THE IIOAB JOURNAL

VOLUME 1 : NO 3 : OCTOBER 2010 : ISSN 0976-3104

Institute of Integrative Omics and Applied Biotechnology Journal

Dear Esteemed Readers, Authors, and Colleagues,

I hope this letter finds you in good health and high spirits. It is my distinct pleasure to address you as the Editor-in-Chief of Integrative Omics and Applied Biotechnology (IIOAB) Journal, a multidisciplinary scientific journal that has always placed a profound emphasis on nurturing the involvement of young scientists and championing the significance of an interdisciplinary approach.

At Integrative Omics and Applied Biotechnology (IIOAB) Journal, we firmly believe in the transformative power of science and innovation, and we recognize that it is the vigor and enthusiasm of young minds that often drive the most groundbreaking discoveries. We actively encourage students, early-career researchers, and scientists to submit their work and engage in meaningful discourse within the pages of our journal. We take pride in providing a platform for these emerging researchers to share their novel ideas and findings with the broader scientific community.

In today's rapidly evolving scientific landscape, it is increasingly evident that the challenges we face require a collaborative and interdisciplinary approach. The most complex problems demand a diverse set of perspectives and expertise. Integrative Omics and Applied Biotechnology (IIOAB) Journal has consistently promoted and celebrated this multidisciplinary ethos. We believe that by crossing traditional disciplinary boundaries, we can unlock new avenues for discovery, innovation, and progress. This philosophy has been at the heart of our journal's mission, and we remain dedicated to publishing research that exemplifies the power of interdisciplinary collaboration.

Our journal continues to serve as a hub for knowledge exchange, providing a platform for researchers from various fields to come together and share their insights, experiences, and research outcomes. The collaborative spirit within our community is truly inspiring, and I am immensely proud of the role that IIOAB journal plays in fostering such partnerships.

As we move forward, I encourage each and every one of you to continue supporting our mission. Whether you are a seasoned researcher, a young scientist embarking on your career, or a reader with a thirst for knowledge, your involvement in our journal is invaluable. By working together and embracing interdisciplinary perspectives, we can address the most pressing challenges facing humanity, from climate change and public health to technological advancements and social issues.

I would like to extend my gratitude to our authors, reviewers, editorial board members, and readers for their unwavering support. Your dedication is what makes IIOAB Journal the thriving scientific community it is today. Together, we will continue to explore the frontiers of knowledge and pioneer new approaches to solving the world's most complex problems.

Thank you for being a part of our journey, and for your commitment to advancing science through the pages of IIOAB Journal.

Yours sincerely,

Vasco Azevedo

Vasco Azevedo, Editor-in-Chief
Integrative Omics and Applied Biotechnology
(IIOAB) Journal

Prof. Vasco Azevedo
Federal University of Minas Gerais
Brazil

Editor-in-Chief

Integrative Omics and Applied Biotechnology (IIOAB) Journal Editorial Board:

The IIOAB Journal

ISSN: 0976-3104

RESEARCH: BIOPHYSICS

PSEUDO-SECOND-ORDER KINETIC MODEL FOR SORPTION OF MALACHITE GREEN ONTO SEA SHELL: COMPARISON OF LINEAR AND NON-LINEAR METHODS

Shamik Chowdhury and Papita Saha*

Biotechnology Department, National Institute of Technology- Durgapur, Mahatma Gandhi Avenue, Durgapur, WB-713209, INDIA

Received on: 09th -June-2010; **Revised on:** 05th -August-2010; **Accepted on:** 23rd -August-2010; **Published on:** 4th-Sept-2010.

***Corresponding author:** Email: papitasaha@gmail.com Tel: +91-9903739855; Fax: +91343-2547375

ABSTRACT

In this study, the sorption of malachite green, a basic dye onto sea shell was studied by performing batch kinetic sorption experiments. The equilibrium kinetic data were analyzed using the pseudo-second-order kinetic model. A comparison between linear least-squares method and non-linear regression method of estimating the kinetic parameters was examined. Four pseudo-second-order kinetic linear equations were discussed. Kinetic parameters obtained from four kinetic linear equations using the linear method differed. Type 1 pseudo-second-order kinetic model very well represented the kinetic uptake of malachite green by sea shell while Type 4 exhibited the worst fit. Present investigation showed that the non-linear method may be a better way to determine the kinetic parameters.

.

Keywords: *malachite green; sea shell; pseudo-second-order, linear method; non-linear method*

[I] INTRODUCTION

Adsorption often referred to as passive uptake and physico-chemical binding of chemical species or ions to a solid surface, is now widely accepted as an efficient and economically feasible process for the removal of synthetic dyes from industrial effluents. The next real challenge in the adsorption field is to identify the adsorption mechanism. Therefore an extensive study of the adsorption kinetics is important since the kinetics describe the uptake rate of adsorbate which in turn helps to predict the adsorption mechanism. Several researchers have used different kinetic models to predict the mechanism involved in the sorption process. These include pseudo-first-order model, pseudo-second-order model, Weber and Morris sorption kinetic model, first-order reversible reaction model, external mass transfer model, first-order equation of Bhattacharya and Venkobachar, Elovich's model and Ritchies's equation [1]. Literature analysis shows that though several kinetic models are available, except the pseudo-second-order model, no other model represents well the experimental kinetic data for the entire sorption period for most of the systems. In recent years, linear regression is frequently used to determine the best-fitting kinetic equation.

An accuracy of a kinetic model is generally a function of the number of independent parameters, while its popularity in relation to the process application is an indicative of its mathematical simplicity. Likewise, linear regression is frequently used to determine the best-fitting kinetic equation primarily owing to its wide usefulness in a variety of adsorption data and partly reflecting the appealing simplicity of its equations. The linear least-squares method with linearly transformed kinetic rate equation has also been widely applied for confirming the experimental data using coefficients of determination. The kinetic equation giving a coefficient of determination closest to unity is considered to be the best fitting. However, during the last few years, a development interest in the utilization of nonlinear optimization modeling has been noted. This is mainly because such transformation of non-linear equations to linear forms implicitly alters their error structure and may also violate the error variance and normality assumptions of standard least squares [2]. As a result, one may obtain different kinetic parameters when using different forms of a kinetic model for a given sorption process. On the contrary, the non-linear method for analyzing the experimental data provides a more complex mathematical method for determining kinetic parameters and is conducted on the same

IIOAB ISSN: 0976-3104

abscissa and ordinate, thus avoiding the drawbacks of linearization.

With the aforementioned, the present study attempts a comparative analysis between the linear least-square and non-linear regression method of the widely used pseudo-second-order to predict the best sorption kinetics and also to obtain the kinetic parameters using the experimental data of malachite green onto sea shell. The present work is aimed at evaluating the accuracy and consistency in parameter prediction by the linear and non-linear method, and to familiarize the knowledge deficiencies regarding non-linearized adsorption kinetics.

[II] MATERIALS AND METHODS

2.1. Adsorbent

Bivalve type sea shells that most commonly wash up on large sandy beaches were used in this study. The sea shells were collected from the sea beaches of Puri, Orissa, India. It was pretreated before use by washing thoroughly with double distilled water and dried at 110 ± 1 ^{0}C for 24 h in an oven drier. The raw biosorbent was crushed and ground using ball mill and sieved to give a fraction of 80 mesh. The biosorbent was again washed thoroughly with distilled water and dried naturally. The resulted shell particles were stored in sterile, sealed glass containers and used in all the adsorption experiments.

2.2. Adsorbate

Malachite Green used in this study was of commercial quality (CI 42000, FW: 365, MF: $C_{23}H_{25}N_2Cl$) and was used without further purification. Stock solution (500 mg L^{-1}) was prepared by dissolving accurately weighed quantity of the dye in double-distilled water. All working solutions were prepared by diluting the stock solution with suitable volume of double-distilled water.

2.3. Adsorption experiments

Adsorption kinetics experiments were carried out using the batch method for different initial dye concentrations. The experiments were carried out in 250 mL Erlenmeyer flasks containing a fixed amount of adsorbent with 100 mL dye solution. Adsorption kinetics was conducted using the batch method. The initial pH of the solution was adjusted with 0.1 N HCl or NaOH solutions by using a pH meter. The flasks were agitated and incubated in an incubator shaker (Model Innova 42, New Brunswick Scientific, Canada) at 30 ± 1 ^{0}C until reaching equilibrium. Samples were taken from the flasks at regular time intervals for analyzing the concentration of malachite green in the solution. The residual amount of malachite green in each flask was investigated using UV/VIS spectrophotometer (Model Hitachi – 2800) at λ_{max} of 663 nm.

2.3. Pseudo-second-order kinetic model

The pseudo-second order kinetic equation was proposed by Blanchard et al. [3] and is expressed as:

$$q_t = \frac{k_2 q_e^2 t}{1 + k_2 q_e t} \qquad \text{---------- (1)}$$

where q_t and q_e are the amount of dye adsorbed at time t and at equilibrium (mg.g^{-1}) and k_2 (g. mg^{-1} min^{-1}) is the pseudo-second-order rate constant for the adsorption process. Eq. (1) can be linearized to at least four different forms [4]. The different linearized forms of the pseudo-second-order equation are given in **Table-1**. The most popular form used is Type 1 [1].

Table 1: Different linearized forms of the pseudo-second-order equation

Linear Regression	Expression	Plot
Type 1	$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \frac{1}{q_e} t$	t / q_t vs. t
Type 2	$\frac{1}{q_t} = \frac{1}{q_e} + \frac{1}{k_2 q_e^2} \frac{1}{t}$	$1 / q_t$ vs. $1 /$
Type 3	$q_t = q_e - \frac{1}{k_2 q_e} \frac{q_t}{t}$	q_t vs. q_t / t
Type 4	$\frac{q_t}{t} = k_2 q_e^2 - k_2 q_e q_t$	q_t / t vs. q_t

[III] RESULTS

In the present study, the coefficient of determination (r^2) was used to determine the best fit equation:

$$r^2 = \frac{(q_{e,meas} - \overline{q_{e,cal}})^2}{\sum (q_{e,meas} - \overline{q_{e,cal}})^2 + (q_{e,meas} - q_{e,cal})^2} \qquad \text{---- (2)}$$

where $q_{e,meas}$ and $q_{e,cal}$ (mg.g^{-1}) are the measured and calculated adsorbate concentration at equilibrium, and $\overline{q_{e,cal}}$ (mg.g^{-1}) is the average of $q_{e,cal}$ (mg.g^{-1})

Linear regression is frequently used to determine the best fitting kinetic model, and the method of least squares is used for finding the parameters of the kinetic models. The pseudo second-order kinetic constant, q_e and k_2 by a Type 1 pseudo second-order expression were calculated from the plot of t/q_t versus t as shown in **Figure-1**. Similarly the pseudo-second-order kinetic constant, q_e and k_2 were obtained from the plot of $1/ q_{,t}$ and $1/t$, q_t and q_t/t, q_t/t and q_t for a Type 2, Type 3, and Type 4 pseudo second-order expressions respectively. The calculated kinetic constants and their corresponding coefficient of determination (r^2) are given in **Table-2.** The experimental q_e value is also shown in **Table-2**. **Table-2** shows that q_e and k_2 values obtained from the four linear forms of pseudo-second-order expressions were different at all initial dye concentration studied. From **Table-2**, it was observed that except Type 1 pseudo-second-order expression, no other model provided a better fit to the experimental kinetic data. The very low r^2 values for Type 2-4 pseudo-

ISSN: 0976-3104

second-order expressions suggest that it was not appropriate to use these models to represent the experimental data of malachite green onto sea shell. Therefore, by linear method, a theoretical pseudo-second-order model was found to aptly represent the experimental data based on Type 1 pseudo-second-order kinetic expression. In addition, the Type 1 pseudo second order expression predicts reasonably the *q*e values theoretically for all the range of initial dye concentrations studied **Table-1**.

