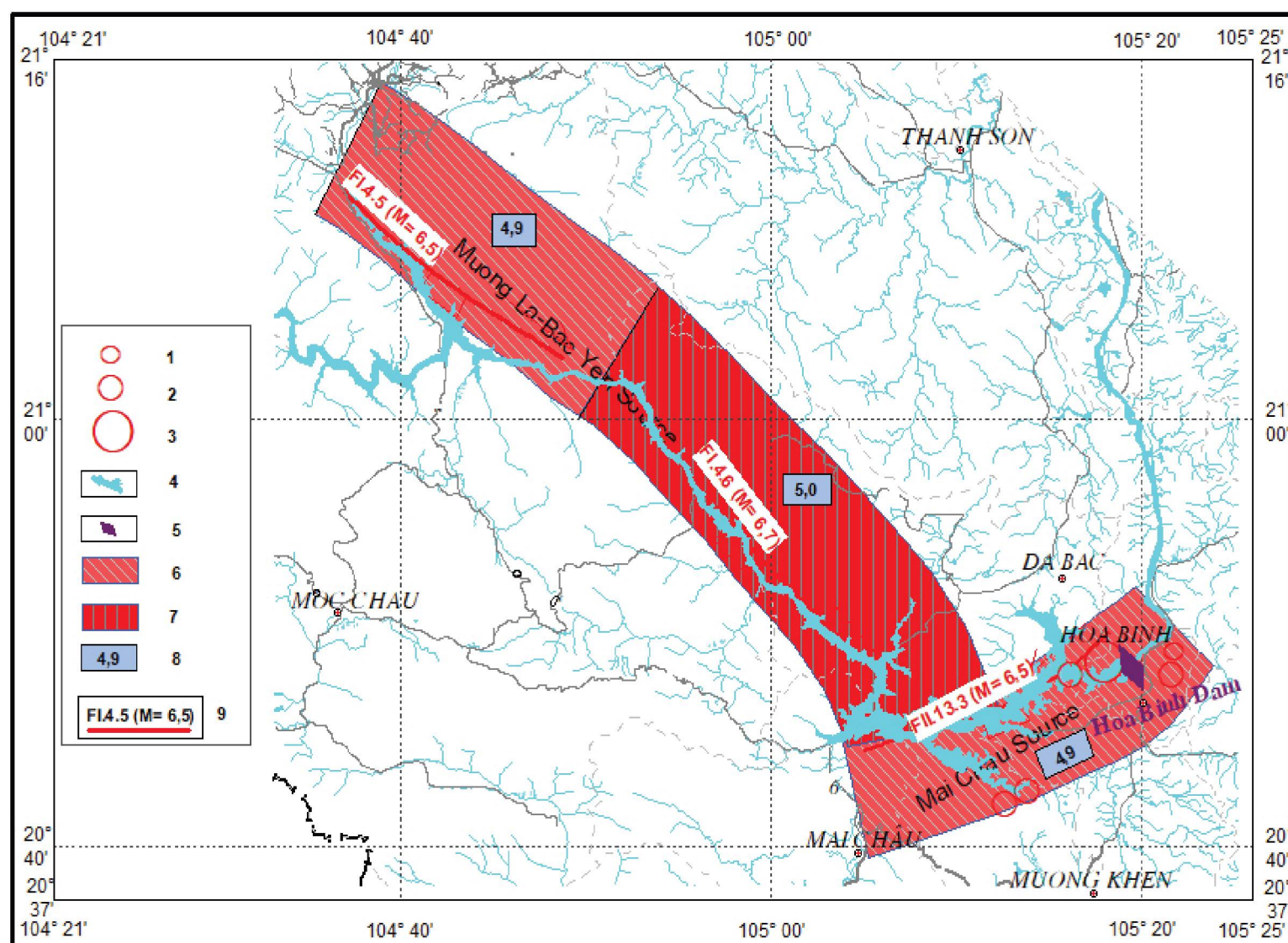




**The characteristics of seismic activity in the ladder zone
of Da river hydro-electric plants**

Vietnam Journal of Science, Technology and Engineering (VJSTE) is a quarterly double-blind peer-reviewed academic journal. VJSTE publishes research articles and review articles in all areas of science, technology and engineering with the focus on the life sciences and environmental sciences including the following topics:

- Agriculture
- Medicine
- Pharmacology
- Biology
- Biotechnology
- Ecology
- Climatology.

CLASSIFICATION SCHEME

1. Mathematics and Computer Science

- 1.1. Mathematics
- 1.2. Computer Science
- 1.3. Computational Science

2. Physical Sciences

- 2.1. Physics
- 2.2. Chemistry
- 2.3. Engineering

3. Life Sciences

- 3.1. Agriculture
- 3.2. Medicine
- 3.3. Pharmacology
- 3.4. Biology
- 3.5. Biotechnology

4. Geosciences

- 4.1. Geography
- 4.2. Geology
- 4.3. Geophysics
- 4.4. Oceanography

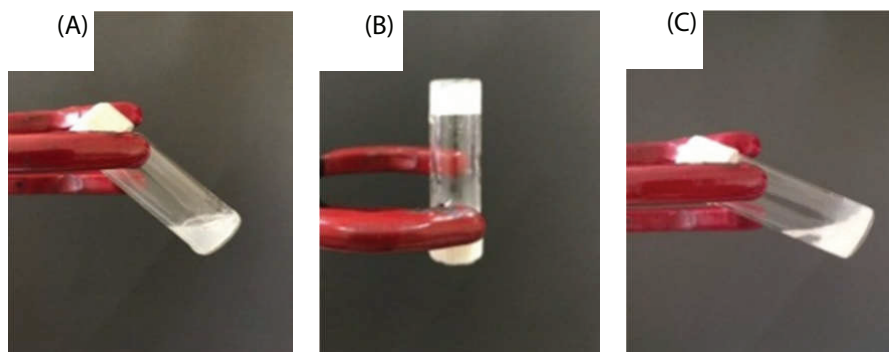
5. Environmental Sciences

- 5.1. Ecology
- 5.2. Climatology

CONTENTS

MARCH 2019 • VOL.61 NUMBER 1

11



MATHEMATICS AND COMPUTER SCIENCE

3 Gaussian Mixture Modeling for Wi-Fi fingerprinting based indoor positioning in the presence of censored data.

Trung Kien Vu, Hung Lan Le

PHYSICAL SCIENCES

9 The effects of temperature, feed ratio, and reaction time on the properties of copolymer PLA-PEG-PLA.

Viet Linh Nguyen-Vu, Mai Anh Pham, Dai Phu Huynh

14 Using a physical model to determine the hydrodynamic dispersion coefficient of a solution through a horizontal sand column.

Le Anh Tuan and Guido Wyseure

LIFE SCIENCES

23 Expression and purification of thioredoxin-his₆-ZmDREB2.7 fusion protein in *Escherichia coli* for raising antibodies.

Thuy Linh Nguyen, Thuy Duong Nguyen, Van Hai Nong, Thi Thu Hue Huynh

30 Identification of cassava bacterial blight-causing *Xanthomonas axonopodis* pv. *manihotis* based on *rpoD* and *gyrB* genes.

Ngo Quang Huy, Mai Van Quan, Le Quang Man, Duong Thi Nguyen, Trinh Xuan Hoat

36 Initial study of learning curves in robot-assisted radical prostatectomy.

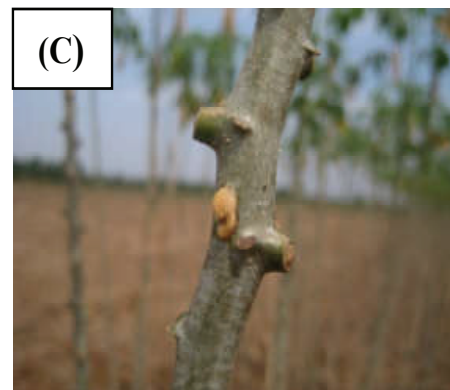
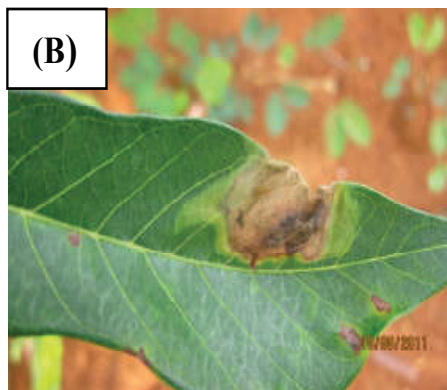
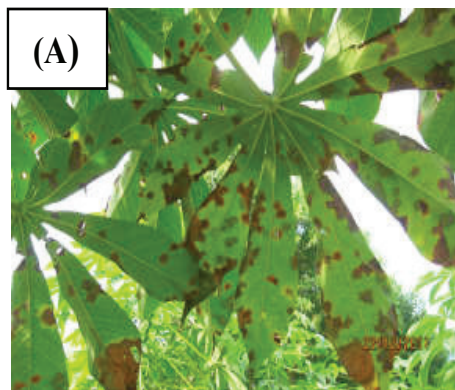
Le Chuyen Vu, Vinh Hung Tran, Phuc Cam Hoang Nguyen, Van An Nguyen, Vu Phuong Do, Ngoc Chau Nguyen, Lenh Hung Do

41 Evaluation of graft function in renal retransplantation at Hue Central Hospital.

Tam Vo, Ngoc Tuan Anh Dang, Nhu Hiep Pham, Gia Khanh Pham

45 Examination of the accuracy of the quantitative in-house kit for determining zinc concentration in seminal fluid.

Thi Trang Nguyen, Thi Minh Phuong Le, Thi Huyen Trang Do, Thi Quynh Dien Tran, Ngoc Thach Pham

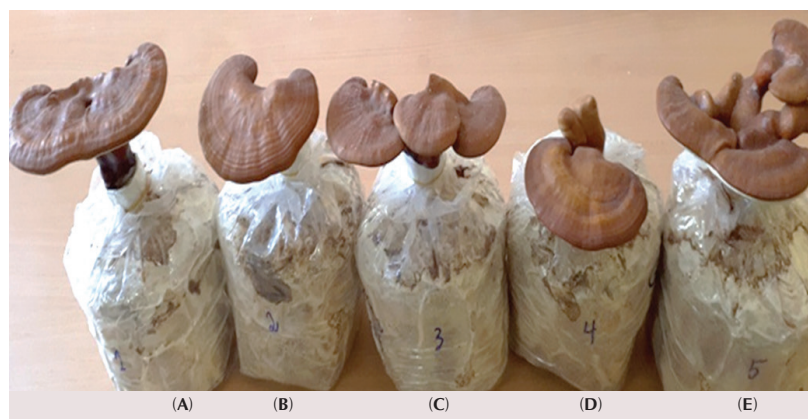


32

CONTENTS

MARCH 2019 • VOL.61 NUMBER 1

66



52 Field trial of beneficial effects of carotenoid-producing *Bacillus aquimaris* SH6 spores to whiteleg shrimps.

Thi Tham Nguyen, Thi Huong Nguyen, Thi Ngoc Que Nguyen, Thi My Tran, Thi Thu Huong Pham, Dinh Thang Nguyen, Thi Van Anh Nguyen

58 Studies of common antibiotic resistance-associated genes of *Acinetobacter baumannii*.

Le Phuong Trinh, Van Son Ta, Thi Ngoc Han Luu, Mai Phuong Nguyen,

Thi Huong Dinh, Van Hung Nguyen, Thi Ha Doan, Bao Yen Pham

62 Optimal culture conditions for mycelial growth and fruiting body formation of Ling Zhi mushroom *Ganoderma lucidum* strain GA3.

Bich Thuy Thi Nguyen, Nghien Xuan Ngo, Ve Van Le, Luyen Thi Nguyen, Ry Kana, Huy Duc Nguyen

68 Characterization of newly isolated thermotolerant yeasts and evaluation of their potential for use in *Cayratia trifolia* wine production.

Doan Thi Kieu Tien, Huynh Xuan Phong, Mamoru Yamada, Ha Thanh Toan, Ngo Thi Phuong Dung

GEOSCIENCES

74 Remote sensing applications for analysing the impacts of land cover changes on the upper part of the Dong Nai river basin.

Hung Pham, Van Trung Le, Le Phu Vo

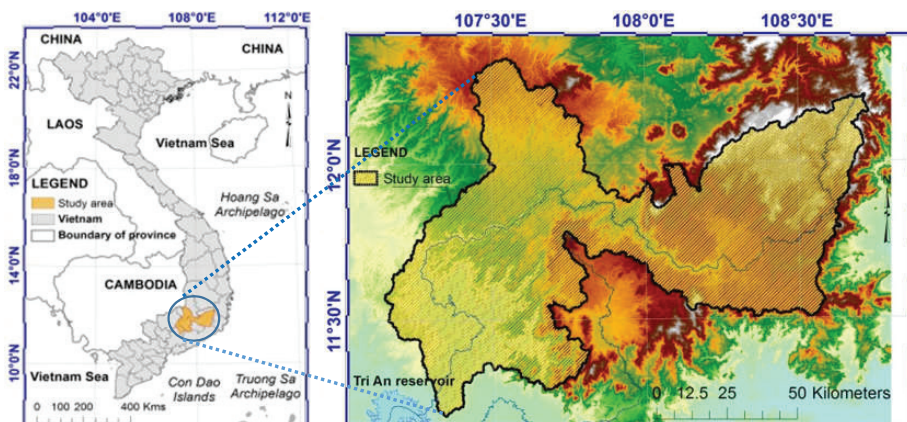
82 The characteristics of seismic activity in the ladder zone of Da river hydro-electric plants.

Dinh Trong Cao, Van Dung Le, Anh Tuan Thai, Thanh Hai Dang, Dinh Trieu Cao

ENVIRONMENTAL SCIENCES

92 River information management and early flood release in response to climate change in Vietnam.

Tran Dinh Hoa, Do Hoai Nam



75

Gaussian Mixture Modeling for Wi-Fi fingerprinting based indoor positioning in the presence of censored data

Trung Kien Vu^{1*}, Hung Lan Le²

¹Faculty of Electronics, Hanoi University of Industry

²National Center for Technological Progress

Received 4 September 2018; accepted 6 December 2018

Abstract:

In complex indoor environments, due to the attenuation of the signal and the changing surrounding environment, the censoring and multi-component problems may be present in the observed data. Censoring refers to the fact that sensors on portable devices cannot measure Received Signal Strength Index (RSSI) values below a specific threshold, such as -100 dBm. The multi-component problem occurs when the measured data varies due to obstacles and user directions, whether the door is closed or open, etc. By accounting for these problems, this paper proposes to model the RSSI probability density distributions using the Censoring Gaussian Mixture Model (C-GMM) and develop the Expectation-Maximization (EM) algorithm to estimate the parameters of this model in the offline phase of the Wi-Fi fingerprinting based Indoor Positioning Systems (IPS). The simulation results demonstrate the effectiveness of the proposed method.

Keywords: censored data, EM algorithm, fingerprinting, Gaussian Mixture Model, IPS.

Classification numbers: 1.3, 2.3

Introduction

With the popularity of wireless local area networks (WLAN), Wi-Fi based indoor positioning techniques are widely used for indoor user localization. Most popular Wi-Fi positioning methods use the received signal strength indication (RSSI). Among available approaches, fingerprinting appears to be the most feasible method for positioning in the indoor environment [1]. This method estimates the position of an object and relies on training data from a set of reference points (RP) with known locations. Fingerprinting-based methods consist of two phases, namely the offline phase and the online phase. In the offline phase, the training data (i.e., RSSI) are collected at the RPs and used to build the database, which is often called the radio map. During the online phase, the online measurements are compared against the training data at every RP. The position of the RP whose training data most closely match the online data can be regarded as the estimated position of the object.

To represent the training data in probabilistic approaches, the parametric model and nonparametric model are two basic categories which are commonly used. The systems which utilized the parametric model had more advantages than the nonparametric model [2].

The probability density function (PDF) of the observed data is assumed to be the single Gaussian in the presence of censoring and dropping problems [3, 4]. Censoring occurs due to the limited sensitivity of Wi-Fi sensors or the sensor driver, which does not intentionally report the overly weak observed signal strengths; in other words, the smart phones do not report the signal strength if it is below a specific threshold, e.g., -100 dBm with typical smart phones.

An EM algorithm was proposed to estimate the

*Corresponding author: Email: vutrongkienfee@gmail.com

parameters of censored and dropped single Gaussian data. Experimental results with real field data can demonstrate the effectiveness of this proposal relative to the others, but the multi-component was not considered.

In [5, 6], the multi-component problem has been noted. In [6], the authors illustrated that human behaviors in the measurement environment (absence, sitting or standing still, moving randomly, and moving specifically) result in the bi-modal phenomena in the experimental data. In this case, using a single Gaussian distribution to model the RSSI histogram is not appropriate. In [5], the Gaussian Mixture Model (GMM) was proposed to model the RSSI measurements. Positioning results were improved relative to the single Gaussian model. However, the censoring problem has not been considered in these studies, although they clearly occurred, as discussed in [3, 4].

In [7, 8], the authors introduced EM algorithms for parameter estimation of the grouped, truncated, and censored data. This proposal can solve the bias of parameter estimation, but the censoring and multi-component problems have not been resolved.

