA, we were also able with help from our previous results in [6] to find two co-mutation networks, one in HA and one in NA, where each mutation in one
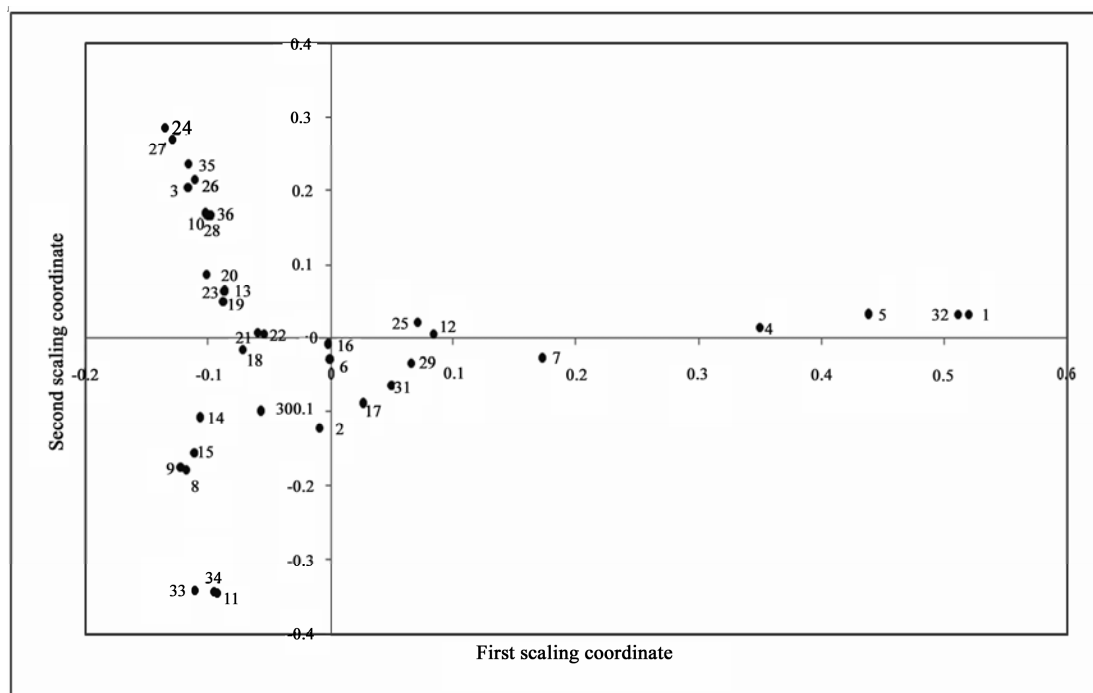
Figure 8. Random Forest-based clusters of the same HA sequences used in **Figure 7**. Here a number from 1 to 36 is used to represent a sequence, which is attached to the start of the corresponding sequence name in **Figure 7**.

network co-mutates with the mutations in the other network across the two proteins HA and NA. These two networks residing in HA and NA separately may provide a link between the mutations that can influence the drug binding sites in NA and those that can affect the host immune response or vaccine efficacy in HA. The distinct entropy responses from HA to the changes of NA stalk motifs were discovered, suggesting the functional dependence between them. Finally, our phylogenetic analysis indicated that the seven representative sequences of HA in 2009 H1N1 were mainly clustered together in the phylogenetic tree made of past representative HA sequences, quite contrary to the NA case [6]. The phylogenetic tree in **Figure 7** was similar in structure to the phylogenetic tree constructed from a collection of HA sequences of 2009 H1N1 on 16 May 2009 in [25], which implied that the diversity of 2009 H1N1 HA sequences remained relatively the same. This view is also supported by the clusters of the same HA sequences created with the Random Forest-based clustering technique (**Figure 8**). Taken together, our results highlighted the importance of conducting timely analysis on the 2009 H1N1 virus and of the integrated approach to studying both surface proteins HA and NA together to reveal their interdependence, which could not be accomplished by studying them individually.

6. ACKNOWLEDGMENTS

REFERENCES

[1] Castrucci, M.R. and Kawaoka, Y. (1993) Biologic importance of neuraminidase stalk length in influenza A virus. *J Virol*, **67**, 759-764.

[2] Els, M.C., Air, G.M., Murti, K.G. *et al*. (1985) An 18-amino acid deletion in an influenza neuraminidase. Virology, **142**, 241-247.

[3] Zhou, H.B., Yu, Z.J., Hu, Y. Tu, J.G. *et al*. (2009) The special neuraminidase stalk-motif responsible for increased virulence and pathogenesis of H5N1 influenza A virus. *PLoS One*, **4(7)**, 6277.

[4] Wagner, R., Matrosovich, M. and Klenk, H.D. (2002) Functional balance between haemagglutinin and neuraminidase in influenza virus infections. Rev. Med. Virol, **12**, 159-166.

[5] Lu, B., Zhou, H.L., Ye, D., Kemble, G. and Jin, H. (2005) Improvement of influenza A/Fujian/411/02 (H3N2) virus growth in embryonated chicken eggs by balancing the hemagglutinin and neuraminidase activities, using reverse genetics. *Journal of Virology*, **79**, 6763-6771.

[6] Hu, W. (2009) Analysis of correlated mutations, stalk motifs, and phylogenetic relationship of the 2009 influenza A virus neuraminidase sequences. *Journal of Biomedical Science and Engineering*, **2**, 550-555

[7] Sebastian, M.S., Ma, J.M., Raphael, T.C.L., Fernanda, L. S. and Frank, E. (2009) Mapping the sequence mutations of the 2009 H1N1 influenza A virus neuraminidase relative to drug and antibody binding sites. *Biol Direct*, **4(18)**.

[8]   Laurel, G., James, S., Dmitriy Z. *et al*. (2005) A single
       amino acid substitution in 1918 influenza virus hemag-
       glutinin changes receptor binding specificity. *Journal of
       Virology,* **79**, 11533-11536.

[9]   Skehel, J.J., Stevens, D.J., Daniels, R.S., Douglas, A.R.,
       Knossow, M. *et al*. (1984) A carbohydrate side chain on
       hemagglutinins of Hong Kong influenza viruses inhibits
       recognition by a monoclonal antibody. *Proc Natl Acad
       Sci U S A*, **81**, 1779-1783.

[10] Wiley, D.C., Wilson, I.A. and Skehel, J.J. (1981) Struc-
       tural identification of the antibody-binding sites of Hong
       Kong influenza haemagglutinin and their involvement in
       antigenic variation. *Nature*, **289**, 373-378.

[11] Caton, A.J., Brownlee, G.G., Yewdell, J.W. and Gerhard,
       W. (1982) The antigenic structure of the influenza virus
       A/PR/8/34 hemagglutinin (H1 subtype). *Cell*, **31**, 417-27.

[12] Tsuchiya, E., Sugawara, K., Hongo, S., Matsuzaki, Y.,
       Muraki, Y. *et al*. (2001) Antigenic structure of the hae-
       magglutinin of human influenza A/H2N2 virus. *J Gen
       Virol*, **82**, 2475-2484.

[13] Michael, W.D. and Pan, K.Y. (2009) The epitope regions
       of H1-subtype influenza A, with application to vaccine
       efficacy. *Protein Engineering, Design & Selection*, **22**,
       543-546.

[14] Katoh, K., Kuma, K., Toh, H., Miyata, T. (2005)
       MAFFT version 5: improvement in accuracy of multiple
       sequence alignment. *Nucleic Acids Res*, **33**, 511-518.

[15] David, M. (2003) Information theory, inference, and
       learning algorithms. Cambridge University Press.

[16] Breiman, L. (2001) Random forests, machine learning,
       45 (1), 5-32.

[17] Cox, T.F. and Cox, M.A.A. (2001), Multidimensional
       scaling, chapman and hall.

[18] Colman, P.M., Hoyne, P.A. and Lawrence, M.C. (1993)
       Sequence and structure alignment of paramyxovirus he-
       magglutinin-neuraminidase with influenza virus neura-
       minidase. *J. Virol*, **67**, 2972-2980.

[19] Andrea, K., Gabriel, R.N. and Ivan, K.H., Sccarontefan, J.
       (2002) Sequence similarities and evolutionary relation-
       ships of influenza virus A hemagglutinins, *Virus Genes*,
       24, 57-63.

[20] Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp,
       M and Barton, G.J. (2009) Jalview version 2 – A multi-
       ple sequence alignment editor and analysis workbench.
       Bioinformatics, **5**, 1189-1191.

[21] Enrique, T. and Mu˜noz, M.W.D. (2005) Epitope analy-
       sis for influenza vaccine design. *Vaccine*, **23**, 1144-1148.

[22] Weis, W., Brown, J.H., Cusack, S., Paulson, J.C., Skehel,
       J.J. and Wiley, D.C. (1988) Structure of the influenza

       virus haemagglutinin complexed with its receptor, sialic
       acid. *Nature*, **333(6172)**, 426-31.

[23] Veljko, V., Henry, L.N., Sanja G. *et al*. (2009) Identifi-
       cation of hemagglutinin structural domain and polymor-
       phisms which may modulate swine H1N1 interactions
       with human receptor. *BMC Structural Biology*, **9(62)**.

[24] Nikolai, V.K., Mikhail, N.M. and Aleksandra, S.G. (2000)
       Intergenic HA–NA interactions in influenza A virus:
       postreassortment substitutions of charged amino acid in
       the hemagglutinin of different subtypes. *Virus Research*,
       66,123-129.

[25] Garten, R.J., Davis, C.T., Russell, C.A. *et al*. (2009)
       Antigenic and genetic characteristics of swine-origin
       2009 A (H1N1) influenza viruses circulating in humans.
       Science, **325**, 197-201.

[26] Kumar, S., Nei, M., Dudley, J. and Tamura. K., (2008)
       MEGA: a biologist-centric software for evolutionary
       analysis of DNA and protein sequences. *Brief Bioinfor-
       matics*, **9**, 299-306.

[27] Lu, B., Zhou, H., Ye, D., Kemble, G. and Jin, H., (2005)
       Improvement of influenza A/Fujian/411/02 (H3N2) virus
       growth in embryonated chicken eggs by balancing the
       hemagglutinin and neuraminidase activities, using re-
       verse genetics. *J. Virol*, **79**, 6763-6771.

[28] Jin, H., Zhou, H., Liu, H., Chan, W.N., Adhikary, L. *et al.*
       (2005) Two residues in the hemagglutinin of A/Fujian/
       411/02-like influenza viruses are responsible for antigenic
       drift from A/Panama/2007/99. *Virology*, **336**, 113- 119.

[29] Elodie G., Naomi A.S., Martin S. *et al*. (2005) Large- scale
       sequencing of human influenza reveals the dynamic nature
       of viral genome evolution. Nature, **437**, 1162-1166.

[30] Du, X.J., Wang, Z., Wu, A.P., Song, L., Cao, Y., Hang,
       H.Y. and Jiang, T.J. (2008) Networks of genomic
       co-occurrence capture characteristics of human influenza
       A (H3N2) evolution. *Genome Res*, **18**, 178-187.

[31] Huang, J.W., King, C.C. and Yang, J.M. (2009)
       Co-evolution positions and rules for antigenic variants of
       human influenza A/H3N2 viruses. *BMC Bioinformatics*,
       10(Suppl 1), S41.

[32] Michel, W., Chen, C.C., Kemp, M.M. and Linhard, R.J.
       (2009) Synthesis and biological evaluation of non- hy-
       drolyzable 1,2,3-triazole-linked sialic acid derivatives as
       neuraminidase inhibitors. *European Journal of Organic
       Chemistry*, **2009(16)**, 2587.

[33] Gocnı´k, M., Fislova´, T., Mucha, V., Sla´dkova´, T. *et al.*
       (2008) Antibodies induced by the HA2 glycopolypeptide
       of influenza virus haemagglutinin improve recovery from
       influenza A virus infection. *Journal of General Virology*,
       89, 958-967.

A new projection method for biological semantic map generation

Hoan N. Nguyen, Nicolas Wicker, David Kieffer, Olivier Poch

Laboratoire de bioinformatique et génomique intégratives, Institut de Génétique et de Biologie Moléculaire et Cellulaire, France
University of Strasbourg, Illkirch Cedex, France.
Email: nguyen@igbmc.fr

ABSTRACT

Low-dimensional representation is a convenient method of obtaining a synthetic view of complex datasets and has been used in various domains for a long time. When the representation is related to words in a document, this kind of representation is also called a semantic map. The two most popular methods are self-organizing maps and generative topographic mapping. The second approach is statistically well-founded but far less computationally efficient than the first. On the other hand, a drawback of self-organizing maps is that they do not project all points, but only map nodes. This paper presents a method of obtaining the projections for all data points complementary to the self-organizing map nodes. The idea is to project points so that their initial distances to some cluster centers are as conserved as possible. The method is tested on an oil flow dataset and then applied to a large protein sequence dataset described by keywords. It has been integrated into an interactive data browser for biological databases.

Keywords: Semantic Map; Dimension Reduction; Biological Database; SOM

1. INTRODUCTION

Thanks to the availability of the human and other genomes and the rapid progress of biotechnologies and information technologies, numerous large biomedical datasets have been generated. Modern biomedical information thus corresponds to a high volume of heterogeneous data that doubles in size every year and that covers very different data types, including phenotypic data, genotypic data as well as standards, processes, protocols or treatments used to generate information from raw data. In this context, systemic approaches are now needed to store, analyze and compare the huge amount of relevant information.

The first and second author contribute equally to this paper.

In addition, the knowledge provided by classical query services on biological data is often unsatisfactory (e.g. a list of proteins or sequences) and there is a need for user-friendly visual representations of the data. Such a representation exists and is called a feature or semantic map. It is used to visualize "land maps" in two or three dimensions that represent, for example, the distribution (similarity and neighborhood) of protein annotations in biological databases. When query results are represented on the map, the repartition of the proteins can be easily observed, as well as their proximity to clusters labeled according to their content. In addition, it is straight-forward to superpose the information obtained from additional requests. Thus, a semantic map can greatly facilitate the interpretation of results from large scale data analyses. To quote a few examples, semantic maps have already been used in fluid mechanics [1], astronomy [2], internet data mining [3,4], scientific literature mining [5] and biology [6].

Many low-dimensional methods have been devised [5,7,8,9] and two of the most popular are the WEBSOM method [9] and the Generative Topographic Mapping (GTM) [1]. These two methods are briefly outlined below.

WEBSOM originates from self-organizing maps [10] which is a classification algorithm where nodes move towards cluster centers. In WEBSOM, the nodes are fixed on a two-dimensional grid and at the same time live in the space of the dataset, typically a R^p space. First, a point y is picked at random from the dataset. Next, the closest node w_i in R^p is selected and then each node w_j moves towards y according to the equation

$$w_j(t+1) = w_j(t) + \eta(t)h_{ij}(t)\left\|y - w_j(t)\right\|$$

where $\eta(t)$ is the learning rate decreasing in time and $h_{ij}(t)$ is a neighborhood function in the two-dimensional grid. These steps are then iterated for all data points. The initialization of the p-dimensional space can be performed randomly, but a more effective method is to select points

along the two first principal axes of the dataset [4]. Finally, the dataset is used again by assigning each point to its closest node in the p-dimensional space using a Euclidean distance. Then, for each node, the number of points it has captured is taken as its density up to a given scaling factor (the size of the dataset).

The generative topographic map (GTM) [1] is a statistical method which is provably (locally) convergent and which does not require a shrinking neighborhood or a decreasing step size. It is a generative model: the data is assumed to arise by probabilistically picking points in a low-dimensional space and mapping them to the observed high-dimensional input space. The statistical model can be described in the following way:

$$p(y|x_i, W, \beta) = (\frac{\beta}{2\pi})^{p/2} \exp\{-\frac{\beta}{2}\|W.\varphi(x_i) - y\|^2\}$$

where x_i is a two-dimensional grid node, β is a scaling parameter, $W.\varphi(x_i)$ a generalized regression model, W a $p \times m$ matrix and the elements of $\varphi(x)$ consist of m basic functions $\varphi_j(x)$ typically equal to radially symmetric Gaussians centered on the nodes of a two-dimensional grid. The parameters W and β of the model are estimated through the expectation-maximization (EM) algorithm [11]. This model can be considered to be the probabilistic counterpart of SOM/WEBSOM. However, the WEBSOM method is quicker than GTM when large amounts of data must be dealt with, especially if the winner selection is optimized so that millions of documents and nodes can be treated [4].

An alternative choice is to follow Flexer's approach [12] which first clusters the points in the data space and then projects cluster centers using Sammon's multidimensional scaling method [13]. However this means that only a subset of points are effectively projected. In this paper, we present a complementary method that projects all points using their distances to the cluster centers.

First this new projection method is presented, then it is evaluated on a benchmark data set and compared to other methods. Finally, it is used in the results section to generate a semantic map in the context of a new integrative navigator for biological databases.

2. METHODS

The principle of the presented method is to project points after they have been clustered and the cluster centers have been projected onto a two-dimensional map. This is done by conserving as much as possible the original distances between the points and the cluster centers. Basically, for each point indexed by i, the two-dimensional coordinates are search such as to minimize the difference between the distances computed in the n-dimensional data space with those computed on the map.

This comes down to finding the point x_i in two dimensions minimizing the following function $E(x_i)$:

$$E(x_i) = \sum_{g=1}^{G} \left(\sum_{k=1}^{2} (x_k - c_{g,k})^2 - d_g^2 \right)^2$$

with d_g denoting the distance between point i and cluster g and $c_{g,k}$ the projection of the kth cluster center. The Newton-Raphson algorithm was used to minimize $E(x_i)$. At each step, $x_i^{t+1} = x_i^t - H^{-1}.\nabla E$ with H the Hessian and ∇E the gradient of E.

$$\nabla E = \begin{pmatrix} \frac{\partial E}{\partial x_1} \\ \frac{\partial E}{\partial x_2} \end{pmatrix}, H = \begin{pmatrix} \frac{\partial^2 E}{\partial x_1^2} & \frac{\partial^2 E}{\partial x_1 \partial x_2} \\ \frac{\partial^2 E}{\partial x_1 \partial x_2} & \frac{\partial^2 E}{\partial x_2^2} \end{pmatrix}$$

The optimizing function is not convex as the Hessian is not always semi-definite positive. To show this, it is sufficient to find a point X verifying $X'HX < 0$. In particular, we show that H_{11} can be negative which is also sufficient. First let us note that

$$\frac{\partial E}{\partial x_l} = \sum_{g=1}^{G} 2 \left(\sum_{k=1}^{2} (x_k - c_{g,k})^2 - d_g^2 \right) 2(x_l - c_{g,l})$$

$$\frac{\partial^2 E}{\partial x_l^2} = 8 \sum_{g=1}^{G} (x_l - c_{g,l})^2 + 4 \sum_{g=1}^{G} \left(\sum_{k=1}^{2} (x_k - c_{g,k})^2 - d_g^2 \right)$$

and then set

$$\begin{cases} x_1 = c_{1,1} = c_{2,1} = c_{3,1} \\ d_1 = \sqrt{(x_1 - c_{1,1})^2 + (x_2 - c_{1,2})^2} \\ d_2 = \sqrt{(x_1 - c_{2,1})^2 + (x_2 - c_{2,2})^2} \\ d_3 > \sqrt{(x_1 - c_{3,1})^2 + (x_2 - c_{3,2})^2} \end{cases}$$

Thus, $H_{11} = 4 \left((x_1 - c_{3,1})^2 + (x_2 - c_{3,2})^2 - d_3^2 \right) < 0$

Consequently, a global optimization process was performed using different initial values. Each cluster center projection was used as an initial value and the best solution after convergence was kept.

3. RESULTS AND DISCUSSION

3.1. Validation Using the Oil Flow Dataset

To validate the new points projection method, a previously established oil flow dataset [14] was used as a benchmark. This training dataset is available at http://www.ncrg.aston.ac.uk/GTM/ and contains 1000
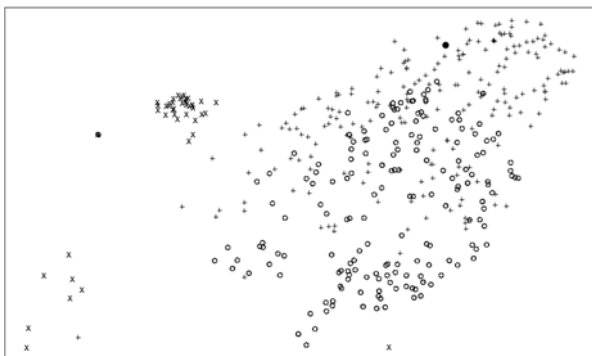
Figure 1. New projection of the dataset. Results of the presented projection on the oil flow dataset. Crosses, circles and plus-signs represent stratified, annular and homogeneous multi-phase configurations respectively. The three group separations are clearly identified.
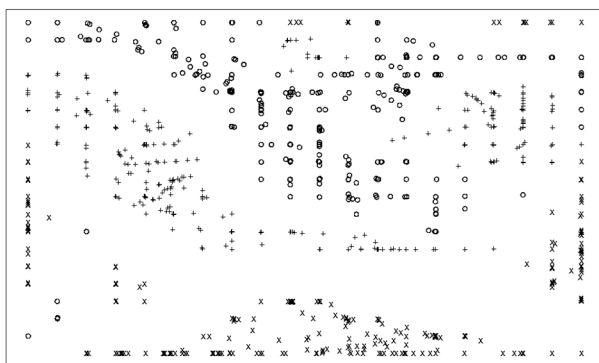
Figure 2. Oil flow dataset after GTM. After projection of the oil flow dataset using the Generative Topographic Mapping, the three group separations are clearly separated, but in a complex way that is far from linear.
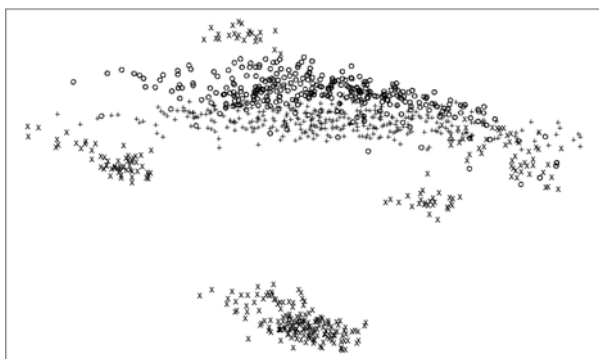
Figure 3. Oil flow dataset after PCA projection. After projection of the oil flow dataset using principal component analysis, the separation of the three groups is not clearly identified. In particular, the crosses are very scattered.

points in 12 dimensions corresponding to 12 measurements on the mixture of oil, water and gas passing through a pipeline. The three phases in the pipe can belong to three different configurations corresponding to laminar, homogeneous and annular flows.

First, the dataset was clustered into 15 clusters and the cluster centers projected according to Sammon's multi-dimensional scaling method [13]. Then the 1000 points were projected in two dimensions using the method described above. The results are shown on **Figure 1**, where it be seen that three different groups are rather well linearly separated. The groups obtained with the GTM and principal component analysis (PCA) methods are shown on **Figures 2** and **3** respectively. In order to objectively measure the quality of these results, we computed the ratio of the between-class inertia and the total inertia for each method. For our method, GTM and the PCA, we obtained a ratio of 0.83, 0.25 and 0.23 respectively, thus confirming the visual impression. Nevertheless, it should be stated that, if only separation is desired and not specifically linear separation, GTM performs better, even though it has the drawback of making the underlying grid very visible.

3.2. Semantic Map Generation for Biological Database

The Laboratory of Genomics and Integrative Bioinformatics (LGBI) at the IGBMC Strasbourg, has developed a new high-performance biomedical information system, called the BIRD System [15,16]. BIRD is able to integrate very quickly heterogeneous data either from the large generalist databases (sequence, structure, function and evolution, etc.) or from specialized databases dedicated to high throughput biology (transcriptomics, interactomic, etc.) in a relational database (IBM DB2). Thus, it allows to organize massive sets of biomedical data according to real world requirements. An original biological query engine, called BIRDQL, has been designed to facilitate access to the heterogeneous databases and to allow pertinent information extraction via a web server. This system has been used in the Decrypthon computing grid [17] in order to provide data to the runtime applications.

To complete the visualization and analyze functionalities of the BIRD System, the new method described above to build semantic maps was integrated in the BIRD query engine (BIRDQL). The maps can be used to explore the data using a combination of high level queries and area selections (**Figure 4**). The method was tested by building a semantic map of the Uniprot database [18] using the keyword descriptions for each protein. After removal of redundant vectors, we obtained 60,000 vectors $z_1, ... z_{60000}$ in a 914-dimensional space corresponding to the 914 keywords extracted from about 6 million proteins. In the following lines, to avoid focusing on the numerical details, we will consider n proteins described by p keywords where n and p stand for 60000 and 914 respectively.

Before projecting the points, some preliminary steps were necessary:
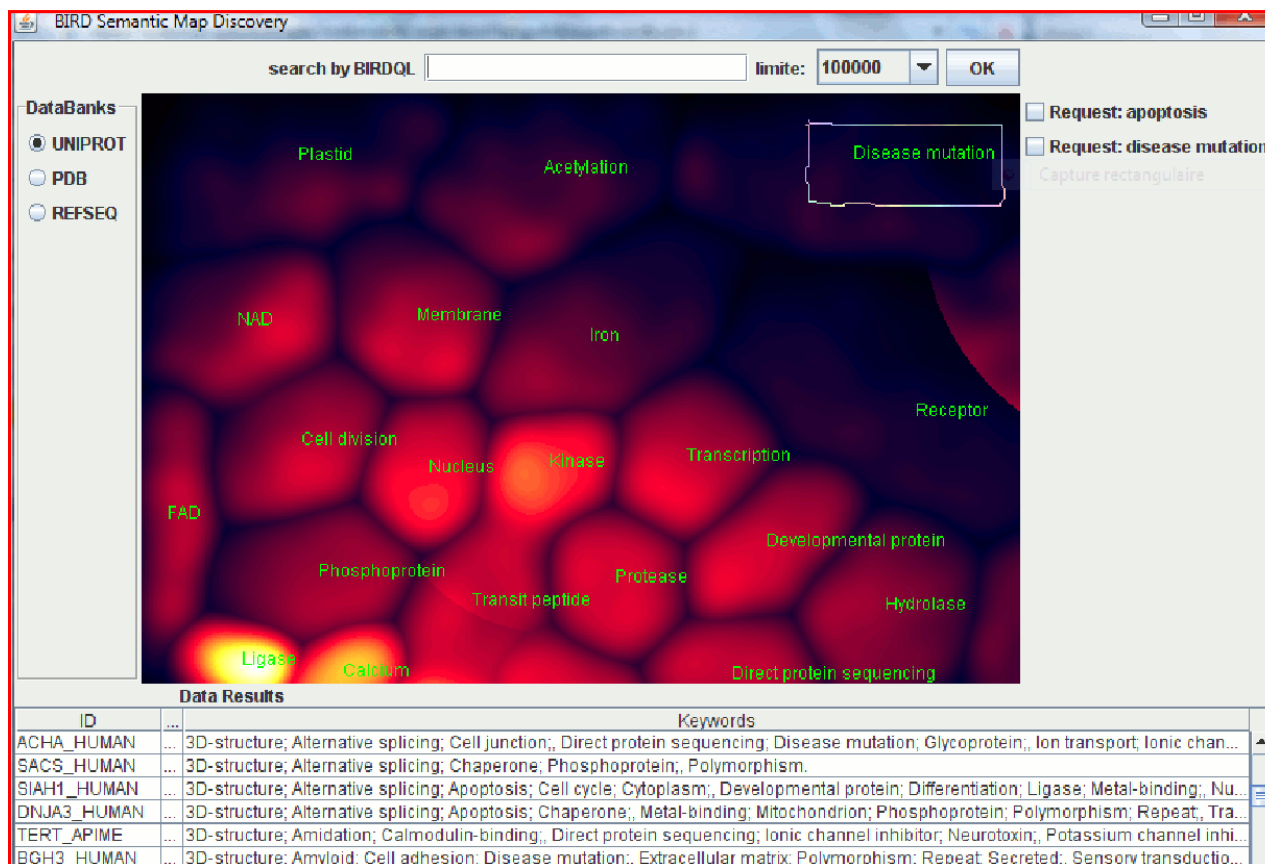
Figure 4. Semantic map with density colours and most frequent keyword labels.

Step 1: dimension reduction

The n proteins were described by p keywords and were thus represented by n points $z_1,...,z_n$ in p dimensions. As in the preprocessing step of WEBSOM [3,4], an initial dimension reduction was performed to reduce p coordinates to p^* using random projection directions. More specifically, random vectors v_j were generated on the p^*-dimensional unit-sphere and then new coordinates were obtained by computing the scalar product $y_{ij} =< v_j, z_i >$ on each document i. Thus, the n proteins were described by n points $y_1...y_n$.

Step 2: mixture models clustering

In a second step, these points were clustered using mixture models. Mixture models are a powerful method to cluster datasets of points described by coordinates. The points are assumed to be independent realizations from a mixture of several distributions. Here the mixture is only briefly described for G components $f_{\alpha_1},...,f_{\alpha_G}$ with parameters $\alpha_1...,\alpha_G$. A general presentation of this method and its applications can be found in [19,20,21,22]. If $\tau_1,...,\tau_G$ indicate the different weights of the components, the likelihood of the model for n points $y_1,...,y_n$ is expressed as:

$$L_M(\tau_1,...,\tau_G,\alpha_1,...\alpha_G \mid y_1,...,y_n) = \prod_{i=1}^{n} \sum_{g=1}^{G} \tau_g f_{\alpha_g}(y_i)$$

The estimation of the different coefficients of the mixture model is commonly performed via the EM (Expectation-Maximization) algorithm of Dempster [11]. Here, in order to simplify the estimation, a variant of the EM algorithm called CEM was used [22]. In this application G was chosen to be equal to 30.

Step 3: cluster centers projection

Once G clusters were obtained, the centers of gravity $c_1^*,...,c_G^*$ were computed in the p-dimensional space. Then, multidimensional scaling (MDS) [23] was applied on the cluster centers to produce two-dimensional coordinates $c_1,...,c_G$. MDS was used because Sammon's method [13] failed on this dataset, since it produced many points with the same coordinates.

After these three preliminary steps, the points were projected on the map using the new projection method. The density $m(x)$ for each point x of the map is given using a kernel method [24]:
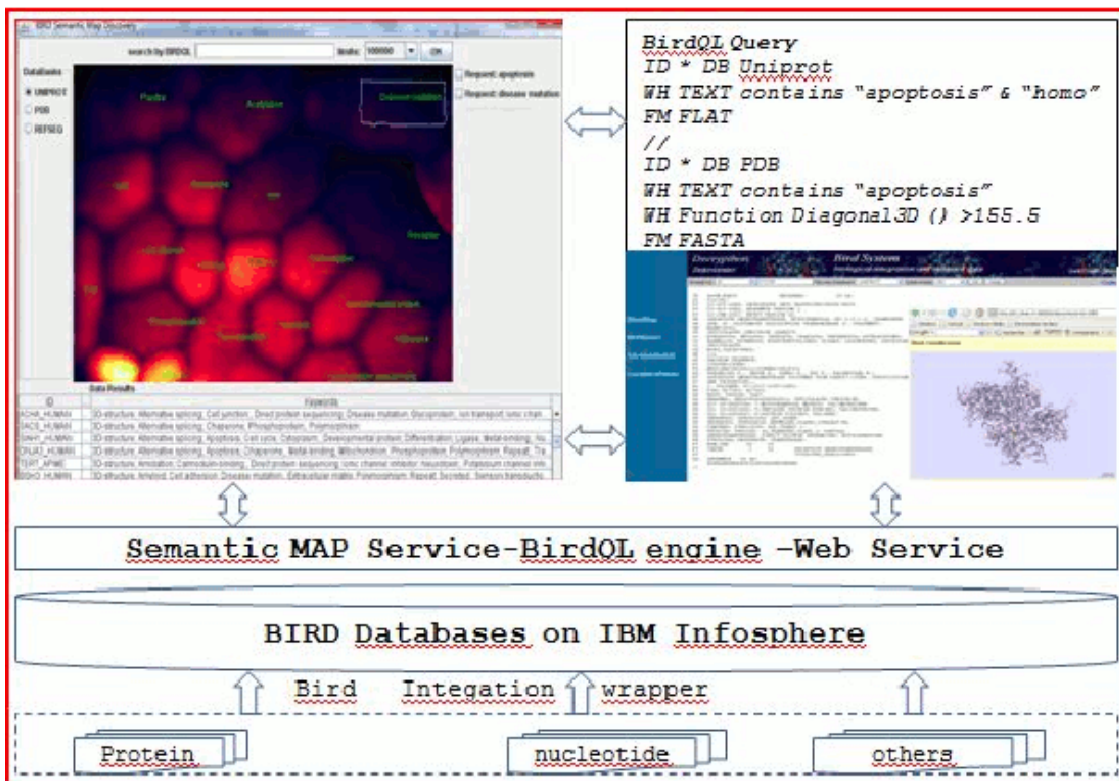
Figure 5. The global architecture of the Semantic Map Discovery prototype coupled with the BIRD System using the BirdQL query engine.
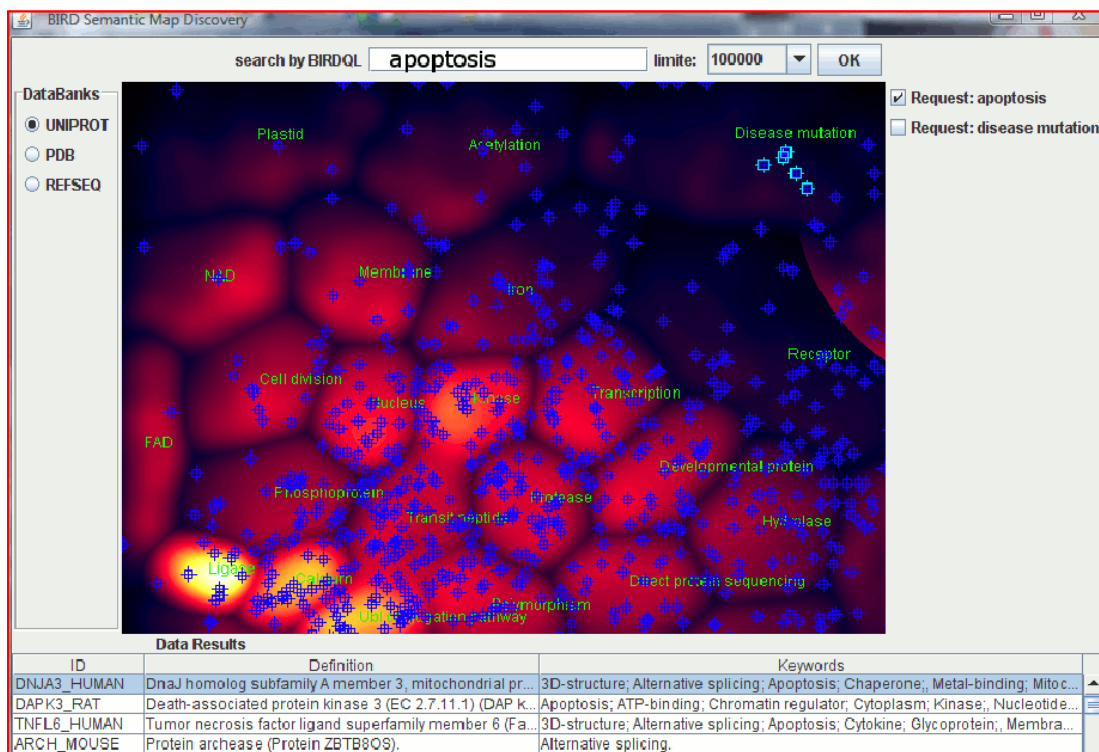
Figure 6. Semantic map with selected proteins. The labels represent the most frequent keywords present inside the cluster points which are not shared between different clusters.

$$m(x)\frac{1}{n}\sum_{i=1}^{n}\frac{1}{2\pi\left|\sum\right|^{1/2}}\exp\{(x-x_i)\sum{}^{-1}(x-x_i)\}$$

Then, a color scale ranging from purple to white, with intermediary colors red, orange and yellow was assigned to each point according to its density. The map is represented in **Figure 4.**

This visual representation allows a global comprehension of the whole database, which is easier to understand than numerical or textual data. Some important keywords shared by many proteins are visible on this map, such as kinase, ligase and protease. At the same time, frequent keywords, such as "complete proteome", that are non-informative, are avoided because they are shared by several clusters. Another observation is that the density is far from being homogeneous, the map being more crowded in the bottom-left corner than elsewhere.

When using the integrated biological query engine BIRD-QL of the BIRD System via a web service or http protocol, as shown in **Figure 5**, the selected proteins are represented on the maps by a plus sign of a given color. If different selections have been performed, different colors are used. An example is shown in **Figure 6**, where proteins selected by a query with the keyword "apoptosis" are shown by blue plus signs. Some of these proteins were selected by the user and are surrounded by a white square. One of the proteins, DNJA3, belongs to the small cluster labeled "disease mutation" but does not possess the "disease mutation" keyword. Interestingly its deficiency implies dilated cardiomyopathy [25] (MIM-608382).

There is still room for improvement in the construction of semantic maps both at the algorithmic level and at the software functionality level. The point's projection is formalized as a global optimization problem and currently, it is resolved simply using different starting points with the Newton-Raphson method. However global optimization methods could also be tested [26,27]. From a practical point of view it would also be useful to determine how many clusters or nodes are necessary to achieve a good projection of the data points.

4. CONCLUSIONS

The main contribution of this work is a new computational solution to the construction of semantic maps. The idea is to project points by locating them according to cluster centers. This method can thus be coupled with other methods such as self-organizing maps or Flexer's approach.

5. ACKNOWLEDGEMENTS

REFERENCES

[1] Bishop, C.M., Svens'en, M. and Williams, C.K.I. (1998) GTM: the generative topographic mapping, *Neural Computation*, **10**, 215-234.

[2] Lesteven, K. (1995) Multivariate data analysis applied to bibliographical information retrieval: SIMBAD quality control. *Vistas in Astronomy*, **39**, 187-193

[3] Kaski, S. (1998) Dimensionality reduction by random mapping: Fast similarity computation for clustering, Proceedings of IJCNN'98, International Joint Conference on Neural Networks, IEEE Service Center, 413-418.

[4] Lagus, K., Kaski, S. and Kohonen, T. (2004) Mining massive document collections by the WEBSOM method. *Information Sciences*, **163**, 135-156.

[5] Chen, C. (2005) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science,* **57**, 359-377.

[6] Grimmelstein, M. and Urfer, W.W. (2005) Analyzing protein data with the generative topographic mapping approach. innovations in classification, data science, and information systems, Baier, D. and Wernecke, K.D. Springer Berlin Heidelberg, 585-592.

[7] Ossorio, P.G, (1966) Classification space: a multivariate procedure for automated document indexing and retrieval. *Multivariate Behavioral Research*, **1**, 479-524.

[8] Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T. K. and Harshman R. (1990) Indexing by latent semantic indexing. *Journal of the American Society for Information Science*, **41**, 391-407.

[9] Kohonen, T. (1997) Self-Organizing Maps, Springer-Verlag.

[10] Kohonen, T. (1982) Analysis of a simple self-organizing process. *Biological Cybernetics*, **44**, 135-140.

[11] Dempster, A., Laird, N. and Rubin, D. (1977) Maximum likelihood from incomplete data via the {EM} algorithm. *Journal of the Royal Statistical Society, Ser. B*, **39**, 249-282.

[12] Flexer, A. (1997) Limitations of self-organizing maps for vector quantization and multi-dimensional scaling. *Advances in neural information processing systems*, **9**, 445-451.

[13] Sammon J.W. (1969) A non-linear mapping for data structure analysis. *IEEE Transactions on Computers*, **18**, 401-409.

[14] Bishop, C.M. and James G.D. (1993) Analysis of multi-phase flows using dual-energy gamma densitometry and neural networks. *Nuclear Instruments and Methods in Physics Research*, Section A, **327**, 580-593.

[15] Nguyen, H., Berthommier, G., Friedrich, A., Poidevin, L., Ripp, R., Moulinier, L. and Poch, O. (2008) Introduction to the new Decrypthon Data Center for biomedical data, *Proc CORIA'*, 32-44.

[16] BIRDQL-Wikili, http://alnitak.u-strasbg.fr/wikili/index.php/BIRDQL.

[17] Décrypthon: le grid-computing au service de la génomique et la protéomique. http://www.decrypthon.fr.

[18] (2008) The uniProt consortium. The Universal Protein Resource (UniProt). *Nucleic Acids Research*, **36**, D190-D195.

[19] Titterington, D., Smith, A. and Makov, U. (1985) Statistical analysis of finite mixture distribution, John Wiley and Sons.

[20] McLachlan, G. and Basford, K. (1988) Mixture models: inference and applications to clustering, Marcel Dekker.

[21] Banfield, J. and Raftery, A. (1993) Model-based Gaussian and non-Gaussian clustering. *Biometrics*, **49**, 803-821.

[22] Celeux, G. and Govaert, G. (1992) A classification EM algorithm for clustering and two stochastic versions, *Journal of Computational Statistics and Data Analysis*, **14**, 315-332.

[23] Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979) Multi-variate Analysis, Academic Press.

[24] Parzen, E., (1962) On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, **33**, 1065-1076.

[25] Hayashi, M., Imanaka-Yoshida, K., Yoshida, T., Wood, M., Fearns, C., Tatake, R. and Lee, J. (2006) A crucial role of mitochondrial Hsp40 in preventing dilated cardiomyopathy, *Nature Medecine*, **12**, 128-132.

[26] Laguna, M. and Marti, R. (2005) Experimental testing of advanced scatter search designs for global optimization of multimodal functions, *Journal of Global Optimization* **33**, 235-255.

[27] Neumaier, A., Shcherbina, O., Huyer, W. and Vinko, T. (2005) A comparison of complete global optimization solvers, *Mathematical Programming*, **103**, 335-356.

Applications of a new *In vivo* tumor spheroid based shell-less chorioallantoic membrane 3-D model in bioengineering research

Nzola De Magalhães[1], Lih-Huei L. Liaw[2], Michael Berns[2], Vittorio Cristini[3], Zhongping Chen[2], Dwayne Stupack[4], John Lowengrub[5]

[1]Department of Biomedical Engineering, University of California, Irvine, USA;
[2]Beckman Laser Institute and Medical Clinic, University of California, Irvine, USA;
[3]Health Sciences Center, University of Texas, Houston, Houston, USA;
[4]Moores Cancer Center, University of California, San Diego, USA;
[5]Department of Mathematics, University of California, Irvine, USA.
Email: nmmagalh@uci.edu

ABSTRACT

The chicken chorioallantoic membrane (CAM) is a classical *in vivo* biological model in studies of angiogenesis. Combined with the right tumor system and experimental configuration this classical model can offer new approaches to investigating tumor processes. The increase in development of biotechnological devices for cancer diagnosis and treatment, calls for more sophisticated tumor models that can easily adapt to the technology, and provide a more accurate, stable and consistent platform for rapid quantitative and qualitative analysis. As we discuss a variety of applications of this novel *in vivo* tumor spheroid based shell-less CAM model in biomedical engineering research, we will show that it is extremely versatile and easily adaptable to an array of biomedical applications. The model is particularly useful in quantitative studies of the progression of avascular tumors into vascularized tumors in the CAM. Its environment is more stable, flat and has a large working area and wider field of view excellent for imaging and longitudinal studies. Finally, rapid data acquisition, screening and validation of biomedical devices and therapeutics are possible with the short experimental window.

Keywords: CAM; Cancer; Spheroid; Optical Coherence Tomography; Photodynamic Therapy; Computational Modeling; Angiogenesis

1. INTRODUCTION

Effective investigations in cancer research, whether in cancer dynamics, drug delivery, drug and diagnostic tool development, involves the use biological models that closely reflect realistic solid tumors.

While *in vitro* models may not provide a complete assessment of the processes that occur only in the environment of solid tumors [1], scientists have relied on tumor *in vivo* models or the integration of tumor *in vitro* models with *in vivo* models to circumvent the shortcomings in *in vitro* models.

Current integration approaches typically introduce tumor cells to in vivo models as single cell suspensions, multi-layered systems such as biopsies, or embedded in scaffolds such as gels or sponges prior to implantation [2, 3]. Additionally, some investigators add exogenous factors in the culture medium or scaffolds to induce certain functions such as angiogenesis [4] and invasion, that may manifest only in *in vivo* environments [5].

Single cell suspensions and biopsies may not be ideal tumor systems for bioengineering applications requiring quantitative analysis. This is because, biopsies cut directly from the animal host may not have a homogeneous cell population, and single cell suspensions lack a uniform and constant shape.

In contrast, the symmetric 3D configuration of multilayered tumor spheroids, a sphere, and the radial dependency of their proliferative and metabolic properties, can facilitate the generation of boundary conditions and fixed data points useful in quantitative investigations, such as computational modeling of tumor growth [5]. Tumor spheroids have been used extensively in studies of cancer dynamics both as an *in vitro* model [7,8,9,10] and combination with an *in vivo* model [9] because they manifest similar growth and morphological dynamics of solid tumors. They begin with an avascular growth stage where the tumor grows to a certain size *in vitro* generating a necrotic center. As we demonstrate later in the article, tumor spheroids can induce angiogenesis and pro-
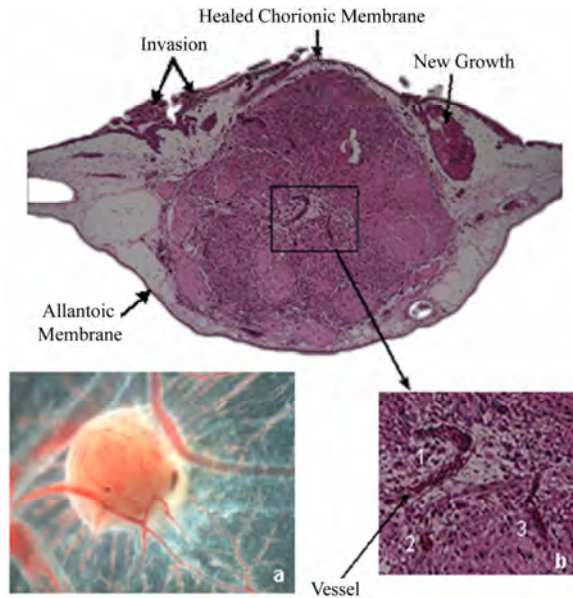
Figure 1. H&E histological section of a 1 mm diameter ACBT glioblastoma spheroid embedded in CAM. Grey area is the CAM; center region in purple denotes the tumor spheroid. Dark purple regions along the top membrane denote tumor cells invading regions of the CAM; (a) Top view of tumor spheroid and CAM interface of the formalin fixed spheroid inside CAM (mesoderm) after removal from the chicken and embryo prior to histological processing; (b) Tumor microvasculature as a result of angiogenesis (1,2,3).

mote vascular growth when introduced to an *in vivo* environment with existing vascular network.

Biomedical imaging systems ideally require a wide working area, and a steady flat environment for consistent analysis. Imaging resolution and light penetration deep into animal tissue can be compromised due to excessive light scattering and absorption from animal tissues. Qualitative and quantitative data acquisition may be facilitated with tissue of greater transparency such as the chicken embryo chorioallantoic membrane (CAM).

The traditional in-shell CAM model [12], although still useful as an alternative system to animal models for studies of tumor angiogenesis[13] and tumor dynamics [14,15], is not optimal for imaging and biomedical engineering investigations due to its restricted and unstable environment. Within the shell environment, the embryo can move freely in the presence of perturbations. Drug injections and imaging studies can be challenging with constant movement of the medium. In addition, the experimental window of the shell model can be very small, restricting access to the entire vessel network and limiting the size of the experimental site. Surface tension may cause the surface of the CAM in the shell model to curve.

The shell-less CAM version was introduced to address the drawbacks of the traditional shell CAM model

[16,17,18]. The shell-less CAM model not only offers a stable and flat environment, it also offers a large experimental area and wider field of view useful for imaging and biomedical engineering applications.

Unlike current shell-less models, our model is the first model to use three-dimensional spheroids as the tumor system of choice. The three-dimensional spheroids are implanted directly into the CAM, and vascularized spheroids are generated endogenously without the use of growth factors to promote angiogenesis or scaffolds to secure and confine the tumor inside the CAM.

This form of vascularization allows the study of the tumorigenic behavior of three dimensional tumors without the application of exogenous signals that could otherwise interfere with the natural processes of tumors. Tumor spheroids combined with the shell-less CAM model can enhance the capabilities of bioengineering modalities by providing a sophisticated tumor model with excellent imaging properties that can facilitate the studies of the angiogenic switch, and model the onset of tumor progression quickly and more effectively.

2. MATERIALS AND METHODS

The *in vivo* shell-less CAM tumor spheroid model was prepared as follows:

Initially, monolayers of ACBT grade IV human glio blastoma were cultured in T-75 culture flasks containing enhanced Dulbecco modified Eagle medium (DMEM), supplemented with 10% heat-inactivated fetal bovine serum, 1% of streptomycin, penicillin, L-glutamine, and non-essential amino acids. The cells were maintained at 37ºC and 8% CO_2 in a tissue culture incubator. After reaching 60% confluence, small tumor aggregates started to form. Using the liquid overlay method [7], ACBT suspension aggregates were collected from the flasks and transferred to culture medium filled square Petri-dishes with the bottom covered with 2% agar. Culture medium was changed three times a week. The tumor spheroids derived from ACBT suspension aggregates grew to 1 mm in diameter in approximately 30 days. For consistency, 1 mm^3 spheroids were selected from the spheroid colony on the day of implantation.

Three days old white fertilized Leghorn chicken eggs (AA Lab Eggs, Inc, Westminster, CA, USA) were disinfected with 70% alcohol wipes, and incubated in a ventilated hatching incubator at 38°C for three hours. After incubation, eggs were removed from the incubator, and under light restricted conditions, egg contents were carefully transferred from the egg shells to a condiment cup. The cups are covered with a breathable polyethylene sheet (fisher brand) and returned to the incubator.

On day seven of embryonic stage (EA), the cups containing the chicken embryos were removed from the incubators. After a small incision was made on the outer membrane of the CAM, one tumor spheroid was im-

planted onto the mesoderm of CAM. Chicken embryos were returned to the incubator for 24 hrs. The implantation was assessed after 24 hours, and angiogenesis was assessed 7 days post implantation (EA 14) using a stereomicroscope (Olympus, model SZH) coupled to a digital camera (Olympus DP 10). After visual assessment, the CAM/spheroid interfaces were fixed with 10% formalin, removed from the CAM, and stored in 10% formalin solution overnight.

The samples were processed, embedded in paraffin and cut in 6 μm serial sections. The sections were stained with hematoxylin and eosin (H&E) to distinguish the tumor, tumor microvasculature and CAM environments.

3. RESULTS AND DISCUSSION

Microvasculature is observed around the spheroid region in **Figure 1(a)**. The middle H&E stained histological section of the CAM/tumor spheroid interface (largest cross-sectional area of tumor spheroid) shows three distinct microvessels at the center of the spheroid (**Figure 1(b)**). New tumor regions are observed, as well as invasion in adjacent CAM areas.