In the case of the non-linear method, the software Origin version 9.0 was used for determining the pseudo-second-order kinetic parameters. Pseudo-second-order kinetic parameters obtained by non-linear method are enlisted in **Table-2**. **Figure-2** shows experimental data and the predicted pseudo-second-order kinetics using the non-linear method. Very high r^2 value suggests that the non-linear pseudo-second-order kinetic model could be used to represent the kinetic uptake of malachite green onto sea shell at all initial dye concentrations studied. By using non-linear method there were no problems with transformation of non-linear pseudo-second-order equation to linear form, and also they were in the same error structures. It is thus logical to use the non-linear method to represent a kinetic model efficiently and effectively.

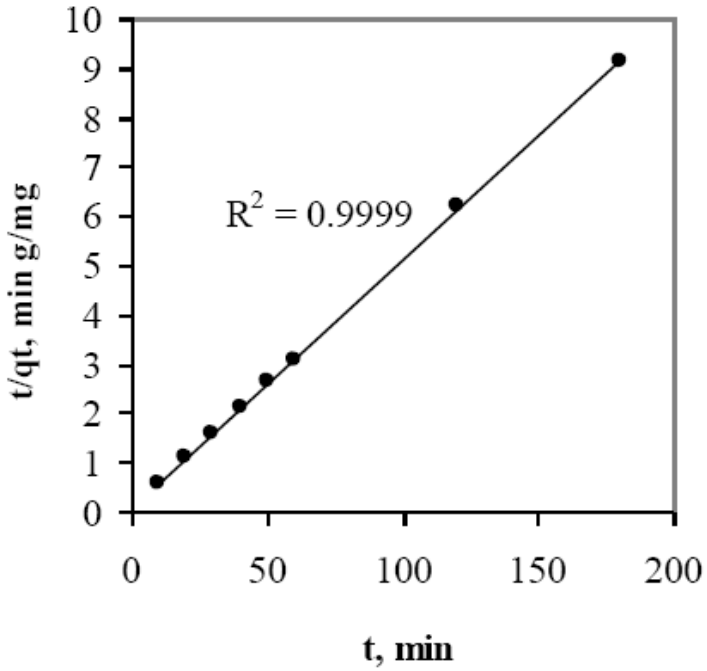

Fig: 1. Type 1 pseudo-second-order kinetics obtained by using the linear method for the sorption of methylene blue onto sea shell (C_0=50 mg L^{-1})

[IV] DISCUSSION

In most adsorption studies, the linear method has been widely used in assaying the quality of fit of a kinetic model to an experimental data primarily due to its simplicity and usefulness. However, the different outcomes obtained by linear regression for the same kinetic model show the real complexities and problems in estimating the kinetic parameters by linearization technique.

The different outcomes for different linearized form of pseudo-second-order models are due to the error alterations while transforming the data that represents a non-linear kinetics to a linearized form. The transformation of a non-linear model to a linear one primarily distorts the normality assumptions of the linear least square method. In addition, a different axis setting alters the regression results, thereby influencing the accuracy as well as consistency, leading to the violation of theories behind the kinetic models. Moreover, the linear method does not test the linearity of the data set. Instead, it assumes that the given data set were linear and gives a straight line that predicts the goodness of fit of the equilibrium experimental data. Furthermore, the linear method is based on the assumption that the scatter vertical points around the line follows a Gaussian distribution, and the error distribution is uniform at every value of X-axis. This is rarely true or practically impossible with kinetics as most of the adsorption kinetic models are non-linear due to different mechanisms.

.
.

IIOAB ISSN: 0976-3104

Table 2: Pseudo-second-order kinetic parameters obtained by using the linear and non-linear methods

Kinetic Model	Parameters	C_0 (mg L^{-1})			
		50	**100**	**150**	**200**
	$q_{e,exp}$ (mg g^{-1})	46.27	82.55	135.89	152.56
Linear-Type 1	q_e (mg g^{-1})	45.67	83.78	137.32	150.98
	k_2(g $mg^{-1}min^{-1}$)	0.0448	0.0540	0.0610	0.0689
	r^2	0.998	0.999	0.990	0.991
Linear- Type 2	q_e (mg g^{-1})	45.14	85.13	132.52	154.39
	k_2(g $mg^{-1}min^{-1}$)	0.0384	0.0516	0.0602	0.0632
	r^2	0.975	0.972	0.969	0.981
Linear-Type 3	q_e (mg g^{-1})	45.02	79.20	130.21	150.47
	k_2(g $mg^{-1}min^{-1}$)	0.0324	0.0578	0.0592	0.0667
	r^2	0.954	0.963	0.923	0.956
Linear-Type 4	q_e (mg g^{-1})	45.36	87.53	129.38	156.62
	k_2(g $mg^{-1}min^{-1}$)	0.0352	0.0521	0.0621	0.0732
	r^2	0.863	0.838	0.762	0.803
Non-linear	q_e (mg g^{-1})	46.17	82.91	135.42	152.87
	k_2(g $mg^{-1}min^{-1}$)	0.0578	0.0621	0.0732	0.0771
	r^2	1.000	0.999	1.000	1.000

The linear method considers error distribution only along the Y-axis irrespective of the corresponding X-axis resulting in the different determined parameters for the four different types of linearized pseudo-second-order kinetic model for the same experimental data. Therefore the linear method is inappropriate in predicting the best-fit kinetics for a particular experimental data set and unable for providing a fundamental understanding of the kinetics of the adsorption systems, resulting in an improper conclusion. On the contrary, the drawbacks of linear method can be avoided by adopting the non-linear method for analyzing the experimental data. This is because in the non-linear method, the experimental equilibrium data and the isotherms are in a fixed x and y axis i.e, the non-linear analysis is conducted on the same abscissa and ordinate resulting in the same error distribution and structure.

The present investigation confirms the non-linear method as an appropriate technique to predict the optimum sorption kinetics. Ho (2006) conducted a similar evaluation using linear and non-linear methods to determine the pseudo-second-order kinetic parameters [5]. He chose cadmium as the adsorbate and tree fern as the adsorbent. The kinetic parameters acquired from the four kinetic linear equations using linear method had discrepancies among themselves. Further, for linear method, the best fit was obtained by using the Type 1 expression because the highest coefficient of determination was calculated from the fitted equation. In contrast to the linear method, kinetic parameters obtained from the four kinetic

ISSN: 0976-3104

linear equations were the same when using the non-linear method. Under such conditions, it would be more rational and reliable to interpret adsorption data through a process of nonlinear regression.

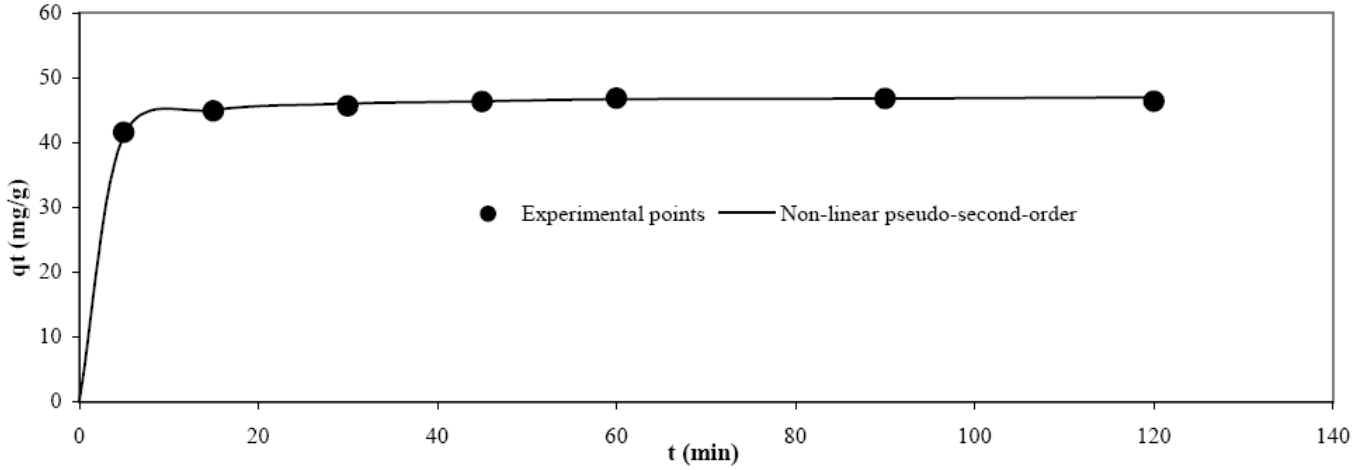

Fig: 2. Pseudo-second-order kinetics obtained by using the non-linear method for the sorption of methylene blue onto shell (C_0=50 mg L^{-1})

[V] CONCLUSION

A comparative analysis between the linear and non-linear method in determining the pseudo-second-order kinetic parameters for sorption of methylene blue onto sea shell was conducted. Present study corroborated that it is not appropriate to use the linear method in determining kinetic parameters of a particular kinetic model. This is mainly because transforming a non-linear kinetic model to a linearized form tends to alter the error distribution, and thus distort the parameters. Non-linear analysis conducted on the same abscissa and ordinate results in the same error distribution and is therefore a better way to obtain the kinetic parameters than linear method.

REFERENCES

[1] Febrianto J, Kosasiha A.N, Sunarsob J, Jua Y, N. Indraswati, Ismadji S. [2009] Equilibrium and kinetic studies in adsorption of heavy metals using biosorbent: A summary of recent studies. *J Hazard Mater* 162:616–645.