This paper accounts for all of the problems discussed and proposes to develop a new extended version of the EM algorithm to enhance the quality of estimated parameters in the offline phase and there by improve the performance of the Wi-Fi fingerprinting-based IPS.

Proposed methods

This section delineates the proposed method, which relies on the characteristics of the collected Wi-Fi RSSI

data for enhancing the accuracy of the fingerprinting-based indoor positioning system (Fig. 1). First, a C-GMM is introduced to model the RSSI distribution in the presence of censored mixture data. Second, an extended EM algorithm is developed to estimate the parameters of this model. This algorithm is employed during the offline phase. Third, in the online phase, the localization and classification procedure is based upon the Maximum a Posteriori (MAP) method.

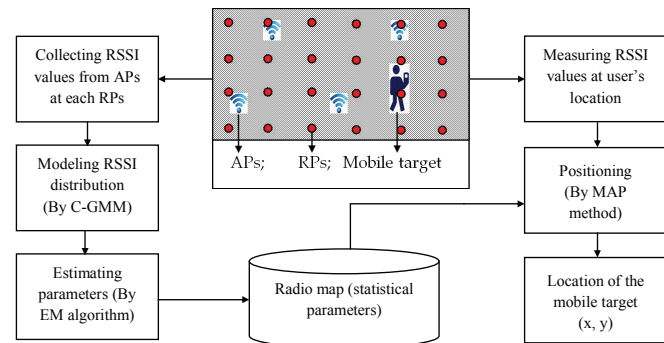


Fig. 1. Block diagram of the proposed Wi-Fi fingerprinting-based IPS.

Modeling RSSI distribution by the C-GMM

Below are several important definitions:

$\vec{y} = [y_1, y_2, \dots, y_N]$; $y_n \in \mathbb{R}$; $n = 1 \div N$ is the set of complete data (non-censored data), y_n are i.i.d. random variables. c is the specific threshold at which a portable device, e.g., smart phone does not report the signal strength. $\vec{x} = [x_1, \dots, x_N]$ is the set of observable data (censored data), $x_n = \begin{cases} y_n, & \text{if } y_n > c. \\ c, & \text{if } y_n \leq c \end{cases}$. Figure 2 illustrates the measurement model.

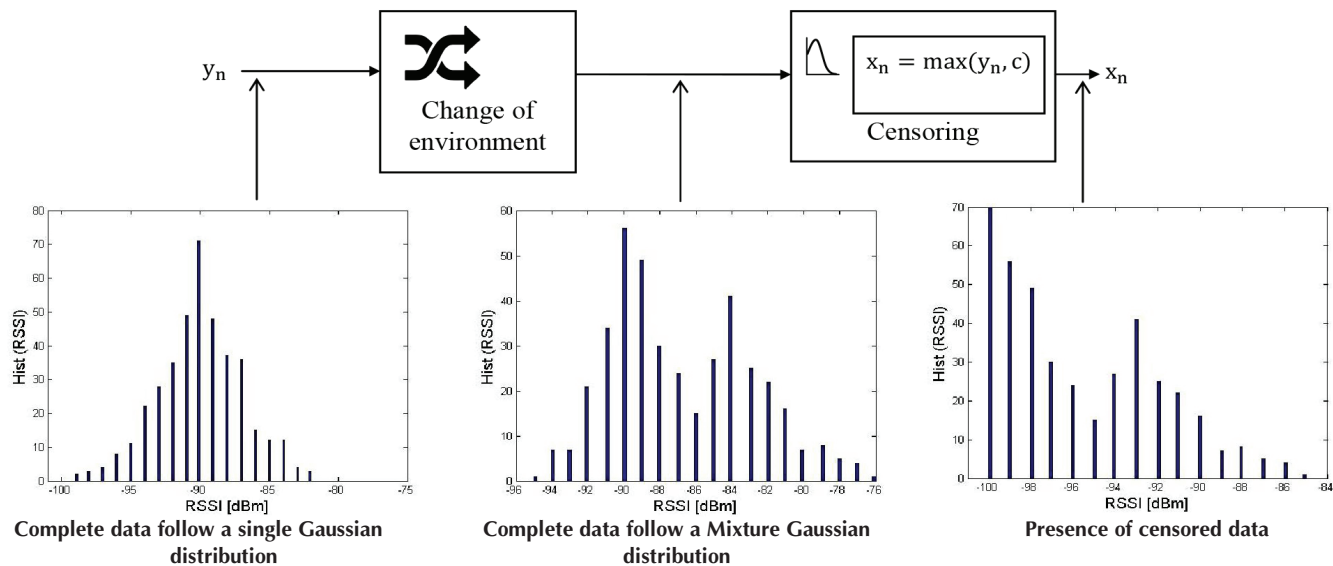


Fig. 2. Proposed measurement model.

Parameter estimation in the offline phase

$\vec{\theta} = [w_1, \dots, w_J; \mu_1, \dots, \mu_J; \sigma_1, \dots, \sigma_J]$ is the set of parameters of the GMM. The GMM includes J Gaussian components; the j^{th} component ($j = 1 \sim J$) is parameterized by $\theta_j = [\mu_j, \sigma_j]$. w_j are positive mixing weights which sum up to one.

The likelihood of $\vec{y}$ following a GMM is as follows:

$$p(\vec{y}; \vec{\theta}) = \prod_{n=1}^N \sum_{j=1}^J w_j p(y_n; \theta_j) \quad (1)$$

To simplify the calculation of the summation in Eq. (1),

a set of auxiliary variables $\vec{\Delta} = \begin{bmatrix} \Delta_{11} & \dots & \Delta_{1J} \\ \vdots & \ddots & \vdots \\ \Delta_{N1} & \dots & \Delta_{NJ} \end{bmatrix}$ were introduced,

where $\Delta_{nj} = 0$ if y_n does not pertain to the j^{th} Gaussian component; otherwise, $\Delta_{nj} = 1$. Then, Eq. (1) becomes the following:

$$p(\vec{y}, \vec{\Delta}; \vec{\theta}) = \prod_{n=1}^N \prod_{j=1}^J [w_j p(y_n; \theta_j)]^{\Delta_{nj}} \quad (2)$$

Then, the log-likelihood is as follows:

$$\ln[p(\vec{y}, \vec{\Delta}; \vec{\theta})] = \sum_{n=1}^N \sum_{j=1}^J \Delta_{nj} \{\ln(w_j) + \ln[p(y_n; \theta_j)]\} \quad (3)$$

E-step:

The expected log-likelihood of the complete data $\vec{y}$ given the observable data $\vec{x}$ is the following:

$$\begin{aligned} Q(\vec{\theta}, \vec{\theta}^{(k)}) &= E\{\ln[p(\vec{y}, \vec{\Delta}; \vec{\theta})] | \vec{x}; \vec{\theta}^{(k)}\} \\ &= \sum_{n=1}^N \sum_{j=1}^J \int_{-\infty}^{+\infty} \Delta_{nj} \{\ln(w_j) + \ln[p(y_n; \theta_j)]\} p(\Delta_{nj}, y_n | x_n; \theta_j^{(k)}) dy_n \end{aligned} \quad (4)$$

In Eq. (4), $\vec{\theta}^{(k)}$ indicates the current estimated parameters, and k is the iteration index. Introducing a set of binary variables $\vec{z} = (z_1, \dots, z_N)$, where $z_n = 0$ when the n^{th} measurement is observable ($x_n = y_n$) and $z_n = 1$ when the n^{th} measurement is not observable ($x_n = c$), the summand in Eq. (4) can be written as follows:

$$Q(\vec{\theta}, \vec{\theta}^{(k)}) = \sum_{n=1}^N \sum_{j=1}^J [(1 - z_n) F_1 + z_n F_2] \quad (5)$$

In Eq. (5),

$$\begin{aligned} F_1 &= \gamma_j(x_n; \theta_j^{(k)}) \{\ln(w_j) + \ln[\mathcal{N}(x_n; \theta_j)]\}; \\ F_2 &= \beta_j(\theta_j^{(k)}) \left\{ \ln(w_j) + \int_{-\infty}^c \ln[\mathcal{N}(y_n; \theta_j)] \frac{\mathcal{N}(y_n; \theta_j)}{I_0(\theta_j^{(k)})} dy_n \right\} \end{aligned}$$

Here, $\mathcal{N}(x, \theta) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ is the Gaussian probability

density function, and

$$\begin{aligned} I_0(\theta_j^{(k)}) &= \int_{-\infty}^c \mathcal{N}(y, \theta_j^{(k)}) dy = \frac{1}{2} \operatorname{erfc}\left(-\frac{c - \mu_j^{(k)}}{\sqrt{2}\sigma_j^{(k)}}\right) \\ I_1(\theta_j^{(k)}) &= \int_{-\infty}^{+\infty} y \mathcal{N}(y, \theta_j^{(k)}) dy = \mu_j^{(k)} I_0(\theta_j^{(k)}) - \frac{1}{\sqrt{2\pi}} \sigma_j^{(k)} \exp\left[-\left(\frac{c - \mu_j^{(k)}}{\sqrt{2}\sigma_j^{(k)}}\right)^2\right] \\ I_2(\theta_j^{(k)}) &= \int_{-\infty}^c y^2 \mathcal{N}(y, \theta_j^{(k)}) dy \\ &= \left[(\sigma_j^{(k)})^2 + (\mu_j^{(k)})^2\right] I_0(\theta_j^{(k)}) - \frac{1}{\sqrt{2\pi}} \sigma_j^{(k)} \mu_j^{(k)} \exp\left[-\left(\frac{c - \mu_j^{(k)}}{\sqrt{2}\sigma_j^{(k)}}\right)^2\right] \\ \gamma_j(x_n; \theta_j^{(k)}) &= \frac{w_j^{(k)} \mathcal{N}(x_n; \theta_j^{(k)})}{\sum_{i=1}^J w_i^{(k)} \mathcal{N}(x_n; \theta_i^{(k)})} \\ \beta_j(\theta_j^{(k)}) &= \frac{w_j^{(k)} I_0(\theta_j^{(k)})}{\sum_{i=1}^J w_i^{(k)} I_0(\theta_i^{(k)})} \end{aligned}$$

M-step:

Computing the derivative of the auxiliary function in Eq. (5), the following iterative parameter estimation formulae can be readily derived as follows:

$$\mu_j^{(k+1)} = \frac{\sum_{n=1}^N (1 - z_n) \gamma_j(x_n; \theta_j^{(k)}) x_n + \beta_j(\theta_j^{(k)}) \frac{I_1(\theta_j^{(k)})}{I_0(\theta_j^{(k)})} \sum_{n=1}^N z_n}{\sum_{n=1}^N (1 - z_n) \gamma_j(x_n; \theta_j^{(k)}) + \beta_j(\theta_j^{(k)}) \sum_{n=1}^N z_n} \quad (6)$$

$$(\sigma_j^2)^{(k+1)} = \frac{\sum_{n=1}^N (1 - z_n) \gamma_j(x_n; \theta_j^{(k)}) (x_n - \mu_j)^2}{\sum_{n=1}^N (1 - z_n) \gamma_j(x_n; \theta_j^{(k)}) + \beta_j(\theta_j^{(k)}) \sum_{n=1}^N z_n} \quad (7)$$

$$\begin{aligned} &+ \frac{\beta_j(\theta_j^{(k)}) \left(\frac{I_2(\theta_j^{(k)})}{I_0(\theta_j^{(k)})} - \frac{2\mu_j I_1(\theta_j^{(k)})}{I_0(\theta_j^{(k)})} + \mu_j^2 \right) \sum_{n=1}^N z_n}{\sum_{n=1}^N (1 - z_n) \gamma_j(x_n; \theta_j^{(k)}) + \beta_j(\theta_j^{(k)}) \sum_{n=1}^N z_n} \\ w_j^{(k+1)} &= \frac{\sum_{n=1}^N (1 - z_n) \gamma_j(x_n; \theta_j^{(k)}) + \beta_j(\theta_j^{(k)}) \sum_{n=1}^N z_n}{N} \end{aligned} \quad (8)$$

The EM algorithm stops when the convergence criterion is satisfied or when the max iteration is achieved. After convergence, the estimated parameters are as follows:

$$\mu_j^{(k+1)} \approx \mu_j^{(k)} := \hat{\mu}_j; \sigma_j^{(k+1)} \approx \sigma_j^{(k)} := \hat{\sigma}_j; w_j^{(k+1)} \approx w_j^{(k)} := \hat{w}_j \quad (9)$$

Given equations (6÷8), both observable and censored mixture data contribute to the estimates. Moreover, if the data

are complete data ($c = -\infty$), then these equations are reduced to the standard EM algorithm for the mixture Gaussian data [5]. On the other hand, if the data have a single Gaussian distribution and suffered from the censoring problem, by setting $J = 1$, over three formulae become those reported in [3]. This means that the proposal can handle both the censoring and multi-component problems presented in the Wi-Fi RSSI data.

The online classification and positioning phase

This sub-section utilizes the Maximum a Posteriori (MAP) method to perform the classification. For each reference position l_k , the parameters of the C-GMM class conditional density $p_Y(y|l_k)$ of RSSI measurements are estimated using equations (6-9). During online classification, the MAP is used to estimate the user's location. First, the posterior is calculated as follows:

$$P(\ell_k|\vec{x}) = \frac{\prod_{i=1}^{N_{AP}} p(x_i|\ell_k)P(\ell_k)}{\sum_{k'=1}^K \prod_{i=1}^{N_{AP}} p(x_i|\ell_{k'})P(\ell_{k'})} \quad (10)$$

In Eq. (10), K and N_{AP} represent the total number of RPs and APs, respectively. x_i is the online measurement from i^{th} AP, and $\vec{x}$ is the set of x_i ($i=1 \div N_{AP}$). It has been considered that the RSSI measurements of different APs are independent, and the prior $P(\ell_k)$ is equal for all locations.

The likelihood $p(x_i | l_k)$ can be calculated as follows:

$$p(x_i | \ell_k) = \begin{cases} \sum_{j=1}^J \hat{w}_{k,i,j} \mathcal{N}(x_i; \hat{\theta}_{k,i,j}), & \text{if } x_i > c \\ \sum_{j=1}^J \hat{w}_{k,i,j} I_0(c; \hat{\theta}_{k,i,j}), & \text{if } x_i = c \end{cases} \quad (11)$$

In Eq. (11), $\hat{\theta}_{k,i,j}$, $\hat{w}_{k,i,j}$ are estimated parameters at the k^{th} RP of the i^{th} AP in the offline phase.

The estimated position of the mobile object is obtained by the following:

$$\hat{\ell}(\vec{x}) = \frac{\sum_{k \in P} \ell_k p(\ell_k|\vec{x})}{\sum_{k \in P} p(\ell_k|\vec{x})} \quad (12)$$

Simulation results and discussion

Parameter estimation

To evaluate the effectiveness of the proposed EM algorithm, complete data $\vec{y}$ with the following parameters has been generated (true parameters):

$$N = 1000; J = 2; [w_1, w_2] = [0.5, 0.5]; [\sigma_1, \sigma_2] = [3, 4];$$

$$[\mu_1, \mu_2] = [-90, -80].$$

Observable data $\vec{x}$ are performed censoring as follows: $x_n = \max(y_n, c)$. The censoring threshold c was changed from $\mu_1 - 2\sigma_1$ to $\mu_1 + 2\sigma_1$. Table 1 indicates the mean of Kullback Leibler (KL) divergence [9] between true parameters and estimated parameters after 1,000 experiments.

Table 1. Parameter estimation compared by mean of KL using the Monte Carlo sampling method.

C (dBm)	Standard EM algorithm for GMM [5]	After [3]	Proposed EM algorithm for C-GMM
-96	0.0018	0.0664	0.0016
-93	0.0329	0.0679	0.0031
-90	3.1491	0.0798	0.0092
-87	5.6358	0.0886	0.0124
-84	7.2847	0.0972	0.0473

As is evident, when $c = \mu_1 - 2\sigma_1 = -96$, data nearly do not suffer from censoring (almost complete); the proposal and the standard EM algorithm for GMM produced the same results. However, when c changes from -93 to -84, the proposed EM algorithm introduces improved results.

Positioning accuracy

To evaluate the effectiveness of the proposed approach in the Wi-Fi fingerprinting-based IPS, a floor plan with 100 RPs (small red circles) and 10 APs (green circles) has been generated, as illustrated in Fig. 3. The first experiment was setup as follows: In the offline phase, 400 measurements are collected for each RP. The measured data at 50% of the training positions (RPs) follow the single Gaussians, randomly; the rest follows the GMMs, and the number of components is $J = 2, 3, 4, 5, 6$, respectively (10% for each model). Measured data at RPs were computed by the log-distance path loss model and by adding a Gaussian with a mean of zero and a standard deviation of two for reflecting the fluctuation of the signal [10]. The limited sensitivity of the Wi-Fi sensor was set to -100 dBm ($c = -100$). The radio map was developed by employing equations (6-9) with $J=4$ and methods, which were proposed in [3, 5]. For the online localization phase, 100 simulations were performed. Each simulation, one online measurement per position was generated in the same scenarios with the training data, and the MAP method was used for computing the final position estimate, as presented in sub-section 2.3.