Vascularized tumor spheroids were endogenously generated with this spheroid based shell-less CAM model using spheroids derived from human glioblastoma (U-87 MG and ACBT), human breast cancer (MCF-7), and human pancreatic cancer (BXPC-3) cell lines. Endogenous tumor induced angiogenesis is evident by the penetration of the blood vessels to the center of the ACBT tumor spheroid as shown in **Figure 1(b)**. New tumor regions on the CAM (grey regions) adjacent to the tumor spheroid (middle purple mass) show the invasive capability of the tumor.

Not only is this system adaptable to different types of tumor cell lines to generate vascularized tumor spheroids, it can be used to investigate multiple tumor related processes including, tumor growth, angiogenesis, invasion and metastasis. The experimental configuration of this system allows its integration with biomedical engineering platforms for more sophisticated analysis of these biological processes.

The following section discusses applications of this model in some areas of biomedical research, including therapeutic, computational, and optical imaging studies.

4. APPLICATIONS IN BIOMEDICAL ENGINEERING RESEARCH

4.1. Application in Therapeutic Studies: Photodynamic Therapy

The CAM system has been used previously to study the effects of therapeutic drugs, such as chemotherapy drugs and photodynamic therapy (PDT) photosensitizers [19] on tumor cells and microvasculature. The advantage of
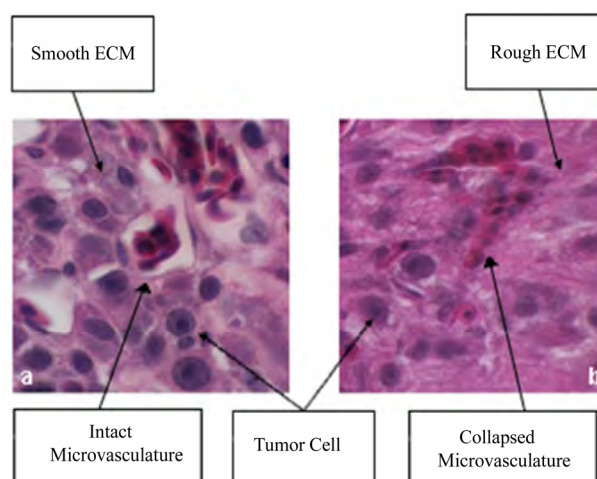
Figure 2. ALA mediated Photodynamic Therapy (PDT) in vascularized tumor spheroids implanted on the CAM of a chicken embryo. (a) Normal vascularized ACBT glioma spheroid pre-PDT; (b) Vascularized ACBT glioma spheroid post-PDT.

using this CAM-spheroid model to study therapeutic efficacy is that, multiple spheroids can be implanted on the same system for dosimetry analysis andmulti-parameter evaluations.

In a previous Photodynamic Therapy study, the shell-less CAM-tumor spheroid model was used to examine the effects of combined Amino-levulinic Acid (ALA) mediated PDT on tumor growth and microvasculature [9]. Damage to the tumor cells, extracellular matrix (ECM) and microvasculature (occluding) after acute ALA-mediated PDT was observed in **Figure 2(b)**. Individual blood cells are no longer distinct, and microvasculature lining was no longer visible (**Figure 2(b)**), compared to normal tumor vasculature observed in **Figure 2(a)**. This shows that PDT was effective in causing vascular damage inside the tumor and on the CAM. The tumor cell shape and nucleus is no longer round. Previous studies have reported similar findings [19,20]. Additionally, the extracellular matrix (ECM) of the tumor environment was observed to have a rough characteristic after PDT treatment.

Standard assays have used tumor cell suspensions embedded in the CAM to investigate the effect of PDT on tumor growth and angiogenesis [5]. Because cells suspensions tend to be more diffuse than spheroids, it is possible that in cell suspension, the PDT effect could be excessive since solid tumors do not appear in nature as cell suspensions. Furthermore, the use of exogenous angiogenic factors to induce angiogenesis may introduce differences in PDT efficacy. This *in vivo* shell-less CAM – tumor spheroid model would be a great model to conduct comparative studies since no exogenous factors are used to induce angiogenesis, and the morphology of the spheroid and its natural growth in the CAM represent a
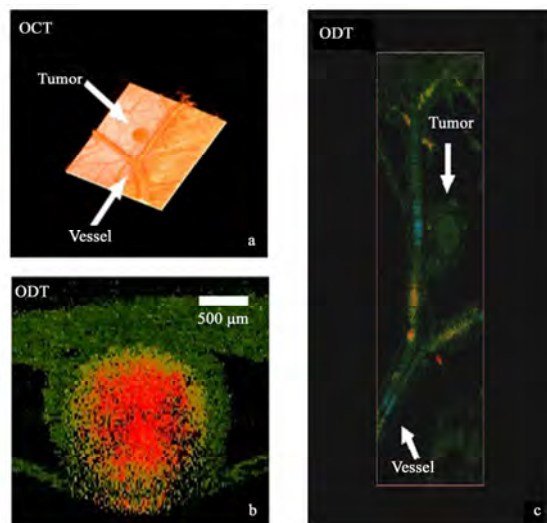
Figure 3. (a) 3-D Optical Coherence Tomography (OCT) and (c) Optical Doppler Tomography (ODT) imaging of CAM-spheroid section of a live chicken embryo (Image Acquisition Time ~2 min). (b) This figure shows real time ODT imaging of blood flow in a major blood vessel in the CAM next to the spheroid. The different colors represent the velocity of blood flow. Black color represents zero velocity and red color represents maximum velocity. Velocity increases from black, green, yellow to red.

more realistic model of *in vivo* tumors.

The use of this model is not limited to photodynamic therapy. Nanotherapy has been a growing field of investigation in recent years. Scientists are interested in developing effective delivery mechanisms for therapeutic drugs of the same nanoscale as their biological targets. Scientists have used the traditional CAM system to test the delivery capabilities and target selectivity of therapeutic drugs and hydrophobic nanoparticles carrying therapeutic drugs [19] for photodynamic therapy based [21] and chemotherapy based [22] cancer treatments.

Nanoparticles for drug delivery and treatment purposes could be introduced into the CAM vascular network for analysis and validation of its effectiveness to reach the tumor spheroid region penetrate the tumor microvasculature and release the therapeutic drug. When used in conjunction with a color coded marker or visual tag, the nanoparticles can be tracked by means of *in vivo* imaging systems as they progress from the CAM vascular network to the tumor spheroid microvasculature. As discussed in the following section, Doppler studies [23,24], such as rate of nanoparticle delivery to the target site could also be quickly assessed by coupling our tumor system with a Doppler based imaging system.

4.2. Application in Optical Imaging Studies

Sophisticated imaging devices with cancer diagnostic implications can be validated and optimized using this fast and simple *in vivo* shell-less CAM – tumor spheroid model. The large aperture and flat surface of the model could facilitate imaging, system alignment and diagnostic experiments. This shell-less CAM tumor spheroid model was used to validate a combined Optical Coherence Tomography (OCT) and Optical Doppler tomography (ODT) system developed by Zhongping Chen et al [23,25,26].

OCT has been used extensively in the clinical arena as an alternative to conventional systems such as MRI and CT to image the morphology of soft tissue [27]. In addition, ODT has been used to investigate and measure flow of blood and other fluids in tissue [24,27,28,29,30]. Both systems can be combined to provide both morphological and functional information of the tissue of interest [24,28].

The structure and morphology of the CAM-spheroid interface (**Figure 3(a)** and **3(c)**), as well as the blood flow in the vessels (**Figures 3(b)** and **3(c)**) were analyzed simultaneously using this combined system.

The morphology of CAM – tumor system was successfully assessed using this combined OCT-ODT system. The shape of the tumor spheroid as well as of the blood vessels along the surface of the CAM is clearly evident in **Figures 3(a)** and **3(c)**. However, any branching of vessels into the tumor spheroid cannot be detected from the OCT and ODT images.

In addition, **Figure 3(b)** shows the real-time imaging of a functional major blood vessel in the CAM with large blood flow in the center, and less blood flow in the periphery. The ODT channel of the system can detect velocity changes based on color mapping, with black color representing zero velocity and red color representing maximum velocity. Thus, velocity increases from black, blue, and green, yellow to red. **Figure 3(c)** illustrates the variation of blood velocity along major vessels. Yellow regions depict centers of high velocity, while blue regions depict areas of low velocity. There is a blue region on the major blood vessel adjacent to the tumor spheroid. One can speculate that low blood flow in that region may be an indication of branching of vessels near the tumor causing a reduction of flow in main vessel. Further work is necessary to investigate this phenomena and any impact that tumor induced angiogenesis may have on blood flow in major vessels adjacent to tumor masses. Moreover, future advancements in the technology of the ODT system may allow the direction of blood flow to be detected, leading to the identification of different types of vessels in tissue such as capillaries and veins.

4.3. Application and Integration in *In Silico* and Computational Studies

Mathematical modeling and multi-scale computer simulations of tumor dynamics have a promising future as innovative diagnostic tools for treatment of cancer in addition and complementary to experimental and clinical
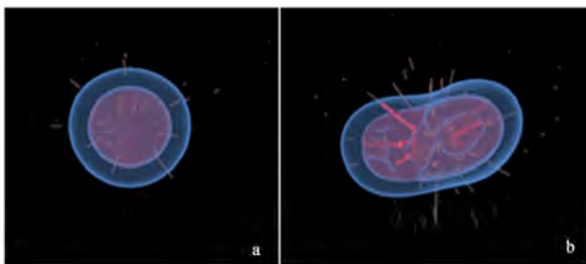
Figure 4. 3-D In-silico simulation of Tumor Spheroid induced Angiogenesis: (a) Early time; (b) Later time. The thin curves show vessel sprouts, the thick red curves describe blood-carrying vessels. The inner surface bounds the perinecrotic region. Figure courtesy of Dr. Fang Jin using methods described in [36,39].

investigations [31].

Various mathematical models have been designed to perform *in silico* (on the computer) experiments or simulations to investigate and predict tumor behavior and response to therapy both *in vivo* and *in vitro* [32,33]. These include modeling of tumor growth [34,35,36], invasion [37], angiogenesis and vascular growth [36,38, 39,40], drug delivery [5,41,42] and therapeutic response [43,44]. The in-silico results are compared with experimental and histological *in vitro* and *in vivo* data to validate and optimize the model's predictions. In addition to providing more insight into the cancer dynamics, the predictive capability of the computer tumor simulator may offer many future clinical applications.

To simulate tumor processes, these models require the acquisition of parameters such as tumor and necrotic core sizes before and after angiogenesis, density of blood vessels, speed of vessel migration towards the tumor, and gradients of growth factors involved in angiogenesis such as VEGF. The predictive capabilities of these models can be improved and optimized with the use of *in vivo* biological models that represent the dynamics of realistic tumors.

Current computational models often use tumor spheroids to integrate *in vitro* experimental data with vascular tumor growth simulations and vascular extrapolations [36,39]. An example of *in silico* tumor spheroid inducing angiogenesis is shown in **Figure 4**. The *in vivo* shell-less CAM 3-dimensional tumor spheroid model has the potential to enhance the capabilities of computational models by providing data for an *in vivo* and three-dimensional component representative of solid tumors. Thus, by integrating this data, tumor simulations will reflect more realistic representation of tumor vasculature and behavior.

The shell-less CAM tumor spheroid system may be used to acquire the parameters described above and validate computational models because it can yield fast results due to its short experimental period. Furthermore, the transparency of the CAM may facilitate data processing using imaging systems. In addition to the CAM's flat surface and closed system, the spheroid's initial symmetrical geometry may facilitate the formulation of equations and boundary conditions that represent the processes such as growth and angiogenesis that govern tumor proliferation. Thus, quantitative information can be extracted from tissue samples of the shell-less CAM tumor spheroid system. This model would be easily adaptable to existing mathematical models already using the tumor spheroid as an in vitro system, as the spheroid is simply being transferred to an *in vivo* platform (the CAM), enabling the incorporation of the *in vivo* component to existing mathematical principles.

Once these models are optimized they have the potential to be used to predict the aggressiveness of a patient's tumor. In addition, they could also predict a patient's tumor response prior to treatment.

That is, the *in silico* model could be used as a diagnostic tool to recommend the most effective individualized treatment based on the patient's prognosis. This can prevent the development of resistance to treatment by eliminating trial and error treatment sampling. Other advantages of the *in silico* model include providing fast and non-invasive diagnosis thus minimizing the discomfort to the patient, and improving overall quality of life. Finally, while the application of *in-silico* models in cancer therapy require improved understanding of cancer behavior and mechanisms, their versatility allows optimization and calibration of parameters to match with new developments and knowledge of cancer dynamics.

5. CONCLUSIONS

The *in vivo* 3-dimensional tumor spheroid based shell-less CAM model presented in this article is a new, attractive and practical system to study multiple mechanisms of tumor biology, including tumor growth, invasion, and angiogenesis. In addition, this system has a vast and dynamic application in many biomedical and bioengineering studies including drug discovery, nanoparticle delivery, therapeutic efficacy, cancer diagnostic imaging device development and validation, and mathematical and computer simulations of cancer dynamics [45].

The development of this system successfully created vascularized spheroids in the CAM in the absence of exogenous factors four days after implantation. Due to the system's analogous representation of the microenvironment found *in vivo* solid tumors, this model can be used as an alternative or complement to animal models, and the research findings can provide preliminary implications to clinical studies.

6. ACKNOWLEDGMENT

REFERENCES

[1] Zietarska, M. *et al*. (2007) Molecular description of a 3D in vitro model for the study of epithelial ovarian cancer (EOC). *Mol Carcinog*, **46(10)**, 872-85.

[2] Nakatsu, M.N. and Hughes, C.C. (2008) An optimized three-dimensional in vitro model for the analysis of angiogenesis. *Methods Enzymol*, **443**, 65-82.

[3] Claudia F. *et al* (2007) Engineering tumors with 3D scaffolds. Nature Methods, **4** (10): 855-860.

[4] Szpaderska, A.M. and DiPietro, L.A. (2003) *In vitro* matrigel angiogenesis model. *Methods Mol Med*, **78**, 311-315.

[5] Duong, H.S. *et al*. (2005) A novel 3-dimensional culture system as an in vitro model for studying oral cancer cell invasion. *Int J Exp Pathol*, **86(6)**, 365-74.

[6] Hammer-Wilson, M.J., Cao, D., Kimel, S. and Berns, M.W. (2002) Photodynamic parameters in the chick chorioallantoic membrane (CAM) bioassay for photosensitizers and administered intraperitoneally (IP) into the chick embryo. Photochem. *Photobiol. Sci.*, **1**, 721-728.

[7] Santini, M. and Rainaldi, G. (1999) Three-dimensional spheroid model in tumor biology. *Pathobiology*, **67**, 148-157.

[8] Liang, Y., Pjesivac-Grbovic, J., Cantrell, C. and Freyer, J.P. (2005) A multiscale model for avascular tumor growth. *Biophysical Journal*, **89**, 3884-3894.

[9] Madsen, S.J, Sun, C.H., Tromberg, B.J., Wallace, V.P. and Hirschberg, H. (2000) Photodynamic therapy of human glioma spheroids using 5-aminolevulinic acid. *Photochemistry and Photobiology*, **72**, 128-134.

[10] Reyes-Aldasoro, C.C. *et al*. (2008) Estimation of apparent tumor vascular permeability from multiphoton fluorescence microscopic images of P22 rat sarcomas in vivo. *Microcirculation*, **15(1)**, 65-79.

[11] De Magalhães, N., Liaw, L.H.L., Li, L., Liogys, A., Madsen, S.J., Hirschberg, H. and Tromberg, B.J. (2006) Investigating the effects of combined photodynamic and anti-angiogenic therapies using a three-dimensional in vivo brain tumor system. *SPIE Proceedings on Photonic Therapeutics and Diagnostics*, **6078**, 503-508.

[12] Knighton, D. *et al*. (1975) Study of avascular and vascular phases of tumor growth in chick embryo. *Clinical Research*, **23(4)**, 557.

[13] Ishiwata, I. *et al*. (1999) Tumor angiogenesis factors produced by cancer cells. *Hum Cell*, **12(1)**, 37-46.

[14] Ribatti, D., Nico, B., Vacca, A., Roncali, L., Burri, P. and Djonov, V. (2001) Chorioallantoic membrane capillary bed: a useful target for studying angiogenesis and anti-angiogenesis *in vivo*. *The Anatomical Record*, **264**, 317-324.

[15] Stewart, J. (1999) Calculus, Brooks/Cole Publishing Company, **6**, 378.

[16] Dunn, B.E. (1974) Technique of shell-less culture of the 72-hour avian embryo. *Poult Sci*, **53(1)**, 409-412.

[17] Jakobson, A.M., Hahnenberger, R. and Magnusson, A. (1989) A simple method for shell-less cultivation of chick embryos. *Pharmacol Toxicol*, **64(2)**, 193-195.

[18] Ono, T. (2000) Exo ovo culture of avian embryos. *Methods Mol Biol*, **135**, 39-46.

[19] Vargas, A., Zeisser-Labouèbe, M., Lange, N., Gurny, R., Delie, F. (2007) The chick embryo and its chorioallantoic membrane (CAM) for the in vivo evaluation of drug delivery systems. *Advanced Drug Delivery Reviews*, **59(11)**, 1162-1176.

[20] Hammer-Wilson, M.J., Akian, L., Espinoza, J., Kimel, S. and Berns, M.W. (1999) Photodynamic parameters in the chick chorioallantoic membrane (CAM) bioassay for topically applied photosensitizers. *J. Photochem. Photobiol*, **53**, 44-52.

[21] Vargas, A., Eid, M., Fanchaouy, M., Gurny, R., Delie, F. (2008) *In vivo* photodynamic activity of photosensitizer-loaded nanoparticles: formulation properties, administration parameters and biological issues involved in PDT outcome. *European Journal of Pharmaceutics and Biopharmaceutics*, **69**, 43-53.

[22] Imtiaz, A. *et al* (2009) Introducing Nanochemoprevention as a Novel approach for cancer control: proof of principle with green tea polyphenol epigallocatechin-3-gallate. *Cancer Res*, **69(5)**, 1712-1716.

[23] Chen, Z. (2004) Optical doppler tomography. *In Handbook of Coherent Domain Optical Methods*, Tuchin V.V. (ed), Kluwer Academic Publishers, Boston, **2**, 315-342.

[24] Ding, Z, Zhao, Y, Ren, H, Nelson, J. and Chen, Z. (2002) Real-time phase-resolved optical coherence tomography and optical Doppler tomography. Optics Express, **10**: 236-245.

[25] Chen, Z., Milner, T., Srinivas, S., Malekafzali, A., Wang, X., Van Gemert, M. and Nelson, J. (1997) Imaging in vivo blood flow velocity using optical doppler tomography. *Opt. Lett.*, **22**, 1119-1121.

[26] Chen, Z. Milner and T.E *et al* (2000) Noninvasive imaging of in vivo blood flow velocity using optical doppler tomography. *Optical Low Coherence Reflectometry and Tomography*, SPIE Milestone Series Book of Selected Papers.

[27] Baumal, C.R. (1999) Clinical applications of optical coherence tomography. *Curr Opin Ophthalmol*, **10(3)**, 182-188.

[28] Donald, I., MacVicar, J. and Brown, T.G. (1958) Investigation of abdominal masses by pulsed ultrasound. *Lancet*, **1(7032)**, 1188-1195.

[29] Kehlet-Barton, J. *et al*. (1999) Three-dimensional reconstruction of blood vessels from in vivo color Doppler optical coherence tomography images. *Dermatology*, **198(4)**, 355-361.

[30] Szkulmowska, A. *et al*. (2008) Phase-resolved doppler optical coherence tomography--limitations and improvements. *Opt Lett*, **33(13)**, 1425-1427.

[31] Gomez-Lopez, G. and Valencia, A. (2008) Bioinformatics and cancer research: building bridges for translational research. *Clin Transl Oncol*, **10(2)**, 85-95.

[32] Roose, T., Chapman, S.J. and Maini, P.K. (2007) Mathematical models of avascular tumor growth, *SIAM Review*, **49**, 179-208.

[33] Cristini Cristini, V., Frieboes, H.B., Li, X., Lowengrub, J. S., Macklin, P., Sanga, S., Wise, S.M. and Zheng, X. (2008) Nonlinear modeling and simulation of tumor growth. *In Modeling and simulation in science, engineering and technology*, ed. Bellomo, N., Chaplain, M.A.J., De Angeles, E., Birkaeuser, Boston.

[34] Macklin, P. and Lowengrub, J. (2007) Nonlinear simulation of the effect of microenvironment on tumor growth. *J Theor Biol*, **245(4)**, 677-704.

[35] Wise, S.M., Lowengrub, J.S., Frieboes, H.B., and Cristini, V. (2008) Three-dimensional multispecies nonlinear tumor growth—I Model and numerical method. *J. Theor. Biol.*, **253**, 524-543.

[36] Bearer, E S., Lowengrub, J S., Frieboes, H.B., Chuang, Y.L., Jin, F., Wise, S.M., Ferrari, M., Agus, D.B., Cristini, V. (2009) Multiparameter computational modeling of tumor invasion. *Cancer Res.*, **69**, 4493-4501.

[37] Cristini, V., Li, X., Lowengrub, J.S. and Wise, S.M. (2009), Nonlinear simulations of solid tumor growth using a mixture model: invasion and branching. *J Math Biol*, **58**, 723-63.

[38] Mantzaris, N.V, Steve, S. and Othmer, H.G. (2004) Mathematical modeling of tumor-induced angiogenesis. *J. Math. Biol.*, **49**, 111-187.

[39] Frieboes, H.B. *et al.* (2007) Computer simulation of glioma growth and morphology. Neuroimage, **37** (Suppl 1), 59-70.

[40] Macklin, P., McDougall, S., Anderson, A.R.A., Chaplain, M.A.J., Cristini, V. and Lowengrub, J.S. (2009) Multiscale modeling and nonlinear simulation of vascular tumor growth. *J. Math. Biol.*, **58**, 765-798.

[41] Sinek, J.P., Sanga, S., Zheng, X., Frieboes, H.B., Ferrari, M. and Cristini, V. (2009) Predicting drug pharmacokinetics and effect in vascularized tumors using computer simulation. *J. Math Biol*, **58**, 485-510

[42] Frieboes, H.B., *et al.* (2006) An integrated computational/experimental model of tumor invasion. *Cancer Res*, **66(3)**, 1597-604.

[43] Madsen, S.J., Sun, C.H., Tromberg, B.J., Cristini, V., De Magalhães, N. and Hirschberg, H. (2006) Multicell tumor spheroids in photodynamic therapy. *Lasers Surg Med*, **38(5)**, 555-564.

[44] Sinek, J. *et al* (2004) Two-dimensional chemotherapy simulations demonstrate fundamental transport and tumor response limitations involving nanoparticles. *Biomed Microdevices*, **6(4)**, 297-309.

[45] Herzenberg, L. *et al.* (2004) American cancer society. *Clinical Oncology* (Blackwell Publishing), 254-260.

Fiber lenses for ultra-small probes used in optical coherent tomography

Youxin Mao, Shoude Chang, Costel Flueraru

Institute for Microstructural Sciences, National Research Council Canada, Ottawa, Canada.
Email: linda.mao@nrc-cnrc.gc.ca

ABSTRACT

We present a design, construction and characterization of different variations of GRIN and ball fiber lenses, which were recently proposed for ultra-small biomedical imaging probes. Those fiber lens modules are made of a single mode fiber and a GRIN or ball fiber lens with or without a fiber spacer between them. The lens diameters are smaller than 0.3 mm. We discuss design methods, fabrication techniques, and measuring performance of the fiber lenses. The experimental results are compared to their modeling results. The fabrication of a high quality beam director for both lens types is presented as well. These fiber integrated beam directors could be added on the tips of the fiber lenses for side-view probes. A needle probe made by these fiber lenses is demonstrated as a sample of the ultra-small probe for biomedical imaging application. *In vivo* human finger images acquired by a swept source optical coherence tomography using the fiber lenses with different beam profiles were shown, which indicates the important impact of fiber lens on the image quality.

Keywords: Optical Fiber Probe; Optical Coherence Tomography; Bio-Medical Imaging

1. INTRODUCTION

Optical biomedical imaging techniques, such as optical coherence tomography (OCT) [1] and Doppler OCT [2,3], are becoming increasingly important tools for both diagnosis and guided surgery because of their high image resolutions. OCT can provide images on the cellular level while Doppler OCT can detect blood flow with velocity sensitivities approaching a few micrometers per second [4,5]. However, in most optically nontransparent tissues, OCT has a typical imaging depth limitation of 1-3 mm. Similarly, Doppler OCT systems suffer from limitations where blood flow can rarely be detected beyond 1–2 mm from the tissue surface without *a priori* velocity profile information and digital extrapolation

algorithms. As a result, the earliest *in vivo* OCT imaging of tissue microstructure and microvasculature was restricted to a few transparent or superficial organ sites, such as the retina [6,7] and skin [8,9]. To overcome this depth limitation, optical probes, such as endoscopes, catheters, and needles have been investigated for *in vivo* OCT imaging in mucosal layers of the gastrointestinal tract [10,11], deep organs and tissues [12,13], and inter-arterial and intra-vascular [14,15]. However, for the imaging of small lumen, narrow space, and deep tissue and organ of humans and small animals, a key concern is the possible damage from the mechanical insertion of the optical probe. Therefore it is critical to develop an ultra-small optical probe that is compatible with the OCT systems, which results in minimum tissue damage.

In vivo optical imaging of internal tissue is generally performed using a fiber-optic probe, since an optical fiber can be easily and cheaply produced with a diameter of less than 0.15 mm. The key components of such optical fiber probe include a small lens and a beam director, where both provide a focused optical beam directing it to a location of interest through a guide-wire. Traditionally, this type of small optical probe has been implemented by attaching a small glass GRIN or SELFOC lens (0.25-1.0 mm) and a glass micro-prism to a single mode fiber (SMF) with optical adhesive or optical epoxy [12]. However, the gluing of a separate small lens and a tiny prism to a fiber is a complex fabrication process that results in a low quality optical interface. A new probe design that uses optical fiber lenses, e.g., fiber GRIN lens or fiber ball lens, has recently been proposed [16,17]. The main advantage of fiber lenses over conventional glass lenses are their small size, ability to auto-align to a fiber, thus creating a fusion-spliced interface with low loss, low back-reflection, and high mechanical integrity. In addition, a beam director can be easily attached to the fiber lenses by the fusion-splice of a polished fiber spacer for GRIN fiber lens and direct polish on the lens surface for ball fiber lens. Swanson *et al.* and Shishkov *et al.* proposed the fiber based optic probes design, but presented the variations of probe

structure instead of the characteristics of their performance [16,17]. Reed *et al.* demonstrated the usage of such probes with emphasis on their insertion loss only [18]. Yang *et. al.* [13], Jafri *et. al.* [19], and Li *et. al.* [20] reported OCT images without detailed characterization of the used fiber lens based probes. Beam quality of a fiber-optic probe is crucial for the imaging system. Ideal characteristics of a fiber-optic probe include a high Gaussian beam intensity profile, an appropriate intensity-distance shape, high flexibility, and low optical aberration and loss. In this presentation, we discuss design methods and fabrication techniques of fiber-lens-based optic probes. We compare in detail measured performance with expected theoretical performance.

2. METHOD OF DESIGN AND FABRICATION

2.1. Design Criteria

For OCT system, image quality and light coupling efficiency from the sample will be directly influenced by the beam quality and profile of the fiber probe. For the best optical performance of a fiber probing lens, its beam profile must be designed to be consistent with the light penetration depth in the sample. In most biomedical imaging systems, light from the probe will be directed into a turbid tissue. Based on interaction properties of light with turbid tissue [21], the range of penetration depth is from 0.5-3 mm at near infrared wavelengths. For example, the penetration depths are 0.7 mm and 3.0 mm in human skin and liver, respectively, at 1300 nm, a conventional wavelength used in OCT systems. Because of these differences of the penetration depth, to design an optical probe, working distance should be in the range of 0.4 – 1.2 mm in the air that depends on the tissues to be tested. There is a tradeoff between the depth of field and beam spot size because the depth of field of a lens is positively related to the square of the spot size according to the theory of Gaussian beam. A large depth of field unavoidable results in a large spot size. Thus, the optimal depth of field is in the range of 0.8 - 1.5 mm in the air; this keeps the spot size in the range of 26 - 35 μm at the 1300 nm wavelength. For an ultra-small optical lens, it is not possible to achieve a large working distance by directly attaching a lens to a SMF because of the strong focus ability of lens and the small mode field diameter (MFD) of the SMF [22]. A fiber spacer with a homogeneous index of refraction has to be added between the SMF and the lens for beam expansion prior to focusing to obtain a larger working distance. Therefore, theoretical modeling becomes necessary to obtain a lens with optimized optical beam performance for imaging different tissues.

2.2. Simulation Method

In this work, we used the commercially available numerical optical modeling software, ZEMAX (ZEMAX Development Corp., WA, USA), to design both GRIN and ball fiber lenses by choosing an appropriate surface type and analysis method. For a GRIN fiber lens, the gradient surface type used in ZEMAX will depend on the profile of the refractive index. The refractive index profile of the GRIN fiber lens used in this work is a radial index gradient, which is very similar to that of a conventional GRIN (or SELFOC) lens. The index of refraction is highest in the center of the lens and decreases with radial distance from the axis. The following quadratic equation closely describes the refractive index of a GRIN fiber lens [23,24]:

$$n(r) = n_0(1 - \frac{g^2}{2} r^2) \tag{1}$$

where r is the radial position from the axis, n_0 is refractive index on the lens axis, g is the gradient constant. The pitch, $p = 2\pi/g$, is the spatial period of the ray trajectory. For modeling the ball fiber lens, a Standard Surface type in ZEMAX is employed. **Figure 1(a)** and **(b)** show the ZEMAX ray trace layout of the GRIN and ball fiber lenses, respectively. The working distance, depth of field and spot size were calculated in ZEMAX using the Gaussian beam theory. The results will be discussed below in comparison with the experimental data.

2.3. Fabrication Method

The GRIN and ball fiber lens probes were made from a standard Corning SMF-28 single mode fiber as the principal light guide, a fiber spacer and a GRIN or a ball fiber lens as the focusing lens. For the GRIN lens probe, a fiber spacer with same outer diameter (0.125 mm) as SMF-28 was fusion-spliced via arc welds to the Corning SMF-28 and then accurately cleaved to a theoretically-determined length. A GRIN fiber was then fusion-spliced to the cleaved fiber spacer and precisely cleaved at a pre-calculated length to generate a desired beam-distance profile (i.e., working distance, depth of field, and spot size). For a short working distance probe, the section of the fiber spacer was omitted resulting in a simple fabrication process. For the ball lens probe, a fiber spacer with same outer diameter of SMF was fusion-spliced via arc welds to the Corning SMF-28 and then accurately cleaved to a theoretically-determined length plus extra 0.2 mm. The tip of the fiber spacer then was fused via arc welds to a perfect ball shape by inputting an appropriate fusion setting. To ensure minimum back-reflection for both probes, the indexes of the fiber spacer and the center of GRIN fiber were matched to the core index of the SMF.

We used a conventional low cost off-the-shelf optical multi-mode GRIN fibers as the GRIN lens, which has
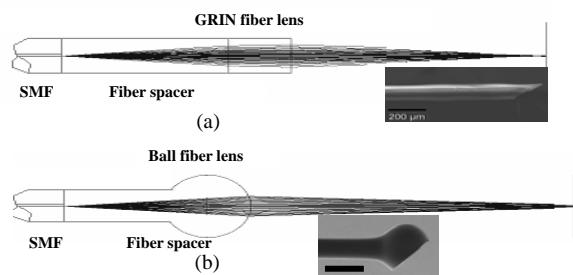
Figure 1. ZEMAX layouts of ray trace for the GRIN (a) and ball (b) fiber lens systems, respectively. Insets: scanning electron micrographs of the GRIN and ball fiber lens tips fused with angle-polished beam director, respectively. Marks are 0.2 mm.

0.1 mm core size, 0.14 mm outer diameter, a core refractive index $n_0 = 1.487$, and a gradient constant $g = 3.76$ at 1300 nm (Prime Optical Fiber Corp., Taiwan). The fiber spacers (Prime Optical Fiber Corp., Taiwan) are made of pure silica without a core. Fusion-splicing was processed using an Ericsson FSU 995 fusion-splicer and an EFC11 fiber cleaver (3SAE Technologies, TN, USA). The spliced interfaces produced minimum back-reflections since the mechanical strength at the interface was similar to that of the untreated fiber. The desired focused beam profile was obtained by tailoring the length of the fiber spacer and parameters of fiber lenses (length of GRIN fiber and diameter of the ball) based on the theoretical results. We fabricated the different variations of the GRIN and ball fiber lens modules with the different length of the fiber spacer and the different lens parameter. All samples were listed in **Table 1** along with detailed descriptions of the samples.

2.4. Characterization Method

A beam profile measurement system (BeamView Analyzer, OR, USA) with an infrared camera (Electrophysics, NJ, USA) and a Super Luminous Diode source (Covega, MD, USA) with 60nm 3dB bandwidth at 1310 nm center wavelength was used to characterize the beam parameters of the lens system. A 40X JIS microscopic objective lens and a related objective tube were attached to the input window of the camera to increase the image resolution. The horizontal and vertical resolutions of 1.0 μm and 1.1 μm were achieved, respectively. The distribution of light intensity at various distances along the direction of propagation after the lens was accurately measured by the beam profile system. Working distance, depth of focus, $1/e^2$ spot size, and Gaussian fitting were analyzed from the measured intensity distribution. The results demonstrated in this work are all in the air medium.

2.5. Fabrication of Fiber Beam Director

After characterization of the lens, a beam director could be attached to the lens for a side-view probe. The differ-

ent attaching methods were used for the two lenses. For the GRIN lens, a fiber spacer was fusion-spliced to the finished lens end as a beam director by polishing the end of the fiber spacer to a 45 degrees angle and coating the polished surface with a total reflection film. This then allowed the beam to be reflected at a 90 degrees angle creating a side-view probe. For the ball lens probe, the beam can be totally internal reflected by a 50 degree polished face on the ball lens.

Insets in the **Figure 1(a)** and **(b)** show the typical scan electron microscope (SEM) pictures of the GRIN and ball fiber lens tip fused with beam directors, respectively. The fiber lens tip together with a tubing system and a connected linearly scanning or 360 degrees rotated motor could be built as an endoscope, or catheter, or a needle probes. The diameter of these probes could be as small as 0.4 mm, which is best suitable for internal *in situ* and *in vivo* biomedical imaging, diagnostic, guided surgery, and treatment with a minimal invasion.

2.6. Needle probe

As a sample, a needle probe designed for the OCT imaging in this work is shown in **Figure 2(a)**. The lens and the uncoated portion of the SMF are protected in a transparent inner catheter (OD 0.49 mm) shown in **Figure 2(b)**. The buffered portion of the fiber is attached to an outer flexible catheter after the syringe (OD 1.4 mm), which is fastened onto a modified syringe piston, while the transparent inner catheter is inserted into a 21 G (OD 0.81 mm) echogenic spinal needle (VWR, Mississauga, ON, Canada). After insertion into the tissue, the needle can be drawn back while the optical probe stays inside of the tissue as shown in **Figure 2(c)**. The probe is then scanned axially inside the tissue driven by a linear scanner, such that a two dimensional OCT image is formed. If a fiber GRIN lens is used, the size of the inner catheter could be as small as 0.4 mm because the diameter of the GRIN lens is smaller than the fiber ball lens.

3. RESULTS AND DISCUSSION

3.1. Experimental Results of Beam Profile

For each sample in this study, optical intensity distribution data on the radial (i.e. x and y) planes were collected along the beam propagation (i.e. optical axial z) direction from the plane of the first half peak intensity (beginning-plane), through the maximum intensity plane, i.e. focus plane (center-plane), to the second half peak intensity plane (end-plane). Beam properties including working distance, spot size, and depth of field were analyzed by measured intensity distribution data with distance from the lens surface to the focal plane, $1/e^2$ beam diameter at the focal plane, and the distance between the begin-plane and the end-plane, respectively. The measured results of the beam properties are listed in **Table 1**

Table 1. Structures of the various samples with measured beam properties.

Samples	Length of Fiber Spacer (mm)	Fiber Lens		Measured Beam Properties		
		Type	Length/ Diameter (mm)	Working Distance (mm)	Depth of Field (mm)	Spot Size (μm)
1	0.52	Ball Lens	0.15	1.00	3.6	50
2	0.55	Ball Lens	0.15	1.40	2.1	45
3	0.62	Ball Lens	0.15	1.20	1.1	27
4	0.70	Ball Lens	0.15	1.00	0.5	20
5	0.75	Ball Lens	0.15	0.90	0.48	18
6	0.00	GRIN Fiber	0.6	0.18	0.16	13
7	0.00	GRIN Fiber	0.55	0.20	0.30	16
8	0.00	GRIN Fiber	0.52	0.28	0.50	22
9	0.00	GRIN Fiber	0.50	0.38	0.60	23
10	0.00	GRIN Fiber	0.48	0.41	0.85	25
11	0.00	GRIN Fiber	0.46	0.40	1.30	30
12	0.00	GRIN Fiber	0.45	0.38	1.45	32
13	0.48	GRIN Fiber	0.17	1.00	0.95	28
14	0.48	GRIN Fiber	0.16	1.10	1.5	35
15	0.48	GRIN Fiber	0.145	1.20	1.8	41
16	0.48	GRIN Fiber	0.14	1.05	2.0	45

along with detailed descriptions of the samples. The theoretical and experimental results of working distance, depth of focus, and spot size of different variations vs. length of GRIN fiber or diameter of the ball lens (bottom x-axis), and length of fiber spacer (top x-axis) are shown in **Figure 3(a)**, **(b)**, and **(c)**, respectively, where, lines represent the theoretical results from ZEMAX at 1300 nm, amount them, dark doted line represent GRIN fiber lens without a fiber spacer, dark and light solid lines represent GRIN fiber lens with a constant length of fiber spacer (0.48 mm) and a constant length of GRIN fiber (0.17 mm), respectively; dark and light dash lines represent ball fiber lens with a constant length of fiber spacer (0.62 mm) and a constant diameter of the ball (0.30 mm), respectively. The related experimental results were represented by points, amount them, triangle points represent GRIN fiber lens and square points represent the ball lens.

3.2. Discussion of Beam Profile Results

From the theoretical result shown in **Figure 3,** short working distance (<0.4 mm) could be obtained by the GRIN fiber lens without fiber spacer shown as the dark dotted lines. To obtain larger working distance, a fiber spacer has to be fusion-spliced between SMF and fiber lens. In these cases, the working distance varies sharply with the length of GRIN fiber for the GRIN fiber lens and with the diameter of the ball for the ball fiber lens, but it varies less sharply with the length of the fiber spacer. The working distances have saturated values for each case. By compensating the working distance with the depth of field and the spot size, the optimized parameters (i.e. 0.9 – 1.2 mm working distance, 0.9–1.1

mm depth of filed, and <30 μm spot size) are not at the position of the largest working distance, instead, the optimized positions are around 0.17 mm length of the GRIN fiber with 0.48 mm length of fiber spacer for GRIN fiber lens and 0.3 mm diameter of ball and 0.62 mm length of fiber spacer for ball fiber lens.

From the experimental results shown in **Figure 3** and **Table 1**, we obtained the working distance of 1.0 mm, the depth of field of 0.95 mm, and the spot size of 28 μm from a GRIN fiber lens module (sample #13) and the working distance of 1.2 mm, the depth of field of 1.1 mm, and the spot size of 27 μm from a ball fiber lens module (sample #3). The results from the ZEMAX numerical optical design software were in a good agreement with the experimental results.

Considering chromatic aberrations, from ZEMAX simulation for the ball fiber lens in the wavelength range of 1260 – 1370 nm, the relative variations of the working distance, depth of field and spot size were calculated all smaller than 4.0%. For the GRIN fiber lens, the range
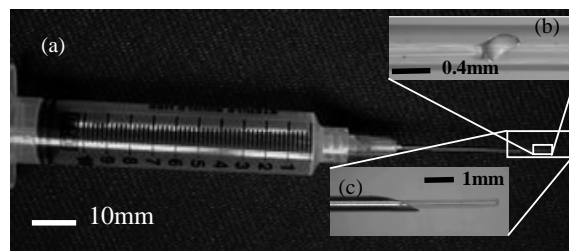
Figure 2. OCT side-view needle probe showing the tubing and angle-polished ball lens. (a) needle probe; (b) protective tubing and exposing the lens; (c) retracted needle tip with protective tubing.
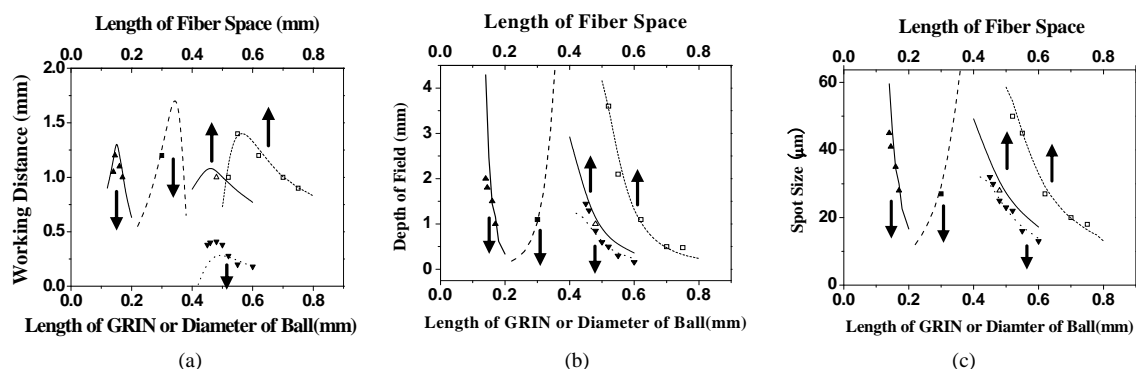
Figure 3. Theoretical and experimental results of working distance (a), depth of field (b), and spot size (c) vs. length of GRIN fiber or diameter of the ball lens (bottom x axis), and length of fiber spacer (top x axis), where, lines represent the theoretical results from ZEMAX at 1300 nm, amount them, dark doted line represent GRIN fiber lens without a fiber spacer, dark and light solid lines represent GRIN fiber lens with a constant length of fiber spacer (0.48 mm) and a constant length of GRIN fiber (0.17 mm), respectively; dark and light dash lines represent ball fiber lens with a constant length of fiber spacer (0.62 mm) and a constant diameter of the ball (0.30 mm), respectively. The related experimental results were represented by the points.

of the zero-dispersion wavelengths, $\{\lambda_0\}$ is 1297-1316 nm. The zero-dispersion slope, S_0, equal to or smaller than 0.101 ps/nm^2-km. Using the standard formula offiber dispersion, $D(\lambda) = S_0[\lambda - \lambda_0^4 / \lambda^3]/4$ (ps/nm-km), we calculated the changes of refractive index in the 1260 – 1370 nm wavelength range. By using these values in ZEMAX, we calculated the relative changes of the working distance, depth of field and spot size were all smaller than 3%. Based on our results, the desired beam profile for the application of optical biomedical imaging systems can be obtained by the GRIN and ball fiber lens with or without fiber spacers. The technique described here possesses a high degree of flexibility for designing ultra-small optical probes with different beam shapes for the different tissue imaging.

For GRIN fiber lens, a beam profile of 15 μm spot size and 1mm working distance was reported [19], although no detailed GRIN fiber and lens structure description in the paper. The 15 μm spot size only could provide less than 0.3 mm depth of field in their OCT system. The patent [16] proposed a beam profile of 30 μm spot size and 2 mm working distance by using a special drawn GRIN fiber with low focus ability. This large working distance could not be achieved by using the conventional low cost off-the-shelf optical multi-mode GRIN fibers, like the GRIN fiber used in this work. Increase working distance will be our future work.

3.3. Experimental Results and Discussion of Beam Quality

We found quality of the beam depends very much on the quality of the surface cleaving and the alignment of the fusion-splicing between the fiber spacer and the fiber lens. The high quality beam is easier to obtain for the probe with the ball fiber lens because the ball is made

from the fiber spacer and there is no interface between the fiber spacer and the ball lens. By well controlling the cleaving and the fusion-splicing, we obtained high quality of beam for the probe of the GRIN fiber lens as well.

Figure 4 shows measured and Gaussian-fitted 1/e^2 intensity beam diameters along the axial distance z (zero is the position of the lens surface) at x (horizontal) and y (vertical) radial coordination in the distance range of depth of field for the samples #8 and #13 with the GRIN fiber lenses. In **Figure 4**, the smallest beam diameter value indicates spot size, x-coordinate value at the pole point indicates the working distance, and the distance range of the curve indicates the depth of field. The working distance of 0.28 mm, depth of field of 0.5 mm, and spot size of 22 μm were obtained for the sample #8 with 0.52 mm length of 100/140 GRIN fiber lens and without fiber spacer. The working distance of 1.0 mm, depth of field of 0.95 mm, and spot size of 28 μm were
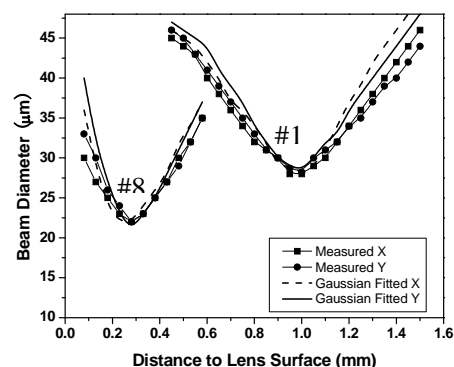
Figure 4. Measured and Gaussian-fitted 1/e^2 intensity beam diameters along the axial distance (zero is the position of the lens surface) at x (horizontal) and y (vertical) radial coordination in the distance range of depth of field of the samples #8 and #13, which was made from the GRIN fiber lens.

Table 2. Measured beam profile images and normalized intensity distributions with Gaussian-fittings at x (horizontal) and y (vertical) radial coordination for sample #13.
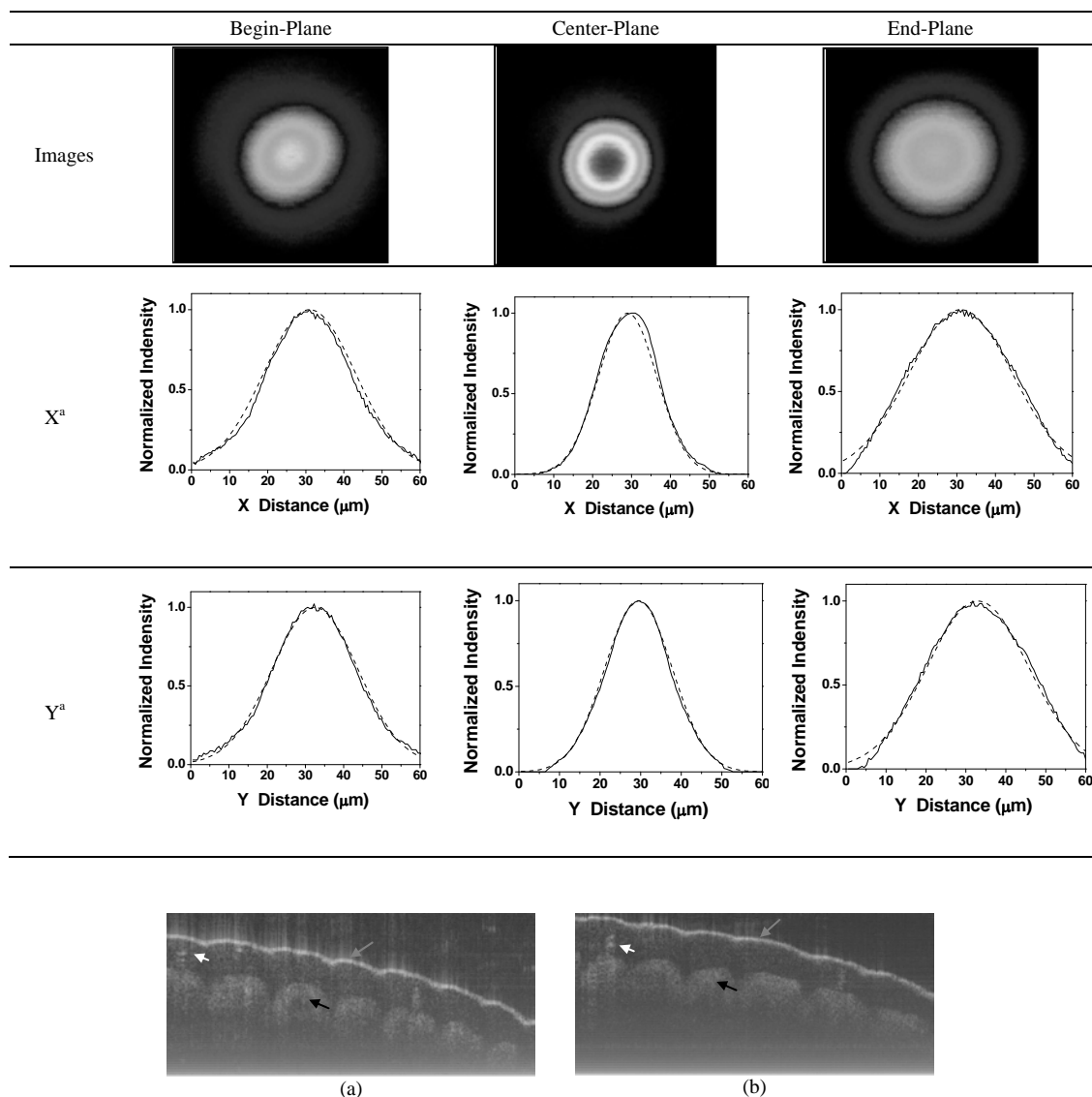
Figure 5. *In vivo* human finger OCT images taken with probe #3 (working distance, depth of field and spot diameter of 1.2 mm 1.1 mm, and 27 μm) and # 5 (working distance, depth of field and spot diameter of 0.9 mm 0.48 mm, and 18 μm).

obtained for sample #13 with 0.17 mm length of 100/140 GRIN fiber lens and 0.48 mm fiber spacer. The x and y symmetry of the beam diameter is very good for both samples. The measured beam diameters are well matched to Gaussian-fitted values in the center (focused) regions, but have small deviations on either side of the center regions.

To further examine the beam quality of the GRIN fiber lens system, **Table 2** shows measured beam profile images and measured normalized intensity distributions with Gaussian-fitted results at the x and y directions on the three typical planes (i.e. begin-plane, center-plane, and end-plane) of the sample #13. From the profile images and distributions shown in **Table 2**, the measured beam profiles match very well with Gaussian distributions at the beginning-plane and center-plane. On the end-planes, the measured and Gaussian-fitted intensity distributions generally match very well despite slight

deviations on both tail ends of the distributions leading to discrepancies between the measured and Gaussian-fitted beam diameters as was shown in **Figure 4**. In addition, the circular shapes in the profile images as shown in **Table 2** indicate high x and y symmetry of the beam profiles through all range of depth of field.