[2] Kumar K.V, Sivanesan S. [2006] Isotherm parameters for basic dyes onto activated carbon: Comparison of linear and non-linear method.*J Hazard Mater* B129:147–150

[3] Blanchard G, Maunaye M, Martin G. [1984] Removal of heavy metals from waters by means of natural zeolites. *Water Res* 18:1501–1507.

[4] Lin J, Wang L. [2009] Comparison between linear and non-linear forms of pseudo-first-order and pseudo-second-order adsorption kinetic models for the removal of methylene blue by activated carbon. *Environ Sci Engin.*3(3):320–324.

[5] Ho YS. [2006] Second-order kinetic model for the sorption of cadmium onto tree fern: A comparison of linear and non-linear methods. *Water Res* 40(1):119–125.

ABOUT AUTHORS

Shamik Chowdhury *is M.Tech, 2nd year student of Biotechnology Department, National Institute of Technology, Durgapur, India. Email:chowdhuryshamik@gmail.com*

Papita Saha *is Assistant Professor at Biotechnology Department, National Institute of Technology, Durgapur, India. Email: papitasaha@gmail.com*

The IIOAB Journal

ISSN: 0976-3104

RESEARCH: BIOINFORMATICS

IN SILICO ANALYSIS OF HEMAGGLUTININ, NEURAMINIDASE, AND MATRIX2 OF H5N1 VIRUS INDONESIAN STRAIN RELATED TO ITS HIGH PATHOGENICITY

Usman Sumo Friend Tambunan*, Agus Limanto, and Arli Aditya Parikesit

Department of Chemistry, Faculty of Mathematics and Science, University of Indonesia, Depok 16424, INDONESIA

Received on: 27th-June-2010; **Revised on:** 7th-Sept-2010; **Accepted on:** 15th-Sept-2010; **Published on:** 1st-Oct-2010.

***Corresponding author:** Email: usman@ui.ac.id Tel: +6221 727 00 27; Fax: +6221 786 34 32

ABSTRACT

In the year of 2007, avian influenza outbreak which occurred in Indonesia caused mortality of almost 85% from detected avian influenza cases. Comparing the mortality rate in Indonesia to other countries with avian influenza outbreak, WHO announced that HPAI H5N1 Indonesia has the highest pathogenicity. Mutations with either antigenic shift or antigenic drift can influence the pathogenicity of influenza virus. Studies on hemagglutinin (HA), neuraminidase (NA), and matrix2 (M2) have been carried out because these three proteins have important roles in the infection process of avian influenza virus. In silico analysis was done by multiple alignment and phylogenetic tree construction. Hemagglutinin mutation was observed at the cleavage site and at the active site, while neuraminidase mutation and matrix2 mutation was observed at the active site. The amino acid character shift from hydrophilic to hydrophobic influenced the virus pathogenicity. The mutation analysis result was utilized for hemagglutinin cleavage by pro-P prediction, 3-D structure prediction, molecular docking simulation, and molecular dynamics simulation. Based on mutation analysis on hemagglutinin cleavage site, a R-X-K/R-R pattern was obtained for H5N1 Indonesia and H5N1 HongKong. Pro-P prediction results showed that the pattern which causes hemagglutinin HPAI H5N1 could be easily cut by Furin. 3-D structure analysis using molecular docking and molecular dynamics also showed that hemagglutinin and neuraminidase H5N1 Indonesia bind better with human sialic acid receptor. Meanwhile H5N1 virus's matrix2 protein gave resistance to amantadine and rimantadine. Results from the analysis revealed a relation between hemagglutinin, neuraminidase, and matrix2 mutation with the pathogenicity of H5N1 in Indonesia.

.

Keywords: *H5N1; hemagglutinin; neuraminidase; matrix2; molecular docking*

[I] INTRODUCTION

Influenza virus is contagious for humans and a number of animals with specific contagiousness towards certain species. That means, if the virus infects one species, it would rarely infect another species. The general symptoms of this disease are fever, headache, throat-ache, and cough. Some influenza cases further caused pneumonia resulting in a number of deaths [1].

Influenza virus is a part of Mononegavirales order, Orthomyxoviridae family, which has single segmented genome. Based on its genus, there are three types of Influenza virus. They are type A, B, and C. Influenza virus A and B have 8 RNA segments, while Influenza virus C has 7 RNA segments. The nucleic acid of influenza virus is translated into approximately 10 proteins, namely hemagglutinin (HA), neuraminidase (NA), matrix protein (M1 and M2), non structural protein (NS1 and NS2), nucleocapsid protein (NP), polymerase basic (PB1 and PB2), and polymerase acidic (PA) [2].

Influenza A virus is classified based on its hemagglutinin and neuraminidase antigens, which are located on the viral coats. Until today, scientists have found 16 types of HA and 9 types of NA. Influenza virus A H5N1, which is widely known as avian influenza, is one of the influenza A subtype that could cause infection of poultry. However, over the course of time, it could infect humans as well. Only four strains of avian

IIOAB ISSN: 0976-3104

influenza A could cause infection in humans. They are H5N1, H7N7, H7N3, and H9N2 [2].

Avian influenza A has two types of pathogenicity: Highly Pathogenic Avian Influenza (HPAI) and Low Pathogenic Avian Influenza (LPAI). Pathogenicity means the ability of a virus to cause disease. HPAI H5N1 is called 'Asian' H5N1, which attracted worldwide attention, while LPAI H5N1 is called 'North American' H5N1 [3].

During the 20th century, Influenza A virus became a frightening pandemic disease. Three occurrences of influenza pandemic have caused mortality for millions. First pandemic (Spanish Flu) in 1918-1919 was caused by H1N1 subtype and caused 50 million deaths. Second pandemic (Asian Flu) in 1957-1958 was caused by H2N2 subtype and caused 1 million deaths. Third pandemic (Hong Kong Flu) in 1967-1968 was caused by H3N2 subtype and caused 1 million deaths as well [1].
The HPAI H5N1 was isolated from a swan ranch in China in 1996. Moreover, HPAI H5N1 has occurred in poultry market in Hong Kong. Besides that, H5N1 had caused 6 deaths out of 18 infected patients [2].

Since 1997, the HPAI H5N1 virus has caused massive mortality on poultry and human. The Asian pandemic area of this virus comprises of Japan (north) and Indonesia (south). Until now, research has proved that H5N1 infection to human occurred because of direct contact between human and infected poultry. Although a few possible human-to-human transmissions of H5N1 influenza have been reported, there is still no evidence of efficient person-to-person spread [4, 5].

In the year of 2004, H5N1 virus reached certain proportion as an Asian pandemic. There are HPAI epidemics in China, Japan, South Korea, Thailand, Vietnam, Indonesia, Cambodia, and Laos. The H5N1 epidemic in Indonesia occurred in 2005 for the first time. The epidemiological data have shown that Avian Influenza A cases have resulted in 141 human infections in the period 2005-2009. Among them, 115 lead to certain deaths. Henceforth, WHO has declared that H5N1 virus Indonesian strain is the most pathogenic avian influenza A virus [6].

The changing infection specificity of H5N1 from poultry to human was caused by single amino acid substitution on position of 226 and 228 at hemagglutinin [5]. It could change the poultry receptor binding site from its specific position of α-2,3 linked sialic acid for poultry, into specific position on human of α-2,6 linked sialic acid [4]. This change was caused by the ability of influenza A virus to mutate by antigenic drift and shift means. This feature made the virus more pathogenic and increased its ability to infect human effectively [7].

Intensive research on this disease has been done, especially for identification, diagnostic development, and prevention. The results from in silico study on H5N1 virus from Banten Province clearly show that there are amino acid substitutions and modification of secondary structure. This has been determined based on several type of H5N1 virus comparison [8].

This research was conducted for observing whether the mutation on HA, NA, and M2 are related on its high pathogenicity on H5N1 virus in Banten, Indonesia, in the year of 2007.The HA glycoprotein forms spikes at the surface of virions, mediating attachment to host cell sialoside receptors and subsequent entry by membrane fusion and the cleavage of HA is required for viral infectivity and is a critical determinant of viral pathogenicity [1]. The NA forms knoblike structures on the surface of virus particles and catalyzes their release from infected cells, allowing virus spread [1]. The M2 is a transmembrane protein that forms an ion channel required for the uncoating process that precedes viral gene expression [1]. These three have important roles in H5N1 infection process and this process develops the pathogenicity of a virus.

The objective of this research is to conduct in silico analysis of HA, NA, and M2 mutation on H5N1 virus in Indonesia, which has certain influence on its high pathogenicity towards human. The general steps are the construction of phylogenetic trees, HA cut-out prediction by furin, the search for 3D structure, molecular docking, and molecular dynamics.

[II] MATERIALS AND METHODS

The following steps were conducted using Microsoft Windows XP based PC.

2.1. Search and choose the sequences

The hemagglutinin, neuraminidase, and matrix2 from H5N1 subtype were downloaded from the Influenza Virus Resource database of the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov.html). The 37 H5N1 virus genomes that have HA, NA, and M2 full length sequences were used in this study.

2.2. Multiple alignments

This step was conducted by using ClustalW online program (www.ebi.ac.uk/Tools/clustralw2/index.html). The alignment result was interpreted to pinpoint the position of different amino acid on the region of the receptor binding domain between H5N1 Indonesian virus and H5N1 virus from other countries. The alignment data will be utilized for mutation analysis of H5N1 Indonesian virus.

2.3. Construction of Phylogenetic tree

The construction was meant to find the sequences with high homology with HPAI. The process was conducted by using CLC Main Workbench 5.0 software, with sequence alignment as its input.

IIOAB ISSN: 0976-3104

2.4. HA0 cleave cutting prediction by pro-protein (Furin)

One of the important factors for pathogenicity is the HA cleaving pattern by furin intracellular protein. The prediction was done by using pro-protein (furin) online server (http://cbs.dtu.dk/services/ProP).

2.5. Building 3D model

The SWISS MODEL service was utilized to build 3D models for H5N1 virus protein, by finding the exact templates. Chosen 3D structures were the ones with a sequence homology near 100% compared with the virus sequences. The PDB files were exposed by using Molecular Operating Environment (MOE 2008.10).