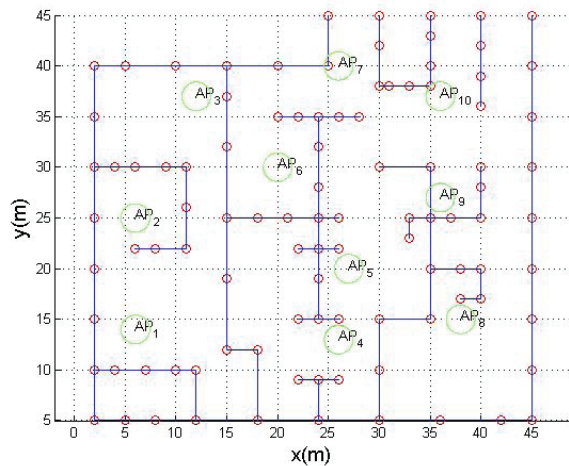


Fig. 3. The computer-generated floor plan.

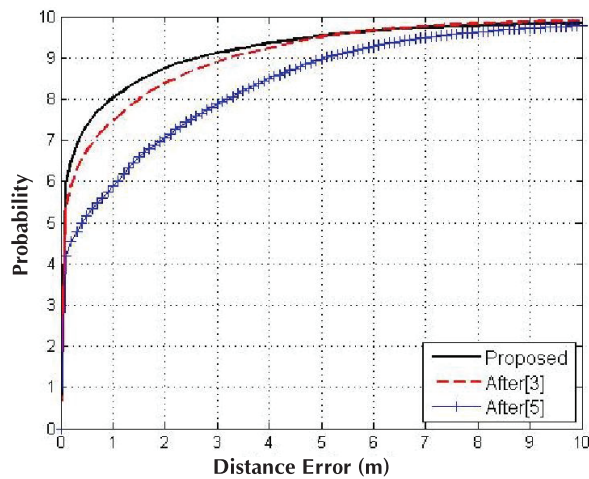


Fig. 4. Comparison of positioning results when the observable training data ratio was 69.77%, the observable online data ratio was 69.82%.

Figure 4 illustrates the probability that the positioning error is lower than a specific distance. The estimated position is the specific position at which the mobile target had collected the online measurements. The plots in the figure are computed by averaging the positioning results of 100 simulations. It is evident that the proposed method outperforms the others, particularly when the error distance is smaller than 2 meters. In term of Wi-Fi fingerprinting-based indoor positioning, while the proposal in [3] is unable to solve the multi-component problem in the observed data, authors of [5] have not considered the censoring problem in their research. This simulation result demonstrates that the proposal can cope with the phenomena presented in the measured Wi-Fi RSSI data.

In the second experiment, data were gathered in the same

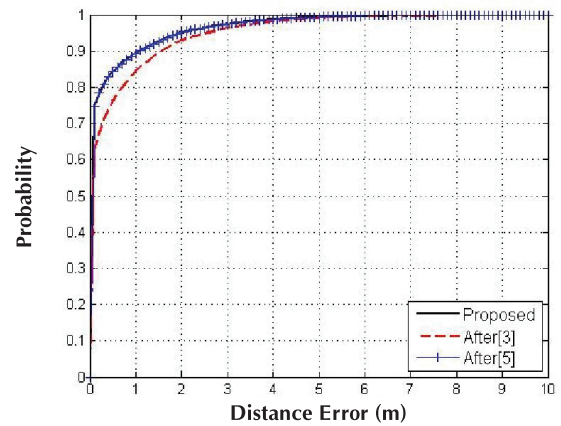


Fig. 5. Comparison of positioning results when the observable training and online data ratio were 100%.

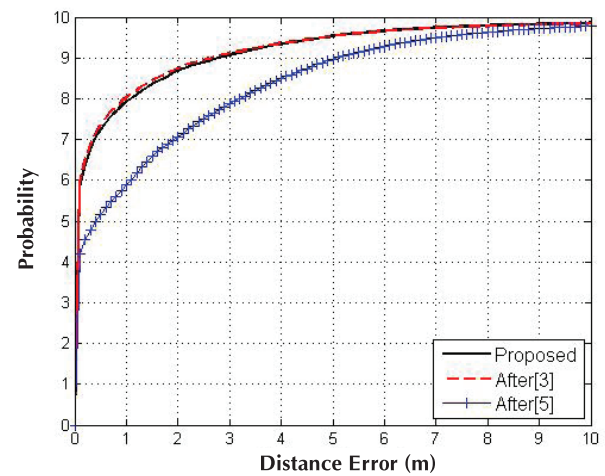


Fig. 6. Comparison of positioning results when the observable training data ratio was 69.24% the observable online data ratio was 69.74%.

manner as in the first experiment, but the limited sensitivity of the Wi-Fi sensor was changed to a value which is smaller than the smallest value of collected Wi-Fi RSSI. This means that collected data at all RPs of all APs are complete. Fig. 5 validated that the proposal and the standard EM algorithm for GMM [5] presented the same results, which means that the C-GMM is still appropriate to model complete mixture data.

The experiment setup for the results in the Fig. 6 is the same as in the first experiment; however, the measured data all RPs of all APs follow the single Gaussian distribution and exert an influence on censoring. It is apparent that the approach nevertheless works as effectively as the method proposed in [3].

Moreover, Table 2 indicates the properties of the Mean Distance Error (MDE) of the three experiments.

Table 2. MDE (m).

	After [3]	After [5]	Proposed
Experiment 1	1.3428	1.6321	1.0452
Experiment 2	1.0402	0.4856	0.4863
Experiment 3	0.9920	1.7395	1.0012

Conclusions

This paper has presented and analyzed an EM algorithm for estimating the parameters of the GMM in the presence of censored mixture data. The results have demonstrated that the algorithm delivers less biased and more efficient estimates relative to existing methods. Further, it has illustrated the enhancement of the Wi-Fi fingerprinting-based indoor positioning system when the novel method was employed. Experimental results on artificial data verify that the proposal produces optimal accuracy of positioning among available approaches. Future research will make substantial use of labor work for gathering real data and evaluate the proposed method. In addition, reducing the computational cost in the online phase and using sensors on the portable devices to predict the current position of the moving objects can significantly enhance the real-time performance of the IPS.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

[1] L. Mainetti, L. Patrono, and I. Sergi (2014), "A survey on indoor positioning systems", *Proceedings of 22nd Int. Conf. on*

Software, Telecommunications and Computer Networks (SoftCOM).

[2] K. Kaemarungsi and P. Krishnamurth (2004), "Modeling of indoor positioning systems based on location fingerprinting", *Proceedings of the INFOCOM*, Hong Kong.

[3] K. Hoang and R. Haeb-Umbach (2013), "Parameter estimation and classification of censored Gaussian data with application to Wi-Fi indoor positioning", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver.

[4] K. Hoang, J. Schmalenstroer, and R. Haeb-Umbach (2015), "Aligning training models with smartphone properties in Wi-Fi fingerprinting based indoor localization", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brisbane.

[5] M. Alfakih, M. Keche, and H. Benoudnine (2015), "Gaussian mixture modeling for indoor positioning Wi-Fi systems", *3rd Int. Conf. on Control, Engineering and Information Technology (CEIT)*, Tlemcen, Algeria.

[6] Jiayou Luo and Xingqun Zhan (2014), "Characterization of smart phone received signal strength indication for WLAN indoor positioning accuracy improvement", *Journal of Networks*, **9(3)**, pp.739-746.

[7] A.P. Dempster, N.M. Laird, and D.B. Rubin (1977), "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society, series B (Methodological)*, pp.1-38.

[8] G. Lee and C. Scott (2012), "EM algorithms for multivariate Gaussian mixture models with truncated and censored data", *Computational Statistics & Data Analysis*, **56(9)**, pp.2816-2829.

[9] J.R. Hershey and P.A. Olsen (2007), "Approximating the Kullback Leibler divergence between Gaussian mixture models", *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Honolulu.

[10] C. Gustafson, T. Abbas, D. Bolin, and F. Tufvesson (2015), "Statistical modeling and estimation of censored pathloss data", *IEEE Wireless Comm. Letters*, **4(5)**, pp.569-572.

The effects of temperature, feed ratio, and reaction time on the properties of copolymer PLA-PEG-PLA

Viet Linh Nguyen-Vu^{1,2*}, Mai Anh Pham¹, Dai Phu Huynh^{1,2}

¹National Key Laboratory of Polymer and Composite Materials, Ho Chi Minh University of Technology

²Faculty of Materials of Technology, Ho Chi Minh University of Technology

Received 30 July 2018; accepted 23 October 2018

Abstract:

Triblock copolymers were polymerised by the ring-opening reaction of D,L-lactide in the presence of poly(ethylene glycol) (PEG), with number-average molecular weight (M_n) of 1500 and 2050 g/mol, using $\text{Sn}(\text{Oct})_2$ as a catalyst. The influences of the reaction time, the ratio of PEG and Poly(D,L-lactic acid) (PLA), and PEG types on structure and sol-gel phase transition of PLA-PEG-PLA triblock copolymers were investigated. Optimal polymerisation parameters were obtained, such as reaction time of 18 hours, a catalyst amount of 1.3%, and PEG/PLA ratio of 1/1.7, PEG ($M_n=1500$); the efficiency of the triblock synthesis was 42.3%. The properties of PLA-PEG-PLA copolymers were evaluated by analytical methods such as proton nuclear magnetic resonance ^1H NMR spectroscopy, gel permeation chromatography (GPC), and the sol-gel state transition at varying temperature. The results show that the triblock was successfully synthesised and its hydrogel had capability of the sol-gel state transition when the temperature changed. The PLA-PEG-PLA copolymer in aqueous solution is a thermo-sensitive hydrogel that can be used for drug and protein delivery systems or triblock denaturation applications for commercial purposes.

Keywords: poly(lactide), state transition, thermo-sensitive hydrogel, triblock.

Classification numbers: 2.1, 2.3

Introduction

Thermo-responsive hydrogels have attracted extensive attention in the field of biodegradable materials because they exist as a solution at low temperature and change to a gel at the human physiological condition. This is convenient for the administration of injections [1-5]. Thermo-responsive PEG hydrogels are one of these hydrogels and are applied in drug delivery and tissue engineering thanks to the sol-gel transition in different temperature conditions and the biocompatibility of PEG [2, 6-8]. Thermo-sensitive hydrogels are synthesised based on two blocks: the hydrophilic block (PEG) and hydrophobic blocks such as poly(lactic-co-glycolic acid) (PLGA), polycaprolactone (PCL), and PLA. The amphiphilic micelles in the aqueous solution contain two parts: hydrophobic blocks form the core and the surrounding micelles are hydrophilic blocks. The sol-gel transition of triblock copolymers in aqueous solution depends on the balance between the hydrophilic/hydrophobic block and others. Gelation was obtained with high temperature due to the generation of the micelles and the hydrophobic bridge between micelles [1-3, 7-10]. These thermo-sensitive hydrogels could encapsulate some types of drugs, such as bovine serum albumin (BSA), paclitaxel, dexamethasone, thymopentine, insulin, and the like, and prolong the release time of the drugs [7, 10-13]. One thermo-sensitive hydrogel that is commercially available is pluronic. It is used for drug delivery systems, implantation, and scaffold or denaturation applications [2].

In this study, we conducted the triblock synthesis by varying the reaction time (14 hours, 18 hours, and 22 hours); the PEG/PLA ratio (1/1.7 and 1/2.2); the proportion of the catalyst $\text{Ti}(\text{II})$ 2-ethylhexanoate (1.3% and 1.6% catalyst); and the PEG types (PEG-2050 and PEG-1500). Finally, we optimised the polymerisation parameters for high efficiency of the triblock. The temperature-sensitive hydrogels of PLA-PEG-PLA copolymers in aqueous solution were in a

*Corresponding author: Email: nguyenvuvielinhh@hcmut.edu.vn

sol state at room temperature, and they changed to a gel state when the temperature increased to human physiological temperature.

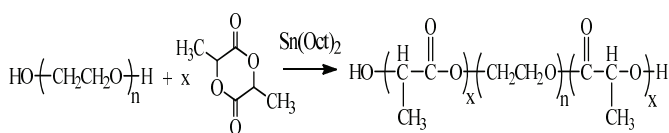
Materials and methods

Materials

D,L-Lactide (synonym: 3,6-Dimethyl-1,4-dioxane-2,5-dione), from Sigma Aldrich, is 99% pure. Poly(ethylene glycol) $M_n=1500$ Da (PEG-1500), $M_n=2050$ Da (PEG-2050), and Tin(II) 2-ethylhexanoate ($\text{Sn}(\text{Oct})_2$) were purchased from Sigma Aldrich.

The PLA-PEG-PLA triblock polymerisation

The PLA-PEG-PLA triblock copolymers were formed by means of a ring-opening polymerisation of D,L-Lactide, initiated by PEG-1500 or PEG-2050, using $\text{Sn}(\text{Oct})_2$ as a catalyst (Scheme 1). The typical bulk polymerization was conducted as follows: PEG and $\text{Sn}(\text{Oct})_2$ were added to a two-neck flask and magnetically stirred at 110°C for 3-4 hours in vacuum environment to remove moisture. D,L-Lactide was added to the flask and then dried at 75°C in vacuum environment for one hour. Next, the flask was evacuated and filled with nitrogen three times to keep moisture away from the reaction. The reaction was maintained at 135°C for 14-22 hours. After reacting, the mixture was cooled to room temperature and dissolved in diethyl ether to precipitate the copolymer. After the evacuation, the purified product was collected at 45°C for 48 hours [5, 12-14].



Scheme 1. Ring-opening polymerisation of D,L-Lactide in the presence of PEG with the catalyst $\text{Sn}(\text{Oct})_2$.

Copolymer characterisation

The structures and composition of the triblock copolymers were determined by means of ^1H NMR spectroscopy. ^1H NMR spectra were recorded at room temperature using a 500 MHz spectrometer (Bruker, USA). The internal standard was tetramethylsilane (TMS) and deuteriochloroform (CDCl_3) was used as the solvent.

The M_n and the polydispersity index (PDI) of the triblock copolymers were determined by GPC (PL-GPC 50 Plus, Agilent Technology, USA). The mobile solvent was chloroform with the flow rate of 1.0 ml/min (30°C , PEG as standard).

Sol-gel phase transition measurement

The sol (flow)-gel (no flow) phase transition of the triblock copolymer in the aqueous solution was determined using the inverting vials method (with 4 ml tightly screw capped vials with a 10 mm inner diameter). The sol (flow) or gel (no flow) condition was determined in 1 minute. Briefly, each sample at a given concentration was absolutely dissolved in phosphate-buffered saline (PBS) (10 mM, pH 7.4) at 0°C overnight and was continuously stabilised at 5°C for 4 hours. The vials were then placed in a water-bath and were heated from 15 to 80°C at the step $2^\circ\text{C}/\text{time}$. The sol-gel transition temperature was determined by inverting the vial after maintaining it at a constant temperature for 10 min [1, 6, 14].

Results

Characteristics of the PLA-PEG-PLA copolymer using ^1H NMR

The PLA-PEG-PLA triblock copolymers were analysed using ^1H NMR spectra, as displayed in Fig. 1. The shift of 4.267 ppm (3) and 5.162 ppm (4) was assigned to the methine hydrogen ($-\text{CH}-\text{CO}$) and ($-\text{CH}-\text{COO}$) of the lactide unit, respectively. The presence of the methyl hydrogen ($-\text{CH}_3$) of the D,L-lactide unit was observed at $\delta=1.533$ ppm (1). The methylene hydrogen ($-\text{CH}_2-$) of PEG was recorded at $\delta=3.504$ ppm (2).

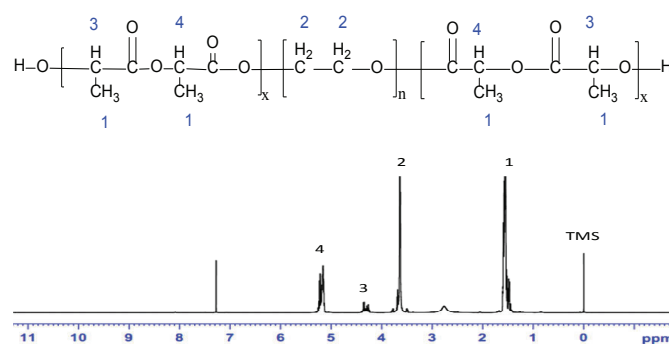


Fig. 1. ^1H NMR spectrum (CDCl_3) of the copolymer.

Effect of the reaction time on triblock structure and sol-gel transition of the hydrogel

The chain length of triblock 2050-14h (reaction time of 14 hours), $M_n=4448$ g/mol, was shorter than the triblock chain 2050-22h (reaction time of 22 hours), $M_n=4826$ g/mol. In addition, when the reaction time was 18 hours, M_n (4844 g/mol) of copolymer 2050-18h was higher than that of triblock 2050-22h (Table 1). Increasing the reaction time from 18 hours to 22 hours not only reduced the length of copolymer (by 18 g/mol) but also slightly

increased the PDI, from 1.18 to 1.19. The PEG/PLA ratio decreased from 1/1.36 to 1/1.35 when the reaction time was prolonged for four more hours (from 18h to 22h).