3.4. OCT Image

Figure 5(a) and **(d)** show *in vivo* images of human finger acquired by a SS-OCT with the fiber probes # 3 (working distance, depth of field and spot diameter of 1.2 mm 1.1 mm, and 27 μm) and # 5 (working distance, depth of field and spot diameter of 0.9 mm 0.33 mm, and 18 μm). The OCT system was described in detail elsewhere [25]. Briefly, the swept source (HSL2000-HL, Santac) used in the system had a central wavelength of 1320 nm and a full scan wavelength range of 110 nm, which was sweeping linearly with optical frequency with a linearity of 0.2%. The average output power and coherence length of the swept source was 12 mW and 10 mm, respectively. A repetition scan rate of 20 kHz was used in our system and the related duty cycle was 68%. The output light from the swept laser source was launched into the first coupler and then divided into the sample arm with 90% power and reference arm with 10% power by two fiber circulators. The reference arm was arranged with a fiber collimator and a mirror. A variable attenuator was inserted between the collimator and mirror for adjusting the optical power on reference arm to achieve the higher sensitivity. The light was illuminated to the fringer through the fiber lensed fiber probe. A galvanometer (Blue Hill Optical technologies) scanner scanned the fiber probe light transversely on the sample up to 4 mm at 20 Hz with 1000 transverse pixels. The total optical power illuminating on the sample was approximately 10 mW. Two polarization controllers (PC) in both reference and sample arms were used for adjustment to match the polarization state of the two arms. The two-pair output signals from the output couplers were detected with two-pair photodiodes to obtain quadrature signals. Two differential photo-detectors (PDB150C, Thorlabs) were used with adjustable bandwidth. A 3 dB bandwidth of 50 MHz was used in our system. The two detector outputs were digitized using a data acquisition card (DAQ) (PCI 5122, National Instruments) with 14-bit resolution and acquired at a sampling speed of *100 MS/s*. The swept source generated a start trigger signal that was used to initiate the function generator for the galvo scanner and initiate the data acquisition process for each A-scan. Because the swept source was linearly swept with wave-number k, A-scans data with resolved complex conjugate artifact were obtained by a direct inverse Fourier transformation (IFT) from direct DAQ sampling data without any re-sampling process.

The image size is 5x2 mm^2 with 900x500 pixels. The image depth shown in **Figure 5(a)** is slightly larger than that in **Figure 5(b)**, but the image is blurrier in **Figure 5(a)** than that in **Figure 5(b)**, which taken by the probe with larger depth of field and spot size. The image shown in **Figure 5(b)** has higher resolution than that in **Figure 5(a)**, which can be seen clearly with finer structures in layer of epidermis (grey arrow), sweat gland (white arrow), and blood vessel in subcutis layer (black arrow).

4. CONCLUSIONS

We presented a design, construction and beam profile characterization of different variations of graded-index (GRIN) and ball fiber lenses, which were recently proposed for ultra-small OCT probes. Those fiber lens modules were made of single mode fibers and GRIN and ball fiber lenses with/without fiber spacers between them. We used fusion-splicing in between the fibers, lenses and spacers to ensure high quality light transmission. We found that beam-distance profiles (i.e. 0.4 - 1.2 mm of focus distance, 0.8 – 1.5 mm of depth of field, and 26 – 35 μm of spot size) can be obtained by precisely adjusting the lengths of the fiber spacer and the GRIN fiber lens or diameter of the ball lens for the different tissue imaging in human body. Using ZEMAX, optical design software, we modeled our optic probes which proved a precise approach. We obtained very high quality focused Gaussian beam profiles with high x and y symmetry using the conventional multi-mode GRIN fibers and home-made fiber ball lenses. The OCT images shown in this paper indicated the important impact of fiber lens on the image quality. The high quality beam and ultra-small size make such fiber lens based probes very valuable for optical coherence tomography systems.

REFERENCES

[1] Huang, D., Swanson, E.A., Lin, C.P., Schuman, J.S., Stinson, W.G., Chang, W., Hee, M.R., Flotte, T., Gregory, K., Puliafito, C.A. and Fujimoto, J.G. (1991) Optical coherence tomography. *Science* **254**, 1178-1181.

[2] Chen, Z.P., Milner, T.E., Dave, D. and Nelson, J.S. (1997) Optical doppler tomographic image of fluid flow velocity in highly scattering media. *Opt. Lett.* **22**, 64-66.

[3] Izatt, J.A, Kulkarni, M.D., Yazdanfar, S., Barton, J.K., and Welch, A.J. (1997) *In vivo* bidirectional color Doppler flow imaging of picoliter blood volumes using optical coherernce tomography. *Opt. Lett.* **22**, 1439-1441.

[4] Yazdanfar, S., Rollins, A.M. and Izatt, J.A. (2001) Ultra-high-velocity resolution imaging of the microcirculation in-vivo using color Doppler optical coherence tomography. *Proc. SPIE*, **4251**, 156-164.

[5] Yang, V.X.D., Gordon, M.L., Qi, B., Pekar, J., Lo, S., Seng-Yue, E., Mok, A., Wilson, B.C. and Vitkin, I.A., (2003) High speed, wide velocity dynamic range doppler optical coherence tomography (Part I): system design,

signal processing, and performance. *Opt. Express,* **11**, 794-809.

[6] White, B.R., Pierce, M.C., Nassif, N., Cense, B., Park, B.H, Tearney, G.J., Bouma, B.E., Chen, T.C. and de Boer J.F., (2003) *In vivo* dynamic human retinal blood flow imaging using ultra-high-speed spectral domain optical coherence tomography. *Opt. Express*, **11**, 3490-3497.

[7] Yazdanfar, S., Rollins, A.M. and Izatt, J.A., (2000) Imaging and velocimetry of the human retinal circulation with color Doppler optical coherence tomography. *Opt. Lett.*, **25**, 1448-1450.

[8] Zhao, Y.H., Chen, Z.P., Saxer, C., Xiang, S.H., de Boer, J.F. and Nelson, J.S. (2000) Phase-resolved optical coherence tomography and optical Doppler tomography for imaging blood flow in human skin with fast scanning speed and high velocity sensitivity. *Opt. Lett.,* **25**, 114-116.

[9] Zhao, Y., Chen, Z.P., Saxer, C., Shen, Q., Xiang, S., de Boer, J.F. and Nelson, J.S. (2000) Doppler standard deviation imaging for clinical monitoring of *in vivo* human skin blood flow. *Opt. Lett.,* **25**, 1358-1360.

[10] Yang, V.X.D., Tang, S., Gordon, M.L., Qi, B., Gardiner, G., Cirocco, M., Kortan, P., Haber, G., Kandel, G., Vitkin, I.A. and Wilson, B.C. (2005) Endoscopic doppler optical coherence tomography in human gastrointestinal tract: initial experience. *Gastrointest. Endosc.,* **61**, 879-890.

[11] Tran, P.H., Mukai, D.S., Brenner, M. and Chen, Z. (2004) *In vivo* endoscopic optical coherence tomography by use of a rotational microelectromechanical system probe. *Opt. Lett.,* **29**, 1236-1238.

[12] Li, X., Chudoba, C., Ko, T., Pitris, C. and Fujimoto, J. G. (2000) Imaging needle for optical coherence tomography. *Opt. Lett.*, **25**, 1520-1522.

[13] Yang, V.X.D., Mao, Y.X., Munce, N., Standish, B., Kucharczyk, W., Marcon, N.E., Wilson, B.C. and Vitkin, I.A. (2005) Interstitial doppler optical coherence tomography. *Opt. Lett.*, **30**, 1791-1793.

[14] Fujimoto, J.G., Bopart, S.A., Tearney, G.J., Bouma, C.B., Pitris, E. and Brezinski, M.E. (1999) High resolution in vivo intra-arterial imaging with optical coherence tomography. *Heart*, **82**, 128-133.

[15] Diaz-Sandoval, L.J., Bouma, B.E., Tearnay, G.J. and Jang, I. (2005) Optical coherence tomography as a tool for percutaneous coronary interventions. *Catheter. And Cardio. Interventions*, **65**, 492-496.

[16] Swanson, E., Petersen, C.L., McNamara, E., Lamport, R.B. and Kelly, D.L. (2002) Ultra-small optical probes, imaging optics, and methods for using same. *U.S. Patent,* 6445939.

[17] Shishkov, M., Bouma, B.E. and Tearney, (2006) G.J. System and method for optical coherence tomography. *U.S. Patent,* 20060067620A1.

[18] Reed, W.A., Yan, M.F. and Schnitzer, M.J. (2002) Gradient-index fiber-optic microprobes for minimally invasive *in vivo* low-coherence interferometry. *Opt. Lett.,* **27**, 1794-1796.

[19] Jafri, M.S., Farhang, S, Tang, R.S., Desai, N., Fishman, P.S., Rohwer, R.G., Tang, C. and Schmitt, J.M. (2005) Optical coherence tomography in the diagnosis and treatment of neurological disorders. *J. Biomed. Opt.* **10(5)**, 051603.

[20] Li, H., Standish, B.A., Mariampillai, A., Munce, N.R., Mao, Y., Chiu, S., Marcon, N.E., Wilson, B.C., Vitkin, A. and Yang, V.X.D. (2006) Feasibility of interstitial doppler optical coherence tomography for *in vivo* detection of microvascular changes during photodynamic therapy. *Lasers in Surgery and Medicine,* **38**, 754-761.

[21] Sainter, A.W., King, T.A. and Dickinson, M.R., (2002) Theoretical comparison of light sources for use in optical coherence tomography. *Proc. SPIE*, **4619**, 289-299.

[22] Mao, Y., Chang, S., Sherif, S. and Flueraru, C. (2007) Graded-index fiber lens proposed for ultrasmall probes used in biomedical imaging. *App. Opt.*, **46**, 5887-5894.

[23] Kogelnik, H.W. and Li, T. (1966) Laser beam and resonators. *Appl. Opt.*, **5**, 1550.

[24] Emkey, W.L. and Jack, C.A. (1987) Analysis and evaluation of graded-index fiber-lenses. *J. of Lightwave Tech.*, LT-**5**, 156-1164.

[25] Mao, Y., Sherif, S., Flueraru, C. and Changm S. (2008) 3x3 mach-zehnder Interferometer with unbalanced differential detection for full range swept-source optical coherence tomography. *Appl. Opt.,* **47**, 2004-2010.

How cross screw length influences the stiffness of intramedullary nail systems

S. V. Karuppiah[1], A. J. Johnstone[1], D. E. T. Shepherd[2]

[1]Orthopaedic Unit, Aberdeen Royal Infirmary, Aberdeen, UK;
[2]School of Mechanical Engineering, University of Birmingham, Edgbaston, Birmingham, UK.
Email: saravanavail@yahoo.com

ABSTRACT

Fractures of long bones are commonly treated with intramedullary (IM) nails and they have been shown to have a very high success rate. Recently we have concerns with the use of the newer IM nailing systems, that uses longer cross screws, which have been developed with variation in implant designs. We believe that the newer implants provide less fracture stability, due to decreased stiffness of the IM nailing system. The aim of this study was to biomechanically determine the influence of the length of cross screw on the stiffness of the IM nailing system, using a composite model. Our test results confirmed our suspicion that the newer IM nailing system using longer cross screw-length is less stiff than traditional nailing systems using shorter cross screw length, during axial loading.

Keywords: Intramedullary Nail; Mechanical Testing; Screw; Stiffness

1. INTRODUCTION

Orthopaedic implants have been popular in the management of fractures for the last six decades. There are various types of implants available for fixation of bone fractures, intramedullary (IM) nails being one of them. These are popular for the fixation of diaphyseal (shaft) fractures of the femur, tibia and humerus with high success rates. It is one of the accepted and most widely used methods of treating, transverse and short oblique, axially stable fractures of the femoral diaphysis [1,2,3,4]. It provides an excellent and prompt return of function with a low rate of complication and non-unions [5].

With the introduction of inter locking screws (cross screws) it has been also possible to provide rotational stability and precise reconstruction of the anatomy in communited fractures using the IM nailing system. The IM nailing system provide stiffness and stability of the bone segments, all factors which are considered to be essential preconditions for its increased success of bone union, while maintaining the function of the joints and soft tissues [6,7]. From time to time further changes in the design of the implant has been made, including the placement of the cross screws, to extend the use of femoral IM nails in all types of femoral fractures.

There is recent concern in our clinical practice regarding the use of newer femoral IM nailing system, using longer cross screw-length, designed to accommodate in extreme distal femoral metaphyseal fractures as compared to the traditional system which were designed to be used in diaphyseal fractures.

The anatomy of the human distal femoral bone is wider in diameter compared to the narrow diaphyseal bone. In order to accommodate the distal wider femoral bone fragment, the newer intramedullary nailing system uses longer cross screw-length that could be threaded from one cortex to the other.

The concept of stability is crucial in fracture surgery. Stability determines the amount of strain (i.e. relative change in fracture gap) at the fracture site, and strain determines the type of healing. Excessive strain could result in delayed or non-union of the fractures [8]. We believe that the longer cross screw-length, used in the newer IM nailing system, provides less stability to the fracture fixation due to decreased stiffness of the IM nailing system. The aim of this study was to determine the influence of the cross screw length on the stiffness of fracture fixation when using IM nailing systems.

2. MATERIALS AND METHODS

2.1. Materials

The intramedullary nails were simulated using stainless steel tubes with a length of 150 mm, outside diameter of 12 mm and inside diameter of 10 mm. These dimensions were similar to intramedullary nails used surgically. The distal hole in the intramedullary nail was 10 mm from the distal end. The distal bone end was simulated using stainless steel cylinders. In the study three different cylinders were used to represent different parts of the femur, with the dimensions comparable to clinical measure
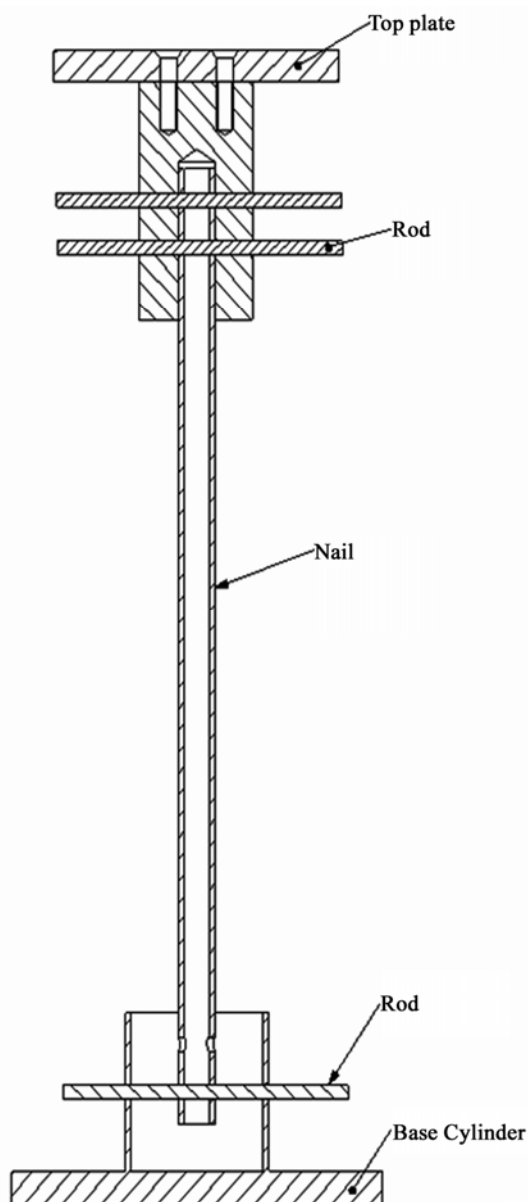
Figure 1. Test set-up.

ments using x-rays:

- outside diameter 50 mm, wall thickness 5 mm, height 50 mm (proximal diaphyses bone with a narrow canal and a thick cortical wall);
- outside diameter 75 mm, wall thickness 3 mm, height 50 mm (metaphyseal junction);
- outside diameter 100 mm, wall thickness 3 mm, height 50 mm (distal condylar bone with a wide canal and a thin cortical wall).

The hole in the cylinder was 25 mm from the top end of the cylinder.

2.2. Methods

An Instron 1822 materials testing machine (Instron Ltd.,

High Wycombe, UK) was used for mechanically testing the intramedullary nails. A customised stainless steel clamp was designed to connect the proximal portion of the intramedullary nail to the cross-head of the testing machine. The stainless steel cylinder, which represented the femur, was attached to the base of the testing machine. The distal end of the intramedullary nail was secured to the cylinders with a single stainless steel rod of diameter 5 mm; this represented the cross screw. The test set-up is shown in (**Figure 1**).

The intramedullary nail was subjected to an axial compressive load, by lowering the cross-head of the testing machine; the load was applied at a rate of 0.05 N/s. During the testing load and displacement was re-
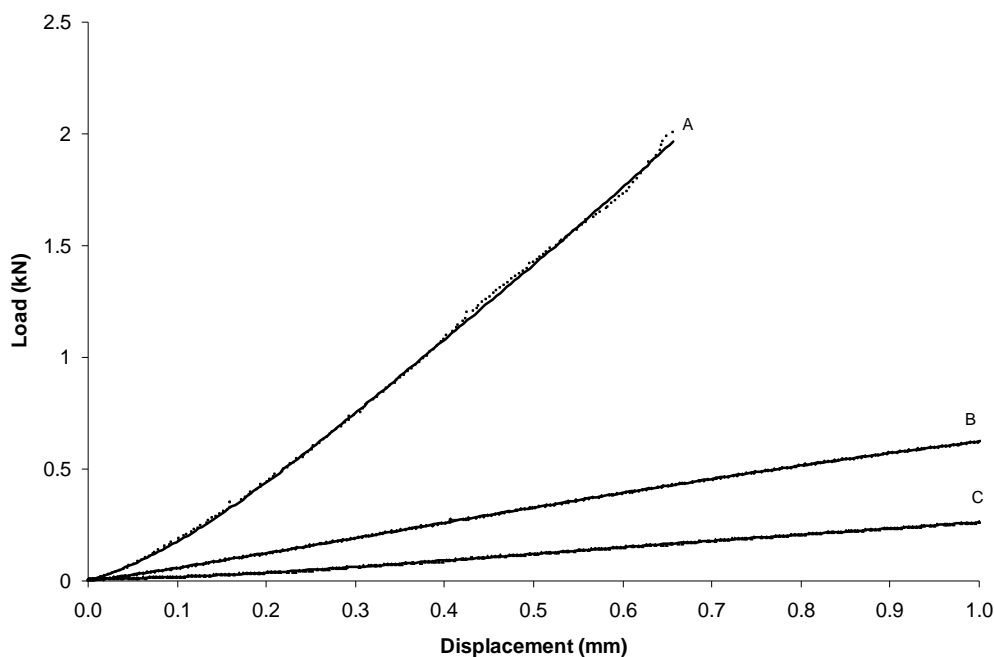
Figure 2. Graphs of load against displacement for a cylinder of outside diameter: A) 50 mm. Curve fit $y=7.1293x^4 - 11.956x^3+7.7119x^2 + 1.0647x+0.0035$, $R^2 = 0.9995$; B) 75 mm. Curve fit $y=0.1341x^4-0.4407x^3+ 0.3588x^2+0.5726x-0.0012$, $R^2 = 0.9999$; C) 100 mm Curve fit $y = 0.2709x^4-0.7248x^3 + 0.6727x^2+0.0352x + 0.0091$, $R^2 = 0.9998$.

corded. Testing continued until a maximum displacement of 1 mm or the maximum force of 2 kN was reached. A displacement of 1 mm was chosen; as the rod and the cylinder underwent mechanical deformation beyond 1mm. A load of 2 kN represents about three times body weight for a 70 kg individual [9]. Each test configuration was tested three times.

Graphs of load against displacement were plotted (**Figure 2** shows a typical example) and a forth-order polynomial fitted to the results. This polynomial fit enables us to use the same fit for all data. The stiffness was then determined from the gradient of the graph at a displacement of 0.5 mm.

3. RESULTS

The mean stiffness of the intramedullary nail system decreases with increasing diameter of the cylinder, which represents the femoral bone. The mean (± standard deviation) stiffness of the system was 3298 ± 144 N/mm, 657 ± 10 N/mm and 297 ± 16 N/mm at diameters of 50 mm, 75 mm and 100 mm, respectively.

4. DISCUSSION

The principle aim of fracture fixation is to provide stability of the bone fragments and restoration of normal anatomy. Intramedullary nails act as a scaffold and are devised to hold together the proximal and distal ends of the long bone for a conducive environment for fracture healing. The placement of the cross screw and the design

of the implant can influence the stability of fracture fixation and hence bone union [10,11].

The proximal (diaphyseal) femoral canal is narrow in which the intramedullary nail has a 'snug fit' against the bone. Hence, the inter bone-nail distance is minimal or none. In the current study it has been shown that with this set-up it has a mean stiffness of 3298 N/mm. With increasing bone diameter, the stiffness of the system decreases, leading to an increased deflection for a given load. At the distal condylar part of the femur, there is a wide canal and a thin cortex. The distance of the cross screw from the bone is increased in this case.

In the clinical situation, when the patient bears weight the greater the diameter of the canal, the greater the deflection will be. As shown in this experiment, there is a proportionate decrease in axial stiffness of the intramedullary nail with increasing cross screw-length. Longer cross screw-length, used in the newer intramedullary nailing system to treat distal femoral fractures, will be far less stiff during weight bearing than the traditional IM nailing system, which uses shorter cross screw-lengths. The newer IM nails hence provide less stability of fracture fixation and may potentially cause delayed union or non-union when used to treat femoral fractures.

REFERENCES

[1] Wolinsky, P.R., McCarty, E., Shyr, Y. and Johnson, K. (1999) Reamed intramedullary nailing of the femur: 551 cases. *Journal of Trauma-Injury Infec-*

tion & Critical Care, **46**, 392-399.

[2] Wiss, D.A. and Stetson, W.B. (1995) Unstable fractures of the tibia treated with a reamed intramedullary interlocking nail. *Clinical Orthopaedics & Related Research*, 56-63.

[3] Rommens, P.M, Verbruggen, J. and Broos, P.L. (1995) Retrograde locked nailing of humeral shaft fractures. A review of 39 patients. see comment. *Journal of Bone & Joint Surgery - British Volume,* **77**, 84-89.

[4] Bick, E.M. (1968) The intramedullary nailing of fractures by G. Kuntscher. Translation of article in Archiv fur Klinische Chirurgie. *Clinical Orthopaedics & Related Research*, **60**, 5-12.

[5] Schatzker, J. (1998) Fractures of the distal femur revisited. *Clinical Orthopaedics & Related Research*, 43-56.

[6] Brumback, R.J., Toal, T.R., Murphy-Zane, M.S., Novak, V.P. and Belkoff, S.M. (1999) Immediate weight-bearing after treatment of a comminuted fracture of the femoral shaft with a statically locked intramedullary nail. *Journal of Bone & Joint Surgery – Am,* **81**, 1538-1544.

[7] Hente, R., Fuchtmeier, B., Schlegel, U., Ernstberger, A. and Perren, S.M. (2004) The influence of cyclic compression and distraction on the healing of experimental tibial fracture. *Journal of Orthopaedic Research*, **22**, 709-715

[8] Taylor, S.J., Walker, P.S., Perry, J.S., Cannon, S.R. and Woledge, R. (1998) The forces in the distal femur and the knee during walking and other activities measured by telemetry. *J.Arthroplasty,* **13**, 428-437.

[9] Schandelmaier, P., Farouk, O., Krettek, C., Reimers, N., Mannss, J. and Tscherne, H. (2000) Biomechanics of femoral interlocking nails. *Injury,* **31**, 437-443.

[10] Henley, M.B., Meier, M. and Tencer, A.F. (1993) Influences of some design parameters on the biomechanics of the unreamed tibial intramedullary nail. *J Orthop Trauma,* **7**, 311-319.

Evaluation of EEG β_2 / θ -ratio and channel locations in measuring anesthesia depth

Zhi-Bin Tan[1], Le-Yi Wang[1], George McKelvey[2], Aliaksei Pustavoitau[2], Guang-Xiang Yu[2], Harold Michael Marsh[2], Hong Wang[2]

[1]Department of Electrical and Computer Engineering, Wayne State University, Detroit, USA;
[2]Department of Anesthesiology, Wayne State University, Detroit, USA.
Email: au6063@wayne.edu; lywang@wayne.edu; gmckelve@med.wayne.edu; yuguangxiang@gmail.com; hmarsh@med.wayne.edu; howang@med.wayne.edu

ABSTRACT

In this paper, the ratio of powers in the frequency bands of β_2 and θ waves in EEG signals (termed as the β_2 / θ -ratio) was introduced as a potential enhancement in measuring anesthesia depth. The β_2 / θ -ratio was compared to the relative β -ratio which had been commercially used in the BIS monitor. Sensitivity and reliability of the β_2 / θ -ratio and EEG measurement locations were analyzed for their effectiveness in measuring anesthesia depth during different stages of propofol induced anesthesia (awake, induction, maintenance, and emergence). The analysis indicated that 1) the relative β -ratio and β_2 / θ -ratio derived from the prefrontal, frontal, and the central cortex EEG signals were of substantial sensitivity for capturing anesthesia depth changes. 2) Certain channel positions in the frontal part of the cortex, such as $F4$, had the combined benefits of substantial sensitivity and noise resistance. 3) The β_2 / θ -ratio captured the initial excitation, while the relative β -ratio did not. 4) In the maintenance and emergence stages, the β_2 / θ -ratio showed improved reliability. Implications: The ratio of powers in EEG frequency bands β_2 and θ derived from the frontal cortex EEG channels has combined benefits of substantial sensitivity and noise resistance in measuring anesthesia depth.

Keywords: Anesthesia Depth; EEG (Electroencephalogram); EEG Channels; β_2/θ -Ratio; Relative β -Ratio

1. INTRODUCTION

In this paper, the ratio of powers in the frequency bands of β_2 and θ waves in EEG signals (termed as the β_2/θ -ratio) was introduced and evaluated as a potential enhancement in measuring anesthesia depth, in comparison to the relative β -ratio which had been commercially used in the BIS (Bi-spectrum Index) monitor (Aspect Medical Systems). Sensitivity and reliability of the β_2/θ -ratio and EEG measurement locations were analyzed for their effectiveness in measuring anesthesia depth during four stages of propofol induced anesthesia (awake, induction, maintenance, and emergence).

Since the physiologic effects of anesthetic agents in ether anesthesia were observed by John Snow in 1847 [1], characterizing, measuring, and continuously monitoring anesthesia depth have been pursued extensively. Accurate monitoring of anesthesia depth can help to avoid overdose of anesthetic agents, prevent intraoperative awareness, and assist the anesthesiologist in anesthesia decisions and management. Case studies have also indicated that objective monitoring of the anesthesia depth can guide more precise administration of anesthesia agents, and consequently can potentially reduce drug costs, expedite post-anesthesia recovery, and shorten hospital stay [2,3].

With the central nervous system (CNS) being the target of anesthesia drugs, the electroencephalogram (EEG) signal processing has naturally become the focus for anesthesia depth monitoring [4,5]. The goal of all these EEG processing methods was to generate some parameters or scales, collectively called quantitative EEG "indices," that were clinically reliable as indicators for anesthesia depth. It was widely believed that anesthetics had effects on the EEG in multiple aspects: such as amplitude, frequency, phase relation, frequency band power transition, etc., [14]. Individual indicators, such as spectral edge frequency, median frequency, band power ratio, etc., demonstrated different levels of capability, but individually did not provide completely reliable descriptors of anesthesia depth [8]. In addition, their sensitivity and reliability were influenced significantly by the EEG signal channels and anesthesia
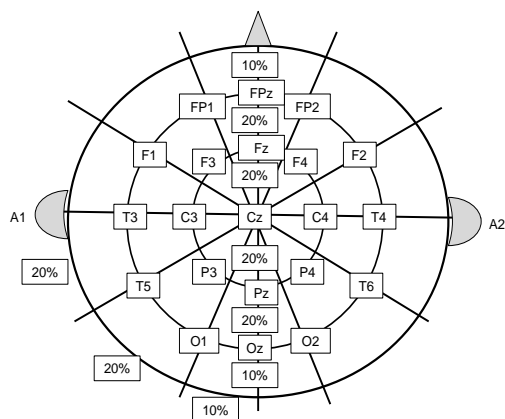
Figure 1. Electrodes placement according to the 10-20 system. Electrode positions are denoted with odd numbers for left electrodes, even numbers for right, z for the midline, F_p the prefrontal, F the frontal, C the central, T the temporal, P the parietal, O the occipital, and A the auricular.
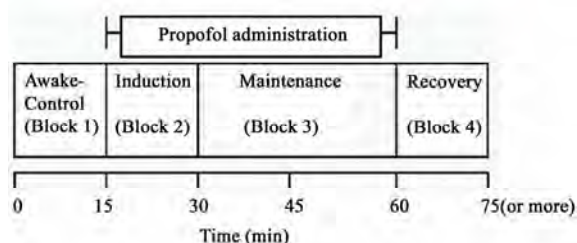
Figure 2. Propofol administration and time interval for each anesthesia stage.

stages. As a result, investigation of EEG parameters and their combined utilities, EEG channel locations, and anesthesia stages can potentially enhance our understanding of anesthesia depth monitoring and achieve a better description of brain activities during anesthesia.

Five frequency bands were frequently identified for the EEG signal: delta band or δ ($0.5-3.5$ Hz), theta band or θ ($3.5-7$ Hz), alpha band or α ($7-13$ Hz), beta-1 band or β_1 ($13-30$ Hz), and beta-2 band or β_2 ($30-50$ Hz) [2,17]. All five bands were influenced by anesthesia agents [17]. The relative β-ratio was one of the main parameters that were used jointly to produce the BIS index in the BIS Monitor. It was defined as $\log(P_{30-47}/P_{11-22})$ [7], where P_{x-y} denoted the average spectral power in the frequency band from x to y in Hz.

In this paper, the β_2/θ-ratio was evaluated as a potential candidate for enhancing depth measurements. Data analysis was performed to evaluate benefits and limitations of the β_2/θ-ratio and EEG electrode locations in relation to reliability in propofol induced

anesthesia depth measurements over different stages.

2. METHODS

In this section, the rationale of introducing the β_2/θ-ratio for anesthesia depth measurements was explained. Then the methods of experiment setup, data acquisition, signal preprocessing, and data analysis were described.

2.1. The β_2/θ-Ratio

It is well understood that the state of human awareness is associated with increased power in the higher frequency bands (β and β_2) and decreased power in the lower frequency bands (θ and δ bands). Consequently, it is a sensible choice of using power ratios of high power ranges to low power ranges as indicators of anesthesia depth.

However, sensitivity of band powers to awareness and alertness varies significantly. Dressler et al. introduced a measure of discriminating capability for awareness and alertness indications [5]. It employed a re-mapped predicting probability, denoted by rP_k, as an indicator for sensitivity of different frequency bins on awareness. The higher the rP_k is, the better the discriminating power of the frequency band has. In particular, the average rP_k value within θ band ($3.5-7$ Hz) was shown to be much higher than that of the band between 15 to 20 Hz, which was a major part of the band ($11-22$ Hz) used in the BIS monitor.

Based on this observation, one objective of this study was to introduce a different band power ratio: the ratio of the β_2 and θ powers, which was defined as $\log(P_{30-47}/P_{3.5-7})$. This is termed as *the β_2/θ-ratio*. There are several potential advantages to this method:

1) The β_2/θ-ratio demonstrated more sensitivity to changes in awakeness.

2) While EMG (electromyography) signals were often considered as the main source of artifacts in EEG signals, during anesthesia depth measurements EMG signals could be a good indicator of awareness. In the frequency band over 30 Hz, the EMG signal became more dominant than EEG signals when the subjects became responsive to environments. As a result, in the BIS system and Datex Entropy Module, the frequency band dominated by EMG was used to enhance sensitivity in the development of the depth indicators [7,18].

One adversary impact of this approach was that EMG frequencies extended to the alpha band, which was covered in the relative β-ratio. This may reduce sensitivity since the EMG increased power concentration in both P_{30-47} and P_{11-22}. In contrast, the EMG had
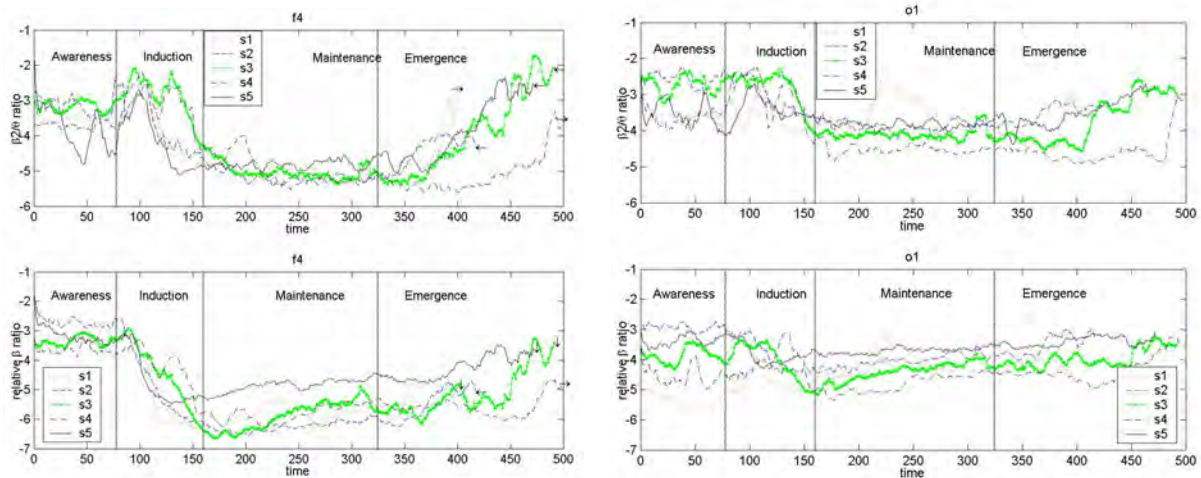
Figure 3. Comparison of the β_2/θ-ratio $\log(P_{30-47}/P_{3.5-7})$ and relative β-ratio $\log(P_{30-47}/P_{11-22})$ between two different channel locations: Left plot for the $F4$ channel and right plot for the $O1$ channel; top for the β_2/θ-ratio and bottom for the relative β-ratio. Five subjects were coded as S1-S5 in the legends. S3 trajectories were highlighted to show trajectories of one subject. The same parameters from $F4$ were more sensitive than those from $O1$.
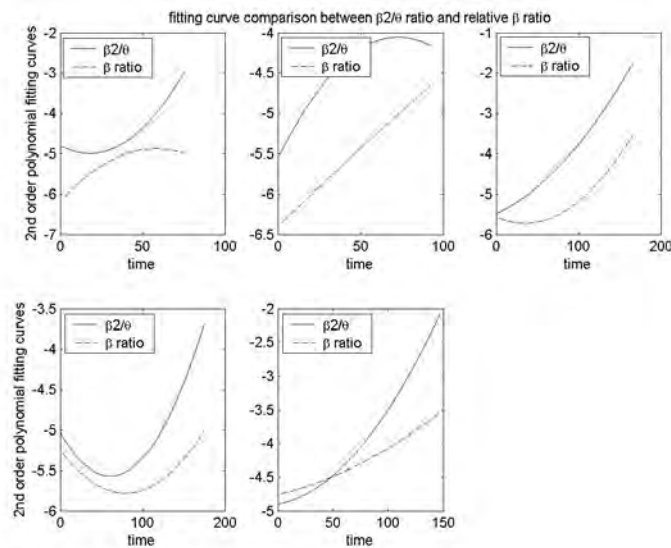
Figure 4. Comparison of the trend patterns between the β_2/θ-ratio $\log(P_{30-47}/P_{3.5-7})$ and relative β-ratio $\log(P_{30-47}/P_{11-22})$ in the emergence stage of the 5 subjects (coded as $S1-S5$). The slopes of the β_2/θ-ratio curves were steeper than those of the relative β-ratio curves in this stage.

little effect in the θ range. As the θ power was used in the β_2/θ-ratio, this could become more effective than the relative β-ratio.

2.2. Data Collection and Analysis

2.2.1. Data Acquisition

EEG signals from 16 channels were recorded by using the 16-channel Nolan Mindset-16 EEG data acquisition equipment (Nolan Computer Systems, L.L.C.). Each subject worn an electrode cap with electrodes arranged according to the international 10-20 system, see **Figure 1**. The subjects were fit with one of the three sizes of the caps. For improved electrode contact and impedance, each electrode adaptor on the cap was injected with a chloride-free gel of electrolyte. The electrodes were then

Table 1. Averaged parameter ranges of different channels during the induction and emergence stages.

	induction	emergence
Fp1	226.09	129.19
Fp2	213.42	126.26
F7	134.28	75.81
F3	102.62	149.19
F4	*289.76*	*153.22*
F8	208.33	116.68
T3	40.27	16.55
C3	120.97	46.85
C4	115.19	47.81
T4	52.45	22.97
T5	22.85	15.97
P3	61.25	27.79
P4	68.13	27.67
T6	39.20	23.54
O1	62.88	26.32
O2	54.83	27.35

Table 2. Averaged SNRs(dB) of different channels in different stages.

	awake	induction	maintenance	emergence
Fp1	6.45	14.03	31.90	22.51
Fp2	8.19	14.41	33.64	19.17
F7	10.32	14.01	30.41	17.14
F3	10.74	13.79	33.38	23.54
F4	*14.54*	*16.44*	*31.36*	*27.86*
F8	10.83	15.86	30.71	18.03
T3	13.71	11.94	29.09	11.52
C3	13.47	14.06	30.32	18.50
C4	16.07	15.42	31.42	18.73
T4	13.29	13.24	33.88	12.43
T5	15.92	12.29	29.07	9.77
P3	17.09	13.87	31.53	15.68
P4	19.68	14.96	31.43	15.15
T6	17.81	13.05	29.78	11.72
O1	16.98	11.20	29.05	11.17
O2	17.53	12.09	27.10	11.10

connected to the adaptor. The machine performed an impedance test first and was ready for recording after the test. Nolan Mindset was connected to a host computer system. The Nolan software runed on the host computer and recorded simultaneously the 16 channel EEG signals with a time reference. Although the Nolan software was capable of performing limited data analysis and display, our data analysis was performed with special programed

algorithms.

The 16 channels were divided equally for the left side ($Fp1$, $F3$, $F7$, $C3$, $T3$, $T5$, $P3$, $O1$) and the right side ($Fp2$, $F4$, $F8$, $C4$, $T4$, $T6$, $P4$, $O2$). The reference montage was used. As a result, the additional reference electrodes were placed with $A1$ (near the left ear) as the reference for the left-side electrodes and $A2$ (near the right ear) as the reference for the right-side electrodes.

2.2.2. Subjects and Anesthesia

The results presented in this paper were based on the EEG recordings from 5 young healthy male volunteers. As a pilot study, the sample size was small. The study was conducted in the Receiving Hospital, Detroit, Michigan, USA, and received institutional approval. All subjects were explained of the nature of the study and consenting participants.

The BIS values were monitored by the BIS monitor. Other physiological vital signs (blood pressures, heart rate, oxygen saturation, etc.) were continuously monitored by the anesthesia monitor (S-5 Anesthesia Monitor by Datex-Ohmeda, Inc.) during the entire process.

Propofol infusion rates ranged from $170-200$ $\mu g/kg/min$. The data collection procedure was divided into four separated stages, which are shown in **Figure 2**:

1) Awake Stage[1] (15 minutes):
The subject was conscious and instructed to close their eyes and be calm and inactive. Facial and body movements were observed in this stage. No anesthesia drugs were administered.

2) Induction Stage (15 minutes):
Propofol infusion started at the beginning of the induction stage. During this stage, propofol infusion rates were adjusted to achieve a BIS value to the desired levels (between 30 and 50). Towards the end of this 15-minute period, the BIS values of all the subjects became stable. This stage was characterized by substantial changes of anesthesia depth towards its steady-state values. Occasional facial and body movements occurred.

3) Maintenance Stage (30 minutes):
Propofol infusion was maintained for 30 minutes to sustain the desired BIS level and depth of anesthesia. This stage was characterized by relatively stable BIS values, no drug rate adjustment, and no body movements.

4) Emergence Stage (15 minutes or longer):
This stage started when propofol administration ceased with the subject gradually recovering to become awake. Due to differences in recovery speed, the duration

---

[1]It was sometimes called the control stage. For terminology consistency it was called awake stage throughout this paper, as shown in **Figure 2.**
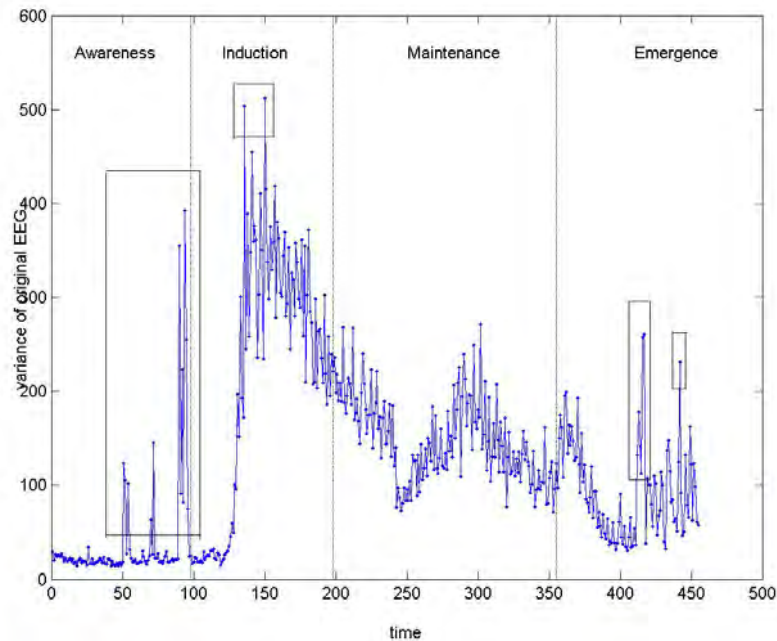
Figure 5. Variances of powers in the original EEG epochs in the Fp1 channel. Variance spikes due to artifact contamination in some epochs are indicated in rectangular frames.

varied. Body movements became gradually apparent in the recovery process of this stage.

2.2.3. Signal Preprocessing

The raw EEG signals were digitized with sampling frequency 256 Hz. The 60 Hz power contamination was visible in the recorded EEG signals. To reduce noise effects, the original EEG data were manually cleared of highly visible artifacts (eye movements, body movements, equipment disturbance, cable movements, etc.). Then, a low-pass filter with cutoff frequency of 47 Hz was designed to filter out the 60 Hz power line disturbances before data analysis.

2.2.4. Data Analysis

EEG epochs of 2560 data points (10 sec.) were used for generating one parameter point of the β_2/θ-ratio (and the relative β-ratio) as follows: the 10-second interval was divided into 4 overlapping segments of 4 seconds each: $[0,4],[2,6],[4,8],[6,10]$. The spectrum of each segment was estimated by Welch's method [19]. The resulting spectra of the four segments were averaged to generate one spectrum. This approach reduced zero-mean independent random sensor noises.

Then, the powers of P_{30-47}, $P_{3.5-7}$ and P_{11-22} were extracted from the resulting spectra to form a value point of the β_2/θ-ratio (and similarly the relative β-ratio) for the EEG epoch. This process was repeated for the entire EEG recording, except for epochs that were removed due to visible artifacts. To further reduce

random fluctuations, the β_2/θ-ratios and relative β-ratios over a moving window of length 10 parameter points were averaged to produce the final data points for analysis.

3. RESULTS

The trajectories of the β_2/θ-ratio $\log(P_{30-47}/P_{3.5-7})$ and the relative β-ratio $\log(P_{30-47}/P_{11-22})$ from the $F4$ EEG electrode for 5 patients were plotted in **Figure 3**, where a patient was only indexed by a code such as $S1$, with data from one patient (code $S3$) highlighted with green color. Both $\log(P_{30-47}/P_{3.5-7})$ and $\log(P_{30-47}/P_{11-22})$ were negative values, as shown in the y-axis of the plots. While the awake, induction, and maintenance stages had fixed lengths, the duration of the emergence stage was variable, with an arrow showing the time when the subject became fully awake.

The trajectories were divided by the four stages marked by "awareness (for awake stage)," "induction," "maintenance," and "emergence." Due to the removal of the corrupted and other unusable data points, the length of each stage was slightly shorter than its designated interval.

The following results were derived from the data.

3.1. Initial Depth Surge Detection

It was common that a patient responded to initial

anesthesia infusion with a surge of excitement for a short time. Capturing this initial phase of response was an indication of causal dependence of the parameter on anesthesia depth. **Figure 3** showed that the β_2/θ - ratio demonstrated an initial surge in each patient in induction, while neither the relative β -ratio nor the BIS did.

3.2. Sensitivity of EEG Parameters to Anesthesia Depth Changes

For the induction and emergence stages in which anesthesia depth changed greatly, the sensitivity of a parameter to depth change was evaluated by either the difference between its maximum and minimum values or by the slope of the parameter trajectory during the stage. This was accomplished by extracting the trend patterns using the least-squares curve fitting with a second-order polynomial model. The trend patterns of the β_2/θ - ratio and relative β -ratio in the emergence stage were showed in **Figure 4**. The larger slopes of the β_2/θ -ratio curves implied that the β_2/θ -ratio was more sensitive to the depth change during the recovery process than the relative β -ratio.

3.3. Sensitivity of EEG Signals from Different Channels to Anesthesia Depth Changes

It was visually apparent from **Figure 3** that the parameters derived from $F4$ responded to depth changes more pronounced than those from $O1$. This was further verified by comparing the sensitivity of EEG parameters from the frontal region to that of the posterior region. **Table 1** listed the parameter ranges in the induction and emergence stages (the awake and maintenance stages are irrelevant since they have nearly constant depths) in all 16 channels. The entries of the table were calculated by $range = maximum - minimum$ of β_2/θ -ratio in the respective stages. Then the ranges were averaged over the five subjects. In both stages, the highest sensitivities occurred at the front channels, especially the $F4$ channel.

3.4. Noise Resistance of EEG Channels

The amplitudes of EEG signals were noticeably lower than noises. When an epoch of the EEG signal was contaminated by artifacts, the variance of the power in this epoch changed markedly from the average of recent previous ones. This artifact detection method was used in the Narcotrend monitor of anesthesia depth [20]. **Figure 5** illustrated the variance spikes caused by artifacts, from one subject and channel $Fp1$.

To quantify the noise resistance capability of different EEG recording channels, the signal-to-noise ratio was used

$$SNR = \frac{Noise\ Free\ Signal\ Power}{Noise\ Power}$$

in which the powers were calculated over one given stage. The larger the SNR, the stronger the ability of the channel was to resist noise. **Table 2** detailed the SNRs of the 16 channels in each stage, averaged over the five subjects.

In the awake stage, patients were alert with eye and facial movements, leading to very small SNRs. In the maintenance stage (in deep anesthesia), SNRs were high across all recording channels. These two stages were not essential in comparison. In the induction and emergence stages, in which anesthesia depth changed most dramatically, the SNRs in the frontal channels, especially the $F4$ channel, had the highest value of all the 16 channels monitored.

4. DISCUSSIONS

A variety of methods and commercial devices for measuring the depth of anesthesia based on EEG signals have been developed [6,7,8,9,10,11,12,13]. At present, the underlying mechanism of effects of anesthesia agents on the CNS is not well understood. The main approaches of EEG signal processing utilized empirical methods to relate EEG parameters to the drug effects by experiments and statistical analysis. Currently, there are a few anesthesia depth monitors in the market, such as the BIS monitor (Aspect Medical Systems, Inc.) and the Entropy Module (GE Medical Systems, Inc.) [14]. Both monitors used Pre-frontal EEG signals. These monitors provided quite reliable monitoring capability in deep anesthesia. On the other hand, due to disturbances from facial and body movements, their reliability during induction and recovery stages and in ICU (intensive care units) applications remained to be enhanced [15,16].

4.1. Summary of Main Findings

This study was focused on potential utility of the β_2/θ -ratio in improvement of anesthesia depth monitoring. Due to its close similarity to the relative β -ratio which was commercially used in the BIS monitor, our study was focused on characteristic comparison between the β_2/θ -ratio and the relative β -ratio. In particular, it highlighted the following preliminary findings: 1) The relative β -ratio and β_2/θ -ratio derived from the prefrontal, frontal, and the central cortex EEG signals were of substantial sensitivity in capturing anesthesia depth changes. However, these parameters from posterior area EEG signals did not provide sufficient sensitivity to measure anesthesia depth variations. 2) Certain channel positions in the frontal part of the cortex had the combined benefits of substantial sensitivity and noise resistance, particularly

in regards to facial and eye movements which were major artifacts in EEG signals. 3) In the induction stage, there was a well-recorded short period (within the initial several minutes of drug administration) of initial excitation in most patients due to initial response to drug, evidenced by patient movements and other responses. The relative β-ratio did not capture this short surge in EEG activity, which may also explain lack of indication of this phenomenon in the BIS monitor. The β_2/θ-ratio captured this in all five subjects. 4) In the maintenance and emergence stages, the β_2/θ-ratio showed smaller sample variances than those of the relative β-ratio, indicating an improved reliability. In fact, in some patients the relative β-ratio showed similar values between a fully awake state (at the end of emergence stage) and deep anesthesia of the same patient. A trend data fitting showed that the β_2/θ-ratio seemed to be more reliable in providing a more consistent trend of anesthesia depth during the maintenance and emergence stages.

4.2. Discussions

Since the BIS monitor was the first anesthesia depth monitor on the market and has been used extensively in operating rooms, the fundamental parameter in the BIS monitor, the relative β-ratio, was used as the main reference standard. On the other hand, there were many modifications in the BIS algorithm that were apart from and in addition to the relative β-ratio. These modifications were needed before a reasonable compa- rison could be made between the (modified) β_2/θ-ratio and BIS measurements. To make such a comparison more relevant, the actual BIS reading was not used, but rather the relative β-ratio was extrated from the raw EEG data in BIS recording. Our findings were in the following key aspects of anesthesia depth monitoring.

The relative β-ratio and the β_2/θ-ratio demonstrated different sensitivities in distinct anesthesia stages. In the induction stage, the relative β-ratio fell down faster than the β_2/θ-ratio. On the other hand, the β_2/θ-ratio captured the initial surge of alertness from anesthesia drugs, but was subject to a delay and less sensitivity in anesthesia depth monitoring. This observation suggested a potential combined utility of the two parameters in the induction stage. During deep anesthesia (in the maintenance stage), both β_2/θ-ratio and relative β-ratio had substantial reliability. In the emergence stage, the β_2/θ-ratio was more responsive to the depth changes during recovery than the β-ratio.

Figures 3 showed the traces of the relative β-ratio

and β_2/θ-ratio that were derived from the $F4$ and Ol channels, respectively. The results demonstrated that both the relative β-ratio and β_2/θ-ratio tracked anesthesia depth changes with substantial sensitivity when they were computed from the frontal channels such as the $F4$ EEG. However, parameters derived from the temporal, parietal and occipital regions were of little utility. For example, the same parameters that were computed from the posterior area channels such as the Ol EEG did not provide sufficient discriminating capability. In particular, in **Table 1** the highest sensi- tivities, both in the induction and emergence stages, occurred in the channel $F4$. This result rendered the $F4$ channel EEG signals most sensitive to the influence of anesthesia agents. On the other hand, the parameters derived from the channels Fp1, Fp2, F3, F7, and F8, were sufficient to make them candidates for depth measurements. From the data collected in this study, it appeared that the EEG signals recorded from the frontal and central channels may best describe the brain activities during anesthesia.