2.6. Molecular Docking

The found 3D structure was docked with its ligand by using molecular docking software. Before the docking, preparation steps must be done. This was done by removing water molecules, addition of hydrogen atoms and charges. Further minimization was done using MOE 2008.10(MMFF94x). The utilized parameters for analyzing the complex between protein and ligand are ΔGo binding and inhibition constant. They are as follows:

The simulation on protein/ligand complex was done after molecular docking steps with MOE 2008.10. Before molecular dynamics was computed preparation steps were done for molecular docking, followed by inserting the ligand in order to form the protein-ligand complex. Then, the complex was minimized by force field MMFF94x and solvable in Born form. The parameters were utilized in accordance with MOE default, which is ensemble NVT (N, total atom; V, Volume; T, Temperature) by using the NPA algorithm.

[III] RESULTS

This research was only focusing on comparing H5N1 Indonesian virus strain (A/Indonesia/CDC1047/2007) as reported to NCBI, with the H5N1 Hong Kong virus strain (A/Hong Kong/482/97). The reason is to determine the molecular difference between the two. The alignment was done twice. First, the Indonesian virus sequence toward the other 36 viruses. Second, the Indonesian virus sequence toward the (A/Hong Kong/482/97) virus.

Table: 1. HA Docking data with alpha-2,6 sialic acid

HA (PDB ID)	ΔG (kkal/mol)	pKi	Ki (μM)
Indonesia (2IBX)	-9,22	6,72	0,19
Hong Kong (2FK0)	-8,4613	6,16	0,69

Table: 2. NA Docking data with alpha-2,6 sialic acid

NA (PDB ID)	ΔG (kkal/mol)	pKi	Ki (μM)
Indonesia (2HTY)	-9,1216	6,64	0,23
Hong Kong (2HU0)	-7,8524	5,72	1,89

Table: 3. M2 Docking data with amantadine

M2 (PDB ID)	ΔG (kkal/mol)	pKi	Ki (mM)
Indonesia (2RLF)	-2,8202	2,05	8,79
Hong Kong (2KIH)	-5,4129	3,94	0,11

Table: 4. M2 Docking data with Rimantadine

M2 (PDB ID)	ΔG (kkal/mol)	pKi	Ki (mM)
Indonesia (2RLF)	-4,6186	3,36	0,42
Hong Kong (2KIH)	-5,1680	3,76	0,17

The protein sequences of HA, NA, and M2 from the influenza virus were the sequences of H5N1 subtype. H5N1 was chosen because this subtype caused avian influenza epidemic in various countries, Indonesia was included from epidemic time range of 2003-2007 and Indonesian H5N1 had the highest pathogenicity. The search was conducted by accessing the National Center of Biotechnology Information website (http://ncbi.nlm.nih.gov/genomes/FLU/Database.html). It has influenza database advanced data search in accordance with the desired influenza virus specification. They are the influenza type, viral host, the country of origin, type of protein sequences, and influenza virus subtype. This research used Influenza A virus with HPAI H5N1 subtype from human host and H5N1 LPAI from avian host. HPAI H5N1 subtype sequences were selected from Hong Kong, China, Vietnam, Thailand, Laos, and Indonesia, whereas LPAI H5N1 subtype sequences were selected from United States and Canada. The downloaded protein sequences were HA, NA, and M2 full length.

The downloaded HA, NA, and M2 FASTA sequences were loaded into ClustalW2 program for Multiple Sequence alignment process (http://www.ebi.ac.uk/Tools/clustalw2/index.html). ClustalW2 was used to align the amino acid sequences, in order to analyze the occurred mutation on HA, NA, and M2 HPAI H5N1 Indonesian virus on the last reported case to WHO (A/Indonesia/CDC1047/2007) with HPAI H5N1(A/Hong Kong/482/97) which has been crystallized and deposited at the Protein Data Bank (PDB). The conducted mutation analysis between A/Indonesia/CDC1047/2007 with A/Hong Kong/482/97 was limited only to the amino acids with important role in active site binding, because this is the determinant factor of the viral intrusion into the host cell. The role of the active site on HA and NA is to bind the Sia (α-2,6) Gal, while on M2 it's acting as ion proton channel and the inhibition target of amantadine and rimantadine. The receptor binding site of HA are in position 190, 225, 226, and 228 **[Table-5]**; on NA are in position 118, 152, 276, 292, and 371 **[Table-6]**; and on M2 are in position 27, 30, 31, 34, 37, and 41 **[Table-7]**. The positions 226 and 228 are specifically affected for HA [9].

ISSN: 0976-3104

Table: 5. HA sequence alignment analysis result

HA Sequence	Positions			
	190	225	226	228
donesia/CDC1047/2007 (2IBX)	V	L	V	K
/Hong Kong/482/97 (2FK0)	E	G	Q	G

Researchers reported that changes from hydrophilic to hydrophobic on HA and NA results in an increase stability in binding to sialic acid. This caused the increasing level of pathogenicity in H5N1 Indonesian virus. Besides making alignments toward H5N1 Indonesian virus and H5N1 PDB, the alignment was done against all the chosen sequences. The alignment results were used for building phylogenetic trees **[Figures 1, 2, and 3]**. It would be useful to determine the homology relation between Indonesian HPAI H5N1 and HPAI H5N1 from other countries [10].

Table: 6. NA sequence alignment analysis result

NA Sequence	Position				
	118	152	276	292	371
A/Indonesia/CDC1047/2007 (2HY)	A	S	W	Y	G
A/Hong Kong/482/97 (2HU0)	R	R	E	R	R

Table: 7. M2 sequence alignment analysis result

M2 Sequence	Position					
	27	30	31	34	37	41
A/Indonesia/CDC1047/2007 (2RLF)	A	A	N	G	H	W
A/Hong Kong/482/97 (2KIH)	V	A	N	G	H	W