Table 1. Characteristics of PLA-PEG-PLA copolymer using ^1H -NMR and GPC.

Sample name	PLA-PEG-PLA M_n^a	M_n^a	PDI ^b (M_n/M_w)	PEG/PLA (wt/wt) ^a
2050-14h	1199-2050-1199	4448	1.21	1/1.17
2050-18h	1397-2050-1397	4844	1.18	1/1.36
2050-22h	1388-2050-1388	4826	1.19	1/1.35

a: determined by ^1H -NMR; b: determined by GPC.

The sol (flow) - gel (no flow) transition and the precipitation (separate two phase, including an aqueous phase and polymer in the bottom of vials) of the copolymer PLA-PEG-PLA in PBS are displayed in Fig. 2.

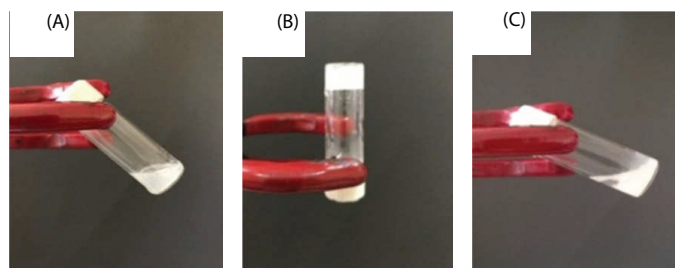


Fig. 2. Sol-gel transition of hydrogel PLA-PEG-PLA in PBS (A) sol, (B) gel, and (C) precipitation.

The sol-gel phase diagram of the copolymer 2050-18h in PBS shows that the hydrogels which were created from 20% and 30% triblock concentration were in a sol state below 38°C. The hydrogel of the 35% copolymer appeared as a gel state at about body temperature and was changed to precipitation state at 74°C. In addition, the 2050-14h hydrogel occurred in a sol state at all given concentrations, from 25% to 40% copolymer (Fig. 3). Therefore, the optimum reaction time was 18 hours; this was retained for the next experiments.

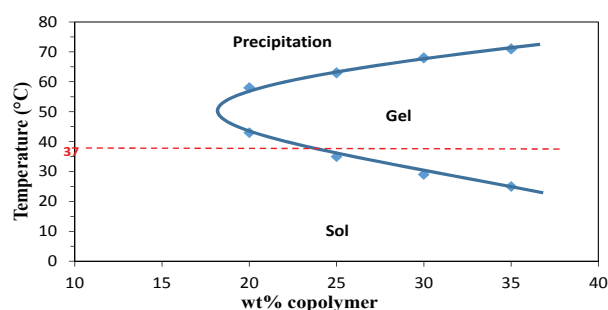


Fig. 3. Sol-gel phase diagram of hydrogel PLA-PEG-PLA (2050-18h) in PBS.

Effect of the PEG/PLA ratio and PEG types on the triblock structure and sol-gel transition of hydrogel

When the PEG/PLA ratio increased from 1/1.7 to 1/2.2, the M_n of triblock 2050-1.7 changed from 4844 Da to 5654 Da (M_n triblock 2050-2.2), especially, the length of PLA chain added to 405 Da. However, the reaction efficiency of triblock 2050-2.2 was reduced to 38.9%. In the case of PEG 1500, the triblock's M_n and reaction efficiency (36.5%) were lower than that of triblock 2050-2.2 (38.9%). The PEG/PLA ratio of triblock 1500-2.2 (1.90) was higher than that of triblock 2050-2.2 (1.76) although their designed PEG/PLA ratios were similar. In addition, the PDI of triblock 2050-2.2 (PDI=1.161) was the lowest while its M_n was the biggest (Table 2).

Table 2. The reaction efficiency, M_n , and PEG/PLA ratios of the PLA-PEG-PLA copolymer using PEG-2050 and PEG-1500.

Sample	PLA-PEG-PLA (M_n) ^a	M_n^a	PDI ^b (M_n/M_w)	PEG/PLA ^a (wt/wt)	Reaction efficiency
2050-1.7	1397-2050-1397	4844	1.184	1/1.36	42.3%
2050-2.2	1802-2050-1802	5654	1.161	1/1.76	38.9%
1500-2.2	1424-1800-1424	4348	1.181	1/1.90	38.5%

a: determined by ^1H -NMR; b: determined by GPC.

According to Fig. 4, the gel phase area of hydrogel using triblock 2050-2.2 was broader than that of hydrogel 2050-1.7, and the gel temperature range was lower than that of the lactide ratio, at 1.7. The critical gel temperature (CGT) of hydrogel triblock 2050-2.2 was lower than the hydrogel 2050-1.7. Furthermore, the hydrogel triblock 2050-1.7 could not convert from sol to gel at 37°C although the CGT of 2050-2.2 could convert from sol to gel. At 30% and 35% copolymer concentration, the hydrogel of triblock 2050-2.2 was converted from sol to gel phase at 37°C and separated to two phases (precipitation) at 72°C.

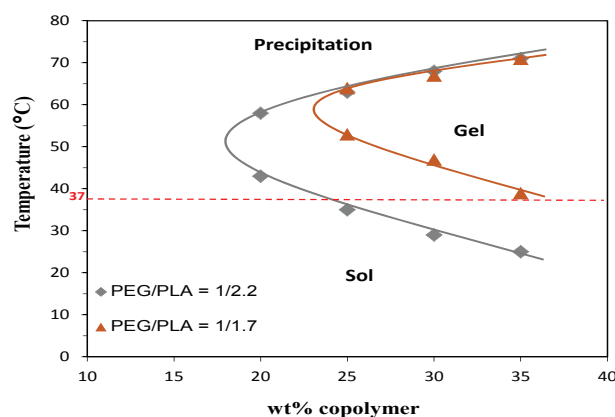


Fig. 4. The sol-gel phase diagram of the PLA-PEG-PLA hydrogel with a comparable ratio of PEG/PLA, 1/2.2 and 1/1.7 (PEG 2050, reaction time: 18 hrs).

As illustrated in Fig. 5, the hydrogel of 1500-2.2 triblock was in a gel state at 25-35%wt copolymer concentration at a low temperature (from 17 to 25°C). In the case of the 2050-2.2 hydrogel (20-35%wt triblock in PBS solution), the gel state was obtained at a higher temperature, more than 25°C for the 35% copolymer, and more than 42°C for the 20% copolymer. At lower copolymer concentration, at 25%wt, the hydrogel of triblock 1500-2.2 was separated to sedimentation at 37°C. Moreover, the ability sol-gel phase transition of triblock 2050-2.2 was better than that of triblock 1500-2.2 and, as a result, the gel area was broadened.

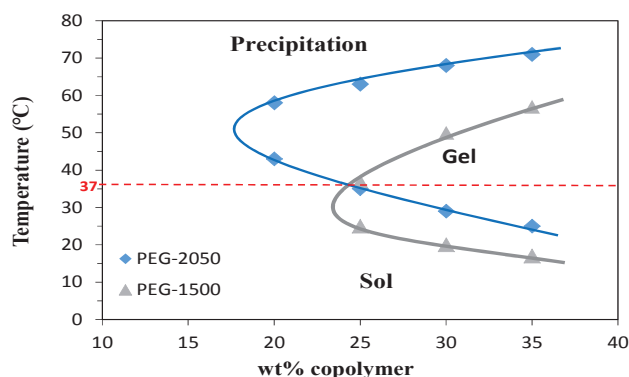


Fig. 5. The sol-gel phase diagram of PLA-PEG-PLA hydrogel with different PEG types, $M_n=1500$ and 2050 g/mol (reaction time: 18h, PEG/PLA ratio: 1/1.2).

Discussion

We succeeded in synthesizing triblock PLA-PEG-PLA using D,L-lactide, PEG and $\text{Sn}(\text{Oct})_2$ catalyst. The ^1H NMR spectrum showed the specific peak (3) and peak (4) of methine hydrogen (CH-CO) and (CH-COO). The PLA-PEG-PLA hydrogel underwent the sol-gel transition at determined temperatures and copolymer concentrations. Furthermore, we investigated the effect of reaction time, PEG/PLA ratios, and PEG types on the properties and the sol-gel transition of the hydrogel of triblock. Time influenced the number of D,L-Lactides reacting with PEG, in the presence of $\text{Sn}(\text{Oct})_2$ catalyst. The reason for this is that the length of the triblock prevented lactides from attaching on the PEG backbone and generated the branch on the main polymer chain [5, 8, 10, 12, 14]. Moreover, with a short reaction time (14 hours), D,L-Lactide could not add to the PEG backbone more effectively. When reaction time was increased, the number of D,L-Lactides attached to PLA-PEG-PLA increased. The reaction temperature caused the differences in the PEG/PLA ratio and the length of PLA blocks [8]. In the hydrogel triblock structure, PLA blocks were hydrophobic and PEG blocks were hydrophilic, similar to the experiment based on the hydrophobic PLGA or PCL block, therefore the micelles

and bridging micelles were formed in aqueous solution. At a high temperature, the number of micelles and bridging micelles increased, so that the hydrogel changed from a sol state to gel. With continuously increasing temperature, the thermo-vibration of H_2O molecules was strong and caused the precipitation of the hydrogel PLA-PEG-PLA copolymer [12, 15, 16]. When the reaction time was less than 14 hours, the length of PLA block was shortened, leading to a decrease in the hydrophobicity of the hydrogel copolymer. As a result, the sample 2050-14h occurred in a sol state at all given concentrations and temperature.

In the second part of the survey, we changed the PEG/PLA ratio, increasing it from 1/1.7 to 1/2.2, and the triblock was synthesized with a low polydispersity index. However, the length of triblock caused the lack of space to lactide to connect with the PEG block and reduced the reaction efficiency. Moreover, the hydrogel of triblock 2050-2.2 had a gel phase area broader than that of triblock 2050-1.7 due to development of the length of hydrophobic block (PLA) in the copolymer structure. The reason for this phenomenon is that the establishment of bridging connections between micelles and high number of micelles in the triblock 2050-2.2 hydrogel structure [16]. Increasing the amount of lactide for polymerisation, resulted in the M_n increasing and PDI decreasing. This proved that the PEG/PLA ratio was 1/2.2 - the triblock product had fewer impurities and had chemical homogeneity. On the other hand, reaction efficiency decreased as the lactide ratio increased to 2.2 because the polymer chain entanglements of the long triblock were formed, leading to the PLA block being difficult to connect with the PEG block. The amount of lactide for polymerisation was increased, the M_n increased and PDI decreased. In addition, the higher lactide ratio increased sublimation and reduced reaction efficiency.

Finally, we changed the molecular weight of PEG block (M_n 2050 g/mol and 1500 g/mol) and fixed the designed PEG/PLA ratio. Triblock PLA-PEG(1500)-PLA had a PEG/PLA ratio that was greater than that of triblock PLA-PEG(2050)-PLA. The reason for this is that the small molecular chains of PEG were effectively added to PLA or lactide. The hydrogel 1500-2.2 triblock had a lower CGT than the hydrogel 2050-2.2 triblock, with the consequences that it became a gel and could not be injected at room temperature (25-32°C). The length of the hydrophilic block (PEG) influenced the balance of the hydrophobic (PLA) and hydrophilic parts (PEG) in the hydrogel structure. When the hydrophobic part was larger the hydrophilic part, the ability to form micelles in an aqueous solution was easier [12, 15]. At low M_n of PEG (1500 g/mol) the micelles were formed more easily than with PEG (2050 g/mol), so that the gel state of hydrogel triblock 1500-2.2 was obtained at a

lower temperature than that of hydrogel triblock 2050-2.2. The sol-gel transition area depended on the total molecular weight of the block copolymer. This effect related to the amphiphilic copolymers with low molecular weight more easily forming micelles than those with high molecular weight [15].

Conclusions

In this study, the influences on the phase diagrams of the reaction times, the PEG/PLA ratio, and the PEG types were investigated. The optimal reaction time was 18 to 22 hours for the synthesis of triblock PLA-PEG-PLA. A lower critical gel concentration was found when we increased PLA/PEG ratios. The sol-gel transition diagrams were shifted to higher temperatures by increasing the molecular weight of the triblock without changing the composition of the copolymers. The results of this study suggest that PLA-PEG block copolymers can be the next step in the development of polymeric drug delivery.

ACKNOWLEDGEMENTS

This research was funded by Vietnam Government according to the co-project between the Ministry of Science and Technology (Vietnam) and the Ministry of Science, ICT and Future Planning (South Korea) under grant number NDT.27.KR/17.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] B. Jeong, Y.H. Bae, and S.W. Kim (1999), "Thermoreversible gelation of PEG-PLGA-PEG triblock copolymer aqueous solutions", *Macromolecules*, **32**, pp.7064-7069.
- [2] C.T. Huynh, M.K. Nguyen, and D.S. Lee (2011), "Injectable block copolymer hydrogels: achievements and future challenges for biomedical applications", *Macromolecules*, **44**, pp.6629-6636.
- [3] H. Nouailhas, A. El Ghzaoui, S. Li, and J. Coudane (2011), "Stereocomplex-induced gelation properties of polylactide/poly(ethylene glycol) diblock and triblock copolymers", *Journal of Applied Polymer Science*, **122**, pp.1599-1606.
- [4] H. Cui, J. Shao, Y. Wang, P. Zhang, X. Chen, and Y. Wei (2013), "PLA-PEG-PLA and its electroactive tetraaniline copolymer as multi-interactive injectable hydrogels for tissue engineering", *Biomacromolecules*, **14**, pp.1904-1912.
- [5] H. Danafar, K. Rostamizadeh, S. Davaran, and M. Hamidi (2014), "PLA-PEG-PLA copolymer-based polymersomes as nanocarriers for delivery of hydrophilic and hydrophobic drugs: preparation and evaluation with atorvastatin and lisinopril", *Drug Development and Industrial Pharmacy*, **40**, pp.1411-1420.
- [6] N.A. Peppas, K.B. Keys, M. Torres-Lugo, and A.M. Lowman (1999), "Poly(ethylene glycol)-containing hydrogels in drug delivery", *J. Control Release*, **62**, pp.81-87.
- [7] C.-C. Lin and K.S. Anseth (2009), "PEG hydrogels for the controlled release of biomolecules in regenerative medicine", *Pharmaceutical Research*, **26**, pp.631-643.
- [8] E.M. Saffer, G.N. Tew, and S.R. Bhatia (2011), "Poly(lactic acid)-poly(ethylene oxide) block copolymers: new directions in self-assembly and biomedical applications", *Current Medicinal Chemistry*, **18**, pp.5676-5686.
- [9] M.K. Nguyen and D.S. Lee (2010), "Injectable biodegradable hydrogels", *Macromolecular Bioscience*, **10**, pp.563-579.
- [10] G.M. Zentner, R. Rathi, C. Shih, J.C. McRea, M.-H. Seo, H. Oh, et al. (2001), "Biodegradable block copolymers for delivery of proteins and water-insoluble drugs", *Journal of Controlled Release*, **72**, pp.203-215.
- [11] Y.J. Kim, S. Choi, J.J. Koh, M. Lee, K.S. Ko, and S.W. Kim (2001), "Controlled release of insulin from injectable biodegradable triblock copolymer", *Pharmaceutical Research*, **18**, pp.548-550.
- [12] G. He, L.L. Ma, J. Pan, and S. Venkatraman (2007), "ABA and BAB type triblock copolymers of PEG and PLA: a comparative study of drug release properties and "stealth" particle characteristics", *International Journal of Pharmaceutics*, **334**, pp.48-55.
- [13] Y. Zhang, X. Wu, Y. Han, F. Mo, Y. Duan, and S. Li (2010), "Novel thymopentin release systems prepared from bioresorbable PLA-PEG-PLA hydrogels", *International Journal of Pharmaceutics*, **386**, pp.15-22.
- [14] D.G. Abebe and T. Fujiwara (2012), "Controlled thermoresponsive hydrogels by stereocomplexed PLA-PEG-PLA prepared via hybrid micelles of pre-mixed copolymers with different PEG lengths", *Biomacromolecules*, **13**, pp.1828-1836.
- [15] H.T. Lee and D.S. Lee (2002), "Thermoresponsive phase transitions of PLA-block-PEO-block-PLA triblock stereo-copolymers in aqueous solution", *Macromolecular Research*, **10**, pp.359-364.
- [16] M.S. Shim, H.T. Lee, W.S. Shim, I. Park, H. Lee, T. Chang, et al. (2002), "Poly(D,L-lactic acid-co-glycolic acid)-b-poly(ethylene glycol)-b-poly (D,L-lactic acid-co-glycolic acid) triblock copolymer and thermoreversible phase transition in water", *Journal of Biomedical Materials Research*, **61**, pp.188-196.