The noise resistance capability was also distinct among different channels. An analysis of data from **Table 2** revealed that in the awareness and induction stages, due to facial, eye, and body movements, the EEG signals suffered from large artifacts, represented by lower SNRs. Within the front and central channels (that provide substantial sensitivity for depth measurements), the $F4$ channel had nearly the largest SNR. In the maintenance stage, as artifacts became very small all channels displayed similar SNRs. During the emergence stage, the front and central channels were relatively noise resistant. Still the $F4$ channel demonstrated nearly the largest SNR.

This analysis suggests that the non-prefrontal channels such as $F4$ may be a sound candidate for a better tradeoff between signal sensitivity to the depth changes and noise resistance capability. This may be especially useful as a potential remedy to the typical cases of BIS reliability in ICU (Intensive Care Unit) settings where noise artifacts make the BIS index far less reliable than in deep anesthesia patients.

It was cautioned that the above discussions and findings were based on a very small sample of five subjects, and as a result, they should be viewed as initial investigation and promising potential benefits of the β_2/θ-ratio. To substantiate the findings, a much larger sample of subjects must be conducted. This was the goal of our next study. The ultimate objective, if the findings were substantiated in the subsequent studies, was to integrate the improved EEG signal processing technique to enhance the existing anesthesia depth monitoring accuracy and reliability.

The BIS monitor used prefrontal EEG channels. This

had the advantage of easiness in applications since the prefrontal EEG electrodes did not contact hairs. However, the BIS monitor had the issue of reliability. This problem was particularly acute in ICU applications since the patients were under low anesthesia sedation. Consequently, muscle movements were far more frequent than in deep anesthesia. The frontal region was affected much less by EMG signals, offering a more reliable location for EEG measurements. It was possible that by using the frontal channels, one may be able to enhance reliability substantially. The tradeoff between monitor performance and easiness in usage needs to be further studied.

5. CONCLUSIONS

The above findings implied some possibilities in improving anesthesia depth monitoring: 1) Use better EEG channels at the front locations, rather than prefrontal locations; 2) Use a combined parameter, such as a weighted sum of the β_2/θ -ratio and relative β -ratio, in the induction stage: Initially more weight on the β_2/θ -ratio to capture the initial surge of awareness, then changes gradually to the relative β -ratio to take advantages of larger response sensitivity (hence a larger signal). 3) Use the β_2/θ -ratio to replace the relative β -ratio in the maintenance and emergence stages.

6. ACKNOWLEDGEMENT

REFERENCES

[1] Snow, J. (1847) On the inhalation of the vapour of ether. *LMG (London Medical Gazette)*, **39**, 498-502, 539-542.

[2] Lalkman, C.J. and Drummond, J.C, (2002) Monitors of depth of anesthesia, Qua Vadis? *Anesthesiology*, **96**, 784-787.

[3] Drummond, J.C. (2000) Monitoring depth of anesthesia: with emphasis on the application of bispectral index and the middle latency auditory evoked response to the prevention of recall. *Anesthesiology*, **93**, 876-82.

[4] Fisch, B.J. (1999) Fisch and Spehlmann's EEG primer: basic principles of digital and analog EEG. 3rd Edition. *Elsevier Science B.V.*.

[5] Dressler, O., Schneider, G., Stockmanns, G. and Kochs, E.F., (2004) Awareness and the EEG power spectrum: analysis of frequencies. *British Journal of Anaesthesia*, **93**, 806-809.

[6] Anier, A., Lipping, T., Melto, S. and Hovilehto, S. (2004) Higuchi fractal dimension and spectral entropy as mea-

sures of depth of sedation in intensive care unit. *The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 526-529.

[7] Rampil, I.J, "A primer for EEG signal processing in anesthesia". Anesthesiology, **89**, 980-1002, 1998.

[8] Tonner, P.H. and Bein, B. (2006) Classic electroencephalographic parameters: Median frequency, spectral edge frequency etc. *Best Practice & Research Clinical Anesthesiology*, **20**, 147-159.

[9] Zhang, X.S., Roy, R.J. and Jensen, E.W. (2001) EEG complexity as a measure of depth of anesthesia for patients. *IEEE Transactions on Biomedical Engineering* **48**, 1424-1433.

[10] Jeleazcov, C., Schneider, G., Daunderer, M., Scheller, B., Schuttler, J. and Schwilden, H. (2006) The discriminant power of simultaneous monitoring of spontaneous electroence- phalogram and evoked potentials as a predicator of different clinical states of general anesthesia. Anesth Analg , **103**, 894-901.

[11] Koskinen, M., Seppanen, T., Tong, S.B, Mustola, S. and Thakor, N.V. (2006) Monotonicity of approximate entropy during transition from awareness to unresponsiveness due to propofol anesthesia induction". IEEE Transactions on Biomedical Engineering, **53**, 669-675.

[12] Schneider, G., Hollweck, R., Ningler, M., Stockmanns, G., Kochs, E.F., (2005) "Detection of consciousness by electroencephalogram and auditory evoked potentials. *Anesthesiology*, **103**, 934-943.

[13] Zikov, T., Bibian, S., Dumont, G.A., Huzmezan, M. and Ries, C.R., Quantifying cortical activity during general anesthesia using wavelet analysis. *IEEE Transactions on Biomedical Engineering*, **53**, 617-632, 2006.

[14] John, E.R. and Prichep, L.S. (2005) The anesthetic cascade: a theory of how anesthesia surpress conciousness. *Anesthesiology*, **102**, 447-471.

[15] Heir, T. and Steen, P.A. (1996) Awareness in anesthesia: incidence, consequences and prevention. *Acta Anaesthesia Scand*, **40**, 1073-1086.

[16] Rampersad, S.E. and Mulroy, M.F. (2005) A case of awareness despitean adequate depth of anesthesia as indicated by a bispectral index monitor. *Anesth Analg*, **100**, 1363-1364.

[17] Jameson, L.C. and Sloan, T.B. (2006) Using EEG to monitor anesthesia drug effects during surgery. *Journal of Clinical Monitoring and Computing*, **20**, 445-472.

[18] Viertio-Oja, H., Laakso, H., Maja, V. and Sarkela, V. (2004) Description of the entropy TM algorithm as applied in the datex-ohmeda S/5 TM entropy module. *Acta Anaesthesiol Scand*, **48**, 154-161.

[19] Welch, P.D. (1967) The use of fast Fourier transforms for the estimation of power spectra: A method based on time averaging over short modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, **15**, 70-73.

[20] Bruhn, J., Myles, P.S., Sneyd, R. and Struys, M.M.R.F. (2006) Depth of anaesthesia monitoring: what's available, what's validated and what's next?" *British Journal of Anaesthesia*, **97**, 85-94.

Mobile and wireless technologies applying on sphygmomanometer and pulsimeter for patients with pacemaker implementation and other cardiovascular complications

Ching-Sung Wang

Department of Electronic Engineering Oriental Institute of Technology, Taipei, Taiwan, China.
Emill: ff020@mail.oit.edu.tw

ABSTRACT

Continuously monitoring pulse is very important for the pacemaker patients., and the continuously observing blood pressure is also a matter of concern for those who have hypertension, coronary heart disease, or other cardiovascular diseases, for example cardiac arrhythmia and hypertension. What we expect is not only to eliminate arrhythmia, also to treat patients as a whole of body. Therefore, how to keeping monitoring blood pressure and pulses rise to a very important issue. This research edits a wrist-able sphygmomanometer and pulsimeter combining with cell phone, to achieve wireless, continuous, and real-time observation, early detects any accident occurring from the patients with pacemaker implementation or with other cardiovascular complications

Keywords: Pacemaker; Hypertension; Coronary Heart Disease; Cardiovascular Disease; Sphygmomanometer and Pulsimeter

1. INTRODUCTION

The population for pacemaker implant is not limited by age, sex, or race. Over 100,000 pacemakers are implanted per year in the United States. Approximately 500,000 Americans have an implantable permanent pacemaker device [1]. Pacemaker can be affected by electromagnetic interference in several different ways, including temporary inhibition of the pacemaker, temporary function at the fixed noise rate, temporary function at the fixed magnet rate, permanent inhibition or malfunction, and random reprogramming. For any of these results to occur, the E field strength must be greater than 200V/m or the magnetic field strength must be greater than 10 Gauss [2].

Cellular telephones can interfere with the function of implanted cardiac pacemakers [3,4,5]. However, when telephones are placed over the ear, the normal position, this interference does not pose a health risk [5]. Barbaro V, *et al.* The research of influence between GSM mobile phone and implanted pacemaker indicates that electromagnetic interference effects were detected at a maximum distance of 10 cm with the pacemaker programmed at its minimum sensing threshold. When the phone antenna was in direct contact with patient's skin over the implant, electromagnetic interference effects occurred at maximum ventricular and a trial sensing thresholds of 4 mV and 2.5 mV, respectively [6]. Therefore, decreasing wireless emission power and rising pacemaker emission distance are the most useful methods of lowering the interference of magnetic to the pacemaker.

In order not to affect users' habits of using mobile and to avoid the electromagnetic interference effects of pacemaker, we design a wrist-able sphygmomanometer and pulsimeter which is using the Bluetooth power class II [7,8] to be the device for short-distance transmission. Furthermore, to combine with a cellular phone, we can do the prolonged distance observation and treatment immediately for the patients with setting pacemaker or with other cardiovascular disease.

2. SYSTEM ARCHITECTURE

This research aims at designing wrist-able sphygmomanometer and pulsimeter. We design a non-invasive sphygmomanometer [9] to remind the user to measure the blood pressure every hour. Pulsimeter is an electronic device which is based on pressure sensor, and can observe pulse continuously. In order to avoid the influence of electromagnetic wave on the pacemaker, this research is using the Bluetooth power class II equipment to shorten the distance which wireless data transmission device of sphygmomanometer transmits to client-end cellular phone, then use the GPRS (General Packet Radio Service)[10,11] to do the long distance data trans-
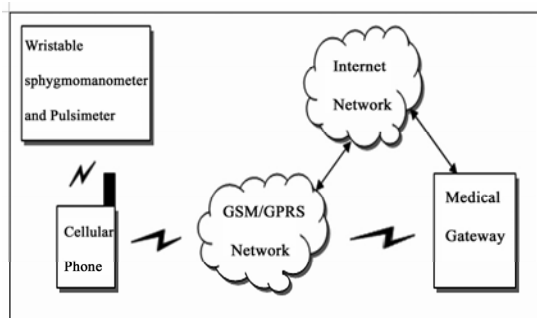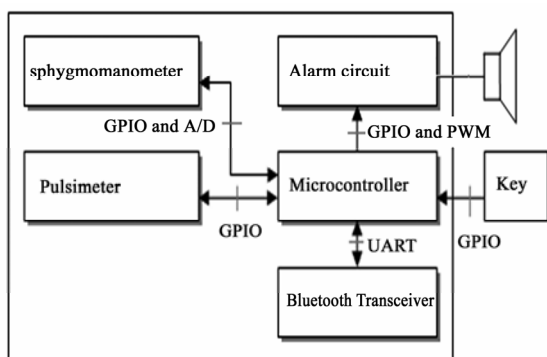
Figure 1. System architecture.

Figure 2. Wristable sphygmomanometer and pulsimeter hardware architecture.
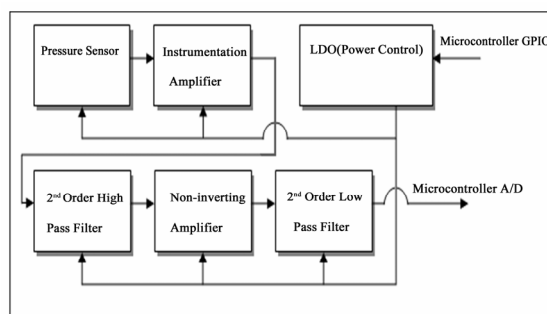

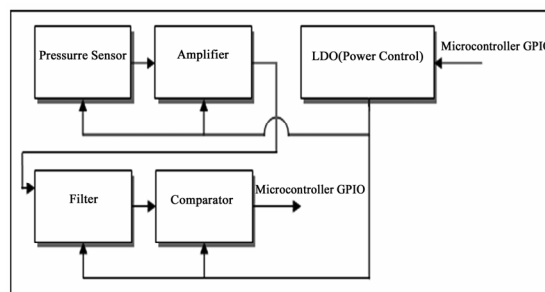**Figure 3**. Sphygmomanometer hardware architecture.

Figure 4. Pulsimeter hardware architecture.

mission. This can transmit blood pressure and pulse rate to Remote Medical Server and then provide them to specialist doctors for reference data.

Because of the connecting method of GPRS is a packet switch [10,11], it doesn't connect with Remote Medical Server continuously. Therefore, this research designs that Remote Medical Server used GSM (Global System for Mobile Communications) circuit connection(voice connection) [10,11,12] to inform wrist-able sphygmomanometer and pulsimeter, and drives sphygmomanometer to measure blood pressure immediately. Then, the sphygmomanometer will send the outcome of blood pressure to Remote Medical Server by GPRS [10,11,13]. **Figure 1** details the block diagram of system architecture.

3. HARDWARE ARCHITECTURE

The hardware system of this research is divided into two parts, one is wrist-able sphygmomanometer and pulsimeter; the other one is client-end cellular phone. The below is the detail of them.

3.1. Wristable Sphygmomanometer and Pulsimeter

Figure 2 details the block diagram of wrist-able sphygmomanometer and pulsimeter hardware architecture. The Controller unit of wrist-able sphygmomanometer

and pulsimeter is microcontroller [14], which controls and reads the peripheral units. Here is the description of each unit.

1) Sphygmomanometer [9,14,15,16]: The sphygmomanometer includes LDO (Low Dropout Regulators) as power controller, pressure sensor , filter, amplifier circuit and accessories. The research is using the pressure sensor of SCC series which is manufactured by HoneyWell company. The pressure sensor will transmit different message according to the changing pressure. The message passes through the procedures of magnifying (magnify the micro-message form sensor), and filtering (remove noise), then all the analog messages will be sent to the microcontroller via A/D (Analog/Digital) interface. **Figure 3** shows the block diagram of sphygmomanometer hardware architecture.

2) Pulsimeter [9,14,15,16] : The pulsimeter includes LDO, pressure sensor , filter, amplifier circuit and accessories. The circuit is using the SCC series of HonyWell company, but it is a different sensitivity pressure sensor. Putting sensor pad on pulse of hand, the pressure sensor will convert the normal pulse beat to a larger voltage output by magnifying, filtering, comparing (transform to digital signal), and then inputs it into microcontroller. **Figure 4** is shown the block diagram of pulsimeter.

3) Bluetooth transceiver [14,17]: The project uses BC3 made by CSR as the Bluetooth transceiver. The microcontroller communicates with the Bluetooth by UART (Universal asynchronous receiver/transmitter) interface and controls all the process by software.

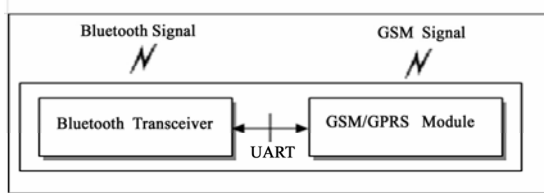4) Alarm circuit: The main purpose of this part is to

Figure 5. Client-end Cellular Phone Hardware Architecture.
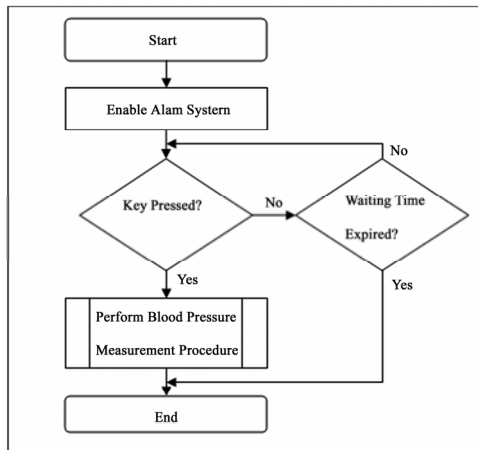
Figure 6. The software algorithm of continuous pulse and blood measuring and monitoring procedure.

send a notice to the users, and then to perform blood pressure measurement. Microcontroller can activate the alarm circuit every hour. However, on the other one end, Remote Medical Server can prescribe the blood pressure measurement, so that the users can measure their blood pressure when they require.

3.2. Client-End Cellular Phone Hardware Architecture

Figure 5 details the block diagram of client-end cellular phone hardware architecture, which includes two main parts, Bluetooth and GSM/GPRS.

1) Bluetooth transceiver [14,17,18]: The hardware framework is similar to the Bluetooth transceiver of wrist-able sphygmomanometer and pulsimeter. It communicates with GSM/GPRS module by UART interface.

2) GSM Module [18]: It includes Baseband part and RF part. The main components are Baseband IC (including DSP), TFT-LCD, NOR Flash, SRAM, Power Management IC and RF IC. We use MT6219 made by MediaTek as the platform of the GSM/GPRS system.

4. SOFTWARE ALGORITHM

4.1. Continuous Pulse and Blood Measuring and Monitoring Procedure

Pulse monitor procedure [14,15,16,19]: Microcontroller

Figure 7. The software algorithm of measure blood pressure procedure.

reads the pulsimeter every minute, then transmits the results to the client-end cellular phone by Bluetooth SPP (Serial Port Profile).

Blood pressure measure procedures [9,14,15,16,19]: Microcontroller enables alarm system to inform user every hour that the time requires to measure the blood pressure. Users need to put hand on the proper position so that sphygmomanometer can have a better detection. While the users put hand on proper position and push trigger button, then the microcontroller controls pulsimeter's LDO to stop action and start to measure the blood pressure. After that, microcontroller transmits blood pressure to the cellular phone by Bluetooth SPP, and to activate pulsimeter. If the user doesn't do the action of blood pressure measurement, the microcontroller would re-activate pulsimeter after a while. The software algorithm of continuous pulse and blood measurement and the monitoring procedure is shown in **Figure 6** and **Figure 7**.

4.2. Client-End Clelular Phone Data Handling Procedure

While the Bluetooth of mobile receives the data from the Bluetooth of wrist-able sphygmomanometer and pulsimeter, it analyzes the data of blood pressure or pulse and to store (includes the measurement time) individu-

Figure 8. The software algorithm of client-end cellular phone data handling procedure.

Figure 9. The software algorithm of remote medical server claim for mearuring blood pressure.

ally at first. Finally, it establishes GPRS connection. After successful connecting, mobile scans stored blood pressure and pulse data, and then it transmits to the Remote Medical Server if the data has not been sent yet.
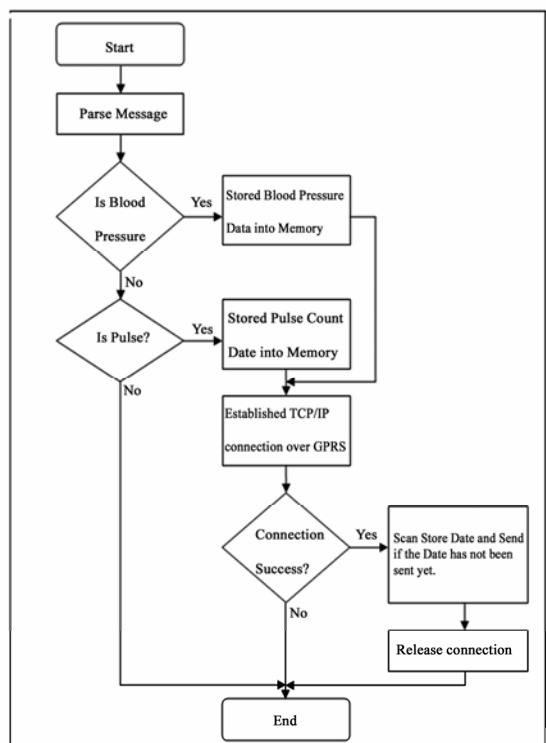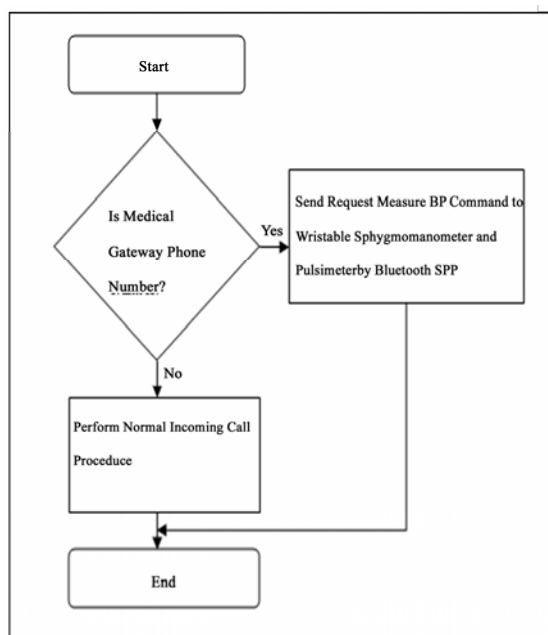
The software algorithm of mobile data handling procedure is shown on **Figure 8** [13,19,20,21].

4.3. Remote Medical Server Claim for Measuring Blood Pressure

If the doctor considers that it is necessary to get the patient's current blood pressure-immediately, he/she can call to client-end cellular phone by GSM module on the Remote Medical Server. While client-end cellular phone receives incoming call form server-end, the Bluetooth SPP transmits command of measuring blood pressure to wrist-able sphygmomanometer and pulsimeter. While wrist-able sphygmomanometer and pulsimeter receive the command from the Remote Medical Server, they activate blood pressure measurement and transmit procedures (**Figure 6**, **7**, **8**). The software algorithm of Remote Medical Gateway claims for measuring blood pressure is shown in **Figure 9** [13,19,20,21].

5. CONCLUSIONS

It is a fatal risk for the patient with artificial pacemaker implementation while the mobile connection is close to the heart. The artificial pacemaker will consider the electromagnetic wave from mobiles is heart signal, then to do the wrong stimulation on muscle fiber of heart ventricle. However, continuous pulse and blood pressure observation is very essential for the patient of arrhythmia with hypertension coronary arterial heart disease or heart failure. Based on this following the Bluetooth power class II RF test specification, we have designed a portable Bluetooth medium. We expect this system will be benefit for the patients of arrhythmia with hypertension coronary arterial heart disease or heart failure and never cause any inconvenience to their daily life.

REFERENCES

[1] http://www.surgeryencyclopedia.com/La-Pa/Pacemakers.html.

[2] Tacey, S. and Ronald, A. (1992) The effects of electromagnetic fields on cardiac pacerkers. *IEEE Transactions on Broadcasting*, **38**.

[3] Barbaro, V., Bartolini, P., Donato, A., *et al*. (1995) Do European GSM mobile cellular phones pose a potential risk to pacemaker patients? *PACE*, **18**, 1218-24.

[4] Okan, E., MD, (2002) Electromagnetic interference on pacemakers. *Indian Pacing Electrophysiol J.*, **2(3)**, 74-78.

[5] Hayes, D.L., Wang, P.J., Reynolds, D.W., *et al.* (1997) Interference with cardiac pacemakers by cellular telephones. *N. Eng. J. Med*, **336**, 1473-9.

[6] Bluetooth, (2004) Specification of the bluetooth system core. Version 2.0.

[7] Bluetooth, (2004) Radio frequency test suite structure (TSS) and test purposes (TP) specification 1.2, Revision 1.2.3.

[8] Nissila, S., Sorvisto, M., Sorvoja, H., Vieri-Gashi, E. and Myllyla, R. (1998) Non-invasive blood pressure measurement based on the electronic palpation method.

Engineering in Medicine and Biology Society, Proceedings of the 20th Annual International Conference of the IEEE, **4**, 1723-1726.

[9] Lin, Y.B. and Imrich, C. (2000) Wireless and mobile network architectures, Wiley Computer Publishing, 1 Edition (October 2).

[10] Michel, M., Marie-Bernadette, P. (1992) The GSM System for Mobile Communications, Telecom Publishing.

[11] Wang, C.S., Lee, J.H. and Chu Y.T. (2007) Mobile telemedicine application and technologies on GSM, bioinformatics and biomedical engineering, *ICBBE*, The 1st International Conference on 6-8 July 2007, 1125-1128.

[12] (2004)MediaTek, AT command set, Revision: 0.03.

[13] (2005) National Semiconductor, SC14440/431/432, SC14435, SC14436/437/438. Baseband processor for PP/FP DECT and WDCT, V1.1, June 16.

[14] (2004) Champion, CM2838, 300mA Low Esr CMOS LDO with Enable, Rev. 1.1.

[15] (2002) HoneyWell, SSC series pressure sensors 0-5 psi through 0-300 psi.

[16] CSR, (2006) BlueCore3-ROM CSP Product Data Book.

[17] MediaTek, (2004) MT6219 GSM/GPRS Baseband Processor Data Sheet, Revision 1.00.

[18] Bluetooth SIG, (2001) Specification of the Bluetooth System Profile, Part K: 5 Serial Port Profile, Version 1.1.

[19] (2003) MediaTek, Customer Device Driver Document, Revision 0.1.

[20] PixTel, (2004) Writing Applications Using PixTel MMI Platform, Version 1.3.

Modelling the inhalation of drug particles in a human nasal cavity

Kiao Inthavong, Jian Wen, Ji-Yuan Tu[*]

School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University, Bundoora, Australia.
Email: jiyuan.tu@rmit.edu.au

ABSTRACT

A human nasal cavity was reconstructed from CT scans to make a Computational Fluid Dynamics (CFD) model. With this model, fluid flow and inhalation of aerosol analysis can be investigated. The surface of the interior nasal cavity is lined with highly vascularised mucosa which provides a means for direct drug delivery into the blood stream. Typical sprayed particles from a nasal spray device produce a particle size distribution with a mean diameter of 50μm, which leads to early deposition due to inertial impaction. In this study low-density drug particles and submicron particles (including nanoparticles) are used to evaluate their deposition patterns. It was found that the low-density particles lightens the particle inertial properties however the particle inertia is more sensitive to the particle size rather than the density. Moreover the deposition pattern for nanoparticles is spread out through the airway. Thus an opportunity may exist to develop low-density and nanoparticles to improve the efficiency of drug delivery to target deposition on the highly vascularised mucosal walls.

Keywords: Nasal Airway; Ultrafine, Fibre; Morphology; CFD; Deposition

1. INTRODUCTION

Nasal drug delivery provides an alternative approach to traditional delivery methods such as oral drug routes that fail in the systemic delivery of compounds due to its dissociation by the digestive system. The nasal airway is dominated by the nasal turbinates that are lined with highly vascularised mucosa opening to the paranasal sinuses. Because of these characteristics it is hypothesised that drug delivery to combat health problems such as lung diseases, cancers, diabetes, sinus infections etc. may be viable if the drug formulation can be deposited in the turbinate region [1].

Despite these advantages, studies have found that tar-geted drug delivery is inefficient [2,3]. The atomisation of the drug formulation produces a mean droplet size of 50μm [4] which exhibits high inertia. This leads to a large proportion of particles impacting in the anterior regions of the nasal cavity. Most drug formulations have close to unit density as they are suspensions in aqueous solutions. Lighter porous drug particles have been developed for pulmonary delivery [5], where the drug particle sizes are in the low micron to sub-micron range and deposition is targeted at the pulmonary airways that exhibit much smaller spaces such as the airway branches in the lungs. Another alternative is the use of engineered nanoparticles which exhibit a large surface area to size ratio leading to greater biologic activity. This increased biologic activity can be exploited for targeted drug and gene delivery, tissue engineering, cell tracking and bio-separation [6,7]. One advantage for nasal drug delivery is its extremely small size which would allow the particles to deposit through diffusion rather than inertial impaction. Thus an opportunity exists for the development of new porous-based particles and/or nano-sized particles for nasal drug delivery.

Computational Fluid Dynamics (CFD) simulations have evolved into a feasible alternative to complement experimental data. For example CFD simulations for airflow patterns [8,9,10] have complemented experimental results [11] by confirming regions of vortices within the nasal vestibule, the olfactory region and posterior to the nasal valve. Simulations for particle deposition however, are fewer in numbers. Spherical particle deposition under conditions related to pharmaceutical nasal spray applications has been studied [3]. Particles in the range of 10 μm to 50 μm subjected to a breathing flow rate of 20L/min found that a large proportion of particles deposited in the anterior third of the nasal cavity which were attributed to the injected particles existing in a high inertial regime. On the other hand nano-sized particles in the nasal cavity under laminar conditions were simulated [12] which found that diffusion was the dominant deposition mechanism for the smallest range of particles (1–30 nm).
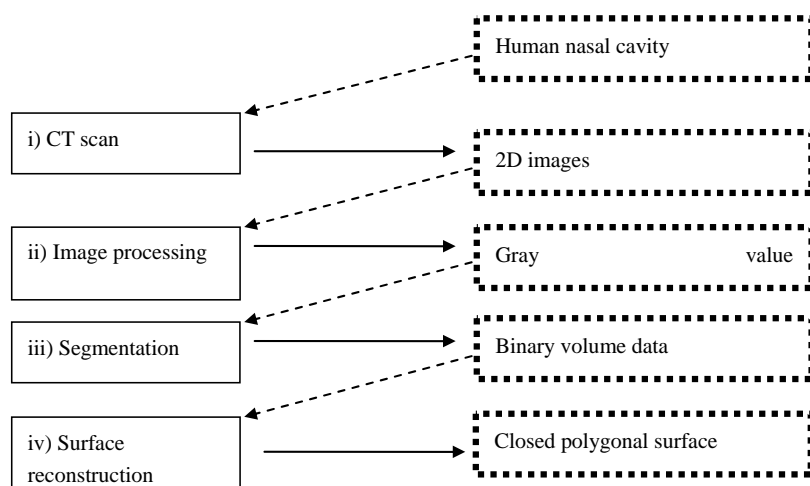
Figure 1. Steps in the construction of the computational model.
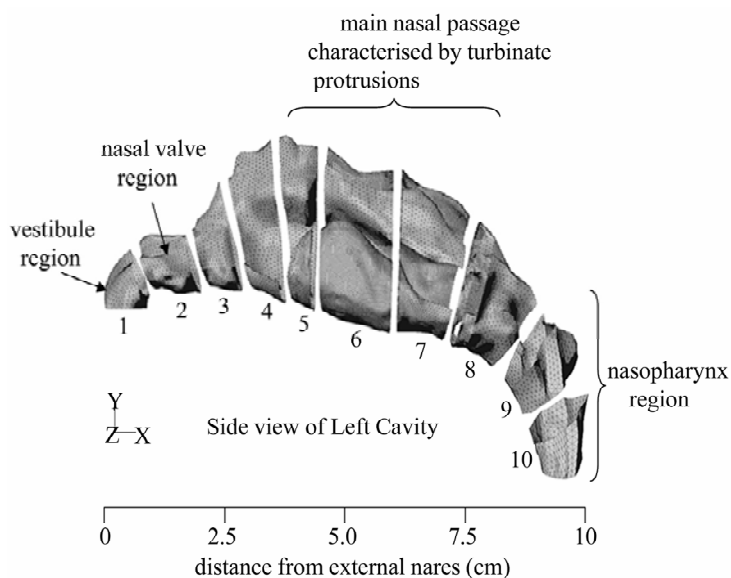
Figure 2. Different views of the computational model. Three coronary slices and the inlet show the internal mesh with a dense region near the walls.

A human nasal cavity was reconstructed from CT scans and a computational model developed for particle flow analysis within the airway. This study presents the use of CFD techniques to investigate the airflow patterns and deposition of low density and nano-sized particles for drug delivery in a human nasal cavity. The nasal airway was chosen for investigation as it is one of the major entries into the respiratory system which can be penetrated to reach the blood streams. With this in mind, it is anticipated that this research will assist in new designs of aerosols and particulates and also help to guide practical clinical tests for toxicological and therapeutic studies.

2. METHOD

The model reconstruction involved four main steps

(**Figure 1**): 1) CT images acquisition; 2) image processing and editing to improve the quality of the image volume; 3) segmentation; and 4) surface reconstruction. CT images of a 25 yr-old Asian male (75kg, 170cm) provided a 3D matrix of volume elements (voxels), in which different tissues and structures having different attenuation characteristics were distinguished from one another by differences in brightness or greyscale.

2.1. Image Processing

The scan was performed using a CTI Whole Body Scanner (General Electric). The single-matrix scanner was used in helical mode with 1-mm collimation, a 40-cm field of view (FOV), 120 kV peak and 200 mA to produce contiguous images (slices) of 1-mm thickness

Left Cavity   Right Cavity

Figure 3. Path streamlines in the left and right cavities.

with voxel size 0.25×0.25×1 mm. The original set of CT images is converted into a file format compatible with the package MegaWave2 (MW2), by means of C language self-developed routines. The conversion program also performs an image enhancement, by rescaling the grey-level histogram to 1-200 and remapping the image volume to an 8-bit/pixel depth file.

A 3D convolution with a Gaussian kernel was used to reduce the background noise present in the images. Because of its isotropic shape, the Gaussian filter has optimal properties such as smoothing mask, removing small-scale texture and noise, which could alter the regional segmentation, without distorting lower spatial frequencies. Filtering was applied in three dimensions in order to obtain a smoothed CT image volume also along the axial direction. Such a procedure attenuates the spatial discontinuities among the slices introduced during acquisition, as an effect of the slice thickness.

2.2. Segmentation & Surface Reconstruction

A 2D segmentation is used to detect and extract, slice by slice, the walls of the airway. For the segmentation process, a region growing algorithm, based on the Mumford and Shah [13] method implemented WM2 is used. The regional segmentation has been included because it allows the tracking only of the domains of interest, even in the presence of noise. A first regional segmentation with a greater number of partitioning regions than necessary is performed on each single slice. This allows the algorithm to detect the walls even in severely disturbed images. A threshold binarisation process is then applied in order to remove sub-regions unrelated to the airway, which typically present a lower intensity value with respect to the signal. In this wok, the threshold has been empirically chosen and represents 45% of the maximum grey-level value of the study. Generation of a surface or solid model from the 2D contour data began with the translation of the segmented, modified and smoothed contour points into a data series that was read into CAD package used in this study: CATIA. The contours were lofted to define surface splines which enclosed the airway volume.

Right cavity   Left Cavity
$x = 2.6$cm from the anterior tip of the nose
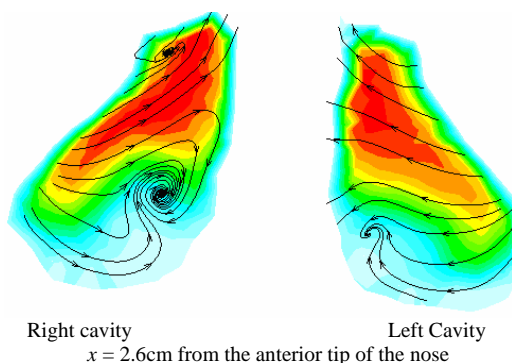
Figure 4. Contour plot of axial velocity (x-velocity) combined with cross flow path streamlines (y-z velocity) in the left and right cavities.
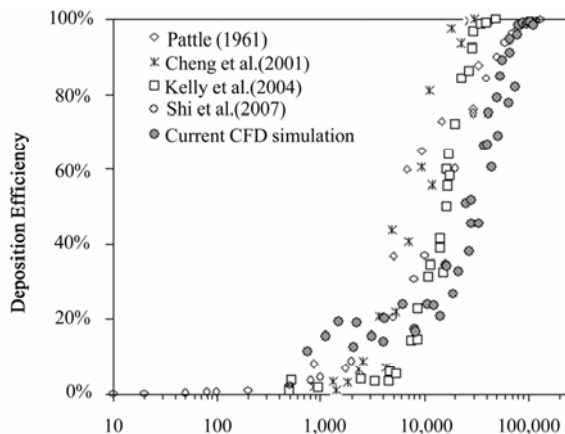
Figure 5. Deposition efficiency vs inertial parameter comparisons for the simulation micron spherical particles.

2.3. Adaptive Meshing

The CATIA models were imported into a 3D modelling program called GAMBIT. An initial model with 82,000 cells was created and used to solve the air flow field at a flow rate of 7.5 L/min. The original model was refined by cell adaptation techniques that included refining large volume cells, cells that displayed high velocity gradients and near wall refinements. This process was repeated twice, with each repeat producing a model with a higher

cell count than the previous model. Subsequently four models were produced, 82000, 586000, 950000 and 1.44 million cells. A grid independence test found that the results for average velocity converged at 950,000 cells. Subsequently the 950,000 cell model was used and is shown in **Figure 2**.

2.4. Numerical Method

Due to the complex geometry of the anatomically real nasal cavity a commercial CFD code, FLUENT, was used to predict the continuum air phase flow through solutions of the conservation equations of mass and momentum. The steady continuity and momentum equations for the gas phase (air) in Cartesian tensor notation are:

$$\frac{\partial}{\partial x_i}\left(\rho_g u_i^g\right) = 0 \tag{1}$$

$$\rho u_j^g \frac{\partial u_i^g}{\partial x_j} = -\frac{\partial p_g}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\mu_g \frac{\partial u_i^g}{\partial x_j}\right) \tag{2}$$

where u_i^g is the *i*-th component of the time averaged velocity vector and ρ_g is the air density. These equations were discretised using the finite volume approach. The third order accurate QUICK scheme was used to approximate the momentum equation whilst the pressure-velocity coupling was realized through the SIMPLE

method. To be consistent with experimental data, a constant flow rate of 7.5 L/min. was used to simulate light breathing. At this flow rate, the flow regime has been determined to be laminar [11,14]. A steady flow rather than a cyclic unsteady flow was used in this case to allow the results to emphasize the effects of particle morphology on deposition sites independent from cyclic conditions. Moreover the effects of a periodic inhalation on the overall flow field are found to be negligible from the Womersley frequency variable which is used to determine the importance of the fluctuating sinusoidal pattern of the inhalation-exhalation breathing cycle. The Womersley frequency variable,

$$W = D/2\left(\omega/v_g\right)^{0.5} \tag{3}$$

was calculated as 0.3 where D is the local cross-sectional distance between the two nasal walls and is about 0.5 cm in this nasal cavity, v_g is the kinematic viscosity of air and ω is the breathing frequency.

2.5. Drug Particles

For spherical particles the drag force is related to the drag coefficient which has been studied quite extensively. The general correlation for smooth spherical particles is given as [15]:

$$C_D = 24/18\,\mathrm{Re}_p\left(18 + 2.367\cdot \mathrm{Re}_p^{\,w}\right) \quad \text{for Re}<20 \tag{4}$$

Table 1. Deposition efficiency based on **Figure 5**.

Density	d_{ae}	Inertial Parameter (IP)	Deposition % present simulation
100	15.8	41,667	67.4%
200	22.4	83,333	95.3%
1000	50.0	416,667	100%

50μm 100 kg/m³   **50μm 1000 kg/m³**

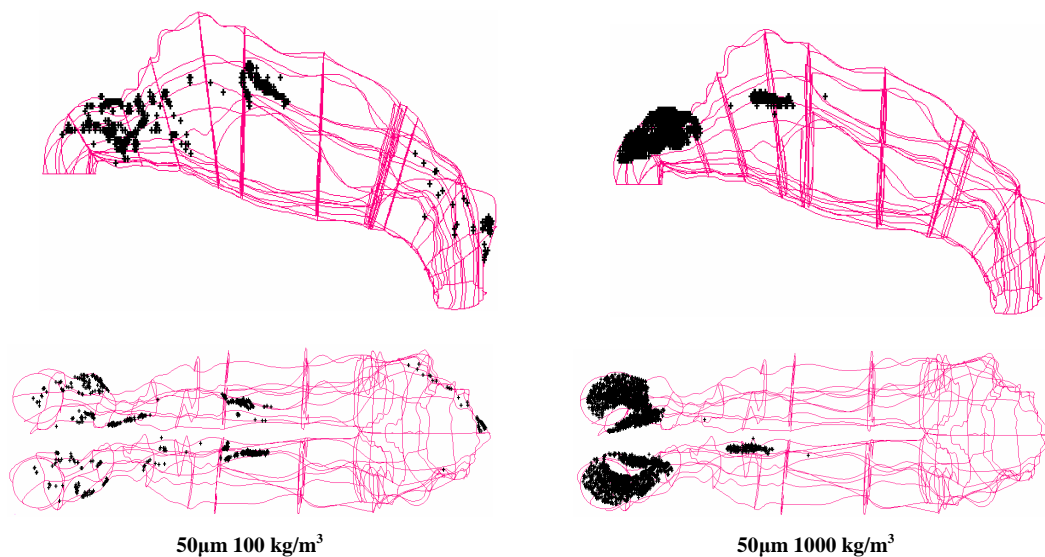**Figure 6.** Deposition pattern for low density particles where ρ = 100kg/m**³** and 1000kg/m**³.**

where $w = 0.82 - 0.05\left(\log_{10} \text{Re}_p\right)$

For submicron particles the drag force per unit particle mass taking the form of Stokes' drag law [16] defined as,

$$F_D = \frac{18\mu}{d_p^2 \rho_p C_c}\left(u_i^g - u_i^p\right) \tag{5}$$

C_c is the Cunningham correction factor to Stokes' drag law. The Brownian force by Li and Ahmadi [17] can be re-arranged to highlight the diffusion coefficient as:

$$F_B = \frac{\zeta}{m_d}\sqrt{\frac{1}{\tilde{D}}\frac{2k_B^2 T^2}{\Delta t}} \tag{6}$$

where m_d is the mass of the particle, T is the absolute temperature of the fluid, v is the kinematic viscosity, k_B is the Boltzmann constant, and $\tilde{D}$ is the diffusion co-efficient. **Eq.6** is inputted into the user-defined-function option in Fluent. Additional forces include Saffman's lift force [18]

$$F_L = \frac{2K v^{1/2}\rho d_{ij}}{\rho_p d_p \left(d_{lk}d_{kl}\right)^{1/4}}\left(\vec{v} - \vec{v}_p\right) \tag{7}$$

and the thermophoretic force [19]

$$F_T = -D_T \frac{1}{m_p T}\frac{\partial T}{\partial i} \tag{8}$$

Particle rebounding from the surfaces was ignored and particle deposition was determined when the distance between the particle centre and a surface was less than or equal to the particle radius.

3. RESULTS AND DISCUSSION

3.1. Airflow Patterns

Path streamlines which act as massless particle tracers to track the flow path of the inhaled air, were released from the nostril inlet to provide visualisation of the flow field (**Figure 3**). The streamlines in the left nasal cavity at a flow rate of 7.5L/min show flow separation and reversed flow in the upper anterior part of the cavity (olfactory region). This low flow characteristic in the olfactory region is important as it is a defence mechanism that prevents particles whose trajectories are heavily de-pendent on flow patterns from being deposited onto the sensitive olfactory nerve fibres, while vapours are al-lowed to diffuse for olfaction. For both cavities, the air flow squeezes through the nasal valve region, before decelerating due to the expansion in the cross-sectional area. The nasal valve region is approximately 2cm from the nostril inlet.

A cross-sectional area located just immediate of the anterior nasal valve at 2.6cm was chosen to reflect the rapid changes in the flow field. The cross-section shown
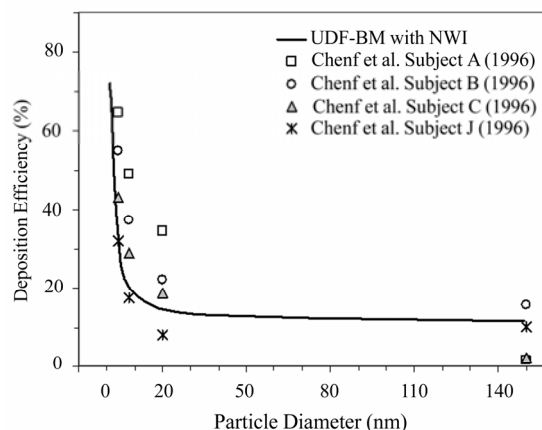
Figure 7. Deposition efficiency of 1nm-150nm particles in a human nasal cavity at a steady inhalation rate of 10L/min.

in **Figure 4** is from a frontal perspective (positive flow into the paper). The contours reflect the axial velocity (*x*-component of velocity) and are combined with streamlines of secondary flow (*y-z* component of veloc-ity). The red contours are maximum values which repre-sent the bulk flow regions. For the cross-section located at 2.6cm from the anterior tip of the nose, two vortices in the right cavity and one in the left are found. The bulk flow is found in the middle and upper regions of the both cavities.

3.2. Deposition of Drug Particles

A parameter used for normalizing impaction-dominant deposition studies is the inertial parameter, IP given by:

$$IP = Q d_{ae}^2 \tag{9}$$

where Q is the air flow rate, given in cm^3/s and d_{ae} is the aerodynamic diameter given in μm. It is a convenient parameter that compares deposition against different flow rates and particle sizes at aerodynamic diameters. Monodispersed particles in the range of 1-30μm were released passively (with the airflow) into the nasal cavity at flow rates of 5, 7.5, 10 and 15L/min. The deposition of particles over a range of the inertial parameter is shown in **Figure 5** and is compared with other experi-mental results.

Since the drug particles exhibit different densities, they can be compared in terms of their aerodynamic properties (inertia, and settling properties) through the equivalent aerodynamic diameter, d_{ae} defined as:

$$d_{ae} = d_p \sqrt{\rho_p / 1000} \tag{10}$$

This means that a small diameter, very dense particle can have the same aerodynamic as a large diameter, but less dense particle if their d_{ae} are the same.

Particles in the micron particle size range exist in the inertial regime where deposition by inertial impaction is

(a) Side view
individual particle deposition

(b) Top view
individual particle deposition

(c) Side view – contours of normalised local regional deposition

(d) Top view – contours of normalised local regional deposition

N_i / N_{max}

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Figure 1. Regional deposition patterns of 1nm particles under a flow rate of 10L/min in a human nasal cavity.

relevant. The idea of low-density porous drug particles will decrease the particle inertial properties, although this decrease is not as significant as changing the particle size. The inertial parameter incorporates both the geometric equivalent spherical diameter along with the particle density to give the particle aerodynamic diameter (**Eq.12**). Thus a larger porous particle will have a smaller equivalent a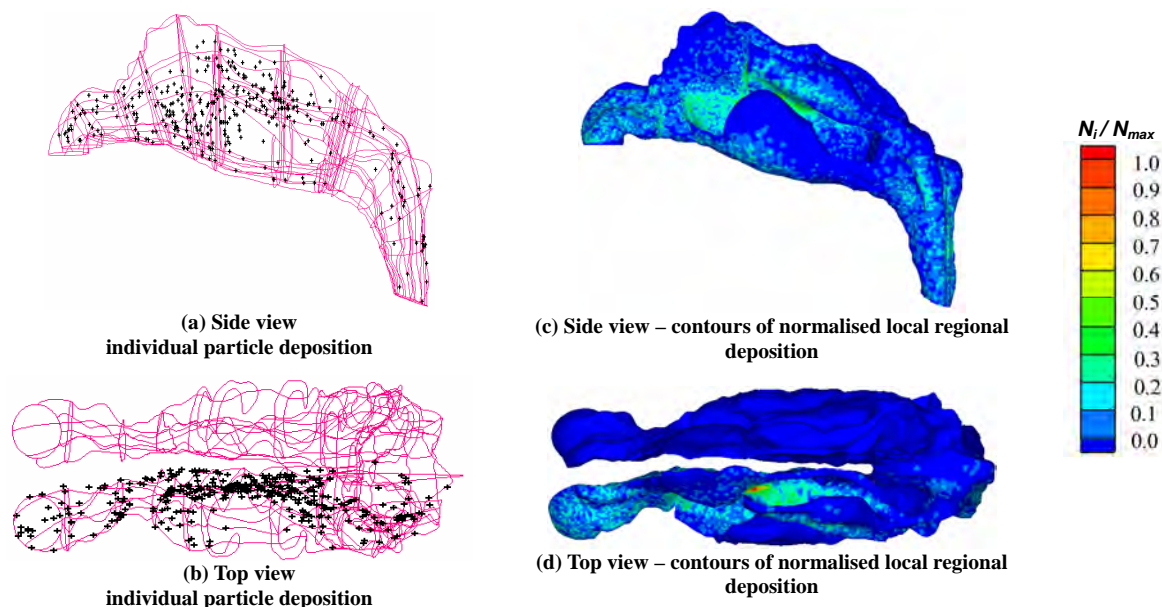erodynamic diameter than a particle of equal size but with a higher density. The effect of particle density on the deposition of a 50μm particle is shown below.

The particle deposition patterns in the nasal cavity in **Figure 6** shows particles with density of 100kg/m^3 and 1000kg/m^3 for brevity as the deposition pattern for ρ = 200kg/m^3 becomes similar to that of ρ = 100kg/m^3. For particles with densities of 100kg/m^3 a portion of the particles deposit superiorly on the septum walls of both sides of the nasal cavity. This suggests that the fluid flow is close to the inner septum walls forcing the particles into this region.

A second concentration of deposited particles occurs at the back of the nasal cavity where the flow changes directions from horizontal to vertically downwards. The change in the flow direction causes the particles to impact in this region. For particles with density of 1000 kg/m^3 deposition is found in the frontal area with only a small proportion of particles passing through the nasal valve region. These particles finally impact onto the superiorly on the septum wall of the left nasal cavity, however this pattern is not found in the right nasal cavity.

Deposition of submicron particles (1 to 150nm) was simulated under a flow rate 10 L/min in order to make

comparisons with experimental data reported by [20] which found deposition efficiencies for a variety of human subjects under a flow rate of 10 L/min. The solid line in **Figure 7** corresponds to the CFD model prediction. The deposition curve is high for very small nanoparticles and the particle diameter range in which the deposition drops from 72% to 18% is between 1nm-10nm. From 10nm-150nm however, there is only a small change in the deposition curve from 18%-15%. This deposition curve profile is characteristic of the Brownian diffusion, where the particles are so small that the fluid may no longer be considered continuous. The trajectory of the nanoparticle is then caused by the collision of the air molecules and concentration gradients to produce the random motion.

Local deposition patterns for a 1nm particle are shown in **Figure 8**. The deposition pattern of the 1nm particle is distributed evenly through the nasal cavity where the diffusion disperses the particles in all directions. The wall contours in **Figure 8(c)**, **(d)** show regions of high concentrations which is determined by the number of particles that deposit onto a wall face divided the maximum number of particles that deposit on any one face. Few particles are able to reach the wider meatus region, and instead the particles remain close to the nasal septum wall (inner regions). High concentrations are found at the upper regions of the cavity with a higher distribution of deposition within that one area.

In general the deposition pattern is spread out through the nasal cavity well. This has interesting applications for drug delivery where traditional nasal sprays are producing micron sized droplets that are prone to inertial deposition. This deposition mechanism leads to high

inertial impaction (up to 100% for a mean atomised particle droplet of 50μm) in the anterior region of the nasal cavity [2,3]. However for high drug efficacy, the delivery of the droplets needs to be deposited in the middle regions of the nasal cavity, where the highly vascularised walls exist. Smaller particles such as 1μm were found to be less affected by inertial properties, which allowed it to bypass the anterior region of the nasal cavity. However because of the particles ability to follow the streamlines more readily, the particles were less likely to deposit in any region of the nasal cavity and instead bypasses it completely, leading to the undesired effects of lung deposition. Delivery of nanoparticles especially 1nm-5nm particles therefore, can provide improved deposition in the middle regions whilst minimising deep lung deposition.