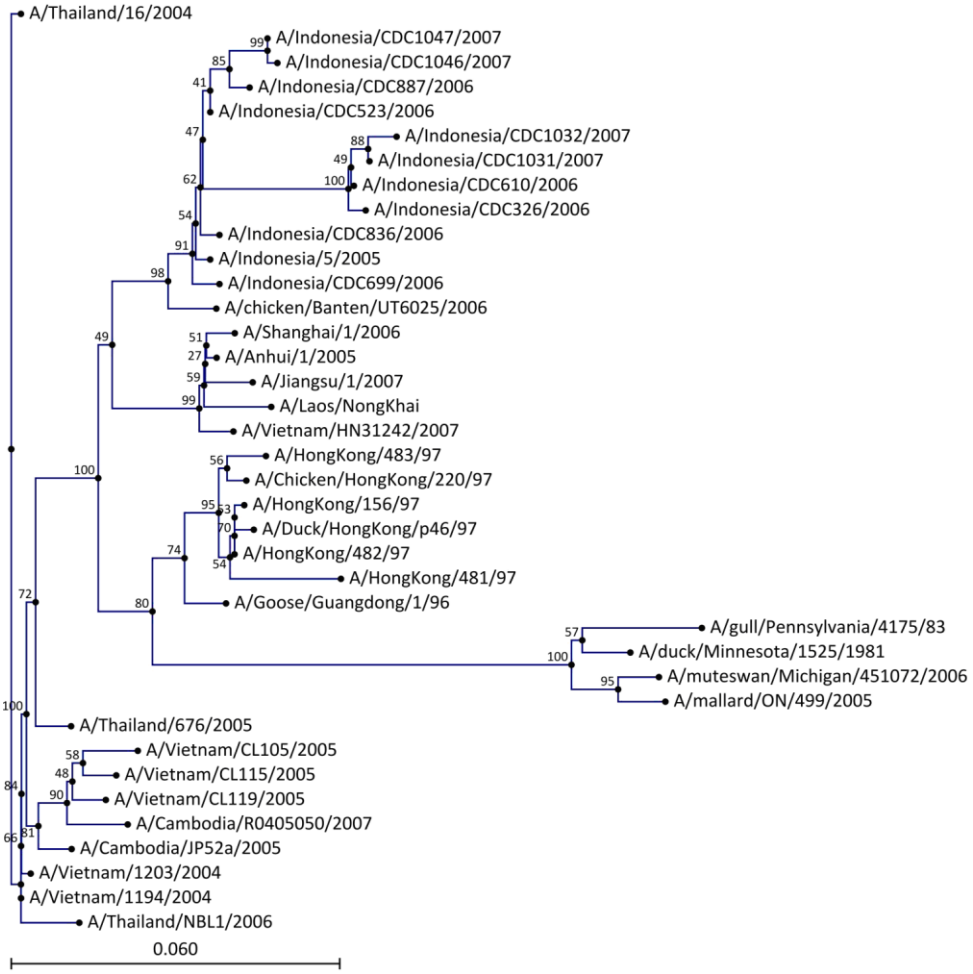

Fig: 1. Phylogenetic tree of HA H5N1

IIOAB ISSN: 0976-3104

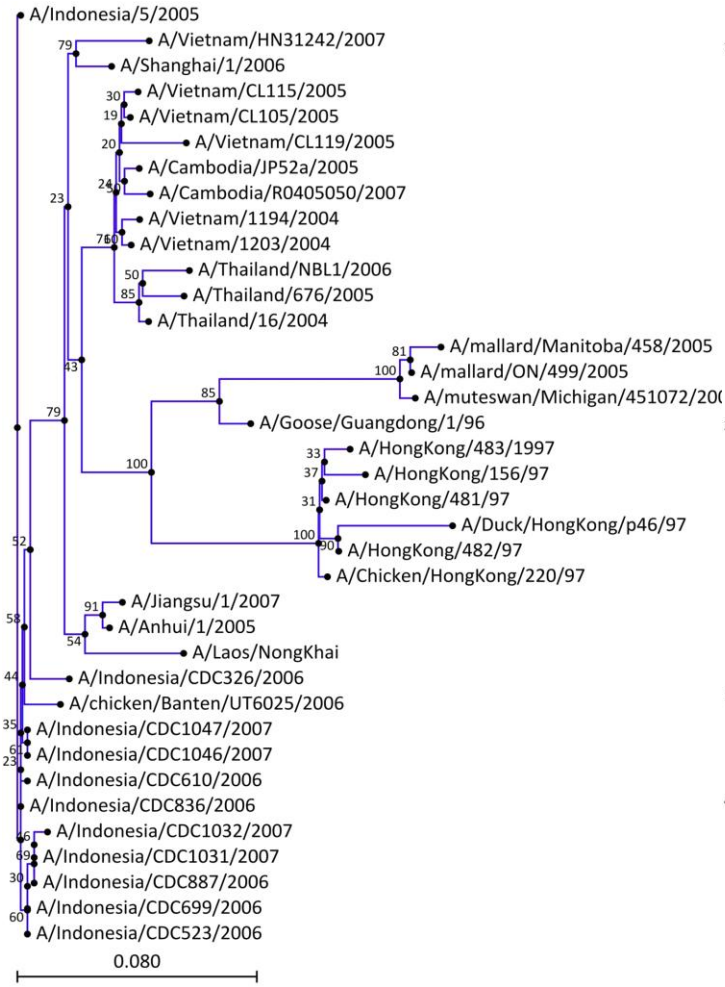

Fig: 2. Phylogenetic tree of NA H5N1

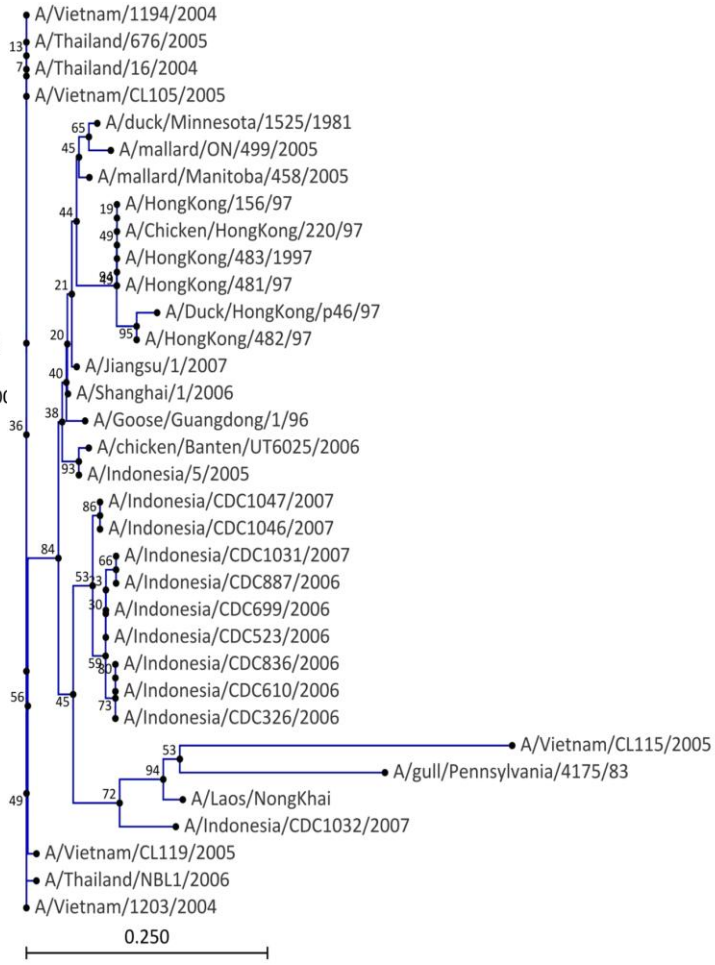

Fig: 3. Phylogenetic tree of M2 H5N1

The HA protein is useful for the binding process toward host cell sialic acid receptor. After proteolitic process activation of HA precursor into HA1 and HA2, the virus starts to fuse with host cell. Previous research had shown that the HA cleaving robustness is an important factor of avian influenza virus virulence. LPAI has one arginine residue at the cleaving area. Henceforth, it could only be cleaved by extra-cellular protease like trypsin. Therefore, viral infection only occurred at the host cell, while HPAI having poly-basic amino acids could be cleaved by intracellular protease like furin. This allowed viral infection to spread systemically to the whole host cell tissue [11].

The computational results from online server furin pro-protein clearly show that A/Indonesia/CDC1047/2007 virus has R-X-K/R-R pattern, which is similar to A/Hong Kong/482/97 **[Table-8]**. This result clearly grouped Indonesian strain along with Highly Pathogenic Virus.

The 3D structure search was done by RSCB server (http://www.rcsb.org/pdb/static.do?p=search/index.html). The method of this research utilized homology modeling on HA, NA, and M2 matrix of H5N1 Indonesian virus by using SWISS-MODEL workspace server. The process was conducted by giving sequence file input for browsing the 3D structure in FASTA format. The chosen mode is automated.

Table: 8. HA cleavage result by Furin

Sample of H5N1	Furin-type Cleavage Site Prediction
A/Indonesia/CDC1047/2007 (ABM90533)	ESRRKKR|GL
A/Hong Kong/482/97 (AAC32100)	ERRRKKR|GL
A/Shanghai/1/2006 (BAH10637)	RERRRKR|GL
A/mallard/ON/499/2005 (ABQ43787)	none

The downloaded 3D structure was utilized to a molecular docking simulation toward Sia(α-2,6)Gal. Molecular docking simulation was conducted by using MOE-dock 2008.10 software. Ligand structure was drawn by using MOE2008.10 builder feature. Before starting the molecular docking, ligand was drawn and prepared. Ligand preparation was conducted

IIOAB ISSN: 0976-3104

by using wash function and gas phase MMFF94x (i.e. no salvation was considered). The default parameters of MOE were utilized. One of the MOE features utilized was protonate-3D. It is a feature to solve the macromolecular protonation state assignment problem by selecting a protonation state for each chemical group that minimizes the total free energy of the system. The important parameter for protonate-3D is the repair of partial charge, which means substitution of the solvation mode while computing the force field of the molecule. The repair of partial charge and hydrogen atom was done to have the optimum state of the ligand, by using default parameter as well. The optimum state of the ligand was reached when the optimum minimization energy of protein-ligand conformation was attained. After the preparation, the same steps were done for the 3D structures of HA, NA, and M2. These steps were necessary to secure the optimum state of HA, NA, and M2 protein. The start configuration of the HA, NA, and M2 was initiated after the protonate-3D procedure [13].

The molecular dynamic simulations on HA, NA, and M2 of H5N1 Indonesian virus from its crystal structure on PDB was done to validate the interaction between protein and its ligand. The dynamic simulation was conducted on initialization step. The protein-ligand complex was optimized with partial charges and minimized with force field MMFF94x. However, this solvation was utilizing Esol calculation on the system. This process was conducted by using solvent.

The utilized statistics for conformation simulation was computed on the ensemble of structures. It was using MOE default, which is ensemble NVT with constant temperature of 300K and 101kPa pressure, while using NPA algorithm for adjusting the whole parameters. The position, velocity, and acceleration results were saved every 0.5 pico second [13].

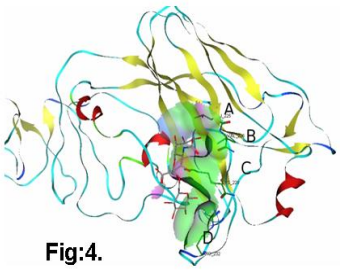

Fig:4.

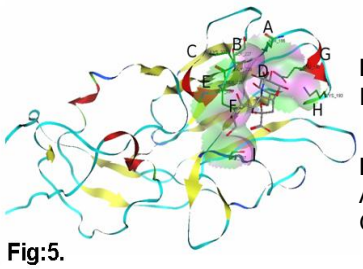

Fig:5.

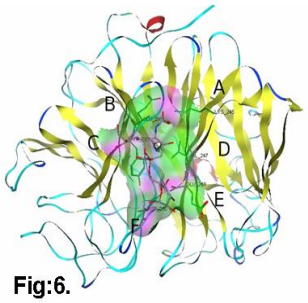

Fig:6.

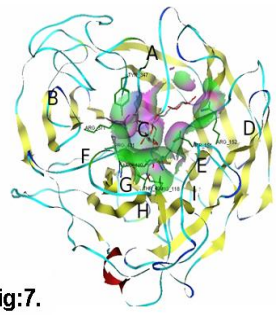

Fig:7.

[IV] DISCUSSION

The alignment of HA H5N1 Indonesian virus toward HA of PDB (2FK0) shows, that mutation is close to H5N1 Indonesian virus receptor binding site. It was observed that in position 190 and 226 a changing amino acid property, from hydrophilic to hydrophobic, occurred. However, on position 225 and 228, the changed amino acid still retains its hydrophobicity. The alignment of NA H5N1 Indonesian virus toward NA of PDB (2HU0) clearly shows that mutation is imminent on receptor binding site. It was observed, that in position 118, 276, 292, and 371 occurs a changing property of amino acid, from hydrophilic to hydrophobic. However, the mutation on position 152 still retains its hydrophilicity. The alignment of M2 H5N1 Indonesian virus with M2 of PDB (2KIH) clearly shows mutation of H5N1 on active site position. It is occurred on active site position 27. However, the mutation still retains its hydrophilicity property.

The phylogenetic tree indicated that Indonesian HPAI H5N1 is closer with a HPAI H5N1 branch from other countries. It could be inferred, that they have a close homology relation. The close relationship is shown in HA, NA, and M2 sequences **[Figures-1 to 3]**.

Information for **[Table 1-4]**: It is a parameter from MOE. The complete formula which is necessary for comprehending the table is as follows:

$$\Delta G = -RT \ln Kp \qquad (1)$$

$$Kp = \frac{1}{Ki} \qquad (2)$$

Fig: 4. Docking visualization of 2ibx with sialic acid: A is Ile_225, B is Val_226, C is Lys_228, D is Arg_232

Fig: 5. Docking visualization of 2fk0 with sialic acid: A is Asn_186, B is Ser_227, C is Lys_232, D is Gly_228, E is Gln_226, F is Gly_225, H is Lys_193, I is Ser_137

Fig: 6. Docking visualization of 2hty with sialic acid: A is Lys_245, B is Glu_291, C is Tyr_292, D is Val_247, E is Glu_248, F is Asp_250

Fig: 7. Docking visualization of 2hu0 with sialic acid: A is Tyr_347, B is Arg_371, C is Pro_431, D is Arg_152, E is Asp_151, F is Pro_431, G is Arg_430, H is Thr_439, I is Arg_118

IIOAB ISSN: 0976-3104

HA docking result with α-2,6 sialic acid **[Figure-4]**, **[Figure-5]** and respective deltaG data **[Table-1]** clearly show that HA H5N1 Indonesian virus has higher binding affinity towards sialic acid.

HA docking results with alpha-2,6 sialic acid **[Figure 6]**, **[Figure 7]** and respective deltaG data **[Table 2]** clearly show that NA H5N1 Indonesian virus has higher binding affinity towards sialic acid.