Using a physical model to determine the hydrodynamic dispersion coefficient of a solution through a horizontal sand column

Le Anh Tuan^{1*} and Guido Wyseure²

¹College of Environment and Natural Resources, Can Tho University, Vietnam

²Laboratory for Land and Water Management, Faculty of Biosciences Engineering, Catholic University of Leuven, Belgium

Received 9 November 2018; accepted 12 January 2019

Abstract:

Miscible displacement can be understood as a physical process in a porous medium whereby two or more fluids fully dissolve into each other when a fluid mixes and goes into the pore space occupied by other fluids without the existence of an interface. A physical model was made in Can Tho University, which included an electrical current system connecting nine groups of four-electrode probes for measuring the electrical conductivity of a potassium chloride solution flowing through a horizontal sand column placed in a firm frame. The experiments were performed with different volumetric flow rates and three types of sand (fine, medium and coarse). The breakthrough curves were analysed, and then the hydrodynamic dispersion coefficients were calculated. The hydrodynamic dispersion coefficient was one of the hydraulic and solute transport parameters used to design a constructed subsurface flow wetland. The research proves that the flows were laminar, and that mechanical dispersions dominated over molecular diffusions and that the dispersions were large enough to cause combined mixing and flowing processes.

Keywords: breakthrough curves, electrical conductivity, four-electrode probes, hydrodynamic dispersion coefficients, physical model.

Classification number: 2.3

Theory

The main mechanisms governing transport in porous media are convection (advection), diffusion, and mechanical dispersion [1]. Partitioning processes and decaying processes also affected to transport mechanisms. Miscible pollutant transport processes are shown in more detail in Fig. 1.

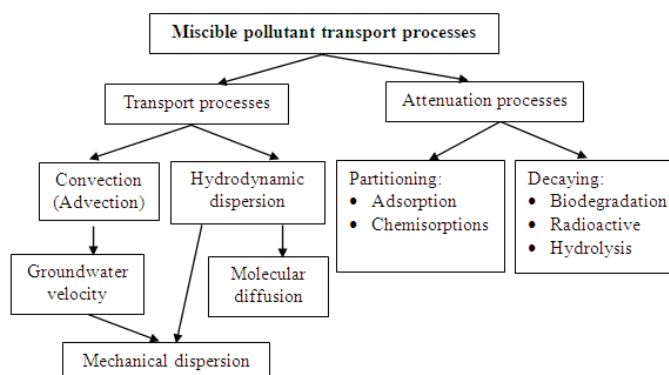


Fig. 1. Flowchart of pollutant transport processes.

The convection-dispersion equation (CDE) describes the transport of solutes through porous media, as in a constructed wetland. Breakthrough experiments with tracers in a horizontal sand column can be used to determine the solute transport parameters for the CDE. The important underlying assumptions for the mathematical analysis are that the sand in the experimental column is homogeneous and that the transport parameters remain constant during the experiment and that, therefore, the solute transport is a linear process. It is necessary to know the transport parameters and the relationship between dispersion and velocity in the solution. The transfer function method is proposed to determine the transport parameters from the solute breakthrough data [2, 3].

*Corresponding author: latuan@ctu.edu.vn

The phenomenon of a solute spreading and occupying an ever-increasing portion of the flow domain in a porous media is called hydrodynamic dispersion. It causes dilution of the solute and is composed of two different processes: mechanical dispersion (or hydraulic dispersion) and molecular diffusion. Hydraulic dispersion refers to the spreading of a tracer due to microscopic velocity variations within individual pores. Molecular diffusion is the net transfer of mass (of a chemical species) by random molecular motion. While these two processes are different in nature, they are in fact completely inseparable because they occur simultaneously. The process of hydrodynamic dispersion is illustrated in Fig. 2.

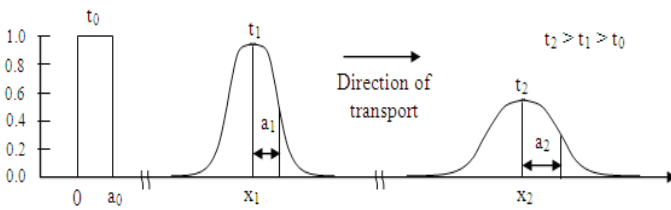


Fig. 2. Spreading of a solute slug with time due to convection and dispersion [4].

The CDE was developed to predict the average concentration of a tracer solute transported in a porous media [5]. It can include adsorption, degradation, and chemical transformation. The CDE for a conservative solute can be expressed in mathematical form as:

$$R \frac{\partial C}{\partial t} = D_h \frac{\partial^2 C}{\partial x^2} - V_{\text{pore}} \frac{\partial C}{\partial x} \quad (1)$$

where the variables t and x represent time and the spatial direction coordinates of the flow, respectively. R is the retardation factor ($R = 1$ means no interaction between the solute and the solid matrix in porous media), C is the solute concentration (mg/l), V_{pore} is the pore water velocity (m/s), and D_h is the coefficient of hydrodynamic dispersion (m^2/s) in the longitudinal direction (i.e. along the x -flow direction). Analytical solutions to the CDE have been developed for a number of specific initial and boundary conditions. Solute transport parameters are estimated by matching analytical solutions to the CDE or alternative models with measured breakthrough curves (BTC) from miscible displacement experiments [6].

By analysing the solution under steady-state flow conditions in the soil column, the initial and boundary conditions for the solute concentration distribution are obtained as follows:

$$\begin{cases} C(x, 0) = C_i \\ C(0, t) = C_0 \\ \frac{\partial C}{\partial t}(\infty, t) = 0 \end{cases} \quad (2)$$

Mojid, et al. [2, 3], following Wakao and Kaguei's [7] use of the Laplace transform of convolution, calculated the estimated response concentration $[C_{\text{r.est}(t)}]$ at time t as:

$$C_{\text{r.est}(t)} = \int_0^\infty C_{i(\alpha)} f(t - \alpha) d\alpha \quad (3)$$

where $C_{i(\alpha)}$ is the time-dependent input concentration of the solute in the soil column, α is the time interval between two consecutive measurements of the input concentration, and $f(t)$, the Laplace inversion of the transfer function, is the impulse response to a Dirac delta input (at $t = 0$) of tracer into the soil column. Equation (2) estimates a set of response concentrations from a set of input concentrations. For a reactive solute, the transfer function $f(t)$ governed by the CDE is calculated as [7]:

$$f(t) = \frac{\left[\pi N \left(\frac{t}{\tau R} \right)^3 \right]^{-1/2}}{2\pi R} \exp \left[- \left(1 - \frac{t}{\tau R} \right)^2 \left(4N \frac{t}{\tau R} \right)^{-1} \right] \quad (4)$$

where N is the mass-dispersion number ($= D_{\text{disp}}/LV_p$), which is the reciprocal of the column Peclet number $P (= LV_p/D_{\text{disp}})$, τ is the mean travel time or the mean residence time of the solute, and L is the distance between the positions where the input and response concentrations were measured.

$$\tau = \frac{L}{V_p} \quad (5)$$

A BTC is a graphical representation of the outflow concentration versus time during an experiment. It shows the concentration of the solute when it breaks through the outflow end [8]. The BTCs should be normalized to identify differences in the areas beneath the peak input and response positions. The mean travel time, the optimal pore velocity V_{opt} , and the optimal hydrodynamic dispersion coefficients D_{opt} are determined for each case. Then, the mean residence time τ is calculated using equation (5) and the dispersivity values λ using the equation $D_{\text{disp}} = \lambda_{\text{disp}} V_{\text{pore}}$. Finally, the column Peclet number is obtained using the

equation $Pe_{col} = V_{pore} L/D_{disp}$, and the mass-dispersion number N is estimated as $N = 1/Pe_{col}$.

Equations (4) and (5) can be used to calculate the estimated response BTCs at any time from the measured BTCs in the input time domain to determine the solute transport parameters. The root-mean-square error (RMSE) between the measured and estimated BTCs is calculated to evaluate the accuracy of fit of the transfer function method. The RMSE is obtained as follows:

$$RMSE = \sqrt{\frac{\int_0^{\infty} [C_{r(t)} - C_{r.est(t)}]^2 dt}{\int_0^{\infty} [C_{r(t)}]^2 dt}} \quad (6)$$

where $C_{r(t)}$ is the time-dependent measured response concentration of the solute.

Method and materials

Method

The objective of this research is to investigate the hydrodynamic characteristics and transport of solutes in a porous media using a physical sand column model. A four-electrode salinity sensor was used to measure the electrical conductivity (EC) of the soil with the purpose of determining the hydraulic characteristics of water movement by conducting tracer tests on a laboratory model of a subsurface wetland. In situ, EC sensors and salinity tracers reduce the amount of time and effort required for sampling and laboratory analysis. They also prevent destructive sampling in experimental column studies. In this last setup, the measurements were taken manually. Breakthrough experiments can take days, so a low-cost data-logging system that measures continuously and automatically throughout the day and night was required. Three grain sizes of sand (coarse, medium and fine) collected from the bottom of the Mekong river in Vietnam were used in the experiments. They are useful materials for domestic wastewater treatment since they can be used to construct a subsurface flow wetland.

Materials

A physical model was made locally in Can Tho University. The model included an electrical multiplexer system connecting nine groups of four-electrode probes. This was fitted into a horizontal sand column placed in a

firm stainless steel frame (Fig. 3). The framework consisted of enclosed transparent Perspex plates of 3 mm thickness covered by a removable lid. The experimental sand column was a long rectangular box with outer dimensions of 2.050 x 0.180 x 0.183 m. A 1 cm-thick polystyrene plate was placed between the lid and the sand column to ensure minimal bypass flow on top of the horizontal column. The whole system was closed watertight.

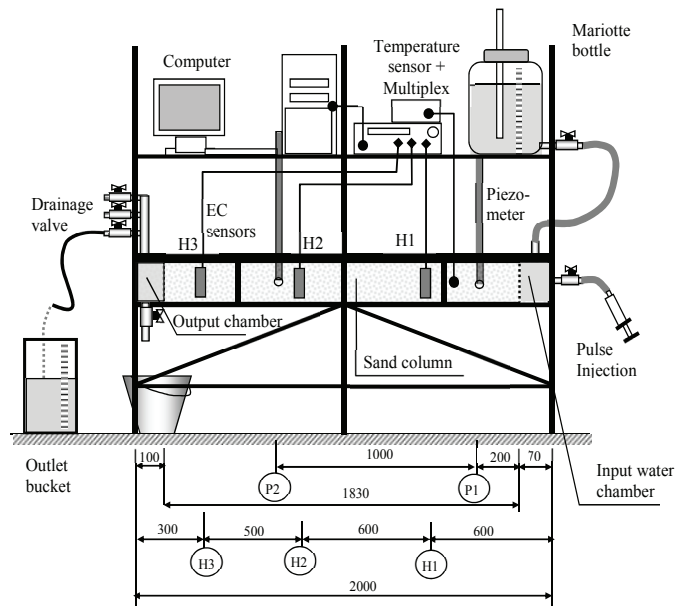


Fig. 3. Sand column system layout (H1, H2 and H3 are groups of three sensors each).

There were three chambers in the rectangular sand column: the input water chamber measuring 0.170 x 0.145 x 0.070 m; the sand column (0.170 x 0.145 x 1.830 m); and the outlet water chamber (0.170 x 0.145 x 0.100 m). The cross-section area of the sand column was 0.02465 m². The input water chamber received water from a 20 l Mariotte bottle. The Mariotte bottle had the function of maintaining constant water pressure and, therefore, constant flux during the experiment. The input chamber was also where the tracer solution was injected. Three groups of three four-electrode sensors were installed and connected to the multiplexer, as shown in Fig. 4. The sensors were 140 mm-long stainless steel rods with an outside diameter of 3 mm. The rods were inserted perpendicularly into the plastic block leaving 8 mm between each rod. The plastic blocks were fastened firmly outside the sand column, and the rods were submerged in the sand to a depth of 137 mm, seen through the Perspex frame.

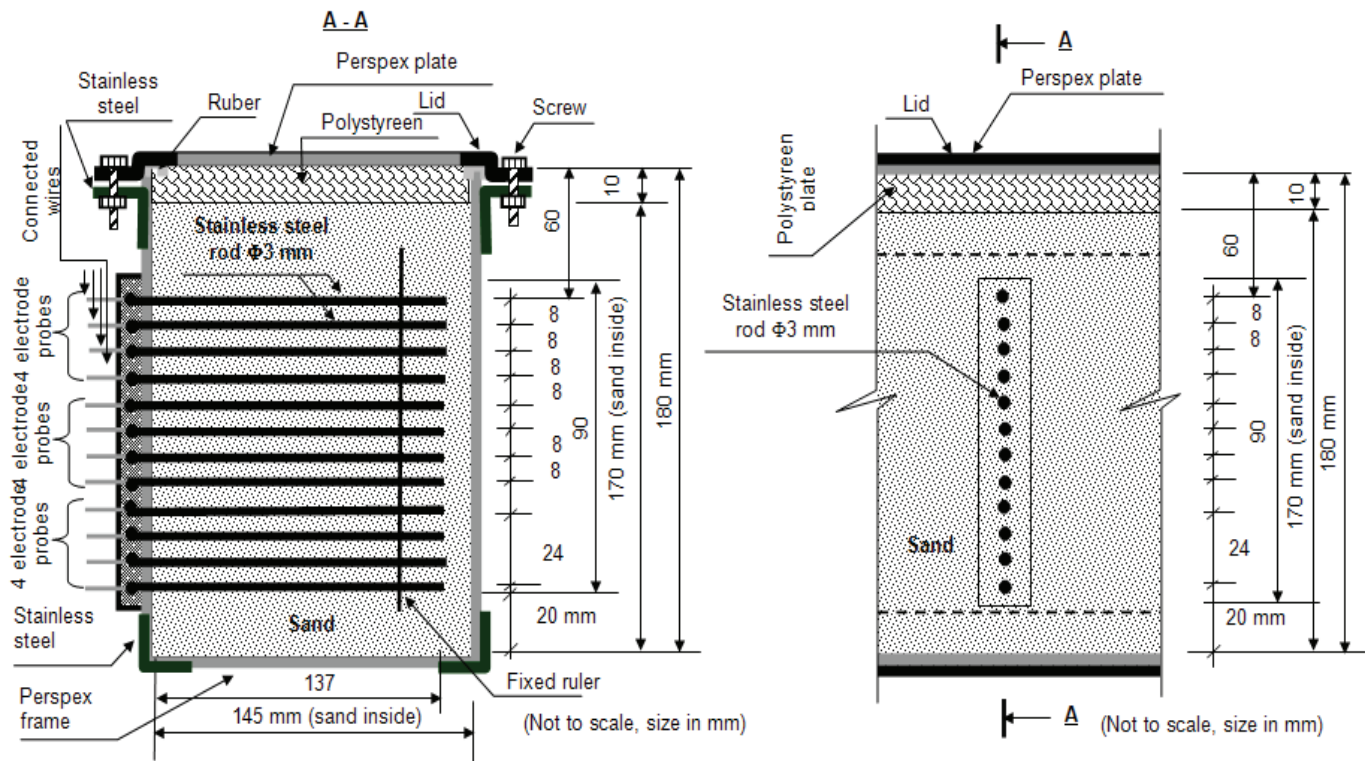


Fig. 4. One vertical group (H1, H2 or H3) of three four-electrode probes each.

The three groups of three four-rod sensors were used to monitor BTCs in the porous horizontal sand column using a saline trace. All sensors were connected to a locally made multiplexing system and a computer. The nine sensors were coded as follows: H1V1, H1V2, H1V3 for group H1; H2V1, H2V2, H2V3 for group H2; and H3V1, H3V2, H3V3 for group H3. H1, H2, and H3 were at a horizontal distance of 53 cm, 113 cm, and 613 cm, respectively, from the start of the sand column. V1, V2 and V3 were 5.6 cm, 4.4 cm, and 3.2 cm, respectively, from the bottom of the sand column. In addition, a thermal sensor was installed and connected to the computer. The codes and distances between the sensor groups are presented in Fig. 5.

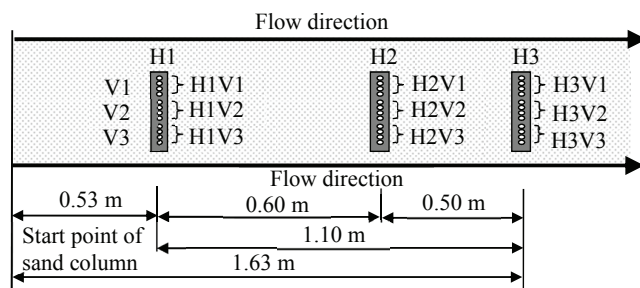


Fig. 5. Distances and coding for groups of sensors.

For each sensor measurement, three values were measured: the current was measured through electrodes 1 and 4; the voltage was measured between electrode 2 and 3, and the temperature was taken. The current through electrodes 1 and 4 was measured by reading the voltage drop over a known resistance R_{cs} . An alternating current (AC) was used, which required amplification and conversion to a direct current (DC), as most data acquisition cards require DC. A type K thermocouple was inserted to measure the temperature.

In order to collect and store data automatically, a measuring system was designed using a commercial personal computer with a data acquisition card. The graphical user interface was developed using the computational language MATLAB and the SIMULINK tool. A cost-effective data acquisition card, HUMUSOFT AD512, with a driver for extended real-time tool box software [9] was installed in a personal computer. The card had eight analogue input channels, two analogue output channels with 12-bit resolution and up to 100 Ks per second data access velocity, which is sufficient for this measurement. In addition, there were eight digital outputs and eight digital inputs which were useful for logical control, as shown Fig. 6.