4. CONCLUSIONS

Simulations of air-particle flows in the nasal cavity found vortices primarily in the upper olfactory region and just posterior to the nasal valve where the geometry begins to expand. This suggests that high inertial particles are unlikely to reach the sensitive olfactory region. Multiple secondary flow regions were found in the lower middle regions within the nasal valve. Low density porous drug particles lightens the particle inertial properties however the particle inertia was more sensitive to the particle diameter rather than its density. The deposition of nanoparticles in the nasal cavity was distributed evenly throughout the airway with a deposition that drops from 72% to 18% for 1nm to 10nm. Because of the evenly distributed deposition pattern for nanoparticles there exists an opportunity to develop low-density and nanoparticles to improve the efficiency of drug delivery to target deposition on the highly vascularised mucosal walls.

5. ACKNOWLEDGEMENTS

REFERENCES

[1] Kimbell, J., Shroeter, J.D., Asgharian, B., Wong, B.A., Segal, R.A., Dickens, C.J., Southall, J.P. and Millerk F. J. (2004) Optimisation of nasal delivery devices using computational models. *Res. Drug Del.*, **9** 233-238.

[2] Inthavong, K., Tian, Z.F., Tu, J.Y., Yang, W. and Xue, C. (2008) Optimising nasal spray parameters for efficient drug delivery using computational fluid dynamics. *Computers in Biology and Medicine*, **38(6)**, 713-726.

[3] Inthavong, K., Tian, Z.F., Li, H. F., Tu, J.Y., Yang, W., Xue, C.L. and Li, C.G. (2006) A numerical study of spray particle deposition in a human nasal cavity. *Aerosol Science Technology*, **40***, 1034-1045.

[4] Cheng, Y.S., Holmes, T.D., Gao, J., Guilmette, R.A., Li, S., Surakitbanharn, Y. and Rowlings, C. (2001) Characterization of nasal spray pumps and deposition pattern in a replica of the human nasal airway. *J. Aerosol Medicine*, **14 (2)**, 267-280.

[5] Edwards, D.A., Hanes, J., Caponetti, G., Hrkach, J., Ben-Jebria, A., Eskew, M.L., Mintzes, J., Deaver, D., Lotan, N. and Langer, R. (1997) Large porous particles for pulmonary drug delivery. *Science*, **276**, 1868-1872.

[6] Gupta, A.K. and Gupta, M. (2005) Synthesis and surface engineering of iron oxide nanoparticles forbiomedical applications. *Biomaterials*, **25(18)**, 3995.

[7] McCarthy, J.R., Kelly, K.A., Sun, E.Y. and Weissleder, R. (2007) Targeted delivery of multifunctional magnetic nanoparticles. *Nanomedicine*, **2(2)**, 153-167.

[8] Keyhani, K., Scherer, P.W. and Mozell, M.M. (1995) Numerical simulation of airflow in the human nasal cavity. *J. Biomechanical Engineering*, **117**, 429-441.

[9] Subramaniam, R.P., Richardson, R.B., Morgan, K.T., Kimbell, J.S. and Guilmette, R.A. (1998) Computational fluid dynamics simulations of inspiratory airflow in the human nose and nasopharynx. *Inhalation Toxicology*, **10,** 91-120.

[10] Finckl M.H. and Wlokas, D.I. (2006) Simulation of nasal flow by lattice Boltzmann methods. *Computers Biology Medicine*, **37(6)**, 739-749.

[11] Hahn, I., Scherer, P.W. and Mozell, M.M. (1993) Velocity profiles measured for airflow through a large-scale model of the human nasal cavity. *J Appl. Physiol.*, **75(5)**, 2273-2287.

[12] Zamankhan, P., Ahmadi, G., Wang, Z., Hopke, P.H., Cheng,Y.S., Su, W.C. and Leonard, D. (2006) Airflow and deposition of nanoparticles in a human nasal cavity. *Aerosol Science Technol.*, **40**, 463-476.

[13] Mumford, D. and Shah, J. (1989) Optimal approximations by piecewise smooth functions and associated variational problems. *Commincations Pure and Applied Mathematics*, **XLII**, 577-685.

[14] Swift, D.L. and Proctor, D.F. (1977) Access of air to the respiratory tract., in *Respiratory Defence Mechanisms.*, Brain, J.D., Proctor, D.F. and Reid, L.M. Editors, Marcel Dekker, New York, NY, 63-93.

[15] Clift, R., Grace, J.R. and Weber, M.E. (1978) Bubbles, Drops, and Particles. London, Academic Press Inc, UK, London, Ltd.

[16] Ounis, H., Ahmadi, G.M.J.B., (1991) Brownian diffusion of submicrometer particles in the viscous sub- layer. *J. Colloid and Interface Science*, **143(1)**, 266-277.

[17] Li, A. and Ahmadi, G. (1992) Dispersion and deposition of spherical particles from point sources in a turbulent channel flow. *Aerosol ScienceTechnology*, **16**, 209-226.

[18] Saffman, P.G. (1965) The lift on a small sphere in a slow shear flow. *J. Fluid Mechanics*, **22**, 385-400.

[19] Talbot, L., Cheng, R.K., Schefer, R.W. and Willis, D. R. (1980) Thermophoresis of particles in a heated boundary layer. *J. Fluid Mech*, **101(4)**, 737-758.

[20] Cheng, K.H., Cheng, Y.S., Yeh, H.C., Guilmette, A., Simpson, S.Q., Yang, Y.H. and Swift, D.L. (1996) *In-vivo* measurements of nasal airway dimensions and ultrafine aerosol deposition in the human nasal and oral airways. *J. Aerosol Science*, **27(5)**, 785-801.

Identifying predictive markers of chemosensitivity of breast cancer with random forests

Wei Hu

Department of Computer Science, Houghton College, Houghton, NY, USA.
Email: wei.hu@houghton.edu

ABSTRACT

Several gene signatures have been identified to build predictors of chemosensitivity for breast cancer. It is crucial to understand how each gene in a signature contributes to the prediction, i.e., to make the prediction model interpretable instead of using it as a black box. We utilized Random Forests (RFs) to build two interpretable predictors of pathologic complete response (pCR) based on two gene signatures. One signature consisted of the top 31 probe sets (27 genes) differentially expressed between pCR and residual disease (RD) chosen from a previous study, and the other consisted of the genes involved in Notch singling pathway (113 genes). Both predictors had a higher accuracy (82% v 76% & 79% v 76%), a higher specificity (91% v 71% & 98% v 71%), and a higher positive predictive value (PPV) (68% v 52% & 73% v 52%)) than the predictor in the previous study. Furthermore, Random Forests were employed to calculate the importance of each gene in the two signatures. Findings of our functional annotation suggested that the important genes identified by the feature selection scheme of Random Forests are of biological significance.

Keywords: Random Forests; Breast Cancer; Chemosensitivity; Gene Signature; Notch Signaling Pathway; Pathologic Complete response; Predictor

1. INTRODUCTION

Breast cancer is a clinically heterogeneous disease that demonstrates a wide variation in its clinical courses and response to chemotherapy. This complexity is a reflection of the molecular oncogenic aberration in DNA repair, cell cycle control, cell survival, and signal transduction in breast tumors. Microarray analysis has identified breast cancer subtypes with distinct gene expression profiles and clinical behavior [1,2,3]. There are several major molecular classes of breast cancers indentified by different research groups. Some studies [2,3] suggested five major classes of breast cancer: normal breast-like, luminal-A, luminal-B, basal-like, and human epidermal growth factor receptor 2 (HER2)-positive cancers. Another study [4] proposed three major classes: ER+/HER2-, ER-/HER2-, and HER2+. The heterogeneity of breast cancer characterized by these subtypes brings great challenge to its research. In a significant proportion of breast cancer patients, chemotherapy does not result in response, but can induce significant side effects and financial costs. The ability to identify predictors of response or resistance to cancer drugs will provide better treatment to the individual patient.

Several studies have suggested that the gene-expression profiles of chemo sensitive tumors are different from those of chemo resistant ones [5]. Gene expression profiling with a measurement of thousands of mRNA transcripts in a single experiment is widely used in human cancer research. Due to the high dimensionality of microarray data, a feature selection step to find a subset of discriminative genes, referred to as a signature, is often necessary for building robust predictors [6,7].

Ayers *et al.* [8] developed a multigene predictor of pCR to sequential weekly paclitaxel and FAC (T/FAC) neoadjuvant chemotherapy for breast cancer patients. The study involved 42 patients: 24 patients were used in the training set and 18 patients in the validation set. pCR was obtained in 13 patients (31%). A gene set of 74 markers ($P < 0.09$) was built using data from the training set and tested on the validation set. Overall, a 78% predictive accuracy was achieved, with a 100% positive predictive value for pCR, a 73% negative predictive value, a sensitivity of 43%, and a specificity of 100%. Later, a follow-up study [9] included 133 patients with stage I-III breast cancer, with a pCR rate of 26% (n=34). A 30-probe set Diagonal Linear Discriminant Analysis (DLDA-30) classifier was selected for independent validation. It showed a significantly higher sensitivity (92% v 61%) than a clinical predictor including age, grade, and estrogen receptor status. This 30-probe set pharmacogenomic predictor correctly identified all but one of the patients who achieved pCR (12 of 13 patients) and all but one of those who were predicted to have residual disease had residual cancer (27 of 28 patients).

Chemosensitivity is better predicted by multigene signatures than by a single molecular discrimination because biological phenomena occur through the concerted expression of multiple genes [10,11,12]. However, within a signature of genes, the important question of how each individual gene contributes to the prediction has not been studied. We attempted in this work to identify predictors and gene signatures that have better prediction performance than the DLDA-30 and to quantify the importance of each gene in a signature in the prediction of pCR.

In [9], an exhaustive search of a good predictor of pCR was conducted. Different machine learning techniques were tested including support vector machines with linear, radial, and polynomial kernels (SVM), Diagnal Linear Discriminant Analysis (DLDA), and K-nearest neighbor (KNN) using Euclidean distance. One interesting discovery was that SVM provided the worst performance of pCR prediction among all these different techniques in this particular data set. Random Forest has demonstrated its comparable performances to SVM in many bioinformatics applications. In the current study, we sought to explore the utility of Random Forests were utilized to construct two predictors based on two signatures, the top 31 probe sets and the Notch signature, and take advantage of the feature selection capability of Random Forests to measure the importance of each gene in these signatures.

2. MATERIALS AND METHODS

2.1. Patient Cohorts and Clinical Information

One breast cancer patient cohort was obtained from a previous publication [9] (n=133). Needle-biopsy samples were collected from 133 patients with stage I, II, or III breast cancer who received preoperative weekly paclitaxel and a combination of fluorouracil, doxorubicin, and cyclophosphamide (T/FAC). These 133 patients were divided into two subsets, one training set of size 81 and one validation set of size 52. These data contain clinical information including patient age, gender, race, histological classification, stage, nuclear grade, ER (estrogen receptor), PR (progesterone receptor), and HER2 (human epidermal growth factor 2) status, pathologic complete response, and residual disease. These data also contain each patient's genome-scale gene expression profiles generated using Affymetrix U133A chip (Santa Clara, CA). pCR was defined as no residual invasive cancer in the breast or lymph nodes. pCR is presently accepted as a reasonable early indicator for long-term survival.

2.2. Top 31 Probe Set Signature

To build a predictor of pCR, the genes that are highly expressed in either the pCR cases or the RD cases need to be identified. To achieve this goal, t-tests for unequal variances for all the probe sets on Affymetrix U133A chip were carried out. The 31 probe sets (27 genes) with the smallest t-test P values (FDR=0.05%) were selected in [9], which was used as our first signature.

2.3. Notch Signature

Notch genes encode highly conserved cell surface receptors. The Notch signaling pathway consists of Notch receptors, ligands, negative and positive modifiers, and transcription factors. It plays a key role in the normal development of many tissues and cell types, through diverse effects on cell regulation, proliferation, and differentiation. Aberrant Notch signaling has been observed in several human cancers including acute T-cell lymphoblastic leukemia and cervical cancer. Recent evidences implied that it might be associated with breast cancer [13,14].

Selecting a gene signature based on differentially expressed genes between two conditions, such as pCR and RD in our study, is a common strategy nowadays. Here we endeavored to take a quite different approach, i.e., to identify a signature of genes involved in a particular pathway that has a key impact on human cancers.

The Oligo GEArray Human Notch Signaling Pathway Microarray [15] was designed for profiling expression of 113 genes involved in Notch signaling. Our second signature was these 113 genes as shown in **Table 1**.

We were particularly interested in uncovering what genes in these two signatures are important for the prediction of pCR and what biological or medical significance they might have.

2.4. False Discovery Rate

The standard P value was designed for testing individual hypotheses. When applied in a multiple testing problem such as selecting informative genes in microarray data, it may result in many false positives. While there are a number of methods to overcome the problems due to multiple testing, False Discovery Rate (FDR) approach [16,17] was used to help select the top 31 probe sets in [9].

2.5. Random Forests

Random Forest, proposed by Leo Breiman in 1999 [18], is an ensemble classifier based on many decision trees. Each tree is built on a bootstrap sample from the original training set. The variables used for splitting the tree nodes are a random subset of the whole variable set. The classification decision of a new instance is made by majority voting over all trees. About one-third of the instances are left of the bootstrap sample and not used in the construction of the tree. These instances in the training set are called "out-of-bag" instances and are used to evaluate the performance of the classifier.

Table 1. Genes in notch signaling related pathways [15].

Notch Signaling Pathway:
Notch Binding: DLL1 (DELTA1), DTX1, JAG1, JAG2.
Notch Receptor Processing: ADAM10, PSEN1, PSEN2, PSENEN (PEN2).
Notch Signaling Pathway Target Genes:
Apoptosis Genes: CDKN1A, CFLAR (CASH), IL2RA, NFKB1.
Cell Cycle Regulators: CCND1 (Cyclin D1), CDKN1A (P21), IL2RA.
Cell Proliferation: CDKN1A (P21), ERBB2, FOSL1, IL2RA.
Genes Regulating Cell Differentiation: DTX1, PPARG.
Neurogenesis: HES1, HEY1.
Regulation of Transcription: DTX1, FOS, FOSL1, HES1, HEY1, NFKB1, NFKB2, NR4A2, PPARG, STAT6.
Other Target Genes with Unspecified Functions: CD44, CHUK, IFNG, IL17B, KRT1, LOR, MAP2K7, PDPK1, PTCRA.
Other Genes Involved in the Notch Signaling Pathway:
Apoptosis Genes: AXIN1, EP300, HDAC1, NOTCH2, PSEN1, PSEN2.
Cell Cycle Regulators: AXIN1, CCNE1, CDC16, EP300, FIGF, JAG2, NOTCH2, PCAF.
Cell Proliferation: CDC16, FIGF, FZD3, JAG1, JAG2, LRP5, NOTCH2, PCAF, STIL (SIL).
Genes Regulating Cell Differentiation: DLL1, JAG1, JAG2, NOTCH1, NOTCH2, NOTCH3, NOTCH4, PAX5, SHH.
Neurogenesis: DLL1, EP300, HEYL, JAG1, NEURL, NOTCH2, PAX5, RFNG, ZIC2 (HPE5).
Regulation of Transcription: AES, CBL, CTNNB1, EP300, GLI1, HDAC1, HEYL, HOXB4, HR, MYCL1, NCOR2, NOTCH1, NOTCH2, NOTCH3, NOTCH4, PAX5, PCAF, POFUT1, RUNX1, SNW1 (SKIIP), SUFU, TEAD1, TLE1.
Others Genes with Unspecified Functions: ADAM17, GBP2, LFNG, LMO2, MFNG, MMP7, NOTCH2NL, NUMB, SEL1L, SH2D1A.
Other Signaling Pathways that Crosstalk with the Notch Signaling Pathway:
Sonic Hedgehog (Shh) Pathway: GLI1, GSK3B, SHH, SMO, SUFU.
Wnt Receptor Signaling Pathway: AES, AXIN1, CTNNB1, FZD1, FZD2, FZD3, FZD4, FZD6, FZD7, GSK3B, LRP5, TLE1, WISP1, WNT11.
Other Genes Involved in the Immune Response: CXCL9, FAS (TNFRSF6), G1P2, GBP1, IFNG, IL2RA, IL2RG, IL4, IL4R, IL6ST, IRF1, ISGF3G, OAS1, OSM, STAT5A, STUB1.

Table 2. Performance measures of three predictors: DLDA-30, RF-31, and RF-Notch.

Measures	DLDA-30	RF-31	RF-Notch
Accuracy	0.76	0.82	0.79
Sensitivity	0.92	0.55	0.27
Specificity	0.71	0.91	0.98
PPV	0.52	0.68	0.73
NPV	0.96	0.85	0.80

2.6. Feature Selection Using Random Forests

Random Forest calculates several measures of variable importance. The mean decrease in accuracy measure was used in [19] to rank the importance of the feature in prediction. This measure is based on the decrease of classification accuracy when values of a variable in a node of a tree are permuted randomly. In this study, two packages of R, randomForest and varSelRF [19], were to compute the importance of the genes in a given signature.

3. RESULTS

The first predictor, RF-31, was based on the top 31 probe sets, and the second predictor, RF-Notch, was based on the Notch signature. As in the case of DLDA-30, the RF-31 and RF-Notch were trained on the training data (n=82) and the accuracy of the two predictors was tested on a separate validation set (n=51).

Random Forests produce non-deterministic outcomes. To reduce the possible variance of our results, the Random Forests algorithm was run multiple times and then the average of the predictions was taken. The prediction results of the RF-31 and RF-Notch were based on the average of 20 repeated predictions, which are shown in **Tables 2**. The importance of each gene in the two signatures was based on the averaged calculations by using the function randomVarImpsRF in varSelRF repeated 10 times, as shown in **Figure 1** and **Table 3**.

The predictions of RF-31 and RF-Notch and the im-

Table 3. 31 genes of the highest importance in Notch signature.

Importance	Gene Symbol	Probe Set ID	P Value	*t*-Test	Higher Expression in
0.000822	CTNNB1	201533_at	0.46320	0.74521	RD
0.000928	SNW1	201575_at	0.04458	2.07657	RD
0.001295	NOTCH2	202443_x_at	0.03601	-2.23736	pCR
0.00185	NOTCH2	202445_s_at	0.02239	-2.46026	pCR
0.000658	HES1	203395_s_at	0.34906	-0.94767	pCR
0.006243	ISGF3G	203882_at	0.01170	-2.75847	pCR
0.007946	CXCL9	203915_at	0.01268	-2.73059	pCR
0.000725	MFNG	204152_s_at	0.06783	-1.92262	pCR
0.000826	LMO2	204249_s_at	0.01955	-2.44963	pCR
0.002534	IL6ST	204863_s_at	0.01141	2.66874	RD
0.003382	NEURL	204889_s_at	8.33E-05	4.15928	RD
0.007544	ADAM17	205746_s_at	0.00049	-4.00315	pCR
0.002078	IL2RA	206341_at	0.10542	-1.67547	pCR
0.001054	RUNX1	208129_x_at	0.63212	0.48313	RD
0.002165	NCOR2	208889_s_at	0.07027	1.85101	RD
0.003526	NUMB	209073_s_at	0.06476	1.88058	RD
0.000775	MAP2K7	209952_s_at	0.05390	-1.98674	pCR
0.001136	ERBB2	210930_s_at	0.03134	-2.28195	pCR
0.001296	RUNX1	211180_x_at	0.19439	-1.3392	pCR
0.00223	RUNX1	211181_x_at	0.00020	3.93243	RD
0.000769	PSEN2	211373_s_at	0.87130	0.16266	RD
0.003479	NOTCH2	212377_s_at	0.00019	3.98301	RD
0.000862	AXIN1	212849_at	0.00131	3.35556	RD
0.001427	MYCL1	214058_at	0.01948	-2.47946	pCR
0.005243	CFLAR	214618_at	0.00984	-2.76622	pCR
0.000663	CD44	216056_at	0.22213	-1.24048	pCR
0.000748	MAP2K7	216206_x_at	0.43559	0.78876	RD
0.001933	CFLAR	217654_at	0.05371	-2.02278	pCR
0.000858	FZD4	218665_at	0.00180	3.25207	RD
0.001928	IL17B	220273_at	0.08425	-1.76439	pCR

portance of the genes in the two signatures are summarized in **Table 2** and **Figure 1**. The metrics of performance in **Table 2** indicate the different strengths of the DLDA-30 and our RF-31 and RF-Notch. Both RF-31 and RF-Notch had a higher accuracy, a higher specificity, and a higher PPV than the DLDA-30.

3.1. Importance of the Genes in Top 31 Probe Sets

Of these 31 probe sets, five probe sets had a higher ex-

pression value in the pCR cases and 26 probe sets had a higher expression in the RD cases, demonstrating the dominance of the highly expressed genes in the patients with RD.

Figure 1 displays several genes of top importance, including MAPT, BBS4, MGC5370, BTG3, MELK, CA12, FGFR1OP, MTRN, FLJ10916, E2F3, RRM2, and KIF3A. MAPT, microtubule associated protein tau, was discovered as the best single gene discriminator of pCR to preoperative chemotherapy with Paclitaxel, 5-Fluoroutacil,
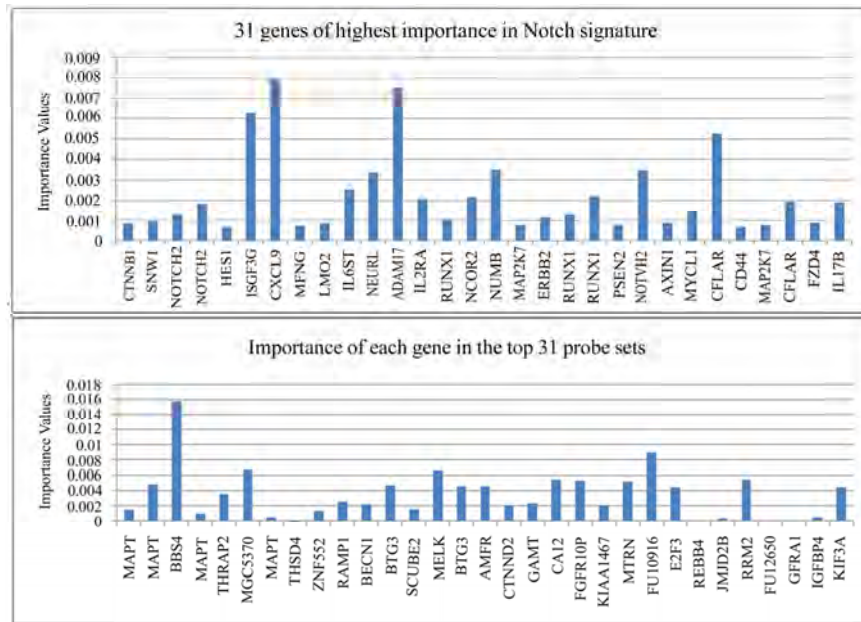
Figure 1. Two plots of the importance of the genes in the top 31 probe set signature and the Notch signature respectively.

Doxorubicin [20]. Its expression correlates closely with ER expression in human breast cancer. In the top 31 probe set signature, there were four probe sets of gene MAPT with very high *t*-test statistic. The multiple selection of MAPT in the signature demonstrates its significance. BBS4, Bardet-Biedl syndrome 4, is a member of the Bardet-Biedl syndrome (BBS) gene family associated with the Bardet-Biedl syndrome. MGC5370 is an alias name of gene MDM2, which is a target gene of the transcription factor tumor protein p53. Over expression of this gene can result in excessive inactivation of tumor protein p53, diminishing its tumor suppressor function. This gene had a very high expression value in our patient with RD. BTG3, a member of the BTG/Tob family, is a transcriptional target of p53. It has a role in DNA damage response. Its antiproliferative action through inhibition of another gene E2F1 was discovered recently [21]. This gene was highly expressed in our patients with pCR, which is consistent with this recent discovery. There were two probe sets of gene BTG3 in the top 31 probe sets, which further illustrates this gene's significance. MELK, maternal embryonic leucine zipper kinase, is a potential marker of proliferating mammary epithelial progenitor cells that are highly expressed in multiple human cancers, including human breast cancer. CA12, Carbonate dehydratase XII, is a member of a large family of zinc metalloenzymes that participate in various biological processes, and was found to be overexpressed in 10% of clear cell renal carcinomas. FGFR1OP is Fibroblast Growth Factor Receptor 1 (FGFR1) Oncogene Partner. Fusing this gene and the FGFR1 gene has been found in cases of myeloproliferative disorder. This gene plays an

important role in normal proliferation and differentiation of the erythroid lineage. MTRN, Meteorin, glial cell differentiation regulator, is a gene clearly involved in cell differentiation. FLJ10916 is an alias name of gene THNSL2, threonine synthase-like 2, which functions in lyase activity, pyridoxal phosphate binding, and metabolic process. E2F3, E2F transcription factor 3, is a member of the E2F family of transcription factors. The E2F family is essential in the control of cell cycle and action of tumor suppressor proteins. *RRM2,* Ribonucleotide reductase M2 polypeptide, provides the precursors necessary for DNA synthesis. During mitosis, Kinesin family member 3A (KIF3A) has a critical function in the equal segregation of chromosomes between two daughter cells.

Based on the above functional annotation, it is evident that these top important genes are not only vital in the prediction of pCR but also strongly implicated in tumourigenesis.

3.2. Importance of the Genes in Notch Signature

Of the 31 genes with highest importance values in Notch signature, 17 probe sets had a higher expression value in the pCR cases and 14 probe sets had a higher expression value in the RD cases as seen in **Table 3**. This somewhat even distribution of the probe sets between the pCR and RD cases was in contrast to the top 31 probe set signature, which could be attributed to the functions of Notch signaling pathway.

The top important genes in Notch signature were the following: CXCL9, ADAM17, ISGF3G, CFLAR, NUMB,

NOTCH2, NEURL, IL6ST, and RUNX1. NOTCH2 is a multifunctional gene involved in apoptosis, cell proliferation, cell differentiation, neurogenesis, and regulation of transcription. There are three probe sets of NOTCH2 and two probe sets of CFLAR in Figure 1, reflecting these genes' importance. CXCL9, ISGF3G, and IL6ST are all involved in immune response. Since the functions of these genes are illustrated through their pathways in **Table 1**, we will not elaborate on them any further here.

There were 15 genes in the top 31 probe set signature with importance values above 0.003, and there were seven such genes in Notch signature. This was expected. Because of their high *t*-test statistics, the top 31 probe sets should be more sensitive to the random permutation employed in the importance calculation than those in the Notch signature. Nonetheless, in **Figure 1** the Notch signature genes displayed their significance.

4. CONLUSIONS

Random Forests were employed to study the prediction of pathologic complete response in breast cancer, and the results improved the predictions of the DLDA-30. Functional annotation demonstrated that the important genes identified by the feature selection scheme of Random Forests are of biological significance.

5. ACKNOWLEDGMENT

REFERENCES

[1] Nguyen, P. L., Taghian, A. G. *et al*. (2008) Breast cancer subtype approximated by estrogen receptor, progesterone receptor, and HER-2 is associated with local and distant recurrence after breast-conserving therapy, *J. Clin Oncol.*, **26(14),** 2373-8.

[2] Perou, C. M., Sorlie, T., *et al*. (2000)Molecular portraits of human breast tumours, *Nature*, **406(6797)**, 747-752.

[3] Sorlie, T., Perou, C. M. *et al*. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications, *Proc Natl Acad Sci U S A,* **98(19)**, 10869-10874.

[4] Kapp, A. V., Jeffrey, S. S. *et al*. (2006) Discovery and validation of breast cancer subtypes. *BMC Genomics*, **7**, 231.

[5] Pusztai, L., (2008) Current status of prognostic profiling in breast cancer, *The Oncologist*, **13**, 350-360.

[6] van 't Veer, L. J., Dai, H., *et al*. (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, **415**, 530-536.

[7] van de Vijver, M. J. *et al*. (2002) A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*, **347**, 1999-2009.

[8] Ayers, M., Symmans, W. F., Stec, J. *et al*. (2004) Gene expression profiles predict complete pathologic response to neoadjuvant paclitaxel/FAC chemotherapy in breast cancer. *J Clin Oncol,* **22**, 2284-2293.

[9] Hess, K. R., Anderson, K., Symmans, W. F. *et al*., (2006) Pharmacogenomic predictor of sensitivity to preoperative chemotherapy with paclitaxel and fluorouracil, doxorubicin, and cyclophosphamide in breast cancer. *J Clin Oncol*, **24**, 4236-4244.

[10] Brenton, J. D., Carey L. A., Ahmed, A. A. *et al*. (2005) Molecular classification and molecular forecasting of breast cancer: ready for clinical application? *J Clin Oncol*, **23**, 7350-7360.

[11] Wang, Y., Klijn, J. G. *et al*. (2005) Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer, *Lancet*, **365**, 671-679.

[12] Ma, X. J., Wang, Z., *et al*. (2004) A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. *Cancer Cell*, **5(6)**, 607-16.

[13] Brennan, K. and Anthony Brown, M. C. (2003) Is there a role for Notch signalling in human breast cancer? *Breast Cancer Res*, **5(2)**, 69-75.

[14] Stylianou, S., Clarke, R. B. *et al*. (2006) Activation of notch signaling in human breast cancer, *Cancer Research*, **66**, 1517-1525.

[15] GEArray, O. Human notch signaling pathway microarray. http://www.sabiosciences.com/gene_array_product/HTML/OHS-059.html

[16] Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol,* **57**, 289-300.

[17] Pounds, S. and Morris, S. W. (2003) Estimating the occurrence of false positive and false negatives in microarray studies by approximating and partitioning the empirical distribution of p-values. *Bioinformatics* **19**, 1236-1242.

[18] Breiman, L. and Random, F. (2001) *Machine Learning*, **45 (1)**, 5-32.

[19] Díaz-Uriarte, R. and Alvarez de Andrés, S. (2006) Gene selection and classification of microarray data using random forest. *BMC Bioinformatics,* **7(3)**.

[20] Rouzier, R., Rajan, R., *et al*. (2005) Microtubule associated protein tau is a predictive marker and modulator of response to paclitaxel-containing preoperative chemotherapy in breast cancer. *Proc Natl Acad Sci U S A*, **102**, 8315-8320.

[21] Ou, Y. H., Chung, P. H., *et al*. (2007)The candidate tumor suppressor BTG3 is a transcriptional target of p53 that inhibits E2F1, *The EMBO Journal*, **26**, 3968-3980.

Phosphatidylinositol transfer proteins: sequence motifs in structural and evolutionary analyses

Gerald J. Wyckoff[1], Ada Solidar[2], Marilyn D. Yoder[3]

[1]Division of Molecular Biology and Biochemistry, University of Missouri-Kansas City, Kansas City, USA;
[2]Vassa Informatics, Kansas City, USA;
[3]Division of Cell Biology and Biophysics, University of Missouri-Kansas City, Kansas City, USA.
Email: wyckoffg@umkc.edu

ABSTRACT

Phosphatidylinositol transfer proteins (PITP) are a family of monomeric proteins that bind and transfer phosphatidylinositol and phosphatidylcholine between membrane compartments. They are required for production of inositol and diacylglycerol second messengers, and are found in most metazoan organisms. While PITPs are known to carry out crucial cell-signaling roles in many organisms, the structure, function and evolution of the majority of family members remains unexplored; primarily because the ubiquity and diversity of the family thwarts traditional methods of global alignment. To surmount this obstacle, we instead took a novel approach, using MEME and a parsimony-based analysis to create a cladogram of conserved sequence motifs in 56 PITP family proteins from 26 species. In keeping with previous functional annotations, three clades were supported within our evolutionary analysis; two classes of soluble proteins and a class of membrane-associated proteins. By, focusing on conserved regions, the analysis allowed for in depth queries regarding possible functional roles of PITP proteins in both intra- and extra- cellular signaling.

Keywords: Protein Evolution; Structural Domain; Phylogenetics; Sequence Motif

1. INTRODUCTION

Phosphatidylinositol transfer proteins (PITP) are monomeric, lipid-binding proteins that bind and transfer phosphatidylinositol (PtdIns) and phosphatidylcholine (PtdCho) between membrane compartments (see reviews by: [1,2,3,4,5,6]. Inositol lipids have specialized functions in the regulation of eukaryotic cells, providing a source of second messengers and acting as signaling molecules. Monomeric phospholipids have extremely low solubilities and negligible spontaneous transfer rates between membranes, necessitating protein factors to shield them in the aqueous environment of the cell. Almost all phospholipid exchange activity within the eukaryotic cytosol is accomplished by three groups of proteins: PITP, phosphatidylcholine transfer protein (PCTP), or sterol carrier protein 2. PITP-domain proteins, which are the focus of this study, are found in five classes of proteins: three are soluble and cytosolic, and two are membrane-associated proteins [1,2] (**Figure 1**).

Soluble PITPs are found in virtually all metazoan organisms. These are approximately 77% identical at the amino-acid level. Two isoforms, PITPα and PITP, exist in mammals; they are highly conserved with about 98% sequence identity at the amino acid level. The proteins bind one molecule of PtdIns or PtdCho and are typically about 32 kDa in mass.

PITP-like domains are also detected in the retinal degeneration B (rdgB) class of proteins (see reviews [6,7]). These 160-170 kDa proteins are membrane-associated proteins first identified in *Drosophila melanogaster*, as mutations in these genes lead to retinal degeneration. The rdgBs contain an
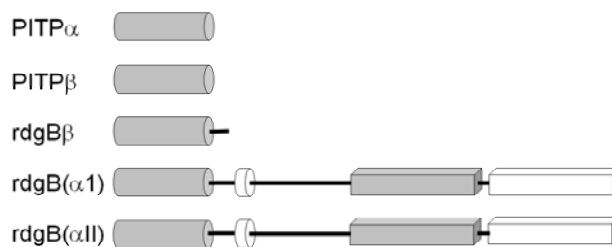

Figure 1. Gene structure of PITP-domain proteins. Grey cylinder is PITP-like domain, white cylinder is the FFAT region of ER binding, the grey box is the DDHD region, and the white box is the LSN2 region. The the human pitpnm1 protein, the FFAT region is residues 360-365, the DDHD region is 686-879, and the LSN2 domain is residues 1022-1152 [6].
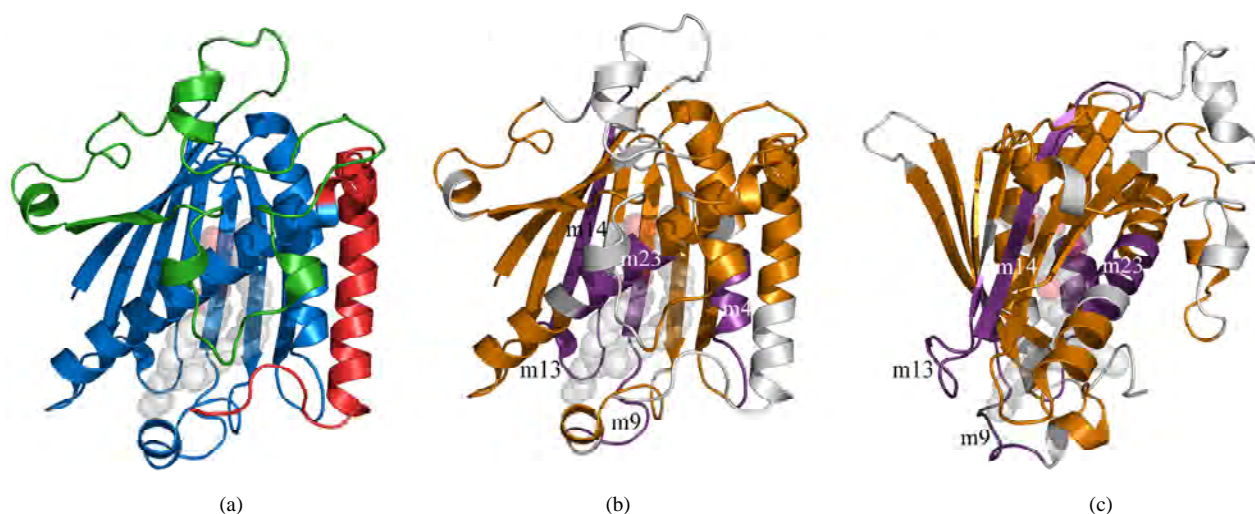
Figure 2. Cartoon representation of rat PITPα complexed to PtdCho. (a). Structural domains of the class I PITPs displayed on a representative PITP (PDB:1T27). The lipid binding core is shown in blue, the regulatory loop in green, and the C-terminal region in red. The PtdCho molecule is in transparent spheres with standard CPK coloring; (b) Conserved sequence motifs mapped onto the same structure. The four conserved sequence motifs; m4, m9, m13, m14, and m23, representing branch points in the clado-gram shown in **Figure 3** are in purple and are annotated. The remaining sequence motifs are in gold, regions not present in conserved sequence motifs are colored grey; (c). Image B is rotated 90º counterclockwise around a vertical axis.

N-terminal PITP-like domain (42% sequence identity to PITP isoforms), a FFAT sequence motif [2,8], a DDHD metal binding domain (which has calcium-binding capabilities in the human proteins [9]), and a LNS2 domain. The FFAT motif is a short sequence containing two phenylalanines in an acidic tract that targets the protein to the endoplasmic reticulum [8]. The DDHD domain of 180 residues may be involved in metal binding, but the function of this domain is unknown [6]. The LNS2 domain is believed to be involved in protein: protein interactions, and in the human homologues is a protein tyrosine kinase 2 (PYK2) binding domain [9]. Two homologues of rdgB are found in most mammalian genomes studied to date and are usually called 1) rdgB(I), rdgBαI, or PITPnm1 and 2) rdgB(II), rdgBαII, or PITPnm2. In *Drosophila,* rdgB is believed to function in the termination of phototransduction and in the establishment and maintenance of rhodopsin levels in photoreceptor cells [10,11]. In humans, PITPnm1 has widespread tissue distribution and can rescue fly rdgB mutant phenotypes [1]; whereas PITPnm2 has a neuronal-specific expression pattern and is unable to functionally rescue fly rdgB mutants [12]. Although *Drosophila* rdgB possesses the capability to transfer PtdCho and PtdIns *in vitro* [10,11], the PITP-domains of rdgB and PITPα are not functionally interchangeable [10].

An additional class of PITP-like proteins, rdgBβ, has been identified in mice, humans, and *Drosophila*. These are 38 kDa, soluble proteins that have sequence similarity most comparable to the N-terminal region of rdgB-class proteins. It shares approximately 40% sequence

identity with PITPα/. The purified protein has been shown to possess *in vitro* PtdIns transfer capabilities [13,14].

Plants and fungi generally do not contain a sequence homologue to PITP, but do possess a functional analog referred to as Sec14p in yeast systems. Sec14p is approximately the same size as PITP, and although there is no detectable amino acid similarity between the two proteins, temperature-sensitive mutants of yeast Sec14p are rescued by rat PITPα and PITP [15,16]. Likewise, Sec14p can successfully substitute for PITP in the PITP-dependent reconstitutions studied to date [17,18,19,20]. Interestingly, the slime mold *Dicytostelium discoideum* has been shown to contain not only homologues to PITP, called DdPITP1 and DdPITP2, but also a Sec14p relative, called DdSec14 [21].

The experimentally determined three-dimensional structures of rat PITPα-PtdCho [22] and PITP-PtdCho [23], human PITPα-PtdIns [24], and the apo form of mouse PITPα [25] have been reported. The PITP structures share little resemblance to the crystal structure of yeast Sec14p [26]. The PITP structure is composed of three regions (**Figure 2(a)**): a lipid-binding core, a regulatory loop, and the C-terminal region. The lipid-binding core of PITPα-PtdCho shares a fold with the steroidogenic acute regulatory protein-related transfer (START) domain [22] first observed in human MLN64 [27]. It has been proposed that the START domain may be a common fold adapted for binding lipid molecules. PCTP is a START-domain protein based on structure and sequence
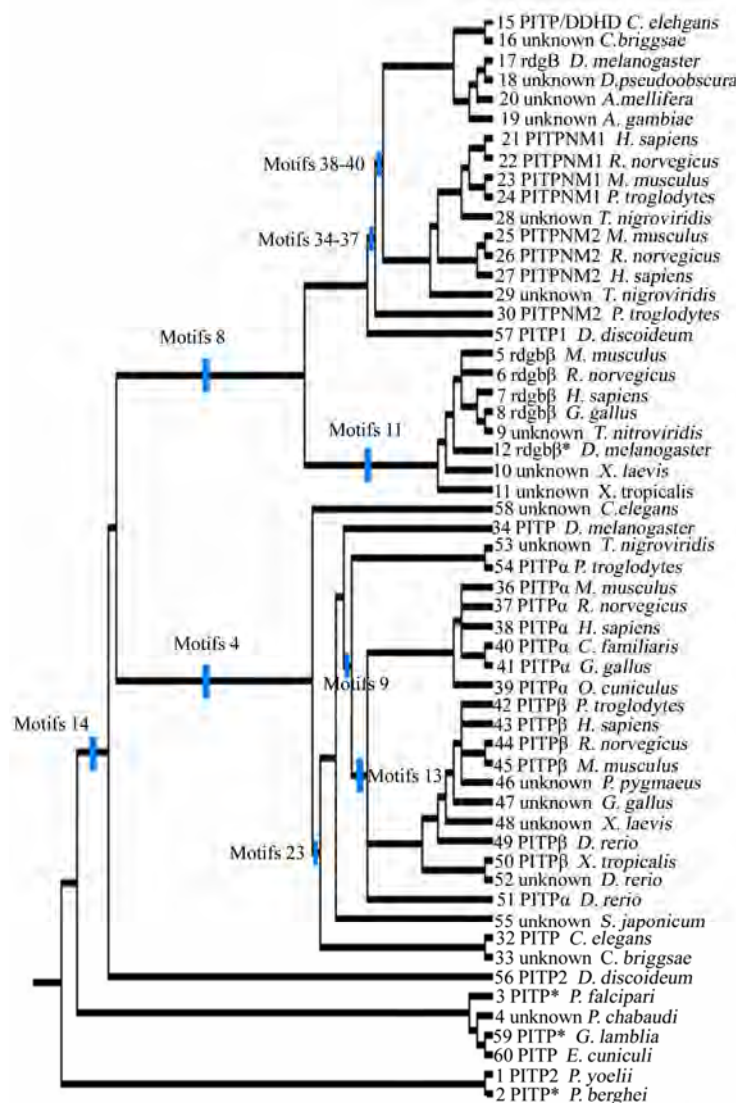
Figure 3. Phylogenetic tree for 56 PITP family proteins. A strict consensus tree was derived from the presence/absence of sequence motifs as well as the motif sequence, analyzed via parsimony as described in the text. The blue lines show which sequence motifs are uniformly gained within each clade on the tree.

identity, whereas PITP is not considered to be a START-domain protein due to a lack of sequence identity.

Mutations and gene-knockout studies provide insight into the functions of PITP. In mice, a mutation in the PITPα gene causes the *vibrator* phenotype, which is characterized by a progressive-action tremor, degeneration of brain stem and spinal cord neurons, and juvenile death. PITPβ does not compensate for loss of PITPα [40]. Furthermore, in mice, embryonic stem cells deficient in PITPα or PITPβ reveal differences in physiological function between the two isoforms. PITPβ deficiency leads to catastrophic failure early in embryonic

development, and the protein is therefore posited to have an essential housekeeping role in the cell [29]. In contrast, PITPα-deficient embryonic stem cells are not compromised in growth or in bulk phospholipid metabolism; however, PITPα is required for neonatal survival. PITPα deficiency affects regulation of phospholipid transport in the ER, endocrine pancreas function, and glycogen metabolism, due to compromised lipid absorption, homeostatic problems, and severe hypoglycemia [30].

Protein similarity is usually detected by conservation of function as revealed by biochemical analysis, sequence
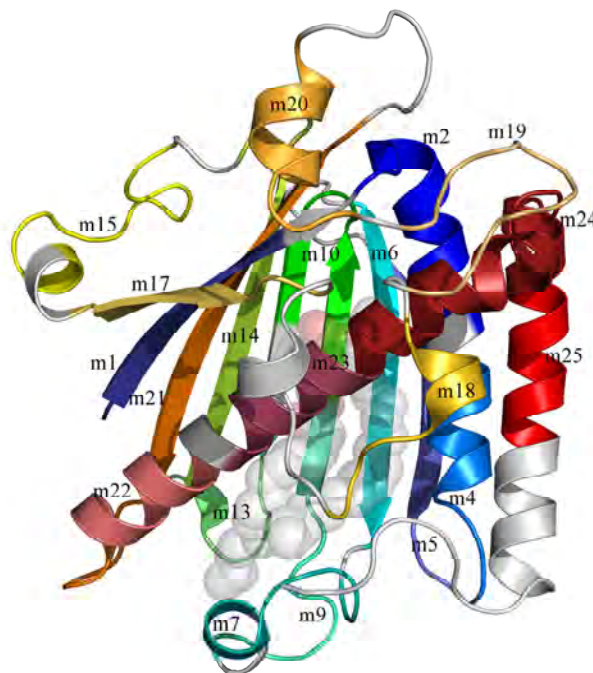
Figure 4. Sequence motifs mapped to the structure of rat PITPα:PtdCho. The sequence motifs present in class I PITPs are shown in a rainbow spectrum of colors from blue to red and from the *N*- to the *C*-terminus. Sequence motifs are annotated 'm1' to m25', the PtdCho molecule is in transparent spheres with standard CPK coloring. The lipid binding core contains sequence motifs 1-2, 4-7, 9-10, 13-15, and 17-25. The regulatory loop contains sequence motifs 15, 17, and 18-20. The C-terminal region contains part of sequence motif 24 and all of motif 25.

similarity as detected by amino acid pattern recognition, or by structural similarity as detected by X-ray crystallography or NMR spectroscopy. The PITP/rdgB proteins possess intriguing patterns of similarity. PITP has functional similarity to Sec14p, but lacks sequence or structure similarity. PITPs have sequence similarity to rdgB and both possess the hallmark ability to transfer phospholipids between the protein and membranes. PITP has been shown to have structural similarity to the START domains, even though it does not share amino acid sequence similarity. Here, we have used a comprehensive evolutionary analysis to synthesize information from sequence analysis and structural comparisons. This approach leads to a cohesive understanding of evolutionary, structural, and functional relationships of the PITP/rdgB protein families, and may suggest that a re-analysis of PITP protein naming conventions is overdue.

2. MATERIALS AND METHODS

2.1. Identification of Sequences

Rat PITPα (gi:8393962) was used as a query sequence in a BLASTp (version 2.2.11, database versions as of May 8, 2005) [31] search. The 134 returned sequences were manually culled to remove duplicated or significantly fragmented sequences, reducing the number to 60 sequences in 26 species (**Table 5**). The final data set contained both cytosolic PITP and rdgBs of approximately 270-330 residues each, as well as the membrane-associated rdgB/PITPnm proteins, of approximately 1250 residues. Several proteins were determined to be splice variants of existing proteins in the tree; they are annotated in all figures and tables.

2.2. Phylogenetic Analysis

Unaligned proteins were processed using the MEME program [32] (implemented in the Wisconsin Package, version 10.4) [33], using the "zero or more" setting, which permits anywhere from zero instances of a block to extreme repetition of blocks, thereby permitting the generation of the most robust matrix of sequence motifs possible. The phylogenetic parsimony tree for the sequence motifs was made by PAUP* (version 4.0b) [34]. The consensus tree was imported into MacClade (v 4.08) [35] for further character-state analysis. Character-state changes were traced along tree clades and unambiguous changes were examined. A database of sequence motifs and proteins was created to facilitate this study and the

Table 1. Conserved sequence motif matrix of PITP and rdgB proteins, PITP domain only.

Track-ing #	Sequence motif identifier																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	1	0
2	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1	0	1	0
3	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	1	0
4	1	1	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
59	0	1	0	0	1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
60	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
56	1	1	0	1	1	1	0	0	1	1	0	0	0	1	1	0	1	0	1	0	1	0	0	1	0
15-16	1	1	1	0	1	1	0	1	0	1	0	0	0	1	0	1	0	1	1	0	1	0	0	1	0
17-19, 25, 28-29	1	1	1	0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	1	0	0	1	0
20	0	0	1	0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	1	0	0	1	0
21-24	1	1	1	0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	1	0	0	1	1
26	0	0	0	0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	0	1	0	0	1	0
27	1	1	1	0	1	1	1	1	0	1	0	1	0	1	0	0	1	1	1	0	1	0	0	1	0
30	1	1	1	0	1	1	1	1	0	1	0	1	0	1	0	0	1	1	1	0	1	0	0	0	0
57	1	1	0	0	1	1	1	1	0	1	0	0	0	1	1	0	0	1	1	0	1	0	0	1	0
5-7	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0	0	1	1	1	0	1	0	0	1	0
8	0	1	1	0	1	1	1	1	0	1	1	0	0	1	0	0	1	1	1	0	1	0	0	1	0
9	0	1	1	0	1	1	1	1	0	1	1	0	0	1	0	0	1	0	1	0	1	0	0	1	0
10	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0	0	1	1	1	0	1	0	0	0	0
11	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0	0	1	0	1	0	1	0	0	1	0
12	1	1	1	0	1	1	1	1	0	1	1	0	0	0	0	0	1	1	1	0	1	0	0	1	0
58	1	1	0	1	1	1	0	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	0	1	0
34	1	1	0	1	1	1	1	0	1	1	0	0	0	1	1	0	1	1	1	0	1	1	1	1	0
53	1	1	0	1	1	1	1	0	1	1	0	0	1	1	1	0	1	1	1	0	0	0	1	1	1
54	0	1	0	1	1	1	0	0	1	1	0	0	1	1	1	0	1	1	1	0	0	0	1	1	1
36-39, 42-51	1	1	0	1	1	1	1	0	1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1
40-41, 52	0	1	0	1	1	1	1	0	1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	1
55	1	1	1	0	1	1	1	0	0	1	0	0	0	1	0	0	1	1	1	0	1	0	0	1	1
32	0	1	0	1	1	1	0	0	1	1	0	0	0	1	0	0	1	0	1	0	1	0	0	1	0
33	1	1	0	1	1	1	0	0	1	1	0	0	0	1	0	0	1	0	1	0	1	0	0	1	0

tables from it are available upon request.

In this phylogenetic analysis, the tree root was desig-nated by the branch that incorporated the most sequence motif additions and that led to the most consistent char-acter-state changes along the tree. That is, the basal prot-eins that rooted the tree had the fewest sequence motifs character-state changes within the tree with this root present. This rooting was borne out by the consistency and was utilized. This rooting suggested the fewest character state changes and least amount of homoplasy within the tree. The tree root consisted of small PITP proteins from *Plasmodium*, *Giardia*, and *Encephalitozoon*. Sequence motif 14 defined this branch point, proteins from these species lacked the motif, and all other proteins possessed

Table 2. Conserved sequence motif matrix of rdgB proteins, C-terminal of the PITP-domain.

Trac-king #	Sequence motif identifier																				
	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
15	0	0	0	0	1	1	0	0	1	1	1	1	1	1	1	0	1	0	0	0	0
16	0	0	0	0	1	1	0	0	1	1	1	1	1	1	1	0	1	0	0	0	1
17,21-22, 26, 28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
18	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	0
20	0	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0
19	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	0	0
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
24	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
25	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0
29	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0
30	1	0	1	1	1	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0

Table 3. Gene and protein nomenclature for proteins utilized in this study.