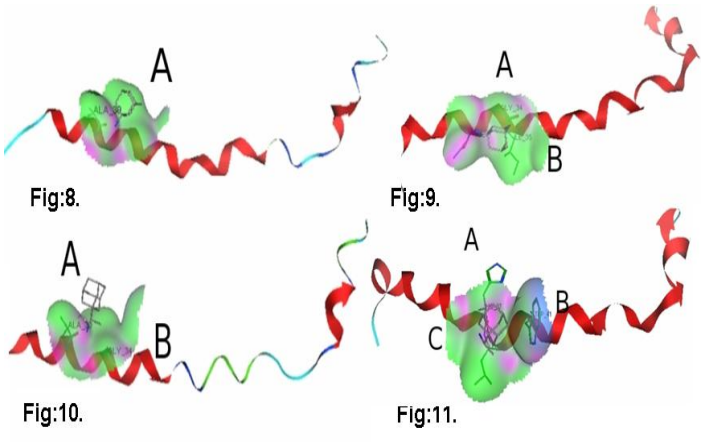

Fig: 8. Visual Simulation Docking of 2RLF toward amantadine: A is Ala_30

Fig: 9. Visual Simulation Docking of 2KIH toward amantadine: A is Gly_34, B is Ile_35

Fig: 10. Visual Docking Simulation of 2RLF toward rimantadine: A is Ala_30, B is Gly_34

Fig: 11. Visual Docking Simulation of 2KIH toward rimantadine: A is His_37, B is Trp_41, C is Leu_38

The visual data from the docking simulation show that HA and NA from H5N1 Indonesian virus binds better than crystallized H5N1 virus from PDB. This clearly shows that H5N1 Indonesian virus has higher infection efficiency compared with other H5N1. It made the pathogenicity level of Indonesian H5N1 higher than the others. The M2 visual docking simulation rimantadine and amantadine show that H5N1 Indonesian virus has higher resistance toward both drugs, compared with crystallized H5N1 virus from Hong Kong.

The dynamic simulation analysis was done by observing the protein-ligand complex interaction among their atoms. The dynamic simulation shows that H5N1 virus has resistance towards both drugs, because they didn't bind with M2 active site. However, this step is not conclusive, and requires further computation (data not shown).

[V] CONCLUSION

The phylogenetic tree analysis shows that H5N1 Indonesian virus has a close relationship with HPAI H5N1 from other countries. However, it belongs to its own cluster, which differs by its pathogenicity.

M2 docking result with Amantadine **[Figure-8]**, **[Figure-9]** and respective deltaG data **[Table-3]** clearly show that M2 H5N1 Indonesian virus has a lower binding affinity toward amantadine.

M2 docking result with Rimantadine **[Figure-10]**, **[Figure-11]** and respective deltaG data **[Table-4]** clearly show that M2 H5N1 Indonesian virus has a lower binding affinity toward Rimantadine.

The sequence alignment analysis has shown that HA, NA, and M2 of H5N1 Indonesian virus has amino acid mutation on its active site, and it is followed with the attribute change from hydrophilic to hydrophobic. It rendered the H5N1 Indonesian virus more pathogenic.

The molecular docking analysis shows that HA and NA H5N1 Indonesian virus has a better ability to bind sialic acid receptor, and the activity of amantadine and rimantadine did not give any inhibition toward active site of M2 H5N1 Indonesian virus. It caused the Indonesian H5N1 to have a higher pathogenicity.

We suggest conducting further research on molecular dynamics of HA, NA, and M2 H5N1 mutation.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Ridla Bakrie, PhD, the head of Chemistry Department, Faculty of Mathematics and Science, University of Indonesia, for his support toward this research. Authors also thank Dr. Lalit Ponnala of Cornell University, USA for his copy editing of the article.

REFERENCES

[1] Kamps SB, Hoffman C, Preiser W. [2006] Influenza report 2006. Flying Publisher. Paris, Cagliari, Sevilla.

IIOAB ISSN: 0976-3104

[2] Peiris JSM, De Jong MD, Guan Yi. [2007] Avian Influenza Virus (H5N1): a Threat to Human Health. *Clin Microbiol* 20: 243–267.

[3] Taha FA. [2007] How Highly Pathogenic Avian Influenza H5N1 has affected World Poultry-Meat Trade. *USDA Economic Research Service*. http://www.ers.usda.gov

[4] Ungchusak K, Auewarakul P, Dowell SF. et al. [2005] Probable Person to Person Transmision of Avian Influenza A (H5N1). *The New England J Medicine* 352 (4): 333–340.

[5] Gambotto A, Barratt-Boyes SM, de Jong MD, Neumann G, Kawaoka Y. [2008] Human Infection with Highly Pathogenic H5N1 Influenza Virus. *The Lancet* 371 (9622): 1464–1475.

[6] WHO [2009] Avian Influenza Disease Outbreak News http://www.who.int/csr/don/2009_09_31/en/index.html.

[7] Bouvier NM, Palese P. [2008] The Biology of Influenza Viruses. *Vaccine* 26: D49–53.

[8] Tambunan SF, Hikmawan O, Theofilus A. [2008] *In Silico* Mutation Study of Hemagglutinin and Neuraminidase on Banten Province Strain Influenza A H5N1 Virus. *Trends in Bioinformatics* 1(1): 18–24.

[9] Larkin MA, Blackshields G, Brown NP, et al. [2007] Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.

[10] Anwar T, Lal SK, Khan AU. [2006] *In Silico* Analysis of Genes Nucleoprotein, Neuraminidase, and Hemagglutinin : A Comparative Study On Different Strains of Influenza A (Bird Flu) Virus Sub-type H5N1. *In Silico* Biology 6: 0015.

[11] Klenk HD. [2007] Molecular Mechanisms of Pathogenicity and Interspecies Transmission of Avian Influenza Virus. Germany: Institut für Virologie Philipps Universität Marburg. Vizier Conference. Marseille, 27.04.2007.

[12] Bock JR, Gough DA. [2002] A New Method to Estimate Ligand-Receptor Energetics. *Mol Cell Proteomics* 1(11): 904–910.

[13] Kitchen DB, Decornez H, Furr JR, Bajorath J. [2004] Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Application. *Nat Rev Drug Discov* 3(11): 935-949.

ABOUT AUTHORS

Prof. Usman Sumo Friend Tambunan *is currently working as a permanent professor in the chair of Bioinformatics, Department of Chemistry, Faculty of Mathematics and Science, University of Indonesia. Previously, he worked at the Indonesian Agency of Technology Assessment and Application, as a senior scientist, and graduated his Phd in Tohoku University, Japan. He was the former vice dean of Faculty of Mathematics and Science, University of Indonesia. He has conducted research on Bioinformatics for biomedics, in the topics of HPV, Avian Influenza, and Dengue Virus. He had secured a copyright patent for HPV vaccine design and published many bioinformatics-related articles in peer-reviewed international scientific journals. He is elected several times as the best lecturer and researcher in University of Indonesia. His research is currently supported by Indonesian ministry of national education grant.*

Agus Limanto*, Bsc is currently working as junior assistant and researcher of Professor Usman Sumo Friend Tambunan at Bioinformatics Laboratory, Department of Chemistry, Faculty of Mathematics and Science, University of Indonesia. Now, he is involved in bioinformatics research, especially in dengue and influenza type A drug design.*

Arli Aditya Parikesit*, Msc is currently working as assistant of Prof Usman Sumo Friend Tambunan, lecturer, and researcher in the chair of Bioinformatics, Department of Chemistry, Faculty of Mathematics and Science, University of Indonesia. He has conducted research on Bioinformatics for biomedics, in the topics of HPV, Avian Influenza, and Dengue Virus. He finished and published his master thesis about HPV, with the support of Indonesian ministry of Education Graduate grant (hibah pasca). Now, he is a doctorate/Phd candidate at the chair of Bioinformatics, Department of Computer Science, Faculty of Computer Science and Mathematics, University of Leipzig, Germany with the support of DAAD fellowship. The theme of his doctorate research is 'Domain Cooccurence Distribution of Genetic Regulators from an Evolutionary Perspective'.*

The IIOAB Journal

ISSN: 0976-3104

RESEARCH: GENOMICS

GENOME SIZE DETERMINATION AND RAPD ANALYSIS OF FOUR EDIBLE AROIDS OF NORTH EAST INDIA

Jyoti P. Saikia[1*], Bolin K. Konwar [2] and Susmita Singh[3]

[1,2] Department of Molecular Biology and Biotechnology (A Department of Biotechnology center, Govt. of India), School of Science and Technology, Tezpur University, Tezpur-784028, INDIA

[3] Department of Biotechnology, North East Institute of Science and Technology (CSIR), Jorhat-785006, INDIA.

Received on: 14th-June-2010; **Revised on:** 25th-Augt-2010; **Accepted on:** 13th-Sept-2010; **Published on:** 15th-Oct-2010.

[*]**Corresponding author:** Email: jyotisaikiazone@gmail.com Tel: +91 03712-267007/8/9; 03712-267045, Extn: 5420; Fax: +91-03712-267005/267006; 03712-267045.

ABSTRACT

Four edible aroid species were selected for the study. The genomic DNA of the plants was isolated and estimated. A part of the genomic DNA was used for analysis using six different primers from Operon Technologies, USA. The genome size determined for the aroids is in the order of Colocasia esculenta> Xanthosoma caracu> Xanthosoma sagittifolium > Amorphophallus paeonifolius. Amorphophallus species was found to be 50% similar to both Xanthosoma caracu and Colocasia esculenta. The analysis will provide a ground for exploring the vast diversified aroid population of the region.

Keywords: *RAPD; aroid; genome size*

[I] NTRODUCTION

Corms were important root crops of Asia [1]. India-Myanmer area is considered as the center of origin for *C. esculenta* [2]. Assessment of genetic diversity of aroids needs immediate attention for the improvement of these crops. Few reports are available on the use of molecular markers to study genetic diversity in taro such as restriction site variation in rDNA, mitochondrial DNA [3] and RAPD markers [4]. Reports are available on unpredictable changes in the chromosomes during cell division resulting in the lack of uniformity in this crop [5]. They reported chromosome number of aroids to be 2n= 22,26,28,38 and 42 in the plants collected from various locations. RAPD profiling of *Xanthosoma caracu* was reported by Schnell *et al.* (1999) [6].

The amount of nuclear DNA and size of the genome (C-value) are two important biodiversity characters, the study provides a strong unifying element in biology with practical and predictive uses. The RAPD analyses of aroid species were important for their identifications and similarity assessment. The letter 'C' of C-value stands for 'constant'; the constant amount of DNA which could be a characteristic of a particular genotype [7]. Hinegardner (1976) described genome size as an important biodiversity parameter [8]. Till date, flow cytometry and fuelgen densiometry (densitometry) are the standard methods for the determination of genome size. But, these established techniques need sophisticated instruments, which is not possible in all situations. Although, the flow cytometry yields almost accurate and reproducible amount of 2C DNA content of plant species, the problem starts with expeditions to more distant areas, when the transport and maintenance of the material become an issue. Moreover, the cost of establishing a flow cytometry laboratory may be prohibitive in certain areas [9]. The north eastern region like many other region of the country, regarded to be the biodiversity hotspot is yet to establish a flow cytometry laboratory. On the basis of the above facts, the present investigation has been undertaken to estimate the genome size using the method described by Konwar *et al.* (2007) [10].