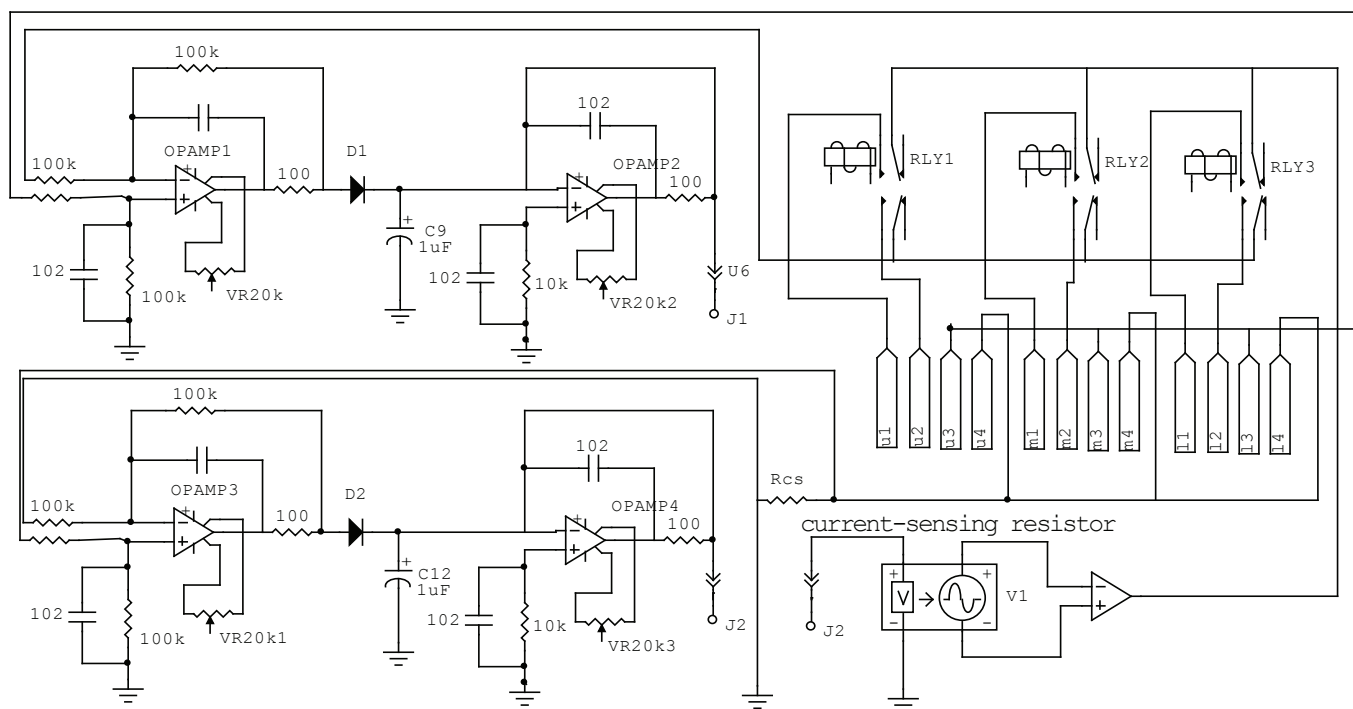


Fig. 6. The signal conditioning circuit for measuring V_{drop} and V_{2-3} .

At a set time interval, the measurement system collected the data at each of the 4-electrode sensors and stored them on the hard drive. Since only one sensor was operated at a time, the multiplexer switched between sensors. The switching circuit was crucial in this design. The ratio of the electric current (I) between the outer electrodes to the voltage difference (V_{drop}) between the two inner electrodes was calculated. The ratio I/V_{drop} was defined as the voltage drop F . First, the different AC frequencies were tested, and it was confirmed that any frequency between 100 and 1,000 Hz was suitable. A constant frequency of 220 Hz was selected. In these experiments, the R_{cs} was 15.8 Ohm. The voltage difference V/V_{drop} was automatically measured using a digital voltmeter. The geometrical factor K_e between the output value V/V_{drop} and the bulk EC depends on the shape and construction of the sensor. The value was calibrated based on the measurements of a laboratory EC meter in water solutions with a prepared concentration and at a known reference temperature, and the F values were measured by the sensor system. The multiplexer recorded EC values in sequence. It began with the sensor H1V1 and switched after 60 seconds to the next sensor, continuing to H1V2, H1V3... until H3V3, after which it returned to H1V1 (Fig. 7). With nine sensor groups, the entire cycle required 540 seconds. The electrical system was designed to record EC values in sequence and display them on a computer monitor.

A program developed in the R programming language was used to calculate the solute transport parameters, and

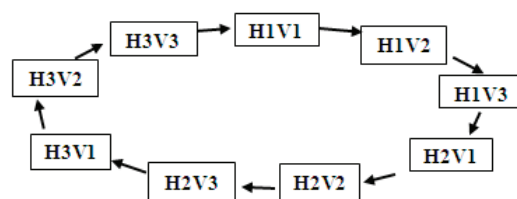


Fig. 7. Sensor group measurement turnover.

the Monte Carlo method was used for the analysis. In the R program, the user can define the random sampling number of the set of transport parameters, i.e. V_{pore} and D_{opt} . The optimised V_{pore} and D_{disp} are expressed as V_{opt} and D_{opt} , respectively. They are determined by searching for the minimal RMSE value in equation (6). In this case, 10,000 sets of $(V_{\text{pore}}, D_{\text{disp}})$ were generated randomly within a sample range of $(V_{\text{opt}}/5, V_{\text{opt}} \times 5)$ for V_{opt} and $(D_{\text{opt}}/5, D_{\text{opt}} \times 5)$ for D_{opt} . The squared correlation coefficient R^2 was determined for each set. Values of $R^2 > 0.5$ were plotted, and the highest R^2 value was identified as the optimized $(V_{\text{pore}}, D_{\text{disp}})$.

Results and discussion

The regression equations and the correlation coefficients (R-square) between the ratios of the measured current to the measured voltage drop (F) over the sensor with the EC measured using an Orion EC-meter (σ_M) are presented in Table 1.

Table 1. Regression equations and R^2 values of F (mA/mV) and σ_M (dS/m).

Sensor groups	Regression equations	R^2
H1V1	$\sigma_M = 13.015F + 0.1557$	0.9930
H1V2	$\sigma_M = 11.453F + 0.2004$	0.9985
H1V3	$\sigma_M = 12.258F + 0.2175$	0.9942
H2V1	$\sigma_M = 12.179F + 0.2116$	0.9970
H2V2	$\sigma_M = 14.400F + 0.0533$	0.9724
H2V3	$\sigma_M = 12.047F + 0.2140$	0.9915
H3V1	$\sigma_M = 11.917F + 0.1799$	0.9940
H3V2	$\sigma_M = 13.010F + 0.2071$	0.9928
H3V3	$\sigma_M = 13.521F + 0.2025$	0.9982

Three kinds of sand, coded as S1, S2 and S3, were used for the sand column experiments. Table 2 shows the sand sieve results and their average porosity. The values of 50% and 10% smaller (d_{50} and d_{10}) were determined by interpolation.

Table 2. Sand sieve analysis.

Sieve size (mm)	% smaller		
	Sand S1	Sand S2	Sand S3
4.000	99.290	98.096	100.000
2.000	98.300	93.205	99.975
1.000	95.662	75.896	99.873
0.500	74.967	40.873	88.005
0.250	7.895	9.716	53.003
0.125	1.039	1.955	0.923
0.075	0.409	0.662	0.840
Pan	0.079	0.000	0.000
d_{50} (mm)	0.407	0.573	0.242
Sand classification	Medium	Coarse	Fine
d_{10} (mm)	0.258	0.252	0.147
d_{60} (mm)	0.444	0.773	0.299
d_{60}/d_{10}	1.723	3.060	2.074
Uniformity	Uniform	Uniform	Uniform
Average porosity n (%)	46.3	49.7	45.7

For each tracer experiment using a particular sand class, the volumetric flow rate was changed. Each experiment was coded with the general identifier QiSj, with i ($i = 1, 2, 3, 4$) representing the flow rates which varies across sand classes j ($j = 1$ for medium sand, $j = 2$ for coarse sand, and $j = 3$ for fine sand). Table 3 summarises the flow rates corresponding to the three different sand types.

Table 3. Flow rates (m^3/s) in the sand column experiments.

	S1 (Medium)	S2 (Coarse)	S3 (Fine)
Q1	2.383×10^{-7}	4.383×10^{-7}	3.933×10^{-7}
Q2	3.400×10^{-7}	6.900×10^{-7}	4.483×10^{-7}
Q3	4.383×10^{-7}	7.250×10^{-7}	4.933×10^{-7}
Q4		7.933×10^{-7}	

Considering that the flows are through a finite area, the soil fluxes in sand column experiments are calculated. When the flow is laminar, Darcy's law is valid. Therefore, the Reynolds number is calculated using the mean grain diameter d_{50} . The water temperatures in the experiments are between 25 and 27°C and the density of the solute varies a little with the tracer concentration. However, to simplify the calculation of the Re number, it is assumed that the density of the solute is approximately that of clean water. If $Re < 10$, the saturated hydraulic conductivity K_s for each experiment is determined. Table 4 summarises the results for Re and K_s .

Table 4. Reynolds number and the saturated hydraulic conductivity.

QiSj	Q (m^3/s)	J_w (m/s)	Re	$-(\Delta h/l)$	K_s (m/s)
Q1S1	2.383E-07	9.649E-06	4.408E-03	0.018	5.371E-04
Q2S1	3.400E-07	1.377E-05	6.288E-03	0.021	6.568E-04
Q3S1	4.383E-07	1.775E-05	8.107E-03	0.025	7.113E-04
Q1S2	4.383E-07	1.775E-05	1.142E-02	0.014	1.270E-03
Q2S2	6.900E-07	2.794E-05	1.798E-02	0.021	1.333E-03
Q3S2	7.250E-07	2.935E-05	1.889E-02	0.022	1.337E-03
Q4S2	7.933E-07	3.212E-05	2.067E-02	0.023	1.399E-03
Q1S3	3.933E-07	1.592E-05	4.337E-03	0.021	7.598E-04
Q2S3	4.483E-07	1.856E-05	5.054E-03	0.023	8.084E-04
Q3S3	4.933E-07	1.997E-05	5.440E-03	0.024	8.339E-04

The results in Table 4 show that the Reynolds numbers are below 10, so all the flows in the experiments were laminar, and Darcy's law can be applied to calculate the saturated hydraulic conductivity K_s . If there is no flow ($Q = 0 \text{ m}^3/\text{s}$) in the sand column, the $(\Delta h/L)$ should be zero. The trend lines of water flux versus hydraulic gradient have to go through the zero point, as shown in Fig. 8.

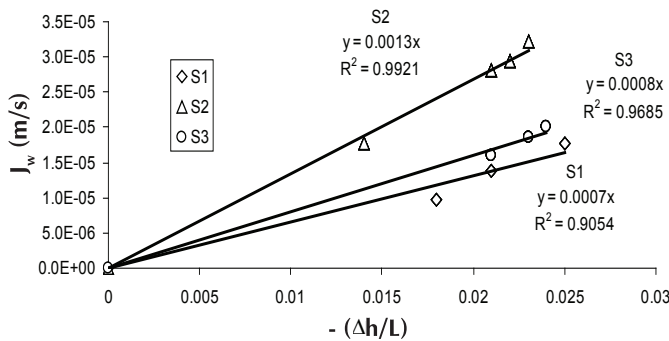


Fig. 8. Water flux versus hydraulic gradient.

The saturated hydraulic conductivity K_s should be constant for each sand class. The standard deviations of the calculated K_s were very small, lower than 5%. Fig. 9 shows the trend lines of the water flux versus the saturated hydraulic conductivity. The slopes of these lines are very small, so the values of K_s can be accepted as having the same order of magnitude.

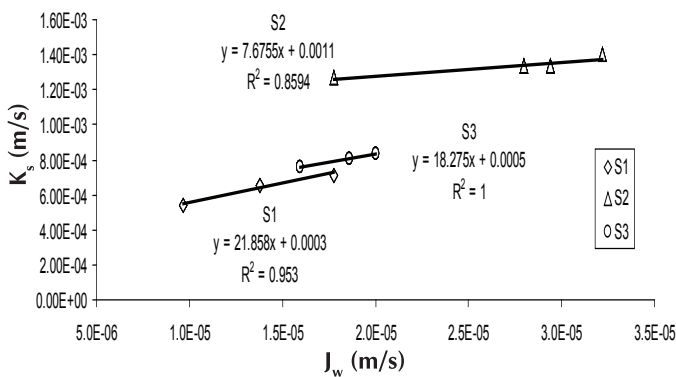


Fig. 9. Water flux versus the saturated hydraulic conductivity.

Figure 10 shows two examples of BTCs measured in the experiments and the normalised BTCs. Based on the results of the transfer function method, the solute transport parameters, which are average residence time (or breakthrough time) τ , dispersivity λ , the column Peclet number Pe_{col} and mass dispersion number N , were estimated for each transport case. The average residence time

decreased with as the water pore velocity increased. This can be seen in Fig. 11, which shows results of the transport from sensor H1V3 to sensor H3V3 for each sand type.

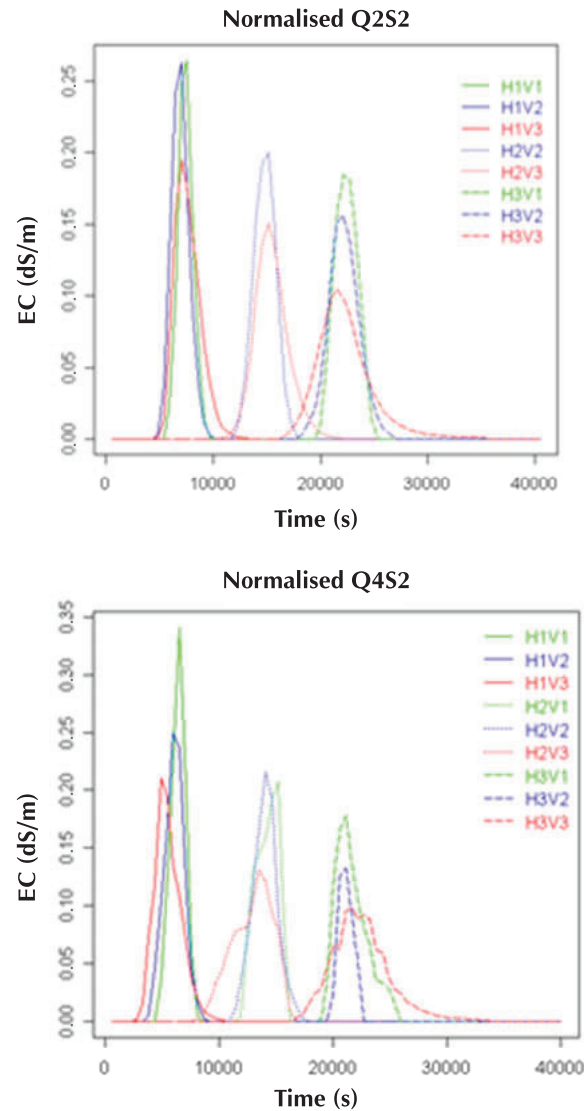


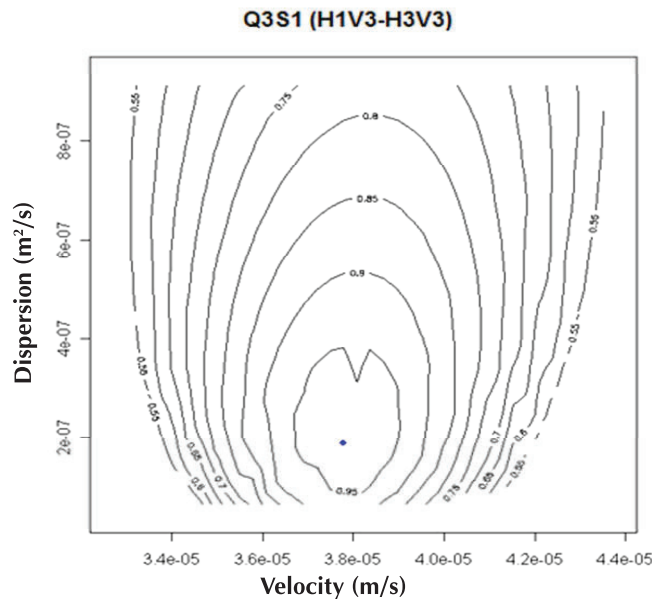
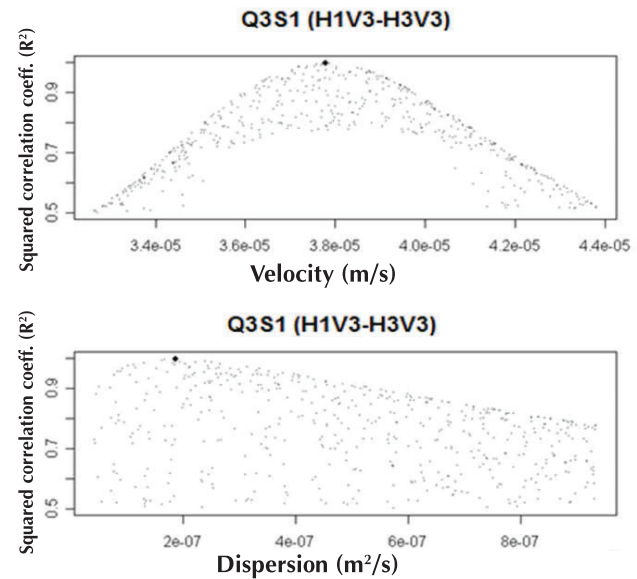
Fig. 10. Normalised BTCs plotted at location Q2S2 and Q4S2.