	PITPα	PITP	rdgBβ	rdgB(1)	rdgB(2)
Gene name	Pitpna (mouse) Pitpn (rat) Pitpna (human)	Pitpnb (mouse) Pitpnb (rat) Pitpnb (human)	pitpnc1 (human)	pitpnm1 (mouse) NIR2 rdgBα1 (fly) DRES9	pitpnm2 (mouse) NIR3 rdgBα2 (fly)
Gene symbols	Pitpna (mouse) Pitpn (rat) Pitpna (human)	Pitpnb (mouse) Pitpnb (rat) Pitpnb (human)	pitpnc1 (human)	pitpnm1 (mouse)	pitpnm2 (mouse)
Gene aliases	Pitpn, vib1A	Vib1B		Rd9; RdgB; DRES9; mpt-1; Pitpnm	NIR3; rdgB2; rdgB 2; mKIAA1457
Protein name	PITPα	PITP	rdgBβ ; rdgB 1; PITP, cyto-plasmic 1*	PITP mem-brane-associated 1	PITP mem-brane-associated 2
Approved gene symbol	PITPNA	PITPNB	PITPNC1	PITPNM1	PITPNM2

the conserved sequence motif.

3. RESULTS

3.1. Characterization of PITP-Like Proteins

An initial approach to characterizing the PITP family was relatively standard. Rat PITPα was used as the seed for a BLAST search targeting relatively distant protein relatives. BLAST results were imported to Pileup [33] to generate a "first pass" global alignment. Under a variety of different conditions, including weighting end gaps, not weighting end gaps, utilizing low gap opening and extension penalties, high-road and low-road alignment options, alignments were very weak, with notable cases where locally homologous segments identified by BLAST were not aligned in the global matrix. The derived nature

of many of the proteins in the PITP family was already apparent; sequence motifs appeared in some classes but not others, rendering standard alignment and analysis methods ineffective. A novel approach was required to identify and analyze the large, divergent PITP family.

3.2. Identification of Conserved Sequence Motifs

When a high-confidence global alignment in Pileup proved impossible, unaligned proteins were processed using MEME [32] which identifies conserved sequence motifs in proteins. The program makes identifications based on prior probabilities of amino acid occurrence. Because it performs ungapped determinations, no preceding global alignment of sequence motifs was necessary, families with highly derived members. Of the 46 conserved sequence motifs detected by MEME, 25 fell

Table 4. Conserved sequence motifs in the PITP-domain.

Motif Identi-fier	secondary structure element, PITPα[a]	Sequence range[b]: PITPα PITPnm1 PITPNC1	Motif sequence[b]: PITPα PITPnm1 PITPNC1
1	sheet 1 (3-11)	2-9	2-VLLKEYRV-9
1		1-8	1-MLIKEYHI-8
1		1-8	1-MLLKEYRI-8
2	helix A (14-33)	12-24	PVSVDEYQVGQLY
2		11-23	PMSLDEYQVAQLY
2		11-23	PLTVDEYKIGQLY
3		-	--------
3		24-31	MIQKKSRE
3		24-31	MISKHSHE
4	helix A (14-33)	26-36	VAEASKNETGG
4		-	-----------
4		-	-----------
5	sheet 2 (39-49)	37-48	GEGVEVLVNEPY
5		37-48	GSGVEILANRPY
5		36-47	GEGVEVVQNEPF
6	sheet 3 (55-64)	55-64	GQYTHKIYHL
6		56-65	GQYTHKVYHV
6		55-64	GQFTEKRVYL
7	helix B (70-75)	66-73	SKVPTFVR
7		204-211	AKIEQFIH
7		198-205	TRVEQFVH
8		-	-----------
8		69-79	IPGWFRALLPK
8		68-78	LPSWARAVVPK
9		76-86	APEGALNIHEK
9		-	-----------
9		-	-----------
10	sheet 4 (84-91)	87-96	AWNAYPYCRT
10		88-97	SWNAYPYTRT
10		86-95	AWNYYPYTIT
11		-	--------
11		-	--------
11		97-104	YTCSFLPK
12		-	--------
12		98-105	RYTCPFVE
12		-	--------
13	sheet 5 (94-100)	99-106	TNEYMKED
13		-	--------
13		-	--------
14	sheet 6 (108-117)	107-120	FLIKIETWHKPDLG
14		107-120	FSIEIETYYLPDGG
14		105-118	FSIHIETKYEDNKG
15	helix C (132-137)	123-134	ENVHKLEPEAWK
15		-	-----------
15		-	-----------
16		-	-----------------------
16		124-146	NVFNLSGAERRQRILDTIDIVRD
16		-	-----------------------
17	sheet 7 (138-142)	137-144	EAVYIDIA
17		-	--------
17		134-141	EVCFIDIA
18		153-160	DYKAEEDP
18		152-159	EYKAEEDP
18		149-156	YYKESEDP
19		163-173	FKSIKTGRGPL
19		162-172	YHSVKTGRGPL
19		159-169	FKSEKTGRGQL
20	helix E (177-183)	175-183	PNWKQELVN
20		-	---------
20		-	---------
21	sheet 8 (191-201)	190-206	MCAYKLVTVKFKWWGLQ
21		187-203	MCAYKLCKVEFRYWGMQ
21		181-197	MCSYKLVTVKFEVWGLQ

22	helix F (206-232)	207-214	NKVENFIH
22		-	--------
22		-	--------
23	helix F (206-232)	217-224	ERRLFTNF
23		-	--------
23		-	--------
24	helix F (206-232)	225-244	HRQLFCWLDKWVDLTMDDIR
24		223-242	HRQAWCWQDEWTELSMADIR
24		216-235	HRQAFAWVDEWYDMTMDEVR
25	helix G (240-261)	245-253	RMEEETKRQL
25		243-252	ALEEETARML
25		-	----------
26		-	--------
26		294-301	PPGPDASP
26		269-276	RSAPSSAP

[a]Secondary structure nomenclature for rat PITP as defined in [22], parenthesis indicate the complete range of the secondary structure element.

[b]Sequences and residue numbering refer to human proteins, PITPα is gi:5453908, tracking #38; PITPnm1 is gi:18490106, tracking #21; PITPNC1 is gi:32307140, tracking #7. PITPNC1 is also called rdgBβ.

within PITP-like domains.

3.3. Phylogenetic Analysis

With sequence motifs identified, two character-state matrices were created. One contained character-state information on the presence or absence of detected motifs, and one contained aligned amino acid data (**Tables 1** and **2**). These matrices were utilized for the creation of a character state phylogenetic tree via maximum parsimony. The overall shape of the tree was determined by the presence/absence matrix, with the amino acid alignment data utilized to determine internal clade structures. Gaps, i.e., missing sequence motifs were considered a new state, and thus, the difference between having a specific set of amino acids and not having an amino acid could served as an informative character state.

Twelve most parsimonious trees of differing structures were identified by PAUP, but the strict consensus tree structure suggests the likely evolutionary history of this intriguing family of proteins (**Figure 3**). The tree revealed that membrane-associated proteins are derivatives of a group of PITP-like genes, and that many genes previously thought to be related to PITPα and PITP via sequence homology fell outside of clades that contain exclusively PITPα or PITP genes.

3.4. Identification of Clades Within the Phylogenetic Tree

In order to be cladistically sound, nomenclature must be hierarchical; all proteins in the tree were "PITP-like." The tree lends itself to consideration of the PITP-like family consisting of three large divisions (the classification nomenclature proposed by Allen-Baume [1] is adapted here): Class I comprised the soluble PITPs, Class IIA were the membrane-associated rdgB proteins (which contain additional domains C-terminal to the PITP-like domain), and Class IIB were the soluble rdgBβ proteins. Further subdivisions along cladistic lines allowed for the naming of the Alpha clade and the Beta clade within Class I, and an ancestral group of PITP-like proteins, predominantly from protista, that occured in several clades rooting the tree. **Table 3** lists the nomenclature commonly utilized for the genes and proteins in this study, as well as the HUGO-approved nomenclature.

3.5. Examination of Conserved Sequence Motifs

To understand the functional shifts that have occurred, examination of some of the sequence motifs in each proposed class was necessary. The amino acid sequence of each motif in the human proteins PITPα (Class I), PITPNM1 (Class IIA), and PITPNC1, also called rdgBβ (Class IIB) are given in **Table 4**.

Class I: These proteins are cytosolic and contain sequence motifs 1-2, 4-7, 9-10, 13-15, and 17-25 (**Figure 4**). This class was distinguished from the Class II divisions by sequence motif 8, which was unique to the Class IIA and IIB clades, and sequence motif 4, which was unique to the non-protista Class I PITPs. A striking observation about sequence motif 8 was that it appears to be an amino acid sequence overlap with sequence motif 9, which defines subnodes within Class I. In the structure of rat PITPα, the residues defined by sequence motif 9 incorporate helix B, which is a putative membrane insertion helix [25]. There was no unique conserved sequence motif that distinguishes the functionally divergent alpha and the beta clades.

Class IIA: These proteins are membrane-associated and typically are 160 kDa. In this analysis they were distinguished by the unique presence of sequence motifs 34-40, which consisted of amino acids residues 831-844, 856-866, 899-920, 937-951, 993-1015, 1022-1038, and 1040-1086 in the human PITPnm1 protein. The division had two major subnodes, one consisting of non-mamma-

Table 5. Proteins used in phylogenetic analysis.

gi #	Tracking #	Species	common	Protein designations
55243798	19	*Anopheles gambiae*	mosquito	
48135931	20	*Apis mellifera*	honeybee	
39590557	16	*Caenorhabditis briggsaae*	nematode	
39590636	33	*Caenorhabditis briggsaae*	nematode	CB09751 [b] (NCBI-COG: PITP)
17556182	32	*Caenorhabditis elegans*		
17554244	15	*Caenorhabditis elegans*	nematode	PITP DDHD (NCBI-COG: PITP)
48059855	58	*Caenorhabditis elegans*	nematode	Y71G12B.17 [b] (NCBI-COG: PITP)
57091327	40	*Canus familiaris*	dog	similar to PITPα [a]
41055500	49	*Danio rerio*	zebrafish	PITP
41055576	51	*Danio rerio*	zebrafish	similar to PITP
28422482	52	*Danio rerio*	zebrafish	
8307957	56	*Dictyostelium discoideum*	slime mold	PITP 2
8307955	57	*Dictyostelium discoideum*	slime mold	PITP 1
24641869	17	*Drosophila melanogaster*	fruit fly	CG11111-PB, isoform B, rdgB
62484257	12	*Drosophila melanogaster*	fruit fly	CG17818-PA, rdgBβ
7300495	35	*Drosophila melanogaster*	fruit fly	CG5269-PA PITP, vib
20151901	34	*Drosophila melanogaster*	fruit fly	SD01527p, vib
54642914	18	*D.pseudoobscura*	fly	GA10766-PA, *Dpse\GA10766*
19170839	60	*Encephalitozoon cuniculi*		PITP
50758480	41	*Gallus gallus*	chicken	similar to PITPα [a]
50756343	31	*Gallus gallus*	chicken	similar to PITP, membrane-associated 2 [a]
50755397	13	*Gallus gallus*	chicken	similar to rdgBβ [a]
53134209	47	*Gallus gallus*	chicken	
50757849	8	*Gallus gallus*	chicken	similar splicing variant rdgBβ [a]
29250063	59	*Giardia lamblia*		similiar to *D. discoideum* PITP1
18490106	21	*Homo sapiens*	human	PITPNM1
24308237	27	*Homo sapiens*	human	PITPNM2
5453908	38	*Homo sapiens*	human	PITPα
6912594	43	*Homo sapiens*	human	PITPβ
32307140	7	*Homo sapiens*	human	PITP, rdgBβ 1
6679337	36	*Mus musculus*	mouse	PITPα
9790159	45	*Mus musculus*	mouse	PITPβ
22003862	5	*Mus musculus*	mouse	rdgB
6679339	23	*Mus musculus*	mouse	PITPNM1
47124324	25	*Mus musculus*	mouse	PITPNM2
2137007	39	*Oryctolagus cuniculus*	rabbit	PITPα
55660937	42	*Pan troglodytes*	chimp	PITP [a]
55644759	54	*Pan troglodytes*	chimp	similar to PITPα [a]
55639157	30	*Pan troglodytes*	chimp	similar to PITP membrane-associated 2 [a]
55636463	24	*Pan troglodytes*	chimp	similar to PITP membrane-associated 1 [a]
56493706	2	*Plasmodium berghei*		PITP [b]
56509884	4	*Plasmodium chabaudi*		protein [b]
23619433	3	*Plasmodium falciparum*		PITP [b]
23485539	1	*Plasmodium yoelii yoelii*		PITP 2
55731914	46	*Pongo pygmaeus*	orangutan	protein [b]
56605814	22	*Rattus norvegicus*	rat	PITPnm, PI membrane-associated [a]
62658898	26	*Rattus norvegicus*	rat	similar KIAA1457, PITPnm2 [a]
8393962	37	*Rattus norvegicus*	rat	PITP, PITPn
16758568	44	*Rattus norvegicus*	rat	PITP
62657241	6	*Rattus norvegicus*	rat	rdgB
56756428	55	*Schistosoma japonicum*		Unknown, similiar to *vib* (fly)
47228496	53	*Tetraodon nigroviridis*	pufferfish	
47228470	28	*Tetraodon nigroviridis*	pufferfish	
47223953	9	*Tetraodon nigroviridis*	pufferfish	
47206861	14	*Tetraodon nigroviridis*	pufferfish	
47226436	29	*Tetraodon nigroviridis*	pufferfish	
47937798	48	*Xenopus laevis*	frog	MGC84500 protein [b]
49256286	10	*Xenopus laevis*	frog	MGC84224 protein
62860160	11	*Xenopus tropicalis*	frog	protein LOC549393 [b]
38512077	50	*Xenopus tropicalis*	frog	PITPβ

[a] indicated as a predicted protein, [b] indicated as a hypothetical or a putative protein

lian proteins and containing the canonical rdgB protein from *D. melanogaster*, and the other subnode containing the mammalian sequences PITPNM1 and PITPNM2. There were 21 conserved sequence motifs outside the

PITP domain, with sequence motifs 34-37 observed in all analyzed proteins in this class. In human pitpnm1, the FFAT domain was found at the end of sequence motif 27, with a derived motif sequence of EFFDCLD. The DDHD region contained sequence motifs 30-35, which comprises 35% of the defined DDHD range, and the LSN2 region contained sequence motifs 39 and the N-terminal portion of sequence motif 40, which is 23% of this region. Lev, *et al*. defined six, short hydrophobic regions based on hydrophathy plots that were originally postulated to be transmembrane domains [9] then later to mediate membrane association [7]. None of the six regions fell in a conserved sequence motif identified in this study On the other hand, sequence motif 26 appeared in multiple repeats throughout the non-PITP domain in sequence motifs 29, 32-33, and 44-45. This is a Ser rich motif of eight amino acids, the significance of this observation is currently unclear.

Class IIB: The proteins in this class are historically referred to as rdgBβs and are monomeric, cytosolic proteins of ~330 amino acids. They were distinguished by the unique presence of sequence motif 11, which consists of residues 97-104 in the human PITPNC1 sequence. Sequence motif 11 is interesting; although it was only present in proteins in this division, it overlaped in sequence with motif 12 and motif 13. Sequence motif 12 was only observed in class IIA, whereas sequence motif 13 was unique to Class I, alpha and beta clades. Another notable feature of this division was PITP1 from *D. discoideum*. It grouped outside the other rdgB . proteins and rooted the class IIA clade.

4. DISCUSSION

4.1. Phylogenetics

Ocaka *et al* [36] provide an extensive analysis of the chromosomal location of the PITPN genes in humans with a phylogenetic tree rooted with *C. elegans* PITP as the chosen outgroup. The analysis presented here differs in two respects. First, the original gene-duplication event leading to the evolution of soluble and membrane-associated PITP proteins likely occurred not early in animal evolution, but instead was initiated in some protists. *D. discoideum* has been shown to have two genes, *pitA* and *pitB*, and expression has been demonstrated for both proteins, PITP1 and PITP2 [21]. The phylogeny shown here indicates that PITP1 is a class IIB protein, perhaps representative of an ancestral precursor to class IIA proteins. Second, data presented here indicate that the Class IIB proteins are derived from Class I proteins, and the membrane-associated Class IIA proteins are derived from the class IIB proteins. Previous cladograms and dendrograms have indicated different derivization patterns. Ocaka *et al* [36] indicate that mammalian PITPNC1 (Class IIB in the present analysis) shares a most recent common ancestor (MRCA) with

PITPNA and PITPNB; more so than with any of the PITPNMs. Fullwood *et al*. present a dendrogram in which the membrane-associated PITPs and the soluble PITPs split from a shared derivation from Class IIB proteins (rdgBβ). An increased sample size, an unbiased tree root, and the use of conserved sequence motifs for global alignment in the present study lead to a different interpretation of the evolution of this protein family.

The phylogenetic tree produced here is indicative of the evolutionary history of these sequence motifs, and not of the proteins as a whole. This is important, because in some proteins, large sequences with no known sequence motif patterns are observed (such as in the *Plasmodium yoelli* "PITP1" protein) that may be the result of gene conversion, unequal crossing-over, or repeat expansion during the evolution of specific proteins. Such large insertions and deletions in protein sequence are difficult to utilize in any phylogenetic analysis; they are purely derived characters that are phylogenetically uninformative. However, the existence of these derived sequences suggests that either the function or regulation of proteins in this family has changed dramatically, indicative of the type of specific adaptation often seen within protein families.

4.2. Functional Shifts in PITP Proteins

Alignment and subsequent phylogenetic analysis of divergent members of protein families is a problem compounded by shifts in primary, secondary, and tertiary structure that occur during protein evolution and functional diversification. Global alignments of complete protein coding regions often fail when functional units (such as regulatory regions, transmembrane domains, DNA-binding domains, protein-protein interaction domains, etc.) are gained or lost over evolutionary time. This issue complicates the otherwise straightforward phylogenetic analysis of related proteins. Additional complications arise in organismal surveys in which entire genomes are sequenced and annotated by automated algorithms based on homology. BLAST and FASTA searches find protein regions with high local identity, which can suggest functional similarity. However, outside of these identified regions, major changes can occur in protein function that obscure identity and make global protein alignment difficult or impossible. Without corresponding functional data, these genes are often named inconsistently. For example, a gene found in one species with locally high homology to a protein region in another species is often named "Similar to…" or "Species X homologue of Species Y protein". Because BLAST is a local homology search tool, it is inadequate for identification and annotation of species orthologs. For this reason, many of the genes identified in species genome projects are insufficiently annotated; phylogenetic analysis of all proteins labeled as "Phosphatidylinositol transfer proteins, Beta" for example will yield a spurious

alignment and therefore a misleading tree. As demonstrated in this study, analysis of the PITP/rdgB family was complicated by both issues. It should be noted that full-length sequences produce similar phylogenetic trees for alignments of many related sets of proteins (all Class I or Class II proteins, for example).

The resultant phylogenetic tree reveals that the general naming convention for the PITP family needs to be critically examined. Many of the proteins in this family cannot be named exclusively by local sequence identity, as their functions appear to be significantly altered from previously studied proteins. Many annotated genes are named in a way that might not be truly indicative of their function. In other words, sequence homology in local segments has obscured the greater functional diversification within this family. For example, although the membrane-associated proteins (Class IIB) are shown to be derived and monocladistic, the tree root appears to be within the PITP family, suggesting that many other functions are derived characteristics from this single-domain protein. At best, this tree demonstrates the need for better functional and structural classification of PITP-like proteins before final assignment of their names.

An additional feature of the strict consensus tree is the hypothesis it suggests about the evolution of PITP protein family members. First, no PITP proteins are found within prokaryotes or within yeast. Yeast have an analogous enzyme, Sec14p, but it appears to have an independent evolutionary origin. The slime-mold and *Plasmodium* sequences represent the most primitive species in this PITP family tree. In contrast, the multiplication of forms within mammals is impressive. Humans, mice, and rats have well-characterized genomes and each have a variety of PITP family-member proteins.

4.3. Conserved Sequence Motifs

The use of conserved sequence motifs in a cladistic analysis of the PITP family highlights several regions for consideration. Two regions are of particular interest: one containing sequence motifs 7, 8, and/or 9, and the other containing sequence motifs 11, 12, or 13. These two regions form loops at the surface of the protein near the ends of the acyl chains of the bound phospholipid (**Figure 2(b)** and **2(c)**). Sequence motif 7 is the only motif present that is not found in the same primary sequence space in the three main protein classes. In human PITPα, the amino acid sequence is at residues 66-73, while it occurs at residues 204-211 and 198-205 in human PITPNM1 and PITPNC1, respectively. Sequence motif 8 is not present in class I PITPs, and partially overlaps with sequence motif 9, which is exclusive to the alpha/beta clades of class I. Sequence motif 7 in the class I PITP's partially overlaps with sequence motif 8 of class IIA and IIB. Sequence motif 7 contains the putative helix insertion loop proposed to anchor PITPα to the membrane during lipid exchange [25]. It is tempting to speculate that this helix has adapted a different conformation or perhaps has migrated across the lipid-binding cavity to near residues 200-208 relative to the PITPα structures.

The second region of interest contains sequence motifs 11, 12, or 13. These three motifs are present in approximately the same space in primary sequence, but each is unique to the three family classes. Sequence motif 11 is unique to class IIB, whereas Sequence motif 12 is unique to class IIA, and sequence motif 13 is unique to class I. The role of this loop in the structure or the function of any class of PITP is currently unexplored.

A concern with the novel phylogenetic analysis technique used in this study is the proportion of the sequence data actually used in the evolutionary analysis and whether amino acids that have been structurally and functionally shown to affect phospholipid transfer and signal transduction capabilities are included in the alignment matrices. In human PITPα, 214 of the 270 amino acids (79.3%) are included in the conserved sequence motifs. Most site-directed mutagenic analysis of the role of specific residues in the phospholipid transfer, specificity, and PLC reconstitution capabilities have been done on human or rat PITPα and PITPβ. Residues in rat PITPα that have been studied and shown to play a functional role in PITPs include Thr59, Lys61, Glu86, Asn90, Tyr103, and the double Trp pair at 203-204 [37,38]. These residues are all in conserved sequence motifs. Thr59 and Lys 61 map to sequence motif 6, which is the only sequence motif present in all sequences in this study. Interestingly, sequence motif 9, containing Glu 86, is present only in Class I proteins. Tyr103 has been shown to diminish PLC reconstitution without affecting phospholipid transfer. This residue is a Tyr only in the Class I proteins, and is a conservative substitution to a Phe in the class IIA and class IIB proteins. Examination of these amino acids and sequence motifs from an evolutionary perspective enables a more thorough understanding of their importance in shaping PITP function and structure.

4.4. Functional Specialization across Evolution

Functional shifts within the PITP family are difficult to assess in light of the relative paucity of experimental analysis of different family members. In rat, PITPs are broadly expressed and have been detected in at lease 20 tissues [39]. Unigene expression profiles (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene) for mammalian species indicate a slightly more general and ubiquitous expression for the alpha isoform than the beta isoform. Mouse PITPβ -/- embryonic stem cells are embryonically lethal, indicating essential functions in cell viability [29]. In contrast, mice deficient in PITPα develop normally to term, but fail to thrive neonatally [30].

These observations and the EST expression profiles support the argument that PITP isoforms are not redundant, and have different functions [29]. One implication of the consensus tree is that, given that the ancestral proteins have putative roles in cellular function, e.g., PITP2 in *Dictyostelium*, they may be more widely expressed than the more derived family members, which appear to have roles only in specific tissues or at specific developmental times. PITPnm1 in mice, for example, is highly expressed primarily in the retina, with weaker expression in the brain and some suggestion (from rat data) of central nervous system expression. Other studies in humans showed PITPnm2 and PITPnm3 expression in brain and retina.

The phylogenetic analysis described here has demonstrated the utility of the use of conserved sequence motifs in detecting evolutionary relationships among proteins in large and/or diverse functions. This method uses short conserved sequence motifs instead of single character amino acids. The presence and absence of motifs becomes an important data point. The use of sequence motifs in phylogenetic analysis as described here should be applicable to the analysis of a wide range of protein families, particularly large, diverse families where reshuffling of domains occurred over evolutionary time.

5. ACKNOWLEDGEMENTS

REFERENCES

[1] Allen-Baume, V., Segui, B. and Cockcroft, S. (2002) Current thoughts on the phosphatidylinositol transfer protein family. *FEBS Letters,* **531**, 74-80.

[2] Cockcroft, S. and Carvou, N. (2007) Biochemical and biological functions of class I phosphatidylinositol transfer proteins. *Biochim. Biophys. Acta*, **1771**, 677-691.

[3] Hsuan, J. and Cockcroft, S. (2001) The PITP family of phosphatidylinositol transfer proteins. *Genome Biol.,* **2**, 3011.1-3011.8.

[4] Routt, S.M. and Bankaitis, V.A. (2004) Biological functions of phosphatidylinositol transfer proteins. *Biochem. Cell Biology*, **82**, 254-262.

[5] Thomas, G.M.H. and Pinxteren, J.A. (2000) Phosphatidylinositol transfer proteins: One big happy family or strangers with the same name? *Mol. Cell Biol. Res. Comm.,* **4**, 1-9.

[6] Wirtz, K.W.A. (2006) Phospholipid transfer proteins in perspective. *FEBS Lett.,* **580**, 5436-5441.

[7] Trivedi, D. and Padinjat, R. (2007) RdgB proteins: functions in lipid homeostasis and signal transduction. *Biochim. Biophys. Acta,* **1771**, 692-699.

[8] Loewen, C.J., Roy, A. and Levine, T.P. (2003) A conserved ER targeting motif in three families of lipid binding proteins and in Opi1p binds VAP. *EMBO J.* **22**, 2025-2035.

[9] Lev, S., Hernandez, J., Martinez, R., Chen, A., Plowman, G. and Schlessinger, J. (1999) Identification of a novel family of targets of PYK2 related to *Drosophila* retinal degeneration B (rdgB) protein. *Mole. Cell. Biol.,* **19**, 2278-2288.

[10] Milligan, S.C., Alb, J.G. Jr., Elagina, R. B., Bankaitis, V. A. and Hyde, D.R. (1997) The phosphatidylinositol transfer protein domain of *Drosophila* retinal degeneration B protein is essential for photoreceptor cell survival and recovery from light stimulation. *J. Cell Biol.,* **139**, 351-363.

[11] Vihtelic, T.S., Goebl, M., Milligan, S., O'Tousa, J.E. and Hyde, D.R. (1993) Localization of *Drosophila* retinal degeneration B, a membrane associated phosphatidylinositol transfer protein. *J. Cell Biol.,* **122**, 1013-1022.

[12] Lu, C., Vihtelic, T.S., Hyde, D.R. and Li, T. (1999) A neuronal-specific mammalian homolog of the *Drosophila* retinal degeneration B gene with expression restricted to the retina and dentate gyrus. *J. Neurosci.,* **19**, 7317-7325.

[13] Fullwood, Y., dos Santos, M. and Hsuan, J.J. (1999) Cloning and characterization of a novel human phosphatidylinositol transfer protein, rdgBb. *J. Biol. Chem.*, **274**, 31553-31558.

[14] Takano, N., Owada, Y., Suzuki, R., Sakagami, H., Shimosegawa, T. and Kondo, H. (2003) Cloning and characterization of a novel variant (mM-rdgB b1) of mouse M-rdgBs, mammalian homologs of *Drosophila retinal degeneration B gene* proteins, and its mRNA localization in mouse brain in comparison with other M-rdgBs. *J Neurochem*, **84**, 829-839.

[15] Skinner, H.B., Alb ,J.G.Jr., Whitters, E.A., Helmkamp, G. M.Jr. and Bankaiitis V. A. (1993) Phospholipid transfer activity is relevant to but not sufficient for the essential function of the yeast *SEC14* gene product. *EMBO*, **12**, 4775-4784.

[16] Tanaka, S. and Hosaka, K. (1994) Cloning of a cDNA encoding a second phosphatidylinositol transfer protein of rat brain by complementation of the yeast sec14 mutation. *J. Biochem* (Tokyo), **115**, 981-984.

[17] Cunningham, E., Tan, S.K., Swigart, P., Hsuan, J., Bankaitis, V. and Cockcroft, S. (1996) The yeast and mammalian isoforms of phosphatidylinositol transfer protein can all restore phospholipase C-mediated inositol lipid signaling in cytosol-depleted RBL-2H3 and HL-60 cells. *Proc. Natl. Acad. Sci. USA*, **93**, 6589-6593.

[18] Hay, J.C. and Martin, T.F.J. (1993) Phosphatidylinositol transfer protein is required for ATP-dependent priming of Ca^{2+}-activated secretion. *Nature,* **366**, 572-575

[19] Jones, S.M., Alb, J.G.J., Phillips, S. E., Bankaitis and V.A., Howell, K.E. (1998) A phosphatidylinositol 3-kinase and phosphatidylinositol transfer protein act synergistically in formation of constitutive transport vesicles from the trans-Golgi network. *J. Biol. Chem.* **273**, 10349-10354.

[20] Ohashi, M., de Vries, K.J., Frank, R., Snoek, G., Bankaitis, V., Wirtz, K. and Huttner, W.B. (1995) A role for phosphatidylinositol transfer protein in secretory vesicle formation. *Nature*, **377**, 544-547.

[21] Swigart, P., Insall, R., Wilkins, A. and Cockcroft, S. (2000) Purification and cloning of phosphatidylinositol transfer proteins from *Dictyostelium discoideum*: homo-

logues of both mammalian PITPs and Saccharomyces cerevisiae Sec14p are found in the same cell. *Biochem. J.,* **347**, 837-843.

[22] Yoder, M.D., Thomas, L.M., Tremblay, J.M., Oliver, R. L., Yarbrough, L.R. and Helmkamp, G.M.Jr. (2001) Structure of a multifunctional protein: mammalian phosphatidylinositol transfer protein complexed with phosphatidylcholine. *J. Biol. Chem.*, **276**, 9246-9252.

[23] Vordtriede, P.B, Doan, C.N., Tremblay, J.M., Helmkamp, G.M.J. and Yoder, M.D. (2005) Structure of PITPb in complex with phosphatidylcholine: Comparison of structure and lipid transfer to other PITP isoforms. *Biochemistry,* **44**, 14760-14771.

[24] van Tiel, C.M., Schouten, A., Snoek, G.T., Gros, P., Wirtz, K.W.A. (2002) The structure of phosphatidylinositol transfer protein a reveals sites for phospholipid binding and membrane association with major implications for its function. *FEBS*, **531**, 69-73.

[25] Schouten, A., Agianian, B., Westerman, J., Kroon, J., Wirtz, K.W.A. and Gros, P. (2002) Structure of apo-phosphatidylinositol transfer protein a provides insight into membrane association. *EMBO J.*, **21**, 2117-2121.

[26] Sha, B., Phillips, S.E., Bankaitis, V.A. and Luo, M. (1998) Crystal structure of the *Saccharomyces cerevisiae* phosphatidylinositol transfer protein. *Nature*, **391**, 506-510.

[27] Romanowski, M.J., Soccio, R.E., Breslow, J.L. and Burley, S.K. (2002) Crystal structure of the *Mus musculus* cholesterol-regulated START protein 4 (StarD4) containing a StAR-related lipid transfer domain. *Proc. Natl. Acad. Sci. USA*, **99**, 6949-6954.

[28] Hamilton, B.A., Smith, D.J., Mueller, K.L., Kerrebrock, A.W., Bronson, R.T., van Berkel, V., Daly, M.J., Kruglyak, L, Reeve, M.P., Nemhauser, J.L., Hawkins, T. L. Rubin, E.M. and Lander, E.S. (1997) The *vibrator* mutation causes neurodegeneration via reduced expression of PITPa: positional complementation cloning and extragenic suppression. *Neuron,* **18**, 711-722.

[29] Alb, J.G.Jr., Phillips, S.E., Rostand, K., Cui, X., Pinxteren, J., Cotlin, L., Manning, T, Guo, S, York, J.D., Sontheimer, H., Collawn, J.F., Bankaitis, V.A. (2002) Genetic ablation of phosphatidylinositol transfer protein function in murine embryonic stem cells. *Mol. Biol. Cell*, **13**, 739-754.

[30] Alb, J.G.J., Cortese, J.D, Phillips, S.E., Albin, R.L., Nagy, T.R., Hamilton, B.A. and Bankaitis, V.A. (2003) Mice lacking phosphatidylinositol transfer protein-a exhibit spinocerebellar degeneration, intestinal and hepatic steatosis, and hypoglycemia. *J. Biol. Chem.* **278**, 33501-33518.

[31] Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.

[32] Bailey, T.L. and Elkan, C.P. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology. AAAI Press, Menlo Park, CA, 28-36.

[33] Accelrys, I. (2005) Wisconsin Package v.10.3. Accelrys, Inc, San Diego, CA.

[34] Swofford, D.L. (2003) PAUP*. Phylogenetic analysis using parsimony and other methods. Version 4. Sinauer Associates, Sunderland, Massachusetts.

[35] Maddison, D.R. and Maddison, W.P. (2005) MacClade v.4.08. Sinauer Associates, Sunderland, Massachusetts.

[36] Ocaka, L., Spalluto, C., Wilson, D.I, Hunt, D.M., Halford, S. (2005) Chromosomal localization, genomic organization and evolution of the genes encoding human phosphatidylinositol transfer protein membrane- associated (PITPNM) 1, 2 and 3. *Cytogenet Genome Res.*, **108**, 293-202.

[37] Alb,J.G. Jr., Gedvilaite, A., Cartee, R.T., Skinner, H.B., Bankaitis, V.A. (1995) Mutant rat phosphatidylinositol/phosphatidylcholine transfer proteins specifically defective in phosphatidylinositol transfer: Implications for the regulation of phospholipid transfer activity. Proc. Natl. *Acad. Sci. USA*, **92**, 8826-8830.

[38] Tilley, S.J., Skippen, A., Murray-Rust, J., Swigart, P.M., Stewart, A., Morgan, C.P., Cockcroft, S. and McDonald, N.Q. (2004) Structure-function analysis of phosphatidylinositol transfer protein alpha bound to human phosphatidylinositol. *Structure*, **12**, 317-326.

[39] Venuti, S.E. and Helmkamp, G.M.J. (1988) Tissue distribution, purification and characterization of rat phosphatidylinositol transfer protein. *Biochim Biophys Acta,* **946**, 119-128.

[40] Lev, S. (2004) The role of the Nir/rdgB protein family in membrane trafficking and cytoskeleton remodeling. *Exp. Cell Res.,* **297**, 1-10.

A new approach for classification of human brain CT images based on morphological operations

Ali Reza Fallahi[1], Mohammad Pooyan[1], Hassan Khotanlou[2]

[1]Biomedical Engineering Department, Shahed University, Tehran, Iran;
[2]Computer Engineering Department, Bu-Ali Sina University, Hamedan, Iran.
Email: afallahi@shahed.ac.ir; pooyan@shahed.ac.ir; hkh@basu.ac.ir

ABSTRACT

Automatic diagnosis may help to decrease human based diagnosis error and assist physicians to focus on the correct disease and its treatment and to avoid wasting time on diagnosis. In this paper computer aided diagnosis is applied to the brain CT image processing. We compared performance of morphological operations in extracting three types of features, i.e. gray scale, symmetry and texture. Some classifiers were applied to classify normal and abnormal brain CT images. It showed that morphological operations can improve the result of accuracy. Moreover SVM classifier showed better result than other classifiers.

1. INTRODUCTION

Medical CT image has been applied in clinical diagnosis widely. It can assist physicians to detect and locate pathological changes, and determine the property of them. But the diagnosis result is often subjective, different physicians may get different diagnosis result at different time [1]. Computer Aided Diagnosis (CAD) aims to provide a computer output as a second opinion in order to assist physicians in the detection of abnormalities, quantification of disease progress and differential diagnosis of lesions [2].

The typical architecture of a CAD system (**Figure 1**) includes four main modules: image pre-processing, definition of region(s) of interest (ROI), extraction and selection of features and classification of the selected ROI [3].

The aim of image pre-processing is to improve the quality of data through the application of methods for denoising such as mean, median, Laplacian and Gaussian filters and enhancing the edges of image structures such as unsharpening and wavelet transform and enhancing image contrast like histogram equalization.

Feature extraction refers to various quantitative measurements of medical images typically used for decision making related to the pathology of a structure or tissue. When the features have been extracted, selection of a subset of the most robust features is essential, aiming at improving classification accuracy and reducing the overall complexity. Some methods such as PCA, LDA and genetic algorithms can be used for this purpose.

One of the major problems of pattern recognition in medical image analysis is the classification of a set of features into the proper classes. The main methods in the brain CT images classification include RBFNN [1], decision tree, See5, inductive learning [4] and Bayesian.

In this paper we applied morphological operations to the images and then extracted three types of features. We then evaluated performance of these operations and extracted features by six types of classifiers.

2. MATERIALS AND METHOD

There are lots of diseases that originate in the brain. The CT image is the lamination cross-sectional image. Different stratification plane images have different patterns. In order to get the better image analysis, we select one kind of disease to determine whether it is normal or abnormal and to extract features from one plane in the CT of brain image (**Figure 2**). The ordinary utilized features are texture, gray scale, shape and symmetry. Our experiment focuses on texture, symmetry, gray scale and their formulation.

2.1. Gray Scale Feature

In digital image processing, the two-dimensional digitized gray scale image (M×N) can be seen as M×N pixels in two-dimensional surface XOY, each pixel (x,y) can be represented as its gray value. Grayscale features that can be extracted are mean, variance, skewness and kurtosis. Among them, the standard variance reflects the separate degree of gray scale value. The skewness takes the mean value as the central data distribution. Skewness is represented by the following equation:
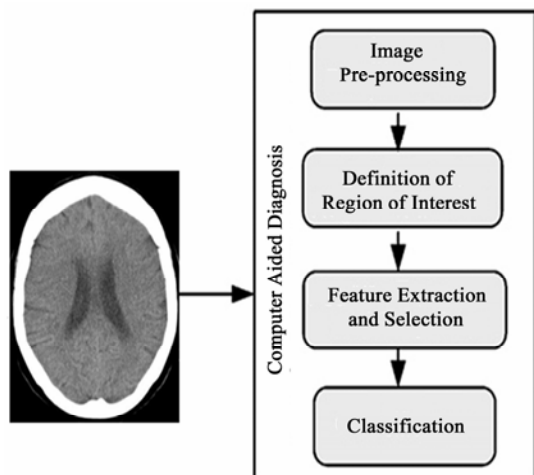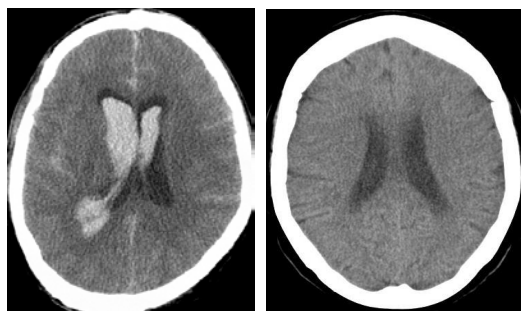
Figure 1. Architecture of a CAD system**.**

Figure 2. Abnormal and normal brain CT images.

$$M_{c3} = (\frac{1}{M_{c2}^{3}})\sum_{x=1}^{m}\sum_{y=1}^{n}(f(x,y) - M_{c1})^{3} \qquad (1)$$

The kurtosis reflects the normal distribution sharpness or smoothness of the compared data that is represented by the following equation:

$$M_{c4} = (\frac{1}{M_{c2}^{4}})\sum_{x=1}^{m}\sum_{y=1}^{n}(f(x,y) - M_{c1})^{4} \qquad (2)$$

Because of symmetry structure of the brain, we divide image in two partitions. The above features are extracted in the left and right sections. We then calculate the Euclidian distance corresponding to four features of two sections as the new features. We applied these new features for classification.

2.2. Texture Features

A co-occurrence matrix (COM) is the square matrices of relative frequencies P (i,j,d,θ) with which two neighboring pixels separated by the distance d at the orientation θ occur in the image (**Figure 3**) , one with gray level i and the other with gray level j [5]. A COM is therefore a square matrix that has the size of the largest pixel value in the image and present the relative frequency distribu-

tions of gray levels and describe how often one gray level will appear in a specified spatial relationship to another gray level within each image region [6]. There are 14 features that may be extracted from COM matrix, but usually four or five features are more interested ones. In this paper four textural features were calculated from the COM for direction h values of $0°$ and a distance d of 1.

The matrix was normalized by the following function:

$$p(i,j,d,\theta) = \frac{p(i,j,d,\theta)}{R} \qquad (3)$$

R is the normalized function, which is usually set as the sum of the matrix.

Energy is also called Angular Second Moment. It is a measure of the homogeneousness of the image and can be calculated from the normalized COM. It is a suitable measure to detect disorder in texture image. Higher values for this feature mean less changes in the image amplitude or intensity result in a much sparser COM.

The energy is formulated by the following equation:

$$J = \sum_{i=1}\sum_{j=1}(p(i,j))^{2} \qquad (4)$$

Contrast is a measure of amount of the local variation in the image. It will have a large value for images which have a large amount of local variation in gray levels and a smaller value for images with uniform gray level distributions and is defined as:

$$G = \sum_{i=1}\sum_{j=1}(i-j)^{2}p(i,j) \qquad (5)$$

The Inverse Difference Moment (IDM) reflects the local texture changes. It is another feature of image contrast and is defined as:

$$IDM = \sum_{i=1}\sum_{j=1}\frac{1}{1+(i-j)^{2}}p(i,j) \qquad (6)$$

Entropy gives a measure of complexity of the image. Complex textures tend to have higher entropy. Entropy is represented by the following equation:

$$S = -\sum_{i=1}\sum_{j=1}p(i,j)\log(p(i,j)) \qquad (7)$$
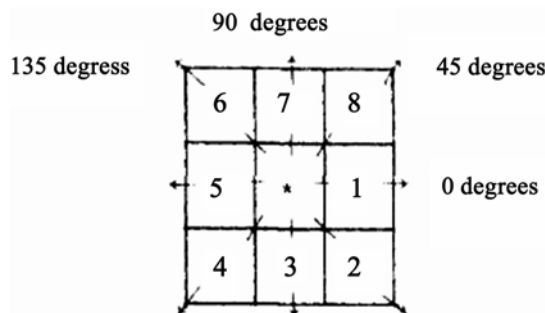
Figure 3. Co-accurance matrix calculation.
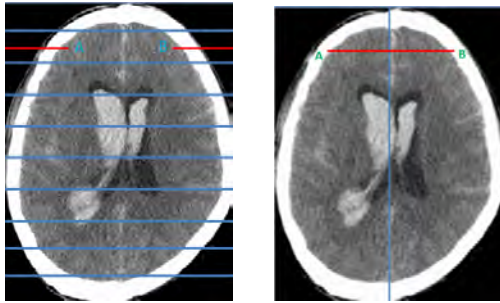
Figure 4. Exterial symmetry (left) and interial symmetry (right) feature calculation.
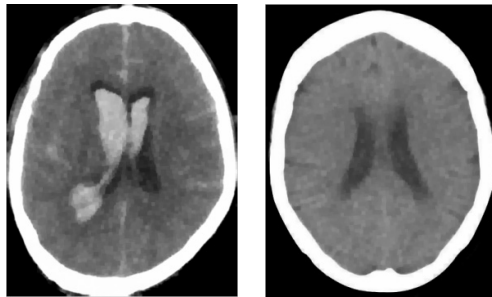
Figure 5. Result of applying morphological operation.
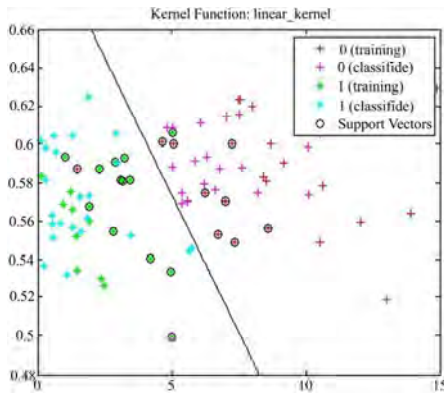
Figure 6. Result of classifing data with SVM.

2.3. Symmetrical Feature

Symmetrical features including two types: Interior symmetry and exterior symmetry.

1) Interior symmetry: We partition one image in ten parts with same length, where p is the part number, S_i is the sum of symmetry pixel points in each part, V_i is the sum of all the pixel points in each part. If $x - y \leq 4$ then point A is symmetrical to point B. x is the gray scale value of point A and y is the gray scale value of point B. From the two edges of each part to the center, for detecting the interior symmetrical pixel point their sum per pixel line is calculated. In a symmetrical images, the value of Sy_p is greater than an unsymmet-

rical images. Therfore, the value of Sy_p in the normal images is greater than the abnormal ones. Interior symmetry is defined as:

$$Sy_p = \left. \sum_{i=1}^{n} s_i \middle/ \sum_{i=1}^{n} v_i \right. \tag{8}$$

2) Exterior symmetry: Let us m_i represents the X coordinate value of midpoint of each pixel line in image, n, the height of the image and M the average X coordinate value of all the pixel line in image (**Figure 4**). The more symmetrical image has higher value of S_E. But, if there are pathological changes in the image, the value of S_E is uncertain. Exterior symmetry is defined as:

$$S_E = \frac{\sum_{i=1}^{n}(m_i - M)^2}{n} \tag{9}$$

2.4. Morphological Operations

Morphological operations are capable of deleting small disconnected regions, filling cavities and smoothing the region-of-interest [7]. Tow types of digital morphological operations, opening followed by closing, are applied to the image to eliminate the small isolated regions. These operations are defined as ordered combinations of fundamental operations, dilation and erosion. If the translation operation is defined as

$$\mathbf{A} + x = \{a + x : a \in \mathbf{A}\} \tag{10}$$

Then the operators' erosion and dilation can be written as:

$$E(\mathbf{A}, \mathbf{B}) = \mathbf{A} \ominus \mathbf{B} = \{x : -\mathbf{B} + x \subset \mathbf{A}\} \tag{11}$$

$$D(\mathbf{A}, \mathbf{B}) = \mathbf{A} \oplus \mathbf{B} = \{x : (-\mathbf{B} + x) \cap \mathbf{A} = \phi\} \tag{12}$$

where A is the image and B is the structuring element. Note that –B denotes the reflection of B with regard to the origin, x denotes a point in space, and a is a point in the image A [8]. Here we use circular element with radius r which can be determined by a compromise between the noise suppression performance and preservation of details.

2.5. Classification

We selected a set of 6 well known classifiers that their computation procedures would be described briefly in the following sections.

1) Linear Bayes Normal Classifier: This computes the linear classifier between the classes of the dataset by assuming normal densities with equal covariance matrices. The joint covariance matrix is the weighted by a priori probabilities average of the class covariance ma-

Table 1. Accuracy of classifiers without morphological operations.

	Classes	Symmetry	Grayscale	Texture	total
	1 : Bayes-Normal-(Quadratic)	97.75	66.75	59	98.5
	2 : Bayes-Normal-(linear)	99.25	67.75	67.5	99.25
	3 : K-NN Classifier	98.5	66.75	37.25	45.2
	4: Decision Tree	89.25	54	62.75	83
5:BP N N	5 hidden neuron	93	74	79	95
	10 hidden neuron	97	73	67	90
	6:SVM	97	88	90	100

Table 2. Accuracy of classifiers with morphological operations.

	Classes	Symmetry	Grayscale	Texture	total
	1 : Bayes-Normal-(Quadratic)	98.5	58.25	61.25	100
	2 : Bayes-Normal-(linear)	99.25	70	64	100
	3 : K-NN Classifier	100	65.75	45.25	45
	4: Decision Tree	91.25	45.75	54	88.75
5:BP NN	5 hidden neuron	100	85	75	97
	10 hidden neuron	96	81	65	95
	6:SVM	97	90	86	100

Table 3. Result of other papers.

Author	classifier	Symmetry	Grayscale	Texture	Total
Zhang et.al	See5	85-95	72-76	75-78	90-94
	RBF NN	85-89	72-75	62-65	83-86
Wang et.al	See5	85-93	-	75-78	-
	RBF NN	85-89	-	67-71	-
	Outlier Detection	91-98.5	-	76-80	-

trices. The covariance matrix of the classes is then decomposed as G = W*W' + sigma^2 * eye (K), where W is a K×M matrix containing the M leading principal components and sigma^2 is the mean of the K-M smallest eigenvalues. Finally, the classification is computed.

2) Quadratic Bayes Normal Classifier: It also computes the quadratic classifier between the classes of the dataset, assuming normal densities. The covariance matrix of the classes is then decomposed as G = W*W' + sigma^2 * eye (K), where W is a K×M matrix, containing the M leading principal components and sigma^2 is the mean of the K-M smallest eigenvalues. The classification is then performed.

3) K-NN Classifier: It computes the common K-nearest neighbor classifier for the dataset.

4) Decision Tree: It computes a decision tree classifier out of a dataset using a binary splitting criterion.

5) Back Propagation Neural Classifier: A feed-forward neural network classifier with five and ten hidden units in the second layer that number of input neurons depends on number of features and hidden neurons are selected manually is computed for the dataset.

6) SVM (Support Vector Machines): SVM is a new classification method for both linear and nonlinear data. It uses a nonlinear mapping to transform the original training data into a higher dimension. With the new dimension, it searches for the linear optimal separating hyperplane. SVM finds this hyperplane using support vectors ("essential" training tuples) and margins that defined by the support vectors (**Figure 5**). With an appropriate nonlinear mapping to a sufficiently high dimension, data from two classes can always be separated by a hyperplane.

3. RESULTS

To evaluate the method, we used 88 images, which 44 images are normal and the remaining is abnormal. We selected 70% of dataset as training data and 30% as testing data. **Table 1** and **2** shows the result of classifiers for two cases: features with morphological operation and features without them. We extracted symmetry, gray scale and texture features with dimension of 10, 4 and 5 respectively. We used a toolbox written for MATLAB named PRTools. This toolbox consists of a complete and useful set of functions for pattern recognition and includes most of the well known classification algorithms [11].

Most of the classifiers showed that the morphological operations improve results in using symmetry and gray scale features and all of them showed reduction results in texture feature that can be caused by reduction of statistical variation in the image because of applying morphological operations. **Table 1** and **2** show that the result of SVM is more accurate than the other classifiers. When a mixture of the features is used, results are better than using the features sole except K-NN and Decision

Tree, that may caused by increasing complexity. Using 5 neurons to run BPNN led better results than using 10 neurons for it. It should be noted that using 10 neurons may result in overfitting problem that reduces accuracy of the test. **Table 3** shows the result of another works in this area and we can see that our method, using SVM and morphological operation can improve these results.

4. CONCLUSIONS

In this study, we applied morphological operations to the images and then extracted the features. Analysis result from classifiers shows that these operations can improve classification result in symmetry and grayscale features but reduce results in texture features. According to **Table 1** and **2**, we conclude that by using SVM classifier, we can classify the patients more accurately into the corresponding groups. Because of the brain symmetrical structure, symmetrical features have best accuracy. Our experiments show that texture features have lower accuracy. In the future works we can apply another feature extraction and feature selection techniques to improve classification accuracy.