[II] MATERIALS AND METHODS

2.1. Plant materials

Four aroid species *A. paeonifolius*, *X. caracu*, *X. sagittifolium* and *C. esculenta* were grown in Tezpur University campus. The young tender shoots of the plants were used for the isolation of genomic DNA. The primers were obtained from Operon technol. Pvt. Ltd. (USA) with sequences as follows: OPW-04: 5-CAGAAGCGGA-3; OPW-05: 5-GGCGGATAAG-3; OPW-08: 5-GACTGCCTCT-3; OPW-10: 5-TCGCATCCCT-3; OPW-15: 5-ACACCGGAAC-3 and OPW-16: 5-CAGCCTACCA-3.

2.2. Isolation of Genomic DNA

Genomic DNA was isolated using the method described by Ronning et al. (1995) [11]. One gram leaf frozen with liquid nitrogen was ground using mortar and pestle. The extraction buffer 10 ml [100mM Tris-HCl, pH 8.0; 250mM NaCl; 25 mM ethylenediaminetetraacetic acid (EDTA); 0.1% (v/v) 2-mercaptoethanol; 100 mM diethyldithiocarbamic acid

ISSN: 0976-3104

(DEDTC); 2% (w/v) polyvinylpyrrolindone (PVPP)] was added and grinding continued. The tissue homogenate was filtered through four layers of cheesecloth. Sodium dodecyl sulfate (SDS) was added to a final concentration of 1.25% the mixture was shaken vigorously and incubated at 65OC for 10 min. Potassium acetate was added to the final concentration of 1.2 M, again shaken vigorously and incubated at 0OC for 20 min, followed by centrifugation at 25,000 g for 20 min at 4OC. The supernatant was filtered through two layers of miracloth, combined with 2/3 volume ice cold isopropanol and precipitated at -20OC for overnight. The DNA was pelleted at 20,000 g for 15 min at 4OC. The pellet was dried and resuspended in 400µl Tris-EDTA (TE) pH 7.4. Each DNA sample was treated with ribonuclease A (RNase A, Bangalore Gene, India) with a concentration of 50 mg ml-1 and incubated at 37OC for 30 min. Organic extraction was performed adding an equal volume of phenol : chloroform : isoamyl alcohol (25:24:1) shaked vigorously and then centrifuged at 12,000 g for 10 min. The process was repeated with chloroform: isoamylalcohol. The DNA was then precipitated by adding NaCl to a final concentration of 0.2 M and two volumes of ice cold 95% ethanol, incubated at -20OC for 30 min and centrifuged for 15 min at 12,000 g to pellet the DNA. The pellet was dried and resuspended in 100 µl TE (pH 7.4).

2.3. Determination of yield and quality of DNA

The yield and quality of the DNA was determined using the method described by Gallagher (1987) [12]. The isolated DNA 50 µl was evaporated to dryness using Maxi Dry Plus under 1mbar atm pressure at 35^{O}C. The dried DNA was dissolved in 1 ml 1X TNE buffer containing 0.01 M Tris base, 1 mM EDTA and 0.2 M NaCl (pH adjusted to 7.4 with concentrated HCl). Both blank and samples were analyzed at 325nm for confirming clean cuvettes. Absorption at 230 nm for detecting any phenol contamination was also measured. The purity level of the DNA isolated was estimated from absorption at 260:280 nm. The quantification of the DNA was done putting the absorption value at 260 nm in the following formulae:

Concentration of dsDNA (µg/ml)=A_{260}/0.020

Where, A_{260}- absorption value of DNA in TNE after subtracting TNE blank absorption

value; ds, double stranded; 0.020- molar extinction coefficient with a unit of M^{-1}

cm^{-1}.

2.4. DNA amplification

PCR amplification was carried out using RAPD decamer primers (Operon Technologies, USA). The reaction mixture (25 µl) consists of 100 ng template DNA, 1X reaction buffer (10 mM Tris-HCl pH 9.0, 50 mM KCl and 1.5 mM MgCl2), 100 µM of each of the dATP, dTTP, dCTP and dGTP, 5 pM primer and 0.5 U of Taq DNA polymerase (Bangalore gene, India). Amplification was performed using the thermal cycler Gene Amp PCR system 9700 (Applied biosystems) with a heated lid to reduce the evaporation under the conditions of an initial 1 min denaturation at 94OC, followed by 45 cycles of 1 min at 94OC, 1 min at 35OC, and 2 min at 72OC, with a final 5 min extension step at 72OC. Approximately, 15 µl of the amplified products were loaded on a 2% agarose gel and separated by electrophoresis in TAE (Tris 1.6 M, acetic acid 0.8 M, EDTA 40 mM) at 100 volts. After electrophoretic separation for 90 min, the gel was stained with ethidium bromide and visualized in UV light. A 100 bp ladder (Bangalore gene, India) was loaded in all gels used as the molecular marker. Following electrophoresis, the gel was photographed using Geldoc 1000 (Biorad).The primers used in the experiment were OPW-04, OPW-05, OPW-08, OPW-10, OPW-15 and OPW-16.

2.5. Genome size determination

Genome size determination was done using the method of Konwar *et al.* (2007) [10]. The method can be divided into two broad categories, (a) determination of leaf tissue and cell volume and (b) genome size determination.

(a) *Determination of leaf tissue and cell volume:* Fine transverse and longitudinal cross sections of the pre-weighted leaf tissue of 1 cm2 size were obtained with a sharp sterile blade and observed under a microscope (Carl Zeiss microlmaging, Germany) at 10X40x magnification. The volume of the whole tissue section (1 cm2) was determined (length×breath×thickness). The length, breath and thickness of rectangular cells; length and radius of the cylindrical cells and radius of spherical cells were measured with a micro scale having 400x magnifications. Data generated from five randomly selected cells of five random sections as well as cell volumes were recorded with the specific formulae. The intercellular space of the leaf tissue of the aroid was measured in five small leaf sections of known dimensions by measuring with a micro scale at 400x magnification.

(b) Genome size determination: The genome size of the plant species was determined using the following calculations:

Average volume of a single cell (l×b×t) = ×µ3 where,

l = length

b= breath

t = thickness

Volume of a tissue (l×b×t) = t µ3

Volume of the intercellular space = s µ3

Actual cell mass = (t-s) = v µ3

Total number of cells in the cell mass = v µ3/×µ3 = y

Now, weight of the tissue section = wg

wg tissue contains = y cells

So, 1g tissue will contain cells = y/w

DNA yield per gram of leaf tissue = d µg

= d×106 pg

So, one cell contains = (d×106)/(y/w) pg

= (d×106)/(y/w) × 978 Mbp

2.6. Data analysis for RAPD

The images were analyzed using Biored software for the estimation of molecular weight of the bands. The pair wise genetic similarity matrix between the genotypes was generated using Jaccard's coefficients.

[III] RESULTS

3.1. Yield and quality of the DNA

The absorption readings obtained of the purified genomic DNA of the four aroids were presented in **Table–1.** From the absorption values it can be concluded that the cuvettes were clean and there is no phenol contamination in the purified DNA [12]. The absorption in 325nm and 230 were taken to confirm phenol free DNA sample and clean cuvette **[Table–1]**. All four DNA samples show 260/280 ratio above 1.7 **[Table–1]**, which signifies the high purity of the DNA. The yield follows the order of X. sagittifolium > A. paeonifolius > C. esculenta > X. caracu. The X. sagittifolium shows not only high yield but also highly pure DNA **[Table–1]**.

3.2. RAPD analysis and dendrogram

IIOAB ISSN: 0976-3104

As shown in the Figure1 the primer OPW-15 sequences are present only in the C. esculenta. Figure 1 also depicts the absence of OPW-15 and OPW-16 sequences in the Amorphophallus. paeoniifolius. Amplification of genomic DNA using six primers yielded 25 reproducible RAPD loci for an average of 4.16 bands per primer; 21 (84%) of them were polymorphic and 4 (16%) were monomorphic. Among Xanthosoma sagittifolium and Xanthosoma caracu only four bands (16%) were polymorphic with the rest of polymorphism existing between Colocasia esculenta and Amorphophallus paeonifolius (figure 1) **[Table 2]**. As seen in the figure 2 the two species of Xanthosoma genus were more similar as compared to the Colocasia esculenta and Amorphophallus paeoniifolius. The Amorphophallus paeoniifolius was more similar to the two Xanthosoma species as compared to the Colocasi a esculenta.

The primers were selected in view of their polymorphism in Indian Colocasia esculenta cultivars [5]. The results showed that the both Xanthosoma sagittifolium and Xanthosoma caracu species are closely related to each other (66%) as compared to their similarity to others **[Table 3]**. Amorphophallus paeoniifolius species was found to be 50% similar to both X. caracu and C. esculenta but dissimilar to X. sagittifolium. Hence, it could be concluded that X. caracu and C. esculenta species are equally distant from A. paeonifolius species **[Table 3]**. C. esculenta used in this study might be the same morphotype M-5 having ID no. TC5 used for RAPD analysis by Lakhanpaul et al. (2003) [5].

3.3. Genome size analysis

The genome size determined for the aroids are in the order of C. esculenta> X. caracu> X. sagittifolium > A. paeonifolius. The genome size of C. esculenta was found to be 14.10 pg and it was in between 2C and 3C value as given by Bennett and Leitch (2005b) for C. antiquorum [13]. They also provided 1C value of X. sagittifolium to be 8.8 pg, which is 0.76 pg higher than that of the present investigation **[Table 4]**.

[IV] DISCUSSION

The study is a preliminary investigation to estimate the genetic similarity and dissimilarity between the edible species of aroids. The primers selected for the study were well established for the *C. esculenta* of India. In the present investigation it was observed that primers are also providing reliable polymorphism in *A. paeonifolius*. These primers selected for the study showed 88.9-100.0% polymorphism in the case of Indian *C. esculenta* morphotypes, which are also reported by Lakhanpaul *et al.* (2003) [5]. The genome diversity analyses of aroids with a limited number of genotype are not enough to prepare a dendrogram. A large scale diverse morphotypes using more number of suitable primers could be needed for the same.