Based on the results of the transfer function method, the solute transport parameters, i.e. the average residence time (or breakthrough time) τ , dispersivity λ , the column Peclet number Pe_{col} , and the mass dispersion number N were estimated for each transport cases of the transport. Table 5 shows the estimated solution transport parameters.

The Monte Carlo method is used to identify the sensitivity of the parameters. The sensitivity analysis evaluates the interactions between the model parameters, i.e. the impact of changes in inputs on the outputs. The dot plots show clearly the optimal point for the (V_{pore}, D_{disp}) set of estimated

Table 5. Estimated solution transport parameters.

Sand type	Flow rate Q (m ³ /s)	Optimal pore velocity V_{opt} (m/s)	Optimal dispers coefficient D_{opt} (m ² /s)	Residence time τ (hr)	Dispersivity λ (m)	Column Pe number Pe	Mass disp. number N
S1	2.383E-07	2.434E-05	1.217E-07	12.554	5.000E-03	2.200E+02	4.545E-03
	3.400E-07	3.779E-05	1.868E-07	8.086	4.943E-03	2.225E+02	4.494E-03
	4.383E-07	3.430E-05	1.362E-07	8.908	3.971E-03	2.770E+02	3.610E-03
S2	4.383E-07	4.202E-05	5.543E-07	7.272	1.319E-02	8.339E+01	1.199E-02
	6.900E-07	6.542E-05	4.750E-07	4.671	7.261E-03	1.515E+02	6.601E-03
	7.250E-07	7.753E-05	6.569E-07	3.941	8.473E-03	1.298E+02	7.703E-03
	7.933E-07	7.633E-05	6.373E-07	4.003	8.349E-03	1.317E+02	7.590E-03
S3	3.933E-07	4.123E-05	1.005E-07	7.411	2.438E-03	4.513E+02	2.216E-03
	4.583E-07	5.705E-05	1.487E-07	5.356	2.606E-03	4.220E+02	2.370E-03
	4.933E-07	5.185E-05	8.608E-08	5.893	1.660E-03	6.626E+02	1.509E-03


 Fig. 11. Contour lines for optimal (V_{pore} , D_{disp}).

 Fig. 12. Dotty plots (.) and highest values (*) of R^2 for optimal V_{pore} and D_{disp} .

transport parameters for the CDE. As an example, Fig. 11 and Fig. 12 illustrates the sensitivity analysis for the case Q3S1 (H1V3 - H3V3) with medium sand and water flux of 1.778×10^{-5} m/s. These two plots represent the response surface between the two parameters V_{pore} and D_{disp} .

Conclusions

This research uses theories on the transport mechanism of a solute in a porous medium. The experiments were performed using sand from the Mekong river. The results were the optimal water pore velocities and the optimal dispersion for three types of sand.

Four-electrode probes were successfully constructed, calibrated and operated using a multiplexing system. The multiplexing system enabled the EC at different locations in the sand column to be continuously monitored. The system was made locally at a low cost and worked well for testing a tracer flowing through a saturated horizontal sand column.

The concentration values of the tracer flowing through the horizontal sand column were measured using a series of sensors and were plotted in the form of BTCs. In each experiment, laminar flow was concluded from the calculated Reynolds number. Laminar flow is necessary for Darcy's law, from which the saturated hydraulic conductivity was

calculated. For the experiments within the same sand class, the values of the saturated hydraulic conductivity had the same order of magnitude. From these curves, the pore water velocity and the mechanical dispersion coefficient were determined using the transfer function method. From these variables, the average residence time, the dispersivity, the column Peclet number and the mass-dispersion number were calculated.

It is possible to conclude that the continuous movement of a solute through sand is governed by the CDE, which is a second-order differential equation. The convection-dispersion equation for inert and non-adsorbing solutes is estimated using measured BTCs and normalised BTCs. The solute transports are identified as mixed-flow processes rather than plug-flow processes. The sensitivity analysis shows that the CDE is highly sensitive to the dispersion parameter.

ACKNOWLEDGEMENTS

The authors thank the VLIR-CTU project for financially supporting this research and all the faculty and staff in the Department of Environmental Engineering, College of Environment and Natural Resources, Can Tho University, Vietnam for their help during the experiments.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] J.A. Cherry and R.A. Freeze (1979), *Groundwater*, Prentice-Hall, Inc., New Jersey, p.604.
- [2] M.A. Mojid, D.A. Rose, and G.C.L. Wyseure (2004), "A transfer-function method for analysing breakthrough data in the time domain of the transport process", *Euro. J. Soil Sci.*, **55**, pp.699-711.
- [3] M.A. Mojid, D.A. Rose, and G.C.L. Wyseure (2006), "A model incorporating the diffuse double layer to predict the electrical conductivity of bulk soil", *Euro. J. Soil Sci.* **58**, pp.560-572.
- [4] C.W. Fetter (1999), *Contaminant hydrogeology*, Prentice-Hall, New Jersey, p.500.
- [5] G. Dagan (1984), "Solute transport in heterogeneous porous formations", *J. Fluid Mech.*, **145**, pp.151-177.
- [6] J.M. Wraith and D. Or (1998), "Nonlinear parameter estimation using spreadsheet software", *J. Nat. Res. Life Sci. Edu.*, **27**, pp.13-19.
- [7] N.S. Wakao and S. Kaguei (1982), *Heat and mass transfer in packed beds*, Gordon & Breach, New York, p.364.
- [8] W.A. Jury and R. Horton (2004), *Soil Physics*, John Wiley & Sons, Inc., New Jersey, p.370.
- [9] Humusoft (2006), *Data Acquisition Products*, available at: <http://www.humusoft.com/datacq/index.htm>.

Expression and purification of thioredoxin-his₆-ZmDREB2.7 fusion protein in *Escherichia coli* for raising antibodies

Thuy Linh Nguyen^{1,2}, Thuy Duong Nguyen¹,
Van Hai Nong¹, Thi Thu Hue Huynh^{1*}

¹Institute of Genome Research, Vietnam Academy of Science and Technology

²University of Science, Vietnam National University, Hanoi

Received 3 August 2018; accepted 23 November 2018

Abstract:

Dehydration-responsive element-binding (DREB) proteins play a critical role in the plant's drought-tolerance mechanism despite their presence in minor amounts in the cell. In this study, a maize-derived transcription factor protein, ZmDREB2.7, was overexpressed in the *Escherichia coli* strain Rosetta 1. The interested gene conjugating with the thioredoxin gene (*TrxA*) and his₆ tag in the pET-32a vector encoded a 55.7 kDa fusion protein. The optimum condition for inducing the thioredoxin-his₆-ZmDREB2.7 expression was five hours of induction with 0.05 mM IPTG at 30°C. The Tris-HCl 20 mM pH 8.0 lysis buffer was harnessed to extract the recombinant protein for the purification process. Using the immobilized-metal affinity chromatography column, the recombinant protein was purified and then injected into rabbits. The antisera containing polyclonal antibodies (pAbs) could specifically recognize the ZmDREB2.7 fusion protein. This study represents updated data on the bacterial expression of the recombinant ZmDREB2.7 protein and the production of anti-ZmDREB2.7 pAbs.

Keywords: *E. coli*, fusion expression, recombinant protein, ZmDREB2.7 protein.

Classification number: 3.1

Introduction

The ZmDREB2.7 protein belongs to the DREBs transcription factor family that involved in the plant abiotic resistance mechanism. The DREB transcription factors can be classified into two groups based on the protein structure: DREB1, and DREB2, despite the fact that they both contain an AP2 DNA-binding domain. In fact, DREB proteins bind specifically to the dehydration-responsive element (DRE) which contains a core motif of A/GCCGAC locating in the promoter region of many genes induced by drought and/or cold [1]. The DREB2 proteins and their coding genes were characterized in different species. In *Arabidopsis thaliana*, DREB2A and DREB2B are induced by osmotic stress and high temperature. Transgenic *A. thaliana* plants overexpressing *AtDREB2A CA*, which was *AtDREB2A* with a deletion of the negative regulatory domain, showed an improved stress tolerance to drought and heat-shock stresses [2, 3]. An *OsDREB2B* gene isolated from rice enhanced drought and cold tolerance in transgenic plants without any phenotypic changes [4]. Meanwhile, a *PeDREB2* gene from the desert-grown tree (*Populus euphratica*) was reported to be induced by cold, drought, and high salinity conditions and *PeDREB2* could specifically bind to the DRE element in the promoter region of many stress-driven genes [5]. In addition, the transient expression of *PeDREB2* in onion epidermis cells showed that the protein localized to the nucleus which confirmed that DREB proteins act as a transcription factor [5]. Pandey and colleagues [6] built a model of a wheat DREB2 protein (*Triticum aestivum* L.) and reported that the protein interacts with the major DNA groove through its β -sheets.

In maize (*Zea mays* L.), a genome-wide analysis [1] successfully identified and cloned 18 *ZmDREB* genes (10 *ZmDREB1* genes and 8 *ZmDREB2* genes). Among them,

*Corresponding author: Email: hthue@igr.ac.vn

ZmDREB2.7 was reported as the most potential gene for crop improvement by marker-assisted breeding and genetic engineering. *ZmDREB2.7-Tv* is a gene isolated from Tevang 1 maize cultivar which exhibits a good tolerance with drought and cold conditions. Originally, the DREB transcription factors are present in a small amount in the plant cells. In addition, the *ZmDREB2.7* protein is induced mainly when a plant confronts osmotic stress [1]. In order to clarify the role of *ZmDREB2.7* protein involvement in drought tolerance, it is necessary to obtain the protein in a high quantity with good quality. Therefore, the heterologous expression of the *DREB2.7* in bacteria brings advantages. A high amount of the *ZmDREB2.7* protein could be used for understanding the characteristics of the protein. In addition, anti-*ZmDREB2.7* polyclonal serum can be employed to detect the presence of the specific *ZmDREB2.7* protein.

Heterologous protein expression in other host systems has been harnessed for the production of many plant proteins [7]. Due to the fact that the protein isolation from the plant is high-cost, labor-intensive, time-consuming, and low-quantity, bacterial expression systems offer a promising alternative. In fact, the plant proteins produced by bacteria were widely employed for research, therapy, and industrial applications. The popularity of using *E. coli* as a workhorse for synthesizing plant protein is a result of its rapid growth at high-cell density on an inexpensive carbon source, well-known genetics, and the commercial availability of enormous expression vectors and strains. However, challenges faced when using bacterial systems to express eukaryotic proteins are lack of post-translation modification and formation of inclusion bodies containing inactive proteins [8]. These causes can be classified into two categories: those that are in the gene sequences and those that are the limitations of the *E. coli* [8]. Fortunately, a number of literature reviews provided comprehensive knowledge to optimize the procedures and parameters involved in the bacterial heterologous protein expression and the purification process [7-11]. In order to troubleshoot the aforementioned problems, the recombinant host, the strain, the expression vector, the inducing conditions, and the approach to modifying the coding sequence of the interested protein should be carefully considered.

The immunization of animals to induce an immune response is a procedure performed routinely worldwide. The process produces antibodies against a specific antigen in laboratory animals such as mice, rabbits, and chickens. Among them, mice and rabbits are the most frequent species used for antibody production. Depending on the desired application and the availability of time and money, scientists may choose between generating monoclonal antibodies (mAbs) or polyclonal antibodies (pAbs). Production of

mAbs is a labor-intensive and time-consuming work. In addition, generation of mAbs comprises cell culture which requires high financial investment, but a low titer of mAbs can be obtained. Meanwhile, the induction of pAbs usually takes 4-8 weeks with high titer. In fact, polyclonal antiserum can be obtained with inexpensive procedures and instruments. Therefore, the pAbs is suitable for many applications and is favored by many scientists [12].

In this study, we introduced the *ZmDREB2.7* gene into the pET-32a expression vector to generate the thioredoxin-*his₆*-*ZmDREB2.7* fusion protein. The recombinant *ZmDREB2.7* fusion protein was overexpressed, purified and used for raising polyclonal antibodies.

Materials and methods

Construction of the recombinant expression vector

The coding sequence of *ZmDREB2.7* originated from maize was cloned into the pJET1.2 vector at Genome Biodiversity Laboratory, Institute of Genome Research. In order to enable cloning the gene into the pET-32a expression vector, two primers (*Zm2.7* BamHI F: 5' - TAGTCGGATCCGATCGGGTGCCGC - 3', *Zm2.7* EcoRI R: 5'- CGACGAGAATTCTAAAGAGGGACGACGA - 3') were designed with *EcoRI* and *BamHI* restriction sites at the 5' and 3' end, respectively. The PCR reaction using the primer pair was conducted with the total volume of 25 µl which contains 12.5 µl 2X Thermo Scientific DreamTaq PCR Master Mix, 1 µl of 10 µM each primer, 1 µl of 10 µg/µl pJET1.2+*ZmDREB2.7* plasmid, 0.8 µl of DMSO, and 8.7 µl ddH₂O. The temperature conditions were as follows: 4 min at 94°C followed by 35 cycles of 45 sec at 94°C, 45 sec at 56°C and 1 min 10 sec at 72°C, then a final extension of 3 min at 72°C. The PCR product of approximately 1.1 kb long was digested with *EcoRI* and *BamHI* restriction enzymes, and the same to the pET-32a expression vector. The two digested fragments, one of *ZmDREB2.7* and one of the linearized pET-32a plasmid in which both flanked by *EcoRI* and *BamHI* restriction sites, were ligated using standard molecular biology techniques [13]. The identity of clones harboring the pET-32a+*ZmDREB2.7* plasmid was identified by restriction enzymes-based screening and confirmed by sequencing.

Expression of thioredoxin-*his₆*-*ZmDREB2.7* in *E. coli*

The pET-32a+*ZmDREB2.7* expression vector was transformed into the *E. coli* strain Rosetta 1. A transformed colony was used to optimize the heterologous protein expression as followed the isopropyl-β-D-thiogalactopyranoside (IPTG)-induce protocol [13]. A colony was inoculated in 3 ml of LB medium supplied with 50 mg/l ampicillin with 200 rpm shaking overnight at 37°C.

The 16-hour culture was transferred into fresh 25 ml of LB medium containing 50 mg/l ampicillin to achieve the final OD₆₀₀ of 0.1. The culture was incubated with 200 rpm shaking at 37°C. When the culture's OD₆₀₀ reached 0.6-0.8, the transformant was induced by adding 0.1M IPTG with an appropriate final concentration. After five hours, cells in 1 ml of the induced culture were collected by centrifuging at 5,000 rpm for 5 min. The cell pellets were suspended with the same volume of lysis buffer and stored at -20°C until further processing.

Cell extract was prepared using the freeze-thaw protocol with Qsonica Q55-220 Sonicator Ultrasonic Processor (Cole-Parmer®) on ice [13]. The condition for sonication step was as follows: five cycles of 1 min 30 sec with a rest period of 2 min between cycles. One hundred µl of lysate was transferred into a new tube as the total protein sample. The cell lysate was separated by centrifuging at 10,000 rpm for 5 min at 4°C. The soluble protein fraction as the supernatant was collected. The bacterial cell debris was resuspended in 900 µl lysis buffer and treated as the insoluble protein fraction.

Purification of the fusion protein

The large-scale soluble protein fraction was prepared as described above then added with 500 mM NaCl and filtered through a 0.45 µm syringe filter. The his₆-tag protein was purified using the 5 ml HisTrap™ HP columns (GE Healthcare, Piscataway, NJ, USA) by following the manufacturer's instruction. The solution flowed through the column at the speed of 0.5 ml/min. The protein sample was loaded on the column and washed with 25 ml washing buffer (20 mM Tris HCl, 100 mM NaCl, 50 mM Imidazole, pH 8.0). The protein was eluted by applying 10 ml elution buffer (20 mM Tris HCl, 100 mM NaCl, 250 mM Imidazole, pH 8.0). All fractions containing the fusion protein were analyzed by SDS-PAGE. The eluted fractions then were applied with Microcon® centrifugal filter (Millipore, MA, USA) for desalting and concentrating, and then used as an antigen for injection into rabbits.