REFERENCES

[1] Zhang, W.L. and Wang, X.Z. (2007) Extraction and classification for human brain CT images. *Proceedings of Sixth International Conference on Machine learning and cybernetics, IEEE.*, 1155-1156.

[2] Stoitsis, J., Valavanis, I., Valavanis, S.G., Golemati, S., Nikita, A. and Nikita K.S. (2006) Computer aided diagnosis based on medical image processing and artificial intelligence methods, *Nuclear Instruments and Methods in Physics Research*, **569**, 591-595.

[3] Haruka, D. and Teruak, A. (2007) Characterization of spatiotemporal stress distribution during food fracture by image texture analysis methods. *Journal of Food Engineering*, **81,** 429-436.

[4] Wang, X.Z. and Lin, W.X. (2007) Application of inductive learning in human brain CT image recognition. *Proceedings of Sixth International Conference on Machine learning and cybernetics, IEEE.*, 1155-1156.

[5] Haralick, R.M., Shanmugam, K. and Dinstein, I. (1973) Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3, **6,** 610-621

[6] Haruka, D. and Teruak, A. (2007) Characterization of spatiotemporal stress distribution during food fracture by image texture analysis methods, *Journal of Food Engineering, 81*, 429-436.

[7] Giardina, C.R. and Dougherty, E.R. (1988) *Morphological Methods in Image and Signal Processing*. Englewood Cliffs, NJ, Prentice–Hal.

[8] Chen, C.W., Luo, J. and Parker, K.J. (1998) Segmentation via adaptive K-Mean clustering and knowledge-based morfological operations with biomedical applications, *IEEE Trans, Image Processing*, Vol. 7, **12**, 1673-1683.

[9] Mitchell, T.M.(2003) Machin learning. China Machine Press, Beijing.

[10] Mandayam, S. and Policar, R. Artificial neural networks. *Lecture Note*.

[11] Duin, R.P.W., Juszczak, P., Paclik, P., Pekalska, E., de Ridder, D. and Tax, D.M.J. (2004) *A* matlab toolbox for pattern recognition. *PRTools4*, Delft University of Technology.

Induced-pluripotent stem cells seeded acellular peripheral nerve graft as "autologous nerve graft"

Jiang Li[1*], **Guo-Dong Gao**[1*], **Ti-Fei Yuan**[*]

[1]Department of Neurosurgery, Tang Du Hospital, The Forth Military Medical University, Xi'an, China.
[*]Email: lijiang19771103@gmail.com; gguodong@fmmu.edu.cn; ytf0707@126.com

ABSTRACT

The hypothesis is that induced pluripotent stem cells (iPSC) derived Schwann cells and/or macrophages can be transplanted into acellular nerve graft in repairing injured nervous system. The efficiency of iPSC seeded acellular nerve graft may mimic the autologous peripheral nerve graft.

Keywords: Induced-Pluripotent Stem Cells, Peripheral Nerve Graft; Axon Regeneration

1. INTRODUCTION

Brain and spinal cord injury conditions are non-reversible and current therapies are quite limited. Biomaterial transplantation, especially when combined with stem cell technology, provides one of the most attractive solutions to repair the injured nervous system, for both central and peripheral. Available biomaterials include natural sourced materials such as nerve, muscle, tendon, vein, fibronectin, collagen as well as fibronectin, and synthetic/engineered materials, such as poly-lactic acid, nanofiber scaffold, biodegradable glass and silicon [1]. However, with more than a century's laboratory and clinical trials, peripheral nerve graft is still the "gold standard", and autologous peripheral nerve graft provided the best axonal regeneration among all these materials.

One existing problem in using peripheral nerve graft to repair the injured nerves is that the obtain of an autologous nerve graft lead to loss of sensory innervation in part of the body, while non-autologous nerve graft though is non-invasive, is of immunogen reactivity and could be expelled. The immunogenic problem could be solved with prior treatment with chemical, thermal, or liquid nitrogen freezing procedures to make the peripheral nerve graft be acellular. However the acellular grafts lost Schwann cells inside and the extracellular matrix suffers from somewhat disruption. Additionally, the acellular graft is not totally free of immunogenic activity due to the existence of non-degraded proteins [2].

2. THE HYPOTHESIS

We propose that the recent developed induced pluripotent stem cell (iPSC) technology combined with acellular nerve graft transplantation can provide a novel approach to repair the injured nervous system. The iPSC-seeded acellular nerve graft from other individuals can be functional similar to the autologous nerve grafts.

The induced pluripotent stem cells, genetically reprogrammed form any somatic cells can be theoretically promoted into any kinds of cells [3,4,5], including Schwann cells and macrophages. We propose that the iPSC could be generated from somatic cells of the patients, and induced into Schwann cells and macrophages before co-transplantation into the acellular allografts or even xenografts. The Schwann cell in the graft has long been found to be neuroprotective and enhancing the axonal remyelination; while macrophage can efficiently remove the remnant proteins and could be helpful in reconstructing the fine structures of extracellular matrix. During the induction of pluripotency, some further genetic modifications could be included, such as the expression of trophic factors to enhance the nerve regrowth.

3. TESTING THE HYPOTHESIS

It is necessary to induce iPSCs into Schwann cells and macrophages. Then the induced cells can be seeded into acellular nerve graft, potentially autologous at first, to examine their efficiency in mimicking natural and autologous Schwann cells when myelinate the regenerated nerves. With these experiences, allografts and xenografts could be tested. Finally this may bring a new avenue in clinical nerve repair procedures with ethical regulations if any are met.

REFERENCES

[1] Schmidt, C.E. and Leach, J.B. (2003) Neural tissue engineering: strategies for repair and regeneration. *Annu Rev Biomed Eng.*, **5**, 293-347.

[2] Gulati, A.K. and Cole, G.P. (1994) Immuno-

genicity and regenerative potential of acellular nerve allografts to repair peripheral nerve in rats and rabbits. *Acta Neurochir (Wien),* **126(2–4)**, 158-64.

[3] Yu, J. and Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., *et al.* (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science,* **318(5858)**, 1917-20.

[4] Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., *et al.* (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, **131(5)**, 861-72.

[5] Takahashi, K. and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell,* **126(4)**, 663-76.

Cepstral and linear prediction techniques for improving intelligibility and audibility of impaired speech

G. Ravindran[1], S. Shenbagadevi[2], V. Salai Selvam[3]

[1]Faculty of Information and Communication Engineering, College of Engineering, Anna University, Chennai, India;
[2]Faculty of Information and Communication Engineering, College of Engineering, Anna University, Chennai, India;
[3]Department of Electronics & Communication Engineering, Sriram Engineering College, Perumalpattu, India.
Email: vsalaiselvam@yahoo.com

ABSTRACT

Human speech becomes impaired i.e., unintelligible due to a variety of reasons that can be either neurological or anatomical. The objective of the research was to improve the intelligibility and audibility of the impaired speech that resulted from a disabled human speech mechanism with impairment in the acoustic system-the supra-laryngeal vocal tract. For this purpose three methods are presented in this paper. Method 1 was to develop an inverse model of the speech degradation using the Cepstral technique. Method 2 was to replace the degraded vocal tract response by a normal vocal tract response using the Cepstral technique. Method 3 was to replace the degraded vocal tract response by a normal vocal tract response using the Linear Prediction technique.

Keywords: Impaired Speech; Speech Disability; Cepstrum; LPC; Vocal Tract

1. INTRODUCTION

Speech impairments or disorders refer to difficulties in producing speech sounds with voice quality [1]. Thus impaired speech is the speech sound that lacks in voice quality. Speech becomes impaired due to a variety of reasons that can be either neurological e.g., aphasia & dysarthria or anatomical e.g., cleft lip & cleft palate [1].

The speech impairment is generally categorized into *Articulation impairment* e.g., omissions, substitutions or distortions of sounds, *Voice impairment* e.g., inappropriate pitch, loudness or voice quality *Fluency impairment* e.g., abnormal rate of speaking, speech interruptions or repetition of sounds, words, phrases or sentences interfering effective communication, *Language impairment* e.g., phonological, morphological, syntactic, semantic or pragmatic use of oral language [2].

The most commonly used techniques to help people with speech impairments are training programmes by speech therapists at home or at hospitals or at a combination of these, sign language like Makaton and electronic aids like text-to-speech conversion unit.

1.1. General Properties of Speech

Though non-stationary the speech signal can be considered as stationary over short periods typically 10-50 msec [3,4,5]. Effective bandwidth of speech is 4-7 kHz [4,5]. The elementary linguistic unit of speech is called a phoneme and its acoustic realization is called a phone [7]. A phoneme is classified as either a vowel or a consonant [3,4,5]. The duration of a vowel does not change much and is 70 ms on an average while that of a consonant varies from 5 to 130 ms [3].

1.2. Speech Production

The diaphragm forces air through the system and the voluntary movements of anatomical structures of this system shape a wide variety of waveforms broadly classified into voiced & unvoiced speech [5]. This is depicted in **Figure 1**.

With voiced speech, the air from the lungs is forced through the glottis (opening between the vocal cords) and the tension of the vocal cords is adjusted so that they vibrate at a frequency, known as ***pitch frequency***, which depends on the shape and size of the vocal cords, resulting in a quasi-periodic train of air pulses that excites the resonances of the rest of the vocal tract. The voluntary movements of the muscles of this vocal tract change its shape and hence resonant frequencies, known as ***formants***, producing different quasi-periodic sounds [7].

Figure 2 shows a sample of voiced speech segment and its spectrum with formant peaks.

With unvoiced speech, the air from the lungs is forced through the glottis and the tension of the vocal cords is adjusted so that they do not vibrate, resulting in a noise-like turbulence that excites normally a constriction
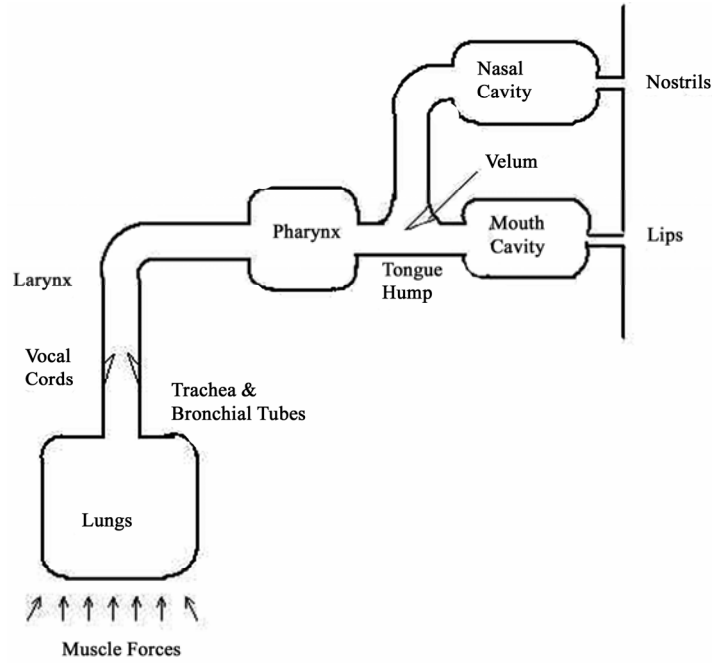
Figure 1. Block diagram of human speech production.
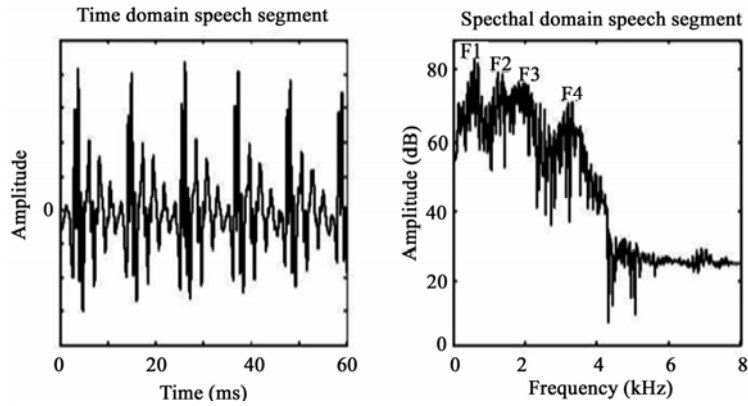
Figure 2. Voiced speech segment and its spectrum exhibiting four resonan frequencies called 'formants'.
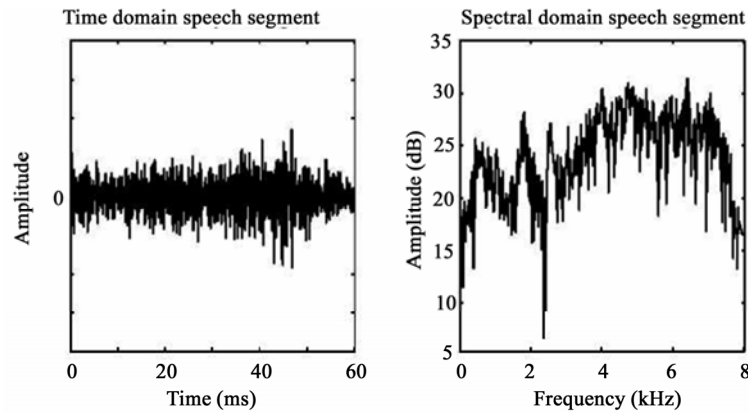

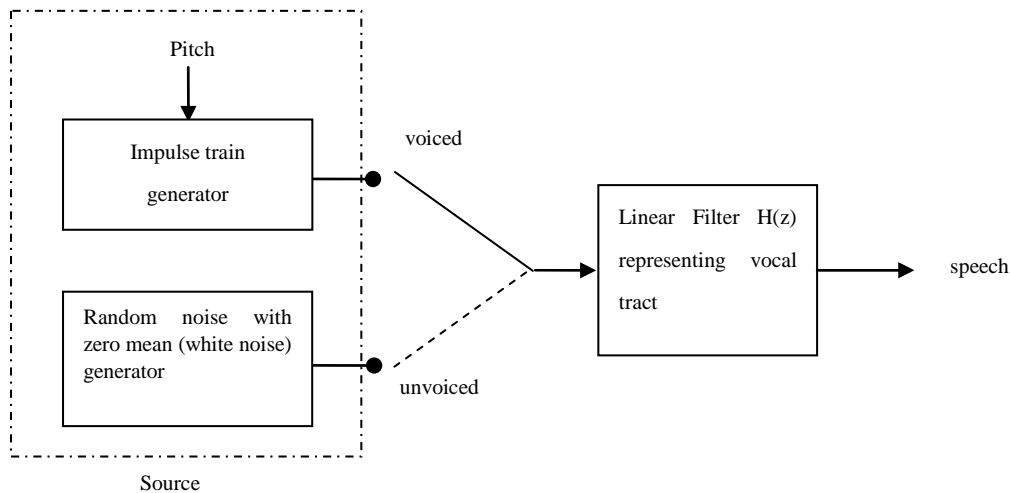**Figure 3.** Unvoiced speech segment and its spectrum.

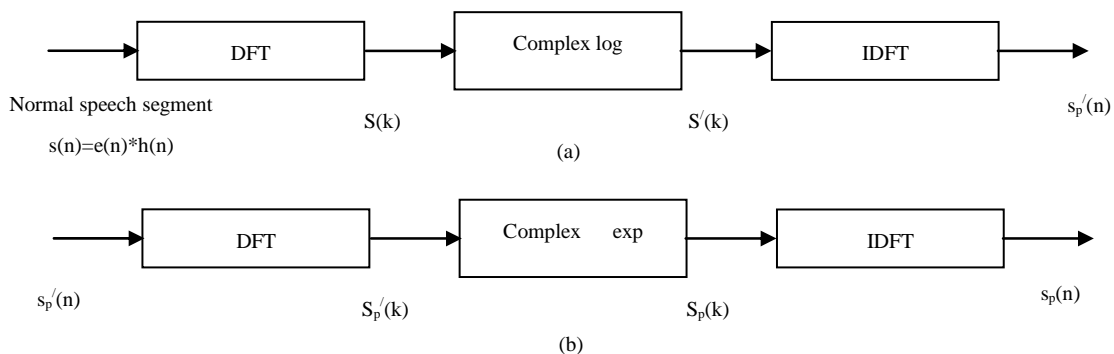**Figure 4.** Source-filter model of a human speech mechanism.

Figure 5. (a) Complex cepstrum and (b) its inverse (after Oppenheim &Schafer).

in the rest of the vocal tract. Depending on the shape and size of the constriction different noise-like sounds are produced.

Figure 3 shows a sample of unvoiced speech segment and its spectrum with no dominant peaks.

Thus a speech signal can be supposed to be a convolution of two signals: 1) a quasi-periodic pulse-like (for voiced speech) or a noise-like (for unvoiced speech) glottal excitation signal generated by a combination of lungs and vocal cords and 2) a system response represented by the shape of the rest of the vocal tract [4].

The excitation signal generally exhibits the speaker characteristics such as pitch and loudness while the vocal tract response determines the sound produced [5].

1.3. Source-Filter Model of Human Speech Mechanism

A speech signal, s(n) is convolution of a fast varying glottal excitation signal, e(n) and a slowly varying vocal tract response, v(n) i.e., s(n)=e(n)*v(n). For a voiced speech, e(n) is a quasi-periodic waveform and v(n) is a combined effect of the glottal wave shape, the vocal tract impulse response and the lip radiation impulse response

while for an unvoiced speech, e(n) is a random noise and v(n) is a combined effect of the vocal tract impulse response and the lip radiation impulse response [4].

A human speech mechanism thus can be viewed as a *source* capable of generating a periodic impulse train at pitch frequency for voiced speech or a white noise for unvoiced speech followed by a *linear filter* having an impulse response that represents the shape of the vocal tract [4]. This is depicted in **Figure 4**.

1.4. Speech Processing Techniques

1.4.1. Cepstral Technique

Complex cepstrum: The complex cepstrum $s'(n)$ of a signal $s(n)$ is defined as the inverse Fourier transform of the logarithm of the signal spectrum $S(e^{j\omega})$ [8].

$$s'(n) = (1/2\pi) \int_{-\pi}^{\pi} \log\left[S(e^{j\omega}|)\right]e^{j\omega}$$

where $S(e^{j\omega})$ is the Fourier transform of $s(n)$. Computation of complex cepstrum requires phase unwrapping, which is difficult due to some theoretical and practical reasons [8]. This is depicted in **Figure 5 (a)** & **(b)**.
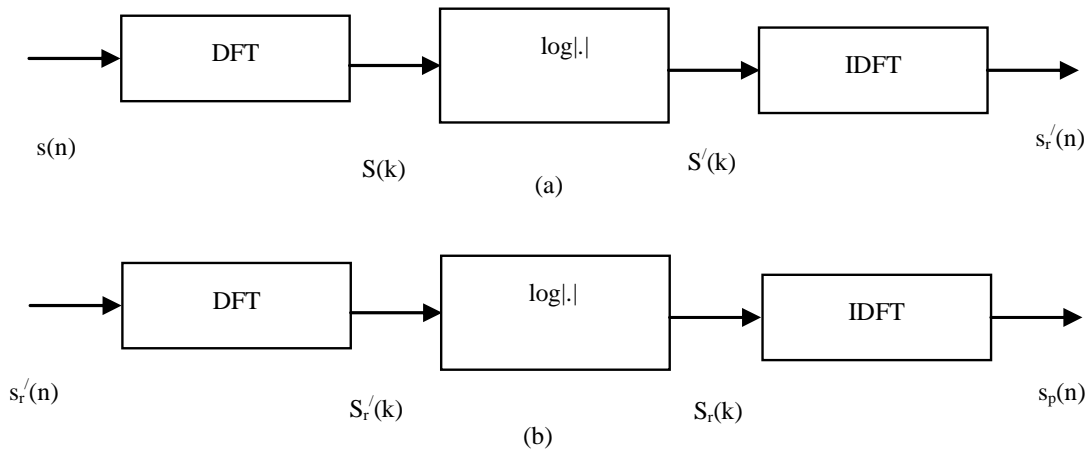
Figure 6. (a) Real cepstrum and (b) its inverse (after Oppenheim & Schafer). Cepstral domain is known as *quefrency* (coined from 'frequency') domain [8].

Real cepstrum: The real cepstrum $s_r^/(n)$ of a signal s(n) is defined as the inverse Fourier transform of the logarithm of the signal magnitude spectrum $|S(e^{j\omega})|$ [8].

$$s_r^/(n)=(1/2\pi)\int_{-\pi}^{\pi}\log(|S(e^{j\omega})|)e^{j\omega}$$

where $S(e^{j\omega})$ is the Fourier transform of s(n).

Real cepstrum is not invertible but provides a minimum phase reconstruction of the signal [8]. This is depicted in **Figure 6 (a)** and **(b)**.

Since speech is convolution of a fast varying glottal excitation signal, e(n) and a slowly varying vocal tract response, h(n), the cepstrum of a speech segment consists of the glottal excitation signal that occupies the low quefrency region and the vocal tract response that occupies the high quefrency region [8].

Since the phase information is not as much important as the magnitude information in a speech spectrum, the real cepstrum is used due to its computational easiness [8].

The first M samples where M is the number of channels allotted to specifying spectral envelope information [9], typically first 2.5 ms to 5 ms duration, of the cepstrum of a speech segment represent the vocal tract response while the remaining samples represent the glottal excitation signal.

A simple windowing process using a rectangular window separates the vocal tract response from the glottal excitation signal in the quefrency domain.

The inverse process of cepstrum involving exponentiation obtains these signals in time domain.

1.4.2. Linear Prediction Technique

Linear prediction technique is a system modeling technique that models the vocal tract response in a given speech segment as an all-pole linear filter with transfer function of the form

$$H(z)=\cfrac{G}{1+\sum\limits_{k=1}^{P}a_p(k)z^{-k}}$$

where G is dc gain of the filter, p is the order of the filter, $a_p(k)$, k=1,2,...p are the filter coefficients, leaving out the glottal excitation as the residual of the process [4]. Thus the LP technique separates the vocal tract response from the glottal excitation.

The various formulations of linear prediction technique are 1) the covariance method 2) the autocorrelation method 3) the lattice method 4) the inverse filter formulation 5) the spectral estimation formulation 6) the maximum likelihood formulation and 7) the inner product formulation [4,14].

In this paper the filter coefficients and the dc gain were estimated from speech samples via the autocorrelation method by solving the so-called Yule-Walker equations with the help of Levinson-Durbin recursive algorithm [11,14].

2. DATA ACQUISITION & SIGNAL PRE-PROCESSING

Adult subjects, 52 subjects (11 females and 41 males) with distorted sound, 41 subjects (3 females and 38 males) with prolonged sound, 12 subjects (all males) with stammering, 9 subjects (1 female and 8 males) with omissions and 5 (all males) with substitutions were selected. They were asked to spell the phonemes "a" as in "male", "ee" as in "speech", "p" as "pet", "aa" as in "Bob" and "o" as in "boat". These speech signals were recorded using a Pentium-IV computer with 2 GB RAM, 160 GB HDD, PC-based microphone, 16-bit sound card, and free audio recording and editing software at a sampling rate of 8 KHz. These signals hereafter will be referred to as
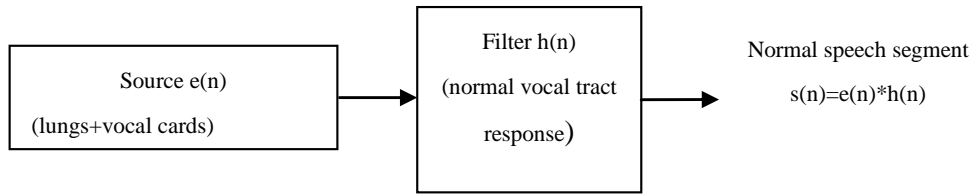
Figure 7. Source-filter model of a normal human speech mechanism.
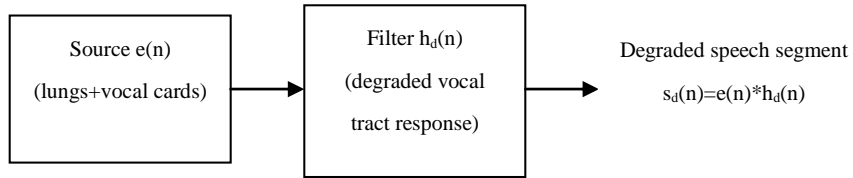
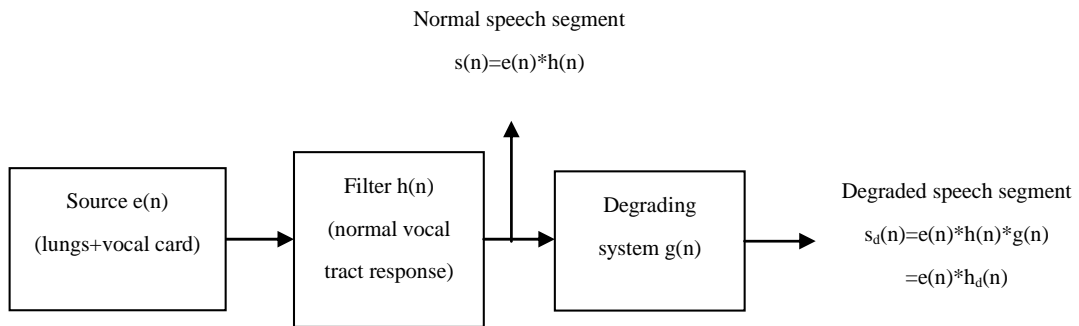**Figure 8.** Source-filter model of a degraded human speech mechanism.

Figure 9. Source-filter model of a degraded human speech mechanism with degradation as separate system.
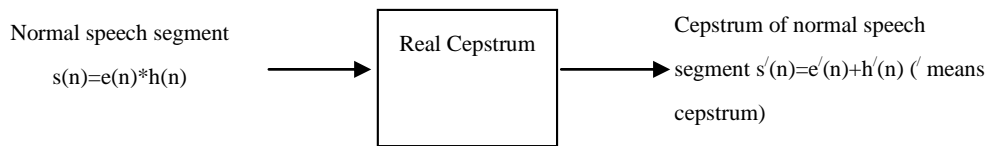

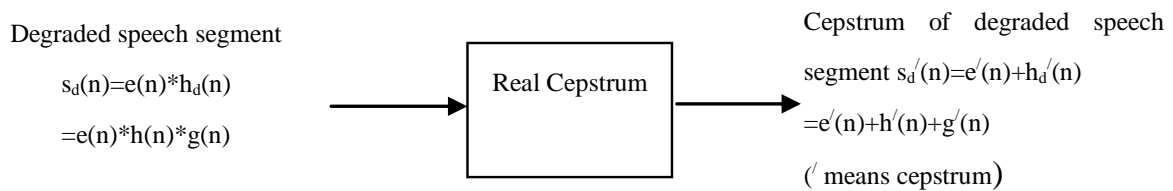
Figure 10. Cepstrum of normal speech segment.
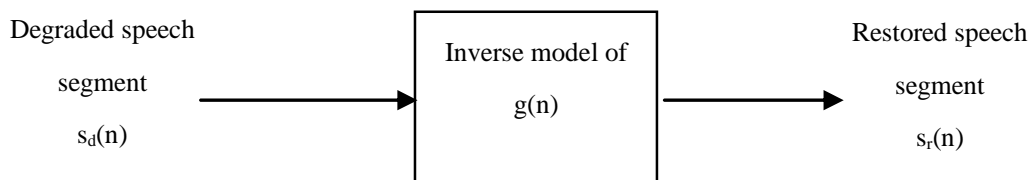

Figure 11. Cepstrum of degraded speech segment.

Figure 12. Restoring normal speech segment from degraded speech segment via inverse model of degradation.

impaired speech signals or utterances in the text.

Same number (119) of normal subjects of similar age and sex were selected and asked to spell the same set of phonemes. These signals were recorded under similar conditions and will hereafter be referred to as normal speech signals or utterances in the text.

These signals were then lowpass-filtered to 4 KHz to avoid any spectral leakage.

The arithmetic mean of each filtered signal was subtracted from it in order to remove the DC offset, an artifact of the recording process [12].

The speech portion of each speech signal was extracted from its background using an endpoint detection algorithm explained in [13].

3. METHODS

Three methods were developed, all being based on the source-filter model of the human speech mechanism. The first two methods were based on the cepstral technique and the third method was based on the Linear Prediction Coding (LPC) technique. In all these three method, the speech was assumed to be the linear convolution of the slowly varying vocal tract response, and the fast varying glottal excitation [4,5,8,9].

3.1. Method 1

This method was based on the following facts: 1) Though non-stationary, speech signal can be considered as stationary for a short period of 20-40 ms [3,4], 2) Speech is a convolution of two signals: the glottal excitation signal and the vocal tract response [4], 3) The excitation signal generally exhibits the speaker characteristics such as pitch and loudness while the vocal tract response determines the sound produced [5] and 4) Cepstrum transforms a convolution process into an addition process [8].

1) and 2) make the short-term analysis of speech signal possible and model a normal human speech mechanism as a linear filter excited by a source as shown in **Figure 7**.

Similarly a disabled human speech mechanism with an impaired vocal tract is modeled as shown in **Figure 8**.

If the degraded vocal tract is modeled as the normal vocal tract followed by a degrading system, then the above source-filter model can be equivalently represented as shown in **Figure 9.**

As suggested by 4), the cepstrum of normal speech segment would be the addition of the cepstrum of normal vocal tract response and the cepstrum of excitation as shown in **Figure 10**.

Similarly for an impaired speech segment the cepstral deconvolution of degraded speech segment is shown in **Figure 11.**

As suggested by 3), if the speech segment, in both cases, represents a similar sound unit (e.g., a similar phoneme), then the $h_d^{/}(n)$ can be represented in term of

$h^{/}(n)$ from **Figure 9** to **Figure 11** as follows

$$h_d(n)=h(n)*g(n)$$
$$h_d^{/}(n)=h^{/}(n)+g^{/}(n)$$

Subtraction of the cepstrum of normal vocal tract from the cepstrum of degraded vocal tract for a similar sound unit yields the cepstrum of degradation as follows

$$g^{/}(n)=h_d^{/}(n)-h^{/}(n)$$

The inverse cepstrum of $g^{/}(n)$ yields the degradation in time domain, $g(n)$. The inverse model of $g(n)$ is obtained as the reciprocal of autoregressive or all-pole model of $g(n)$ obtained via the Levinson-Durbin algorithm.

The speech segment is restored by passing the degraded speech segment through the inverse model of the degradation as shown in **Figure 12**.

Figure 13 shows a complete block diagram representation of the entire method of restoring the speech via the inverse model of the degradation.

3.2. Method 2

Method 2 was based on the same set of facts as Method 1. In this method, the degraded vocal tract response, $h_d(n)$ for a particular phoneme from a disabled speech mechanism is replaced by the normal vocal tract response, $h(n)$ for the same phoneme from a normal speech mechanism. The extraction of the vocal tract responses and reconstructing the speech of improved intelligibility and audibility is done using cepstral technique. The cepstrum of a speech segment consists of the glottal excitation signal that occupies the low quefrency region and the vocal tract response that occupies the high quefrency region [8]. The first M samples where M is the number of channels allotted to specifying spectral envelope information [9], typically first 2.5 ms to 5 ms duration, of the cepstrum of a speech segment represent the vocal tract response while the remaining samples represent the glottal excitation signal. A simple windowing process using a rectangular window separates the vocal tract response from the glottal excitation signal in the quefrency domain [8].

The block diagram representation of the second method has been shown in **Figure 14**.

3.3. Method 3

Method 3 was based on the following facts: 1), 2) and 3) were the same and 4) A speech segment of short duration, e.g., 10-40 ms can be effectively represented by an all-pole filter of order p [14], which is often chosen to be at least 2*f*l/c where f is the sampling frequency, l is the vocal tract length and c is the speed of sound [7,14]. For a typical male utterance with l=17 cm and c=340 m/s, p=f/1000. These filter coefficients were estimated through linear predictive analysis for a speech segment of short duration, e.g., 10-40 ms. The excitation signal is
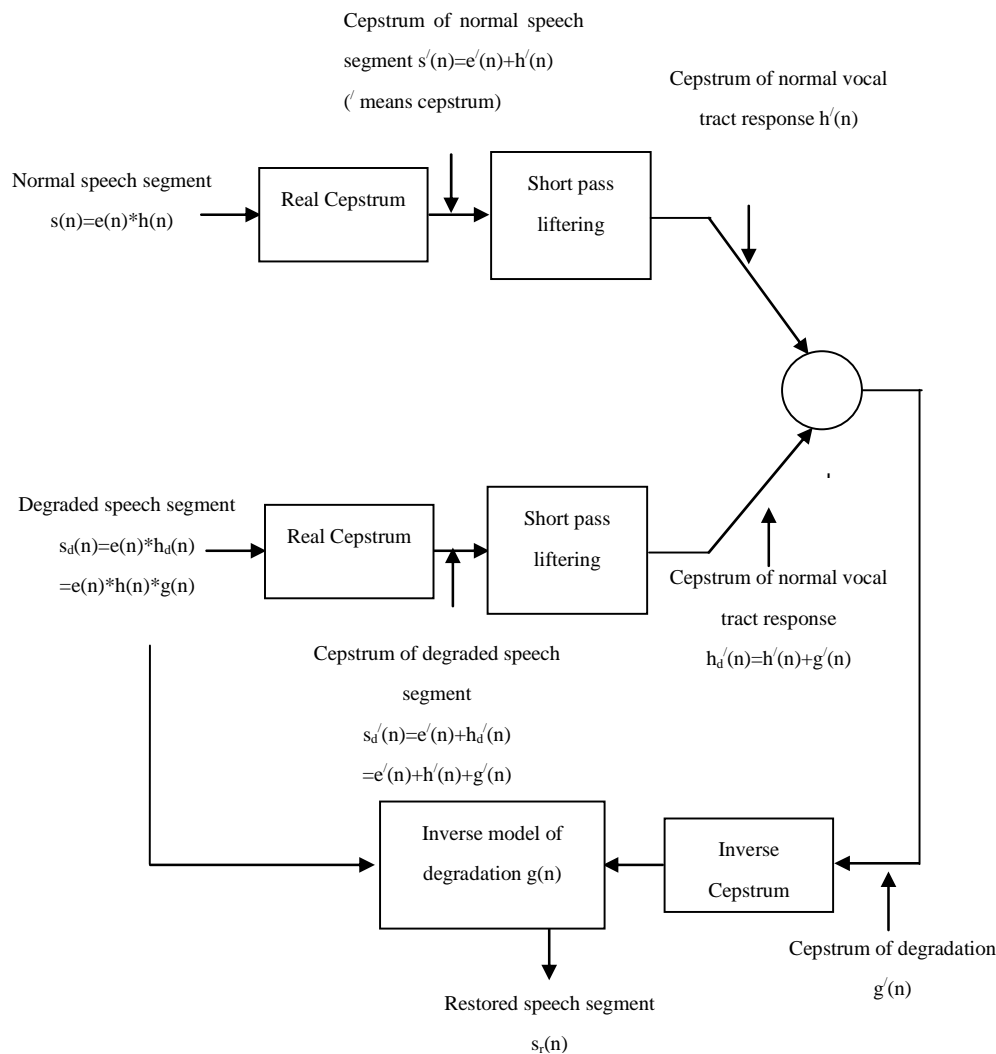
Figure 13. Block diagram representation of Method 1.

obtained either by passing the speech segment through this filter or by synthesizing with estimated pitch period, gain and voicing decision for that segment [14,4]. Here the former is utilized. Thus the linear predictive analysis splits a speech segment into excitation and vocal tract response [14,4].

The first two assumptions make the short-term analysis of speech signal possible and model both the normal and disabled human speech mechanisms as described for Method 1 & 2.

As suggested by 4), both the normal and impaired speech can be split into excitation and vocal tract response.

As suggested by 3), the LP coefficients of the normal speech segment in the place of those of the impaired speech segment are used while the excitation is obtained either from the LP residual for the impaired speech segment or from synthesis from pitch period, gain and voicing decision estimated from the impaired speech

segment. Here the former is utilized

Here the degraded vocal tract response, $h_d(n)$ for a particular phoneme from a disabled speech mechanism, obtained via the LPC technique, is replaced by the normal vocal tract response, $h(n)$ for the same phoneme from a normal speech mechanism obtained via the LPC technique.

The glottal excitation from the degraded (impaired) speech is obtained via the LPC technique as the linear prediction residual [14,4].

The block diagram representation of the method has been depicted in **Figure 15**.

4. IMPLEMENTATION

In all three methods, the speech portions from both the normal and degraded phonemes were extracted using the algorithm in [13] and the normal utterance was time-scaled to match the length of the impaired utterance using the modified phase vocoder [15]. Then each utter-
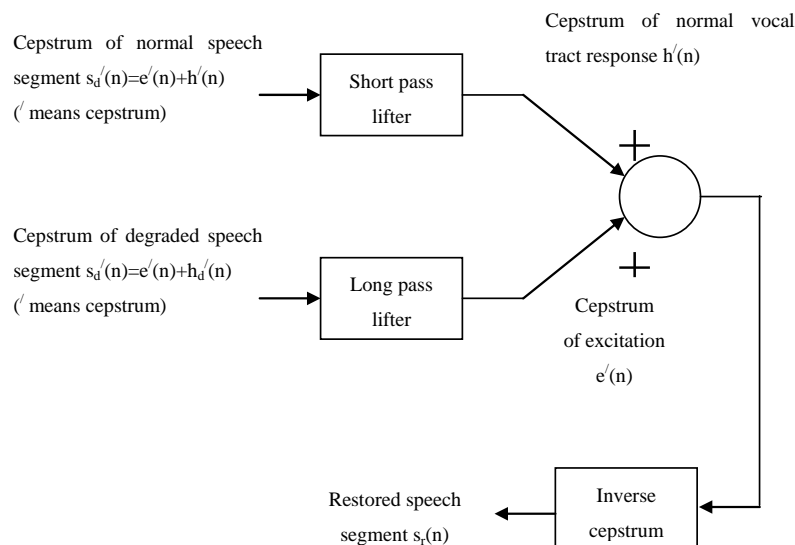
Figure 14. Block diagram representation of Method 2.
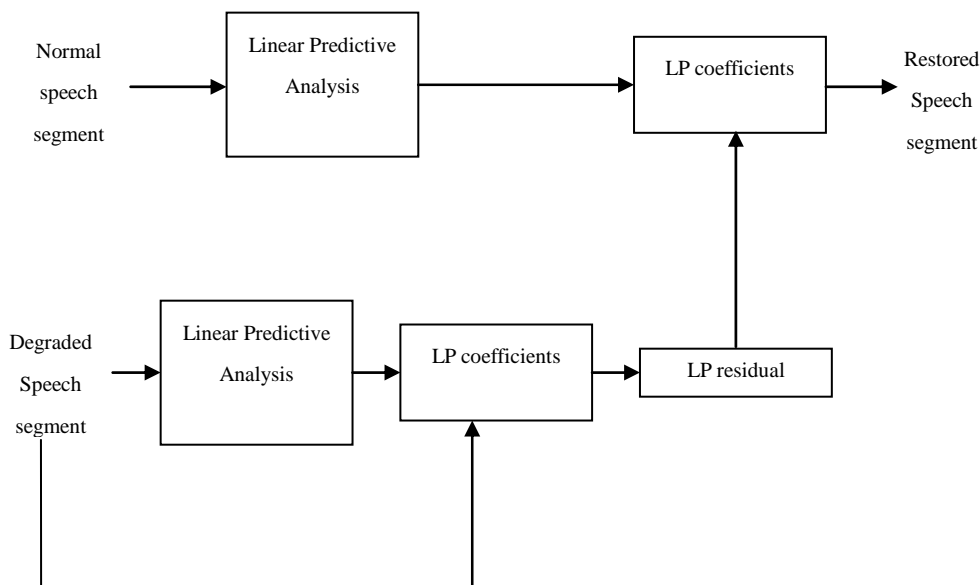
Figure 15. Block diagram representation of Method 3.

ance was segmented into short frames of 20 msec duration overlapping by 5 msec.

In Method 1 & 2, both the frames were preemphasised to cancel the spectral contributions of the larynx and the lips to the speech signal using $H(z) = 1 - \mu z^{-1}$ with $\mu = 0.95$ [7,12]. Then the cepstra of both the frames were computed [8,9].

In Method 1, the first 40 samples of the cepstrum of the normal speech frame were subtracted from those of the cepstrum of the degraded speech frame to extract the cepstrum of the degrading function. The inverse cepstrum of the resultant yielded the degrading function which was then modeled as an all-pole filter. The inverse

of the model was then used to restore the speech.

In Method 2, the first 40 samples of the cepstrum of the degraded speech frame were replaced by those of the cepstrum of the normal speech frame.

In Method 3, after segmentation, the frames were preemphasised and their autocorrelations were computed. The resultant autocorrelations were used to compute the LPC coefficients using Levinson-Durbin recursive algorithm. Then the LPC coefficients computed from the degraded speech frame and the frame itself were used to compute the LPC residues. These LPC residues and The LPC coefficients computed from the normal speech frame were used to synthesize the restores speech frame.

All the above steps were repeated for all frames and for all phonemes. MATLAB 7 was used for programming purposes.

5. COMPARISON OF METHODS

The main advantage of these three methods was that the restored speech had the speaker characteristics since the excitation from the distorted sound was used for restoration as the excitation (the glottal impulse) exhibits mainly the speaker characteristics while the vocal tract response (the articulation) gives rise to various phonetic realization.

All the three methods worked acceptably well with certain articulation impairment such as distorted sound & prolonged sound.

All the three methods suffered from a basic problem, the phonetic mismatching. That is, the process of matching the respective phonemes in the normal and impaired utterances lacks the accuracy due to the fact that the duration of a phoneme in a syllable or a word from two different speakers may not be equal and also its articulation and temporal-spectral shape varies with respect to the preceding and succeeding phones. Moreover, the dynamic time warping techniques used to match two similar time-series may not be used to match the normal and the distorted strings, though they are the same utterances, since they are not 'similar', one being normal and the other, distorted.

All the three methods did not suit for all the speech impairments. For example they did not help solving certain common impairments such as stammering, omissions, substitutions.

Method 1 & 2 suffered from the fact that the real cepstrum is not invertible; only a minimum phase reconstruction is possible. The phase information was lost.

Method 1 suffered from the problem of extracting the degradation exactly since the vocal tract response for a phoneme independently obtained from two speakers do not match sample vice. Hence subtracting the first 25 samples of the cepstrum (representing the normal vocal tract response) of normal speech segment from those (representing the degraded vocal tract response) of impaired speech segment may not exactly give the degradation in the vocal tract response of the impaired subject.

The LP coefficients do not represent the vocal tract response independently of speakers. Hence the restored sound possessed the quality of both of the speakers, the normal and the problem but more towards the problem speaker and less towards the normal speaker.

6. RESULT

In order to assess the result of the above experiments, one thousand observers (500 females and 500 males) of different age group varying from 20 to 40 were selected and requested to listen to the degraded, normal and re-

Table 1. Votes of favour for subjects with distorted sound: Total No. of Votes=52×1000.

	Bad	Good	Excellent
"a" in "male"	2080	46904	3016
"ee" in "speech"	1976	46956	3068
"p" in "pet"	5304	46644	156
"aa" in "Bob"	1144	49452	1404
"o" in "boat"	520	50076	1404

Table 2. Votes of favour for subjects with prolonged sound: Total No. of Votes=41×1000.

	Bad	Good	Excellent
"a" in "male"	820	39360	820
"ee" in "speech"	2747	37761	492
"p" in "pet"	6150	34850	0
"aa" in "Bob"	2132	37392	1517
"o" in "boat"	369	39360	1271

Table 3. Votes of favour for subjects with stammering: Total No. of Votes=12×1000.

	Bad	Good	Excellent
"a" in "male"	9612	1800	588
"ee" in "speech"	9612	1800	588
"p" in "pet"	11856	144	0
"aa" in "Bob"	10680	1200	120
"o" in "boat"	1200	10800	-

Table 4. Votes of favour for subjects with omissions: Total No. of Votes=9×1000.

	Bad	Good	Excellent
"a" in "male"	8982	18	-
"ee" in "speech"	8955	45	-
"p" in "pet"	9000	-	-
"aa" in "Bob"	8982	18	-
"o" in "boat"	-	-	-

Table 5. Votes of favour for subjects with substitutions: Total No. of Votes=5×1000.

	Bad	**Good**	**Excellent**
"a" in "male"	5000	-	-
"ee" in "speech"	5000	-	-
"p" in "pet"	5000	-	-

stored phonemes and to rate them as bad, good or excellent in terms of their intelligibility and audibility. The votes obtained in favour was tabulated as shown in **Tables 1, 2, 3, 4, & 5**.

7. CONCLUSIONS

The future development of this research work, thus, will be focused on developing 1) a formant-based technique and 2) a homomorphic prediction-based technique with complex cepstrum since real cepstrum lacks in phase information [8,9] and on developing a system for continuous speech i.e., words and sentences. For a real-time continuous speech processing, use of dedicated digital signal processor could be an opt suggestion.

REFERENCES

[1] (2004) NICHCY disability fact sheet., Speech & Language Impairments. *NICHCY*. **11**.

[2] (2002) Department of Education, Special education programs and services guide, State of Michigan State.

[3] Shuzo, S. and Kazuo, N. (1985) Fundamental of Speech Signal Processing. Academic Press, London.

[4] Rabiner, L.R. and Schafer, R.W. (1978) Digital processing of speech signal, Prentice-Hall, Engliwood Cliffs, *NJ*.

[5] Rabiner, L.R. and Juang, B.H. (1993) Fundamentals of speech recognition, Prentice-Hall, Engliwood Cliffs, *NJ*.

[6] Rabiner, L.R. and Bernard, G. (1992) Theory and application of digital signal processing, Prentice-Hall of India, New Delhi, Chapter 12.

[7] Thomas, F.Q. (2004) Discrete-time speech signal processing. Pearson Education, Singapore.

[8] Oppenheim, A.V. and Schafer, R.W. (1992) discrete-time signal processing, Prentice-Hall of India, New Delhi.

[9] Oppenheim, A.V. (1969) Speech analysis-synthesis based on homomorphic filtering, *Journal of Acoustic Society of America*, **45**, 458-465.

[10] Oppenheim, A.V. (1976) Signal analysis by homomorphic prediction. *Proc. IEEE, ASSP*, **24**, 327.

[11] Proakis, J. G. and Manolakis, D. G. (2000) Digital Signal Processing, Prentice-Hall of India, New Delhi.

[12] Tony, R. (1998) Speech Analysis Lent Term.

[13] Nipul, B, Sara, M., Slavinskym J.P. and Aamirm V. (2000) A project on speaker recognition' rice university.

[14] Makhoul, J. (1975) Linear prediction: a tutorial review, *Proc. IEEE*, **63**, 561-580.

[15] Jean, L. and Mark, D. (1999) New phase-vocoder techniques for pitch-shifting, harmonizing and other exotic effects, *Proc. IEEE WASPAA*.

Aquaporin 1-expressing MCF-7 mammary carcinoma cells show enhanced migration in vitro

Yong Jiang[1,2], Zhi-Bin Jiang[3]

[1]Key Laboratory of Industrial Microbiology, Ministry of Education, Tianjin University of Science and Technology, Tianjin, China;
[2]College of Biotechnology, Tianjin University of Science and Technology, Tianjin, China;
[3]Jilin Combined Traditional Chinese and Western Medicine Hospital in Jilin Province, Jilin, China.
Email: jiangyong@tust.edu.cn

ABSTRACT

Recent studies have demonstrated that aquaporin (AQP) expression facilitates cell migration and promotes angiogenesis and neutrophil motility. Migration of tumor cells is a crucial step in tumor invasion and metastasis. Here we investigated the expression of AQP in MCF-7 human mammary carcinoma cells and characterized its function in cell migration. Reverse Transcription–Polymerase Chain Reaction, Immunoblot and Immunofluorescence analysis demonstrated two populations of MCF-7 cell clones with low (AQP1L) and high (AQP1H) AQP1 expression and the AQP1 protein expression patterns in the plasma membrane of MCF-7 cells. MCF-7 cell clones (AQP1L and AQP1H) with low and about two-fold higher osmotic water permeability were identified by functional assays with corresponding low and high AQP1 expression. Cell migration rate was remarkably higher in AQP1H cells as compared to AQP1L cells, assessed by wound healing and transwell migration assays. Adenoviral-mediated mRNA and protein expression of AQP1 in AQP1L cells increased their water permeability and migration rate to the level similar to AQP1H cells. The results provided direct evidence that aquaporin-mediated plasma membrane water permeability played an important role in mammary carcinoma cell migration and may be associated with mammary carcinoma invasion and metastasis.

Keywords: Aquaporin; Water Permeability; Cell Migration

1. INTRODUCTION

Since Agre laboratory discovered the first aquaporin [1] from red blood cells, thirteen homologous members (AQP0-AQP12) in mammals and hundreds of AQPs in other species have been identified molecularly [2,3]. Aquaporins (AQPs) are membrane water channels that play pivotal roles in physiological and pathophysiological processes of diverse mammalian organs [4,5,6]. Recent studies indicated a novel role of AQPs in cell migration. Mice lacking AQP1, the endothelial water channel, were found to have impaired endothelial cell migration and tumor angiogenesis [7]. In the same study, over-expression of AQP1 or AQP4 (a structurally related water channel) was also found to promote migration of non-endothelial cells. Subsequent studies demonstrated that AQP1 expression promoted the migration of epithelial cells of kidney proximal tubules and facilitated wound healing of injured proximal tubules [8]. In the brain, AQP4 deletion in astroglial cells was also found to result in impaired migration and remarkably reduced glial scar formation after stab injury [9]. In addition, some other studies provided evidence that aquaglyceroporin AQP9 may be involved in neutrophil cell motility [10,11]. These studies suggested that aquaporin-Mediated plasma membrane water permeability may be a fundamental determinant of cell migration and associated physiological and pathological processes. Migration of tumor cells is a crucial step in tumor invasion and metastasis [12,13]. Previous studies demonstrated the expression of AQPs in some types of tumor cells [14,15,16].

We hypothesized that AQPs also play a critical role in mammary carcinoma cell migration. In the present study, we provided the first evidence for the involvement of AQP1 in mammary carcinoma cell migration.

2. MATERIALS AND METHODS

2.1. Cell Culture

MCF-7 human mammary carcinoma cells (from Cell Bank of Shanghai Institute of Biochemistry and Cell Biology) were maintained in DMEM supplemented with 10% FBS, 100 μg/mL penicillin and 100 μg/mL streptomycin. The cells were plated in 6-well plates

and maintained at 37°C and 5% CO_2. The monoclonal MCF-7 mammary carcinoma cell line was isolated and expanded by limited dilution to investigate the expression and function of AQP1.

2.2. Reverse Transcription-Polymerase Chain Reaction

Total RNA was extracted from the cultured MCF-7 human mammary carcinoma cells with RNeasy micro kit (QIAGEN) and cDNA was reversely transcribed (RT) from the total RNA using a superscript first-strand synthesis system (invitrogen). The cDNA was used as a template for polymerase chain reaction (PCR) amplification (94°C, 5 min and then 30 cycles of 94°C for 30s; 60°C for 30s; 72°C for 60s) using primers flanking a 252 bp fragment of AQP1 (sense, 5'-CTCTCTgTAgCCCTTggACACCTC-3'; antisense, 5'-ggCATCCAggTCATACTCCTCCAC-3') and flanking a 213 bp fragment of GAPDH (sense, 5'-ATT-CAACGGCACAGTCAAGG-3'; antisense, 5'- GCA-GAAGGGGCGGAGATGA-3'). The PCR products were analyzed by agarose gel electrophoresis.