The genome size of the edible species was determined using a comparatively easy and established method than prevalent expensive methods like fuelgen densiometry and flow cytometry. The genome size determination was already established by Konwar *et al.* (2007) [10]. At present, flow cytometry is becoming a popular method in determination of genome size. The method needs the establishment of the sophisticated costly equipment, flow cytometer. The cost is indeed prohibitive for most of the organizations and handling the equipment needs expert man power. On the other hand, discrepancies are observed in flow cytometry data of a single sample analysed in different laboratories. Hence, an effort was made to develop a novel but simple and less expensive method for the determination of genome size of plants without compromising the quality of the work. With the collection of suitable tender leaves from the raised corms of the selected edible aroids for genomic DNA isolation we could obtain most of the genomic DNA. The protocol followed in the investigation could yield quality DNA as evident from the UV-VIS spectrophotometric absorbencies ratio at 260:280 nm being 1.735-1.916. Two critical points were taken into account in this method; firstly, the method for the isolation of genomic DNA was such that it could isolate almost all the DNA from the nuclei, and secondly the accurate determination of the intercellular space in the concerned edible aroids species. The developing countries on one hand have a vast resource of plant biodiversity while on the other hand they face the problem of acquisition and maintenance of flow cytometers [14]. Hence, present method could be great help for the researcher of developing country having limited excess to the facilities like flow-cytometry.

Aroid species	260/280	DNA µg/gm leaf
Amorphophallus paeoniifolius	1.735	183.5
Xanthosoma caracu	1.814	159.21
Colocasia esculenta	1.754	175.21
Xanthosoma sagittifolium	1.916	207.24

Table: 1. DNA purity and content of the aroids

Primers	Base pair length (bp)
OPW-04	327a, 571b, 636a, 1264a.
OPW-05	471c, 542c, 610a, 812c.
OPW-08	903a, 924c, 1244a, 1992c
OPW-10	454a, 460c, 1027c, 1047a, 1153a, 473b.
OPW-15	1224c.
OPW-16	297d, 386c, 461d, 505c, 1149b, 2043b.

NB: a, *A. paeonifolius*; b, *X. caracu*; c, *C. esculenta* and d, *X. sagittifolium*.

Table: 2. Primers and base pair lengths of RAPD generated markers

ISSN: 0976-3104

Aroids	*A. paeonifolius*	*X. caracu*	*C. esculenta*	*X. sagittifolium*
Amorphophallus paeoniifolius	1			
Xanthosoma caracu	0.5	1		
Colocasia esculenta	0.5	0.33	1	
Xanthosoma sagittifolium	0	0.66	0.33	1

Table: 3. Pair-wise genetic similarity matrix between 4 aroids based on Jaccard's coefficient

Aroid species	Genome size in pg (C)	Genome size in Mbp (C)
A. paeonifolius	5.04	4929.12
X. caracu	8.21	8029.38
C. esculenta	14.1	13789.8
X. sagittifolium	8.04	7863.12

Table: 4. Genome size of the aroid species

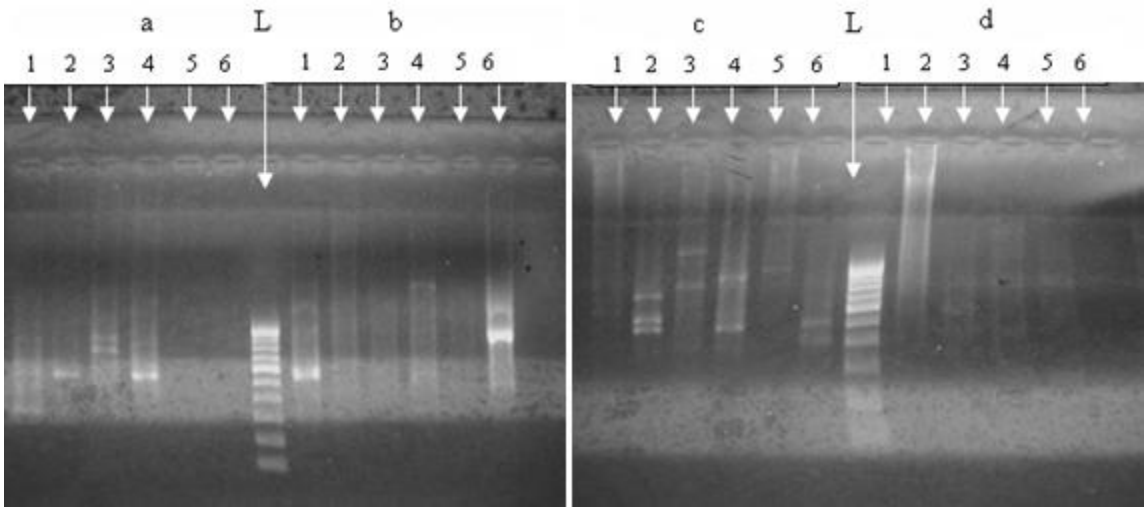

Fig: 1. RAPD profile of aroid species. 1-6: OPW-04, OPW-05, OPW-08, OPW-10, OPW-15 and OPW-16, respectively. (a) *A. paeonifolius*, (b) *X. caracu*, (c) *C. esculenta* and (d) *X. sagittifolium*. L represent kilobase ladder, 100-1000 bp were developed.

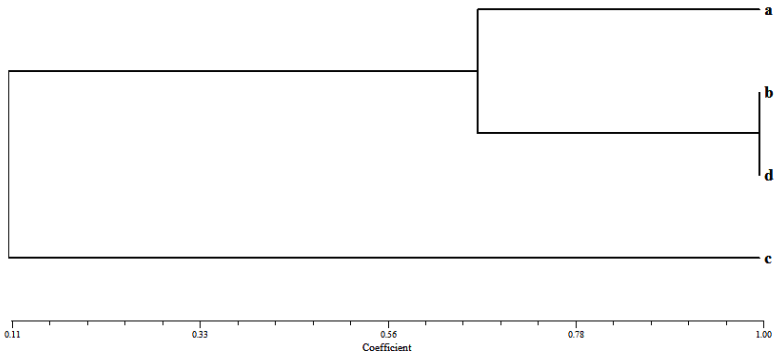

Fig: 2. The dendrogram based on RAPD data (a, *Amorphophallus paeoniifolius*; b, *Xanthosoma caracu*; c, *Colocasia esculenta* and d, *Xanthosoma sagittifolium*).

IIOAB ISSN: 0976-3104

[V] CONCLUSION

The genome size of the species was determined using a new method. The method did not require costly instruments like flowcytometry. As mentioned in the discussion the method provides comparable results with flow cytometry. Reports are available regarding variations among the flowcytometry results [7].The amount of DNA in the nucleous is directly proportional to the chance of survival of a species by not becoming extinct [7]. The contribution to the genome size research from a country like India is very little/not representative [7]. In the light of the discussion above, it can be suggested that popularizing this type of genome size estimating method will allow the developing country researcher to contribute more to the field of genome size. The RAPD analysis of the aroid species provides information on the genetic similarity and phylogenetic relation.

FINANCIAL DISCLOSURE

The present method is suitable especially for comparisons of genome sizes of different species and genus of living forms. The method needs up gradation and future research to establish it in the field of genome side estimation.

ACKNOWLEDGEMENT

We acknowledge the National Medicinal Plant Board for providing us the fund for the present research

REFERENCES

[1] Ochiai T, Nguyen VX, Tahara M, Yoshino H. [2001] Geographical difference in Asian Taro, Colocasia esculenta (L.) Schoot, detected by RAPD and isoenzyme analyses, *Euphytica* 122: 219–234.

[2] Kuruvilla KM, Singh A. [1981] Karyotyping and electrophoretic studies on taro and its origin, *Euphytica* 30: 405–413.

[3] Matthew P, Matsushita Y, Sato T, Hirai M. [1992] Ribosomal and mitochondrial variation in Japanese taro (Colocasia esculenta (L.) Schott). *Jap J Breed.* 42: 825–833.

[4] Irwin SV, Kaufusi P, Banks K, Pena de la R, Cho JJ. [1998] Molecular characterisation of taro (Colocasia esculenta) using RAPD markers. *Euphytica* 99: 183–189.

[5] Lakhanpaul S, Velayudhan KC, Bhat KV. [2003] Analysis of genetic diversity in Indian taro [Colocasia esculenta (L.) Schott] using random amplified polymorphic DNA (RAPD) markers. *Genetic Resources and Crop Evolution* 50: 603–609.

[6] Schnell RJ, Goenaga R, Olano CT. [1999] Genetic similarities among cocoyam cultivars based on randomly amplified polymorphic DNA (RAPD) analysis. *Scientia Horticulturae* 80: 267–276.

[7] Bennetta MD, Leitch IJ. [2005] (a) Nuclear DNA amount in angiosperms: Progress, Problems ad Prospects. *Annals of Botany* 95: 45–90.

[8] Hinegardner R. [1976] The cellular DNA content of sharks, rays and some other fishes. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry* 55(3): 367–370.

[10] Konwar BK, Chowdhury D, Buragohain J, Kandali R. [2007] A new less expensive method for genome size determination of plants. *Asian Journal of Plant Sciences* 6 (3): 565–567.

[11] Ronning CM, Schnell RJ, Gazit S. [1995] Using randomly amplified polymorphic DNA (RAPD) markers to identify Annona cultivars. *American Society For Horticultural Science (USA)* 120 (5): 726–729.

[12] Gallagher SR.[1999] Detection of nucleic acid using absorption spectroscopy. *Short protocol in Molecular Biology*, Wiley.

[13] Bennettb MD, Leitch IJ. [2005] (b) Plant Genome Size Research: A Field In Focus. *Annals of Botany* 95: 1–6.

[14] Temsch EM, Greilhuber J. [2001] Genome size in Arachis duranensis: a critical study. *Genome* 44: 826–830.

ABOUT AUTHORS

Prof. Bolin Kumar Konwar is the Dean, School of Science and Technology, Tezpur University. Along with many different academic positions he is an eminent worker in the field of genetics and plant breeding, biodegradation, bio-energy research (ethanol production from yeast), nanobiotechnology, metagenomics, bioinformatics etc. Along with these polyhydroxy alkanoate from microbes, biocompatibility of nanomaterials, targeted drug delivery and biosurfactant from crude oil degrading microbes, starch characterization and application in bio-film preparation etc. were also carried out by him.

Mr. Jyoti Prasad Saikia is a Ph. D. student under the supervision of Prof. B. K. Konwar from January, 2006. Mr. Saikia expertise in genomic study, biochemical technique, cytotoxicity study, nanobiotechnology, biodegradation, starch research. Mr. Saikia has published eight different papers in peer reviewed international journals.

ISSN: 0976-3104

Ms. Susmita Singh is a Senior Research Fellow (CSIR) pursuing Ph. D under the supervision of Dr. R.L.Bezbaruah, Scientist F, in the department of Biotechnology, North East Institute of Science and Technology (CSIR), Jorhat. Her research interests include industrial microbiology, biochemistry of enzymes and molecular biology. Presently she is working on microbial L-amino acid oxidases.