2.3. Immunoblotting

MCF-7 cells were washed with phosphate-buffered saline (PBS), incubated with 2mL 50 mM Na_2-EDTA solution for 5 to 10 min at room temperature and solubilized at 60°C for 15 min in Laemmli sample buffer. Equal amounts of protein (10μg) were resolved by SDS-PAGE on 12% polyacrylamide gel, blotted to polyvinylidene difluoride (PVDF) membranes, blocked with 5%(w/v) nonfat milk and washed in Tris-buffered saline (TBS)-Tween solution. Primary antibodies (anti-AQP1, anti-Tubulin) (Chemicon, 1:1000) and secondary antibody conjugated to horseradish peroxidase (Sigma, 1:3000) were applied successively and detected by enhanced chemoluminescence (Amersham).

2.4. Immunofluorescence

MCF-7 cells were seeded on coverglasses 24 h before fixation in PBS supplemented with 4% paraformaldehyde (PFA) for 15 min. Cells were washed with PBS, blocked and permeabilized by incubation in PBS containing 2% (w/v) bovine serum albumin (Sigma), and 0.1% (v/v) Triton X-100 for 30 mins. Cells were then exposed to a polyclonal rabbit anti-rat AQP1 antibody (Chemicon, 1:1000) overnight at 4°C. After washing, a Cy-3-conjugated anti-rabbit IgG secondary antibody (Sigma) was applied at 1:500 for 1 h at room temperature. Cells were washed with PBS and visualized by fluorescence microscopy (Olympus).

2.5. Adenovirus-Mediated AQP1 Expression

The recombinant adenovirus expressing human AQP1 (Ad-AQP1) and control β-galactosidase adenovirus (Ad-lacZ) were generated using the ViraPower adenoviral expression system (Invitrogen) according to the manufacturer's instruction and purified/ concentrated with an Adeno-X virus purification kit (BD Biosciences-Clontech). The MCF-7*l* cells were infected with Ad-AQP1 or Ad-lacZ (control adenovirus) at a concentration of 150 PFU/cell respectively. AQP1 expression and function were analyzed at 48 h after infection.

2.6. Water Permeability Measurements

The osmotic water permeability of MCF-7 cell plasma membrane was measured by a calcein fluorescence quenching method described previously [17], with modifications. Briefly, MCF-7 cells were grown on round glass coverslips pre-coated with polylysine for 24 h. The cells were incubated with 10μmol/L calcein-AM (Molecular probes) for 10 min and then mounted in a perfusion chamber designed for rapid solution exchange. The time-course of cytoplasmic calcein fluorescence in response to an osmotic gradient was monitored by exchanging perfusate osmolality between 300 mOsmol (PBS) and 150 mOsmol (diluted with distilled water). The rate of change of cell volume was presented as reciprocal exponential time constant ($1/\tau$), which is proportional to osmotic water permeability, where τ is the time needed from the beginning of osmotic switch to the point when the cytoplasmic calcein fluorescence reaches its maximum.

2.7. In Vitro Wound Healing Assay

Assays were performed as described previously with modification [7]. Briefly, MCF-7 cells were cultured as confluence monolayers on 6-well plates and synchronized in 1% fetal bovine serum for 24 h. The monolayers were wounded by removing a 300-500 μm strip of cells across the well with a 200 μL pipette tip and then washed twice to remove non-adherent cells. Wound healing was quantified as the average linear speed of the wound edges over 24 h.

2.8. Transwell Migration Assays

Invasive cell migration was measured using a modified Boyden chamber chemotaxis assay. MCF-7 cells were seeded on top of a culture plate insert (Corning Costar) containing a polycarbonate filter (6.5 mm diameter, 8μm pores) pre-coated with fibronectin (0.5mg/mL). The upper chamber contained cells in
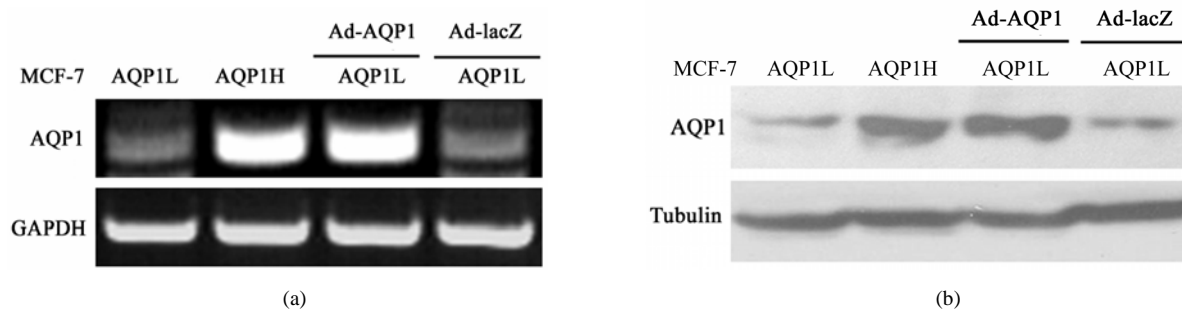
(a)

(b)

Figure 1. The expression of AQP1 in MCF-7 mammary carcinoma cells. (a) Adenovirus-mediated increased mRNA expression of AQP1 in MCF-7 cells. RT-PCR analysis of AQP1 in clones (AQP1L, AQP1H, AQP1L-Ad-AQP1 and AQP1L-Ad-lacZ). (b) Adenovirus-mediated increased protein expression of AQP1 in MCF-7 cells. Immunoblot analysis showing AQP1 protein expression in clones (AQP1L, AQP1H, AQP1L-Ad-AQP1 and AQP1L-Ad-lacZ).
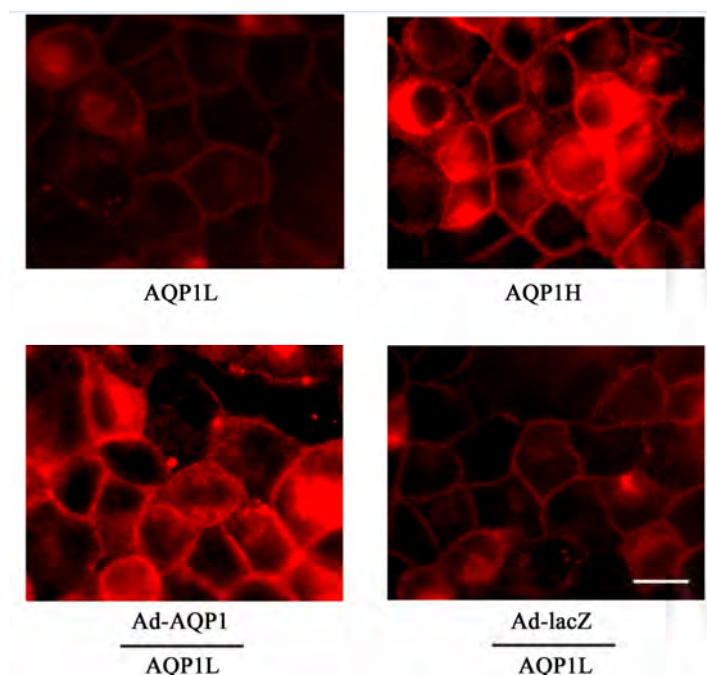

Figure 2. Immunofluorescence showing AQP1 expression in plasma membrane of AQP1L, AQP1H, AQP1L-Ad-AQP1 and AQP1L-Ad-lacZ MCF-7 cells. Scale bar = 10 μm.

(a)

(b)
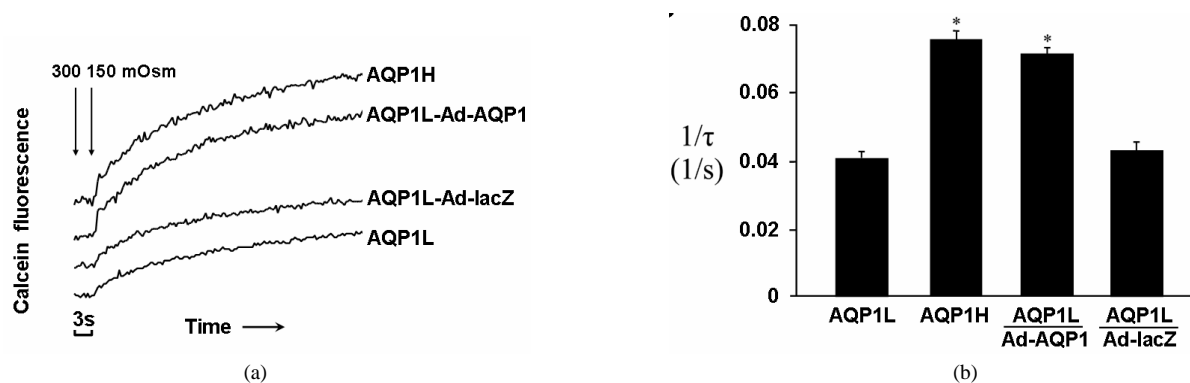
Figure 3. MCF-7 cell water permeability analysis. A, Representative curves showing time-course of calcein fluorescence in response to changes in perfusate osmolality from 300 to 150 mOsmol. B, Cell swelling rates presented by reciprocal exponential time constants ($1/\tau$) in six sets of experiments, (n=6, mean±SE, Asterisks denote $P<0.01$ versus AQP1L cells).

(a)   (b)
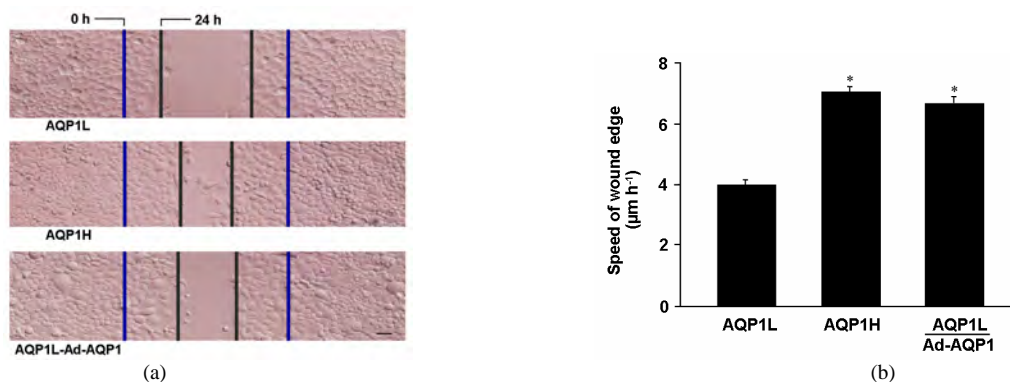
Figure 4. AQP1 facilitated wound healing of MCF-7 cell monolayers. A, Representative images of wound closure showing initial wound edge and wound edge after 24 h. Scale bar = 20 μm. B, The quantified wound edge speed (n=6, mean ± SE, Asterisks denote $P<0.01$ versus AQP1L cells).

(a)   (b)
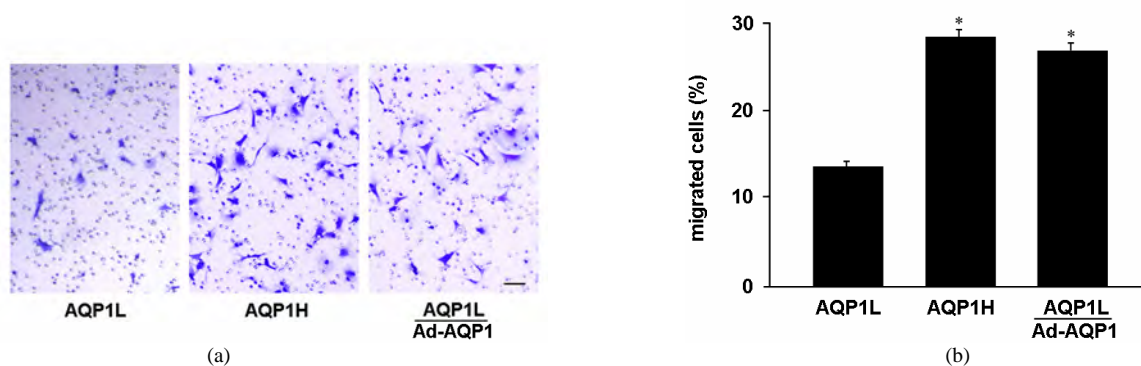
Figure 5. AQP1 promoted invasive migration of MCF-7 cells. A, AQP1L and AQP1H cells migrated through the porous transwell filter. Infection of Ad-AQP1 in AQP1L cells enhanced the transwell migration rate of cells to the level similar to AQP1H cells. Cells were stained with Coomassie blue. Scale bar = 10 μm. B, Percentage of migrated cells at 12 h after plating (n=6, mean ± SE, Asterisks denote $P<0.01$ versus AQP1L cells).

DMEM plus 1% fetal bovine serum, and the lower chamber contained DMEM plus 10% (chemoattractant). Cells were incubated for 12 h at 37°C and 5% CO_2. Cells that did not migrate were wiped away from the top of transwell filter and migrating cells on the bottom surface were counted after staining with Coomassie blue.

2.9. Statistical Analysis

Data were expressed as mean±SE and accompanied by the number of experiments performed independently, and analyzed by t-test. Differences at P<0.05 were considered statistically significant.

3. RESULTS

3.1. Expression of AQP1 in Two Populations of MCF-7 Mammary Carcinoma Cells

We generated monoclonal MCF-7 mammary carcinoma cell line by limited dilution. Selective expression of AQP1 was identified by RT-PCR analysis using primers specific amplifying fragments of human AQP1. Interestingly, two populations of MCF-7 cell

clones with low (AQP1L) and high (AQP1H) AQP1 expression were identified by RT-PCR analysis. The recombinant adenovirus expressing human AQP1 (Ad-AQP1) increased expression of AQP1 in AQP1L cells versus control adenovirus (Ad-lacZ) infected cells (**Figure 1(a)**). Immunoblot and immunofluorescence using a AQP1 polyclonal antibody confirmed the same AQP1 expression patterns as shown in **Figure 1** B and **Figure 2**. Immunofluorescence indicated expression of AQP1 protein in the plasma membrane of MCF-7 cells (**Figure 2**). The AQP1L and AQP1H cells possessed identical protein pattern on SDS-PAGE (data not shown). The identical protein pattern of the two cell populations indicated that they were from the same parental cell type except that the AQP1 expression was different.

3.2. Expression of AQP1 Confered High Water Permeability of MCF-7 Cells

Osmotic water permeability of MCF-7 cell plasma membrane was measured by a calcein fluorescence quenching method. **Figure 3** showed the results of

water permeability measurements of MCF-7 cell lines. The two populations of MCF-7 cell clones (AQP1L and AQP1H) with low and about two-fold higher osmotic water permeability with corresponding low and high AQP1 expression were identified by functional assay (**Figure 3(a), 3(b)**). To determine whether AQP1 expression accounted for the increased water permeability of AQP1L cells, we overexpressed AQP1 in AQP1L cells. Accordingly, osmotic water permeability was increased by about 1.7 fold in AQP1L cells infected with Ad-AQP1. In the same experimental conditions, infection with Ad-lacZ did not increase plasma membrane water permeability of AQP1L cells.

3.3. Expression of AQP1 Facilitated Wound Healing of MCF-7 Cell Monolayers

Wound closure experiments were performed to confirm the role of AQP1 in MCF-7 mammary carcinoma cell migration. **Figure 4** showed significantly accelerated wound closure in AQP1H cells and AQP1L cells infected with Ad-AQP1 at concentration of 150 PFU/cell as compared to AQP1L cells.

3.4. Expression of AQP1 Promoted Invasive Migration of MCF-7 Cells

Transwell migration assay using a modified Boyden chamber was performed as an in vitro model of invasive migration of AQP1L and AQP1H cells, and AQP1L cells infected with Ad-AQP1. As shown in **Figure 5**, migration rate of AQP1H cells towards 10% FBS through the 8 μm pores in the transwell filter was significantly higher (~2.1 fold) than that of AQP1L cells. Infection of Ad-AQP1 at concentration of 150 PFU/cell enhanced the transwell migration rate of AQP1L cells to the level similar to AQP1H cells, suggesting an important role of AQP1 in facilitating cell migration.

4. DISCUSSION

Migration of tumor cells through narrow extracellular spaces requires cell volume regulation and shape change. Previous studies revealed that ion channels and transporters were essential for cell volume regulation and invasive migration of some human tumors [18,19]. Our present study demonstrated for the first time that aquaporin-mediated transmembrane water flux was an important determinant of MCF-7 mammary carcinoma cell migration and may be associated with mammary carcinoma invasion and metastasis. We proposed that transmembrane water efflux through aquaporin water channels facilitated cell volume change mediated by net salt efflux through ion channels and transporters in the mi-

grating mammary carcinoma cells, leading to enhanced motility and invasion. There were evidence suggesting a role of AQPs in tumor cell migration and metastasis. Expression of several AQPs was found in a wide range of tumor cells in addition to the ubiquitous expression of AQP1 in the microvessels of solid tumors [14,15,16]. Warth *et al.* [16] found that AQP4 expression at the end-feet membranes was redistributed in high-grade astrocytomas.

Our experiments provided direct evidence that aquaporin-mediated high plasma membrane water permeability contributed to cell shape-volume change required for mammary carcinoma cell migration. Cell migration rate was remarkably higher in AQP1H cells as compared to AQP1L cells in both wound healing and invasive transwell migration assays. Adenoviral-mediated AQP1 expression in AQP1L cells increased their plasma membrane water permeability and migration rate, further supporting the role of AQP1 in facilitating migration of the mammary carcinoma cells. Future studies are needed to uncover the role of aquaporin water channels in vivo tumor cell migration and metastasis.

5. ACKNOWLEDGEMENT

REFERENCES

[1] Preston, G.M., Carroll, T.P., Guggino, W.B. and Agre P. (1992) Appearance of water channels in Xenopus oocytes expressing red cell CHIP 28 protein. *Science*, **256**, 385-387.

[2] Agre P. (2006) The aquaporin water channels. *Proc Am Thorac Soc*, **3**, 5–13.

[3] Takata, K., Matsuzaki, T. and Tajika, Y. (2004) Aquaporins: water channel proteins of the cell membrane. *Prog Histochem Cytochem*, **39**, 1-83.

[4] Agre, P., King, L.S., Yasui, M., Guggino, W.B. and Ottersen, O.P. (2002) Aquaporin water channels–from atomic structure to clinical medicine. *J Physiol*, **542**, 3-16.

[5] Verkman, A.S. (2002) Physiological importance of aquaporin water channels. *Ann Med*, **34**, 192-200.

[6] Verkman, A.S. (2005) More than just water channels: unexpected cellular roles of aquaporins. *J Cell Sci*, **118**, 3225-3232.

[7] Saadoun, S., Papadopoulos, M.C., Chikuma, M.H. and Verkman, A.S. (2005) Impairment of angiogenesis and cell migration by targeted aquaporin-1 gene disruption *Nature*, **434**, 786-792.

[8] Chikuma, M.H. and Verkman, A.S. (2006) Aquaporin-1 facilitates epithelial cell migration in kidney proximal tubule. *J Am Soc Nephrol*, **17**, 39-45.

[9] Saadoun, S., Papadopoulos, M.C., Watanabe ,H., Yan, D. and Manley G.T. (2005) Involvement of aquaporin-4 in

astroglial cell migration and glial scar formation. *J Cell Sci*, **118**, 5691-5698.

[10] Loitto, V.M., Forslund, T., Sundqvist, T., Magnusson, K. E. and Gustafsson, M. (2002) Neutrophil leukocyte motility requires directed water influx, *J Leukoc Biol*, **71**, 212-222.

[11] Loitto, V.M. and Magnusson, K.E. (2004) Hg^{2+} and small-sized polyethylene glycols have inverse effects on membrane permeability, while both impair neutrophil cell motility. *Biochem Biophys Res Commun*, **316**, 370-378.

[12] Bogenrieder, T. and Herlyn, M. (2003) Axis of evil: molecular mechanisms of cancer metastasis. *Oncogene*, **22**, 6524-6536.

[13] Harlozinska, A. (2005) Progress in molecular mechanisms of tumor metastasis and angiogenesis. *Anticancer Res*, **25**, 3327–3333.

[14] Yang, J.H., Shi, Y.F., Cheng, Q. and Deng, L. (2006) Expression and localization of aquaporin-5 in the epithelial ovarian tumors. *Gynecol Oncol*, **100**, 294-299.

[15] Mazal, P.R., Susani, M., Wrba, F. and Haitel, A. (2005) Diagnostic significance of aquaporin-1 in liver tumors. *Hum Pathol*, **36**, 1226-1231.

[16] Warth, A., Mittelbronn, M. and Wolburg, H. (2005) Redistribution of the water channel protein aquaporin-4 and the K^+ channel protein Kir4.1 differs in low-and high-grade human brain tumors. *Acta Neuropathol*, **109**, 418-426.

[17] Solenov, E., Watanabe, H., Manley, G.T. and Verkman, A. S. (2004) Sevenfold-reduced osmotic water permeability in primary astrocyte cultures from AQP-4-deficient mice, measured by a fluorescence quenching method. *Am J Physiol Cell Physiol*, **286**, 426-432.

[18] Ransom, C.B., O'Neal, J.T. and Sontheimer, H. (2001) Volume-activated chloride currents contribute to the resting conductance and invasive migration of human glioma cells. *J Neurosci*, **21**, 7674-7683.

[19] Soroceanu, L., Manning, T.J. and Sontheimer, H. (1999) Modulation of glioma cell migration and invasion using Cl^- and K^+ ion channel blockers. *J Neurosci*, **19**, 5942-5954.

Automatic detection of multiple oriented blood vessels in retinal images

P. C. Siddalingaswamy[1], K. Gopalakrishna Prabhu[2]

[1]Department of Computer Science & Engineering, Manipal Institute of Technology, Manipal, India;
[2]Department of Biomedical Engineering, Manipal Institute of Technology, Manipal, India.
Email: pcs.swamy@manipal.edu

ABSTRACT

Automatic segmentation of the vasculature in retinal images is important in the detection of diabetic retinopathy that affects the morphology of the blood vessel tree. In this paper, a hybrid method for efficient segmentation of multiple oriented blood vessels in colour retinal images is proposed. Initially, the appearance of the blood vessels are enhanced and background noise is suppressed with the set of real component of a complex Gabor filters. Then the vessel pixels are detected in the vessel enhanced image using entropic thresholding based on gray level co-occurrence matrix as it takes into account the spatial distribution of gray levels and preserving the spatial structures. The performance of the method is illustrated on two sets of retinal images from publicly available DRIVE (Digital Retinal Images for Vessel Extraction) and Hoover's databases. For DRIVE database, the blood vessels are detected with sensitivity of 86.47±3.6 (Mean±SD) and specificity of 96±1.01.

1. INTRODUCTION

In clinical ophthalmology colour retinal images acquired from digital fundus camera are widely used for detection and diagnosis of diseases like diabetic retinopathy, hypertension and various vascular disorders. Retinal images provide information about the blood supply system of the retina. Diabetic retinopathy is a disorder of the retinal vasculature that eventually develops to some degree in nearly all patients with long-standing diabetes mellitus [1]. The timely diagnosis and referral for management of diabetic retinopathy can prevent 98% of severe visual loss, for that, the patient has to undergo regular screening of eye for retinopathy. The process involves dilating the eyes with mydriatic drops and capturing the retinal image using standard digital colour fundus camera. Screening program results in large number of retinal images needed to

be examined by ophthalmologists. Manual diagnosis is usually performed by analyzing the images from a patient, as not all images show signs of diabetic retinopathy, it increases the time and decreases the efficiency of ophthalmologists. Therefore, an automatic segmentation of the vasculature could save workload of the ophthalmologists and may assist in characterizing the detected lesions and to identify false positives [2]. Another important application of automatic retinal vessel segmentation is in the registration of retinal images of the same patient taken over period of time [3]. The registered images are useful in monitoring the progression of disease and to observe the effect of treatment.

Different techniques of segmentation of retinal images have been investigated so far. They are filter based methods, vessel tracking methods, classifier based methods and morphological methods. The techniques utilize the prior knowledge such as, contrast that exists between the blood vessels and surrounding background, origin of vasculature from the same point that is the optic disc and connectivity of the vessels. Filter based methods [4,5] and [6] employ a two dimensional linear structural element that has a Gaussian cross-profile section to identify the blood vessels, which typically has a Gaussian profile. The gaussian kernel is rotated into different orientations to fit into vessels of different configuration to obtain a vessel enhanced image. The image is then thresholded to extract the vessel part from the background. In Hoover *et al.*, 2000 [4] threshold is computed using piece wise threshold probing of matched filter response image. This works well on images of healthy retina, but in diseased states such as diabetic retinopathy it results in detection of many false positives. These methods suffer from problems associated with detecting smaller and tortuous vessels that are prone to changes in background intensity. Vessel tracking methods [7] and [8] use a model to track the vessels starting at given points. Here each vessel segment is defined by three attributes which are direction, width, and center point. Individual segments are identified using a search procedure which keeps track of the center of the vessel and makes some decisions about the future
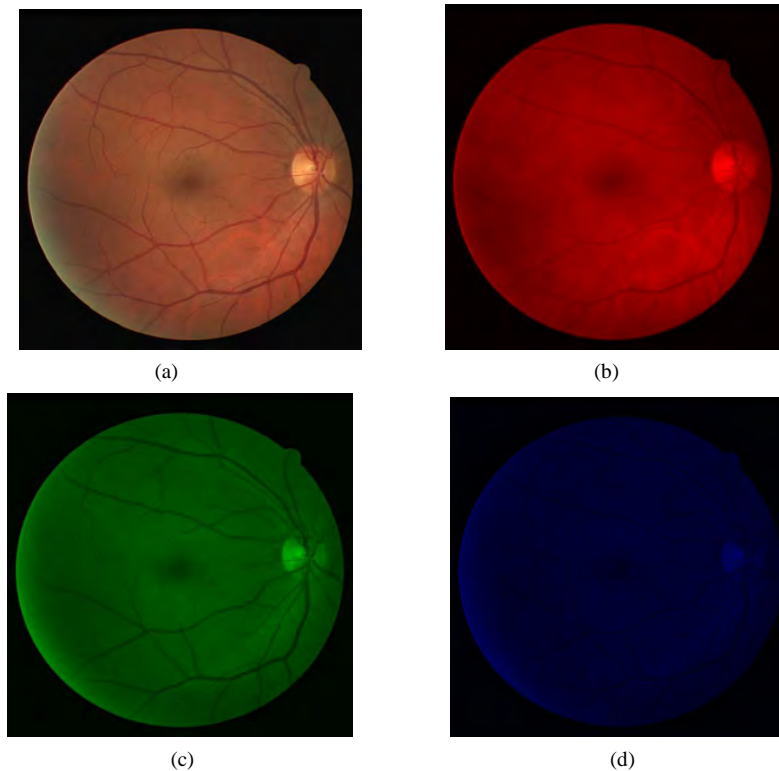
Figure 1. (a) Colour retinal image; (b-d) Red, Green and Blue component images.

path of the vessel based on certain vessel properties. These methods require that beginning and ending search points are manually selected using cursor or by using simple thresholding techniques. Vessel tracking methods provide very accurate measurements of vessel widths but tracking methods often tend to terminate at branch points. Classifier-based method employs two-step approach [9]. They start with a segmentation step often by employing one of the mentioned matched filter-based methods and then the regions are classified according to many features. In the next step neural networks classifier is constructed using selected features by the sequential forward selection method with the training data to detect vessel pixels. Mathematical Morphology is employed for segmentation of blood vessels as reported in [10,11] and [12]. These methods exploits features of the vasculature shape that are known prior, such as it being piecewise linear and connected. They work well on normal retinal images with uniform contrast but suffer when there is a noise due to pathologies within the retina of eye. Many papers have reported work on segmentation of vessels, but still there is scope for improvement as these methods detect vessels along with artifacts. Also detection process becomes much more complicated in presence of lesions and other pathological changes affect the retinal images.

The proposed retinal vessel detection method is comprised of two steps that is the retinal vessel enhancement followed by entropic thresholding. A set of Gabor filters tuned to particular frequency and orientation are used to enhance the blood vessels suppressing the background. Entropy based thresholding based on gray level co-occurrence matrix is employed for the segmentation of the vessels. The following sections elucidate materials and methods for vessel segmentation method, results and discussion.

2. MATERIALS AND METHODS

To develop a retinal vessel segmentation system, the first important thing is to obtain an effective database. To realize this and also for facilitating comparison with the existing methods, two sets of publicly available databases are used.

2.1. Retinal Image Database

The DRIVE database provides forty images [19]. The images are acquired using a Canon CR5 non-mydriatic 3CCD camera with a 45 degree field of view (FOV). Each image was captured using 8 bits per colour plane at 768×584 pixels. The FOV of each image is circular with a diameter of approximately 540 pixels. The Hoover's database [4] consists of twenty digitized slides captured by a TopCon TRV-50 fundus camera at 35 FOV. The slides were digitized to 700× 605 pixels with eight bits per colour channel. Both the databases provide the hand labeled blood vessels identified by experts as gold standard and binary mask image identifying the boundary of the effective portion of each image. Apart from these

(a)   (b)
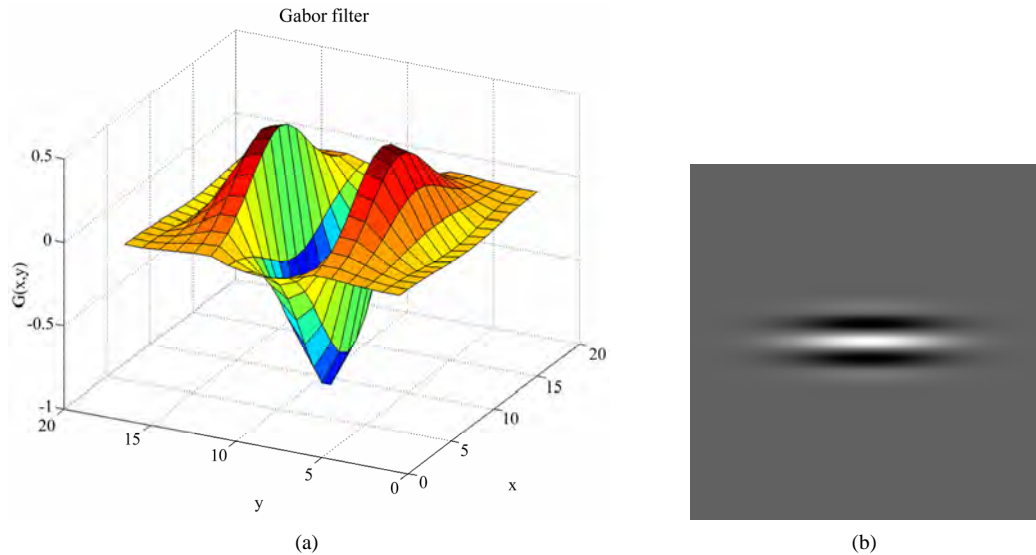
Figure 2. (a) Surface representation of Gabor filter; (b) Real part of Gabor filter.

two databases the images are also acquired from the Department of Ophthalmology, Kasturba medical college (KMC), Manipal using Sony FF450IR digital colour fundus camera with 24 bit colour depth and 768×576 pixel resolution. These images are of very large variability in terms of fundus disease and image quality and were used to test the robustness of proposed retinal vessel detection method.

2.2. Preprocessing

In the colour retinal images, blood vessels appear darker than the background similar to the colour of lesions like microaneurysms and hemorrhages. So it becomes essential to exempt the vessel area during the detection of lesions to avoid false positives. Only one step is involved in the preprocessing of retinal images for segmentation of vessels. It can be seen in the **Figure 1** that the blood vessels appear most contrasted in the green channel compared to red and blue channels in RGB image. Only the green channel image is used for further processing suppressing the other two colour components.

2.3. Vessel Enhancement

The Gabor filters are widely applied to image processing and computer vision application problems such as face recognition and texture segmentation, strokes in character recognition and roads in satellite image analysis [13,14] and [15]. Since, the vessels in the retinal image are connected and piecewise linear, for their segmentation gabor filters are better suited as they are capable of detecting oriented features and can be fine tuned to specific frequencies. Because of their frequency sensitiveness it is possible to filter out the background noise of retinal images. The spatial Gabor filter kernels are sinusoids modulated by a Gaussian window, the real part of

which is expressed by

$$g(x,y) = exp\left[-\pi\left(\frac{x_p^2}{\sigma_x} + \frac{y_p^2}{\sigma_y}\right)\right]\cos\left(2\pi f x_p\right) \quad (1)$$

where,

$$x_p = x cos\theta + y sin\theta$$
$$y_p = -x sin\theta + y cos\theta$$

The Gabor function is defined by five parameters as follows. θ is the orientation of the filter, an angle of zero gives a filter that responds to vertical features. f is the central frequency of pass band. σ_x is the standard deviation of Gaussian in x direction along the filter that determines the bandwidth of the filter. σ_y is the standard deviation of Gaussian across the filter that control the orientation selectivity of the filter. The parameters are to be derived by taking into account the size of the lines or curvilinear structures to be detected. In retinal images the width of the vessels varies along the length of the vessel and in manual segmentation it is found that majority of the vessel diameters are of 4 pixels wide. Therefore to accommodate the all vessels in the detection the thickness parameter t is set to four. The other parameters are derived using the procedure in [15] as $f = 0.5/t$, $\sigma_x = \lambda t/0.75\pi$ and $\sigma_y = 0.85\sigma_x$ and $\lambda = \sqrt{2ln2/\pi}$.

The surface representation and real part of resulting Gabor kernel is shown in **Figure 2**. It can be seen that it is suited for the orientation of directional features to provide good response for pixels associated with retinal blood vessels.

To obtain good response of the vessels oriented along
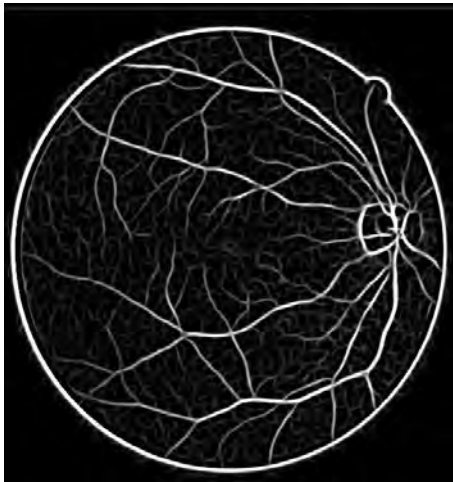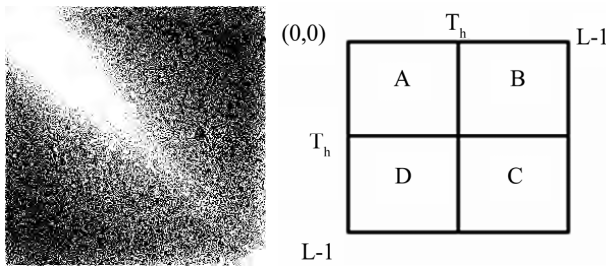
Figure 3. Enhanced vessels in Gabor response.

Figure 4. (a) Gray level distribution in GLCM; (b) Representation of GLCM of 4 quadrants

different directions, the θ of filter is rotated from 0° to 170° in the steps of ten degrees to produce a single peak response on the center of a vessel segment. At each pixel only the maximum response is retained. **Figure 3** shows the result of convolving the image in **Figure 1(c)** with the set of gabor filters. It can be seen that the vessels are enhanced and background is suppressed considerably.

To obtain good response of the vessels oriented along different directions, the θ of filter is rotated from 0° to 170° in the steps of ten degrees to produce a single peak response on the center of a vessel segment. At each pixel only the maximum response is retained. **Figure 3** shows the result of convolving the image in **Figure 1(c)** with the set of gabor filters. It can be seen that the vessels are enhanced and background is suppressed considerably.

2.4. Entropic Thresholding

Thresholding is used to segment the blood vessels from the background. An entropy based thresholding method based on gray level co-occurrence matrix (GLCM) is used to find optimal threshold as it takes into account the spatial distribution of gray levels and preserves the spatial structures in thresholded image [16,17] and [18].

Gray level co-occurrence matrix contains information on the distribution of gray level frequency and edge information, as it is very useful in finding the threshold value.

The gray level co-occurrence matrix $T = [t_{i,j}]$ of the image I with $M \times N$ dimensional matrix gives an idea about the transition of intensities between adjacent pixels, indicating spatial structural information of an image. Depending upon the ways in which the gray level i follows gray level j, different definitions of co-occurrence matrix are possible. The GLCM is obtained as follows:

$$t_{i,j} = \sum_{l=1}^{M} \sum_{k=1}^{N} \delta \qquad (2)$$

where,

$$\delta = 1 \quad if \left\{ \begin{array}{c} f(l,k)=i \ and \ f(l,k+1)=j \\ or \\ f(l,k)=i \ and \ f(l+1,k)=j \end{array} \right.$$

$$\delta = 0 \quad otherwise$$

The probability of co-occurrence P_{ij} of gray levels i and j is written as

$$P_{i,j} = \frac{t_{i,j}}{\sum_i \sum_j t_{i,j}} \qquad (3)$$

Let T_h be the threshold within the range $O_h T_h L$-1, where L is the number of gray levels. Threshold T_h partitions the GLCM into four quadrants, namely A, B, C, and D as in **Figure 4**. Quadrant A represents gray level transition within the object while quadrant D represents gray level transition within the background. The gray level transition between the object and the background or across the object's boundary is placed in quadrant B and quadrant C.

The probabilities of object class and background class are defined as

$$P_A = \sum_{i=0}^{T_h} \sum_{j=0}^{T_h} P_{i,j} \qquad (4)$$

Table 1. Performance of retinal blood vessels segmentation method.

Database	No. of images	Sensitivity (%)	Specificity (%)
DRIVE	40	86.47±3.6	96±1.01

Table 2. Comparison of vessel segmentation results on Hoover's database.

Method	Sensitivity range (%)	Specificity range (%)
Proposed method	79-91	94-98
Hoover *et al.*2000	80-90	92-93

(a)   (b)   (c)

(d)   (e)   (f)
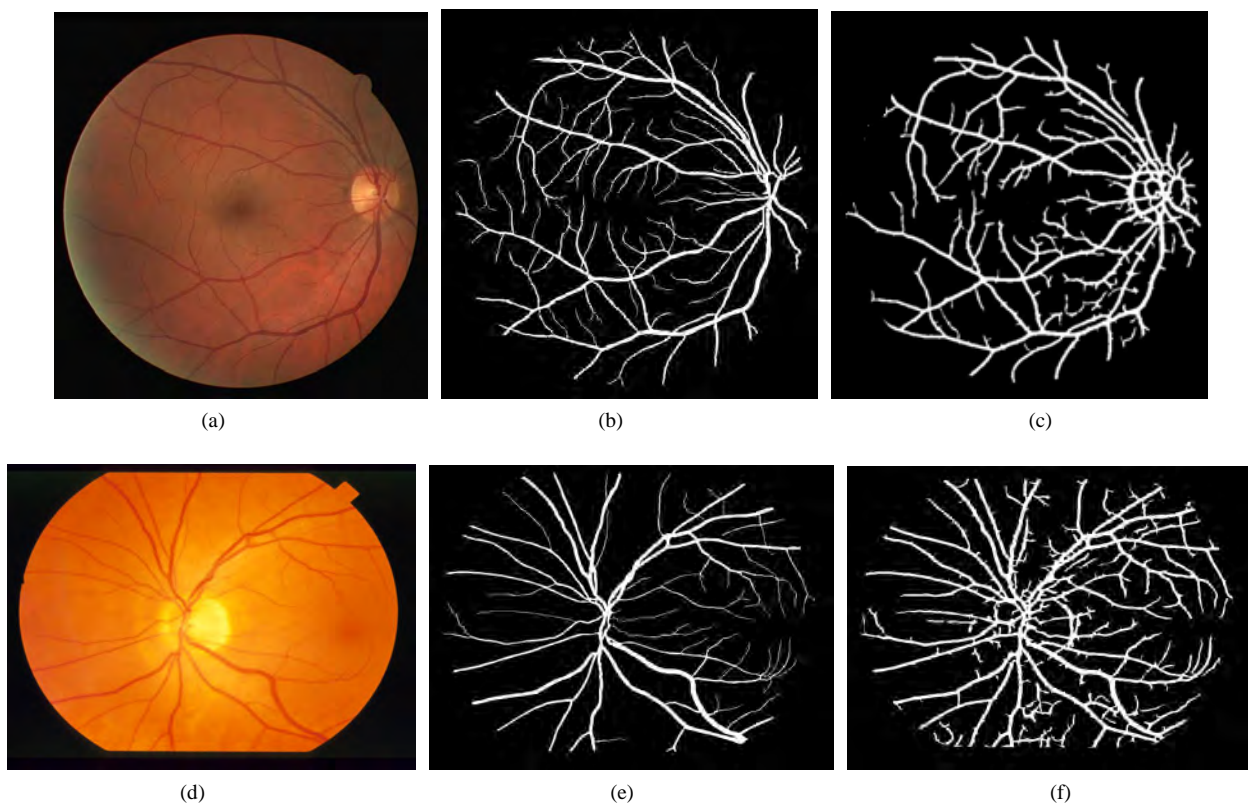
Figure 5. Retinal vessel segmentation. (a) Image from DRIVE database; (b) Corresponding manual segmentation; (c) Vessel segmentation result; (d) Image from Hoover's database; (g) Corresponding manual segmentation; (f) Vessel segmentation result.
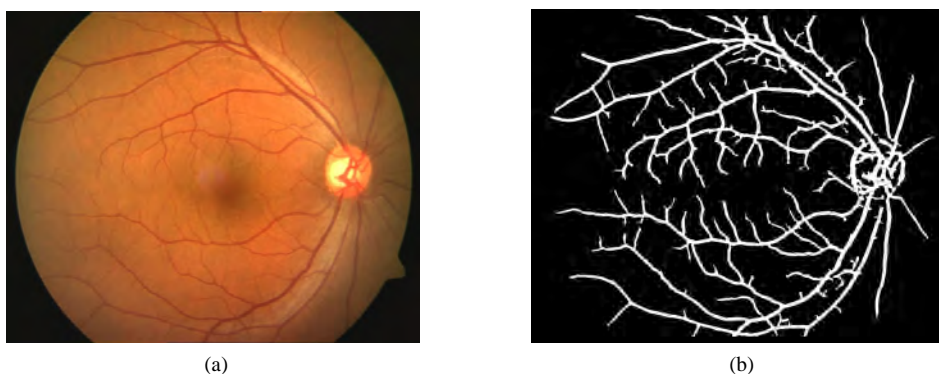

(a)   (b)

Figure 6. (a) Digital color retinal image from KMC database; (b) Segmented vessels.

$$P_C = \sum_{i=T_h+1}^{L_{h1}} \sum_{j=T_h+1}^{L_{h1}} P_{i,j} \qquad (5)$$

Using (4) and (5) as normalization factors, the normalized probabilities of the object class and background class are functions of threshold vector (T_h, T_h) are defined as

$$P_{i,j}^A = \frac{P_{i,j}}{P_A} \qquad (6)$$

$$P_{i,j}^C = \frac{P_{I,J}}{p_c} \qquad (7)$$

The second-order entropy of the object is given by

$$H_A(T_h) = h\frac{1}{2} \sum_{i=0}^{T_h} \sum_{j=0}^{T_h} P_{i,j}^A \log_2 P_{i,j}^A \qquad (8)$$

Similarly, the second-order entropy of the background is given by

$$H_C(T_h) = h\frac{1}{2} \sum_{i=T_h+1}^{L_{h1}} \sum_{j=T_h+1}^{L_{h1}} P_{i,j}^C \log_2 P_{i,j}^C \qquad (9)$$

The total second-order local entropy of the object and the background is given by

$$H_T(T_h) = H_A(T_h) + H_C(T_h) \qquad (10)$$

Finally T_E the gray level corresponding to the maximum of $H_T(T_h)$ gives the optimal threshold for vessel and non vessel classification.

$$T_E = \arg[\,_{T=0L_{h1}}^{\max} H_T(T_h)] \qquad (11)$$

3. RESULTS AND DISCUSSION

Retinal images from the DRIVE database and Hoover's database are used for the automatic segmentation of retinal vessels. **Figure 5** shows the manual segmented image and the final output obtained by the proposed method on the images from both the image databases. On Windows XP, Intel PC with 1.66 GHz CPU and 512MB memory using Matlab 7.0, the method takes about 20 seconds to detect the vessels in retinal image.

The performance of the method is evaluated using sensitivity and specificity at pixel level in comparison with manually segmented vessels by an expert.. Sensitivity gives the percentage of pixels correctly classified as vessels by the method and specificity gives the percentage of non vessels pixels classified as non vessels by the method. given by

$$Sensitivity = \frac{T_p}{T_p + F_n} \qquad (12)$$

$$Specificity = \frac{T_n}{T_n + F_P} \qquad (13)$$

where T_p is true positive, T_n is true negative, F_p is false positive and F_n is false negative at each pixel. **Table 1** show the performance of the proposed method on forty images from DRIVE database.

The results of the proposed method are also compared with those of Hoover *et al.* 2000 [4] on twenty images from the Hoover's database. The database has ten normal and ten abnormal retinal images and the result is depicted in **Table 2**. It can be seen that the proposed method performs better with lower specificity even in the presence of lesions in the abnormal images.

Apart from two standard databases the method is also tested on retinal images obtained from Ophthalmology department, KMC. These images are of large variability in terms of presence of lesions and image quality. These are considered to evaluate the robustness of the method. The result for one of the image is shown in **Figure 6** and is validated by ophthalmologists.

4. CONCLUSIONS

An efficient method for automatic segmentation of retinal blood vessels has been presented. Images from three different datasets are used to evaluate the robustness and accuracy of the method, demonstrating that it may be useful in a wide range of retinal images. Based on a brief

comparison with some other vessel segmentation algorithms, we can conclude that the Gabor filter and entropic threshold provides a better segmentation. The segmented vessels can be used to obtain the control points used in the retinal registration techniques. It is hoped that automated segmentation of vessel technique can detect the signs of diabetic retinopathy in the early stage, monitor the progression of disease, minimize the examination time and assist the ophthalmologist for a better treatment plan.

5. ACKNOWLEDGEMENT

REFERENCES

[1] Emily, Y.C. (2003) Diabetic retinopathy. *Preferred Practice Patterns*, American Academy of Ophthalmology, USA.

[2] Thomas, W., Klein, J.C., Pascale, M. and Ali, E. (2002) A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina. *IEEE Trans. Medical. Imaging*, **21**, 1236-1243.

[3] Laliberté, F., Gagnon, L. and Sheng, Y. (2003) Registration and fusion of retinal images: an evaluation study. *IEEE Trans. Medical. Imaging*, **22**, 661-673.

[4] Hoover, A., Kouznetsoza, V., Goldbaum, M. (2000) Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, **19**, 203-210.

[5] Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M. and Goldbaum, M. (1989) Detection of blood vessels in retinal images using two dimensional matched filters. *IEEE transactions on Medical Imaging,* **8**, 263-269.

[6] Thitiporn, C. and Fan, G.L. (2003) An efficient Algorithm for extraction of anatomical structures in retinal images. *Proc. Of Intl. Conf. on Image Processing*, **1**, 1093-1096.

[7] Wu, D., Zhang, M., Liu, J.C. and Wendall, B. (2006) On the adaptive detection of blood vessels in retinal images. *IEEE Transactions on Biomedical Engineering*, **53**, 341-343.

[8] Pinz, A., Bernogger, S., Datlinger, P. and Kruger, A. (1998) Mapping the human retina. *IEEE Transactions on Medical imaging*, **17**, 606-619.

[9] Sinthanayothin, C., Boyce, J.F., Cook, H.L. and Williamson H.T. (1999) Automated location of the optic disc, fovea, and retinal blood vessels from digital color fundus images. *British Journal of Ophthalmology*, **83**, 902-910.

[10] Zana, F. and Klein, J. (2001) Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Transactions on Image Processing*, **10**, 1010-1019.

[11] Siddalingaswamy, P.C., Prabhu, G.K. and Mithun, D. (2006) Feature extraction of retinal image. *Proc. Of the National Conference for PG and Research scholars,* 24-27.

[12] Yang, Y., Huang, S.Y. and Rao, N. (2008) An automatic hybrid method for retinal blood vessel extraction. J*ournal of Appl. Math. Comp. Sci.*, **18**, 399-407.

[13] Lee, T.S. (1996) Image representation using 2D Gabor wavelets. *IEEE Transactions of Pattern Analysis and Machine Intelligence*. **18**, 959-971.

[14] Chen, J., Sato, Y. and Tamura, S. (2000) Orientation space filtering for multiple orientation line segmentation. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, **22**, 417-429.

[15] Liu, Z.Q., Cai, J. and Buse, R. (2003) Hand-writing recognition: soft computing and probablistic approaches. *Springer Verlag*. Berlin.

[16] Yang, C.W., Chung, P.C. and Chang, C.I. (1996) Hierar-chical fast two dimensional entropic thresholding algorithm using a histogram pyramid. *Optical Engineering*, **35**, 3227-3241.

[17] Mark, L. Althouse, G. and Chang, C. (1995) Target detection in multispectral images using the spectral co- occurrence matrix and entropy thresholding. *Optical Engineering,* **34**, 2135-2148.

[18] Mokji, M.M. and S.A.R and Bakar A. (2007) Adaptive thresholding based on co occurrence matrix edge information. *Journal of Computers*, **2**, 44-52.

[19] Staal, J.J., Abràmoff, M.D., Niemeijer, M., Viergever, M. A. and van Ginneken, B. (2004) Ridge based vessel segmentation in color images of the retina. *IEEE Trans. Med. Imag.*, **23**, 501-509.

Call for Papers

Journal of Biomedical Science and Engineering (JBiSE)

www.scirp.org/journal/jbise

JBiSE, an international journal, publishes research and review articles in all important aspects of biology, medicine, engineering, and their intersection. Both experimental and theoretical papers are acceptable provided they report important findings, novel insights, or useful techniques in these areas. All manuscripts must be prepared in English, and are subject to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by printed in hard copy.

Subject Coverage

- Bioelectrical and neural engineering
- Bioinformatics and Computational Biology
- Biomedical modeling
- Biomedical imaging, image processing and visualization
- Clinical engineering, wearable and real-time health monitoring systems
- Biomechanics and biotransport
- Software, tools and application in medical engineering
- Biomaterials
- Physiological signal processing
- Biomedical devices, sensors, artificial organs and nano technologies
- NMR/CT/ECG technologies and electromagnetic field simulation
- Structure-based drug design

JBiSE

ISSN: 1937-6871
Volume 3 Number 1 January 2010

Journal of Biomedical Science and Engineering

www.scirp.org/journal/jbise

Editor-in-Chief
Kuo-Chen Chou

Editor−in−Chief

Kuo-Chen Chou
Gordon Life Science Institute, San Diego, California, USA

Editorial Board

Notes for Intending Authors

Submitted papers should not have been previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. All papers are refereed through a peer review process. For more details about the submissions, please access the website.

Website and E-Mail

www. scirp. org/journal/jbise
Email: jbise@scirp.org

TABLE OF CONTENTS

Volume 3, Number 1, January 2010

JBiSE

977193768700101 2