Raising of polyclonal antibodies

Two healthy 3-month-old rabbits used for immunization were provided by Vetvaco National Veterinary Joint-Stock Company (VETVACO., JSC), and weighed about 2.5-3.0 kg at the time of acquisition. The pAbs production procedure and laboratory animals care were adopted from the CCAC guidelines on antibody production by the Canadian Council on Animal Care (CCAC) with some modifications (https://www.ccac.ca/Documents/Standards/Guidelines/Antibody_production.pdf). Rabbits were given intramuscular injections at one site on their limbs and

subcutaneous injection at five sites on their backs. The first priming injection was performed with a low dose of 0.25 mg/ml purified recombinant ZmDREB2.7 protein (Antigen-Ag) emulsified in Freund's Complete Adjuvant (FCA). After that, the rabbits received three additional injections with raising concentrations of Ag 0.5 mg/ml, 0.75 mg/ml, and 0.1 mg/ml in Freund's Incomplete Adjuvant (FIA), respectively. Each additional injection was administered at 10-day intervals.

Bleeding was implemented from ear veins three times, seven days after each administered day and the last time at day 10 of the final injection. Rabbit blood was collected into a sterile 15 ml centrifuge tube and placed at room temperature for 30 min followed by incubating at 4°C for one hour. The antiserum was collected by centrifuging the blood tubes at 5,000 rpm for 10 min at 4°C then pipetting the supernatants into new tubes and stored at 4°C.

Agglutination test was conducted by mixing 20 µl of antiserum and Ag on a sterile plate. The plate was placed at room temperature for 10 min. After that, if the collected serum contained pAbs of a specific Ag, white clumps could be observed.

SDS-PAGE and dot blot analysis

The SDS-PAGE analysis was conducted using Tris-Glycine Gel, including a separate gel of 12.6% and a stacking gel of 5%, with the Bio-Rad system according to the manufacturer's instructions. Protein was then electrophoresed using a Bio-Rad PowerPac Basic Mini Electrophoresis system (Bio-Rad), for 35 min at 200 V. Protein was visualized by Coomassie blue staining.

For dot blot analysis, a range of concentration of the purified recombinant ZmDREB2.7 protein (from 1 mg/ml to 5 mg/ml) was loaded onto nitrocellulose membrane by pipetting. The membrane was dried at room temperature for about 20 min and incubated with a 1:8 dilution of rabbit serum containing anti-ZmDREB2.7 antibodies. After that, the primary antibody was recognized by the secondary antibody Goat Anti-Rabbit IgG (whole molecule)-Alkaline Phosphatase (Sigma-Aldrich), and the membrane was exposed to 1-Step™ NBT/BCIP substrate solution (Thermo Fisher Scientific).

Results

Construction of the recombinant expression vector

The recognition sites of restriction enzyme *Eco*RI and *Bam*HI were introduced at the 5' and 3' ends of *ZmDREB2.7* gene, respectively, in order to clone the gene into the pET-32a plasmid (Fig. 1A). The *ZmDREB2.7* gene was designed to be in frame with *TrxA* (thioredoxin) gene and fused with

his₆ tag sequence for a further purification experiment (Fig. 1A). Thus, the expected thioredoxin-his₆-ZmDREB2.7 fusion protein contains 522 amino acids and has a theoretical weight of 55.7 kDa.

By mean of PCR with two specific primers, Zm2.7 BamHIF and Zm2.7 EcoRIR, the approximately 1089 bp-long coding sequence of ZmDREB2.7 was successfully amplified from the cloning vector pJET1.2+ZmDREB2.7 (Fig. 1B). The PCR product was double digested with two mentioned restriction enzymes, and the vector pET-32a was linearized using the same enzymes (Fig. 1C). The ligation product of the two digested samples was transformed into the DH10β competent cells. We isolated plasmids from six random colonies and characterized by electrophoresis on 1% agarose gel. As shown in Fig. 1D, all plasmid bands obtained from putative recombinant clones were higher than the empty vector pET-32a. The recombinant vectors were verified by restriction enzyme-based screening (Fig. 1E) and confirmed by sequencing (data not shown). Taken together, we successfully constructed the pET-32a+ZmDREB2.7 bacterial expression vector.

Expression of thioredoxin-his₆-ZmDREB2.7 in *E. coli*

The pET-32a+ZmDREB2.7 recombinant vector was transformed into the competent *E. coli* strain Rosetta 1, and a number of recombinant colonies were obtained. At first, we accessed the solubility of four different lysis buffers based on Tris buffer and Phosphate buffer to the heterologous protein (data not shown). The result showed that most of the proteins produced by the recombinant strain (about 90-95%) were in the soluble protein fraction. The Tris-HCl pH 8.0 buffer was chosen for subsequent experiments due to the highest solubility to the fusion protein.

Other factors affecting protein expression-including time of induction, temperature, and IPTG's concentration - were examined (Figs. 2A-2C). As expected, the expression of the recombinant protein increased over time and reached the highest level after five hours (Fig. 2A). However, the production of the fusion protein was not influenced by the tested concentration of IPTG as the intensity of bands representing the interested protein was nearly the same in all lanes on the SDS-PAGE gel (Fig. 2C). The same situation was observed when inducing protein expression at 30°C

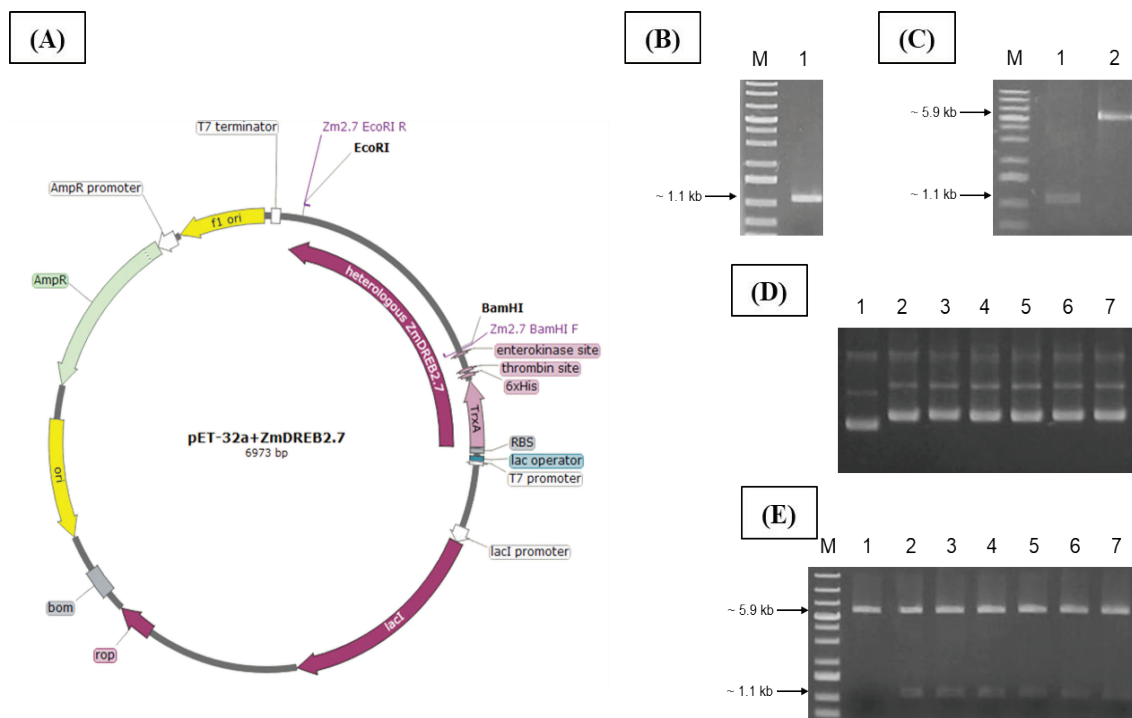


Fig. 1. Construction of the pET-32a+ZmDREB2.7 expression vector. (A) schematic illustration of the pET-32a+ZmDREB2.7 expression vector; (B) PCR amplification of ZmDREB2.7 fragment flanked by EcoRI and BamHI recognition sites. 1: product of PCR using pJET1.2+ZmDREB2.7 as the template; (C) double digestion of DNA with EcoRI and BamHI, 1: PCR products amplifying ZmDREB2.7 gene from pJET1.2+ZmDREB2.7, 2: the pET-32a plasmid; (D) plasmid isolation from bacteria colonies. 1: the pET-32a vector, 2-7: plasmids isolated from six putative recombinant colonies, respectively; (E) restriction enzyme-based screening of putative recombinant colonies. 1: the pET-32a vector, 2-7: plasmids isolated from six putative recombinant colonies, respectively. M: marker 1 kb plus (Thermo Fisher Scientific).

and 37°C (Fig. 2B). Taken together, the optimum condition established to produce the thioredoxin-his₆-ZmDREB2.7 fusion protein was five hours of induction using 0.05 mM IPTG at 30°C (Fig. 2D).

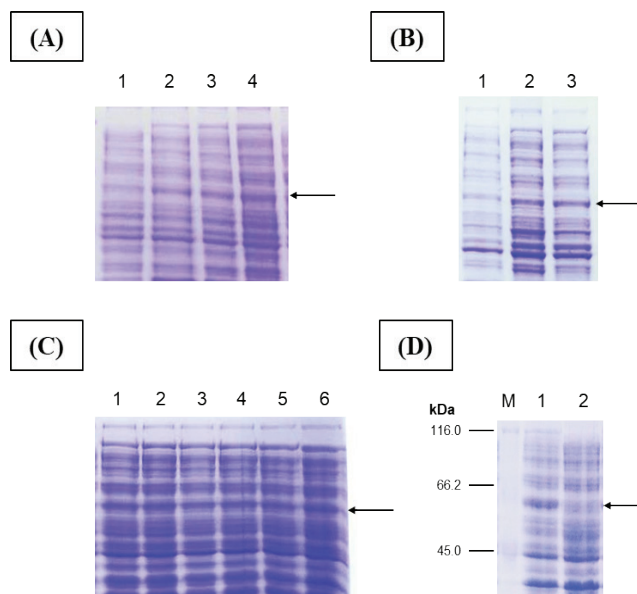


Fig. 2. Expression of the thioredoxin-his₆-ZmDREB2.7 fusion protein. (A) effect of induction period on the expression of the fusion protein. 1-4: 0h, 1h, 3h, 5h after adding IPTG, respectively; (B) effect of temperature on the expression of the fusion protein. 1: 25°C, 2: 30°C, 3: 37°C; (C) effect of IPTG's concentration on the expression of the fusion protein. 1-6: IPTG of 0.05, 0.1, 0.25, 0.5, 0.75, 1.0 mM, respectively; (D) overexpression of the recombinant protein in the *E. coli* strain Rosetta 1 harboring the pET-32a+ZmDREB2.7 vector. 1: optimized induction conditions. 2: without IPTG. M: Thermo Scientific™ Pierce™ Unstained Protein Molecular Weight Marker. The arrow indicates the interested protein.

Purification of the thioredoxin-his₆-ZmDREB2.7 fusion protein

We took advantage of the fact that the DREB2.7 fusion protein contains a his₆ sequence at N-terminal to purify the fusion protein by immobilized-metal affinity chromatography (IMAC). The fusion protein was large-scale overexpressed with optimized conditions and utilized for the purification process. Fig. 3 showed the SDS-PAGE analysis of the recombinant protein purified through the IMAC column. Most of the proteins of the host strain were in the unbound fraction (lane 2) and the wash fraction (lane 3). Lanes 4-10 showed the protein fractions after applying the elution buffer containing 250 mM Imidazole. The arrow pointed to the expected full-length protein. The interested protein eluted with the high amount as judged by Coomassie staining. Thus, we concentrated and desalted the elution fractions for antibodies production.

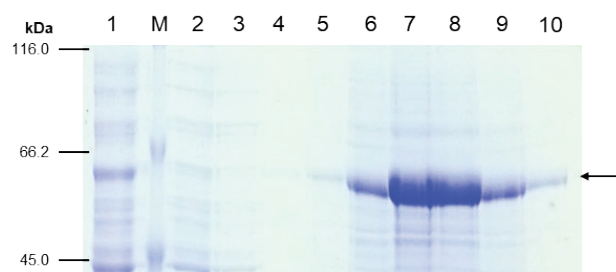


Fig. 3. Purification of the thioredoxin-his₆-ZmDREB2.7 fusion protein using HisTrap™ HP columns. M: Thermo Scientific™ Pierce™ Unstained Protein Molecular Weight Marker. 1: the soluble fraction of the recombinant *E. coli* strain. 2: the flow through from column. 3: elute after washing with 50 mM Imidazole. 4-10: fractions after applying the elution buffer containing 250 mM Imidazole. The arrow indicates the interested protein.

Raising of anti-ZmDREB2.7 fusion protein polyclonal antibodies

The protein after the purification step was used for injection into two rabbits via the procedure described above. The agglutination test was implemented using sera from the two rabbits against the purified recombinant ZmDREB2.7 fusion protein. The assays were conducted seven days after each injection to monitor antibody response during the immunization process. We obtained the positive result of the agglutination test immediately after the priming injection. In addition, the intensity of reactions rose as more injections were given. As shown in Fig. 4A, there were visible white clumps after 30 minutes combining the serum of the last bleeding with the antigen. Moreover, it was obvious that serum originated from the first rabbit exhibited higher

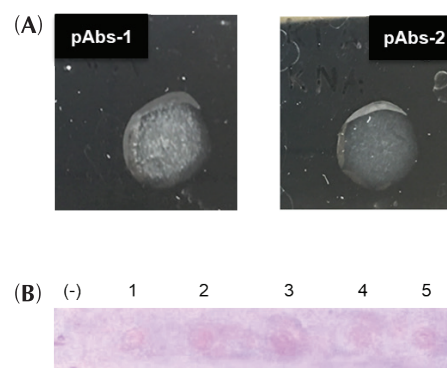


Fig. 4. Agglutination test and dot blot analysis using rabbit anti-fusion protein sera. (A) agglutination test of rabbit sera (the last bleeding) to the antigen. pAbs-1, pAbs2: the antisera from the first and the second immunization rabbit, respectively. (B) dot blot analysis of the rabbit sera to the antigen. (-): H₂O. 1-5: serial dilutions of the ZmDREB2.7 fusion protein ranging from 1 mg/ml to 5 mg/ml, respectively.

response than that from the second one. It was supposed that immunized rabbits produced pAbs to the thioredoxin-his₆-ZmDREB2.7 fusion protein. Therefore, we harnessed the anti-fusion protein serum from the first rabbit for dot blot analysis to confirm the specificity of the serum. As the result shows, the pAbs generated in rabbit serum can efficiently recognize the recombinant ZmDREB2.7 fusion protein (Fig. 4B).

Discussion

The *E. coli* expression system has been exploited for the production of a variety of proteins. Even though there are some drawbacks, such as lack of post-translation modification, the bacterial expression system remains faster and cheaper for producing eukaryote proteins. Therefore, we adopted an *E. coli* expression system and did optimization of components involved in the protein expression process to obtain the high expression level of the ZmDREB2.7 fusion protein.

Even though the *ZmDREB2.7* gene contains few rare codons to *E. coli*, it has a high GC content (70%). Meanwhile, the GC content in the *E. coli* genome was about 50.5% [14]. Additionally, the maize-derived gene shows a low codon adaptation index in *E. coli* which is 0.68 (the acceptable figure is from 0.8 to 1.0). The difference in codon bias between maize and *E. coli* normally causes early termination and produces truncated versions of the heterologous protein. Due to such limitations in the gene sequence, we harnessed the *E. coli* strain Rosetta and the protocol which gradually induces the heterologous protein expression. The Rosetta 1 strain has many advantages for enhanced protein expression [15]. The strain as a BL21 derivation is deficient in protease Lon and OmpT which could increase the stability of expressed recombinant proteins. In addition, Rosetta 1 harbors a compatible plasmid which produces tRNAs for rare codons AUA, AGG, AGA, CGG, CUA, CCC, and GGA. Then, the tRNA pool can compensate for the difference in codon bias between *E. coli* and the original source of the interested gene. Therefore, it was not necessary to optimize codon usage of the *ZmDREB2.7* gene to ensure that the heterologous protein was expressed in full length.

When a eukaryote protein expresses at a high level in the bacteria cell, it may be found in inclusion bodies due to inappropriate folding. To overcome this issue, the *ZmDREB2.7* gene was conjugated with the *TrxA* in the vector pET-32a. TrxA normally located in *E. coli* cytoplasm is a compact, highly soluble, and thermally stable protein. These properties allow *trxA* to serve as a molecular chaperone. Therefore, when ZmDREB2.7 N-terminally fused with *trxA* protein, the recombinant protein could avoid forming an inclusion body [16]. Additionally, theoretically, slowing

down the production rate can help the newly synthesized proteins fold more properly [10]. It is also reported that sometimes inducing at low temperature facilitates soluble thioredoxin-fused protein [17]. In the study, we induced the fusion protein at 30°C for five hours as there was no significant difference of the protein expression level between 30°C and 37°C (Fig. 2B). Because the expression levels of the fusion protein were nearly the same at a range of inducer concentration (Fig. 2C), the lowest concentration of inducer (0.05 mM IPTG) was the optimum choice to increase the protein production.

In addition, the ZmDREB2.7 was fused with the his₆ sequence to enable purification using IMAC system. The his₆ tag at N-terminal guarantees that the translation process initiates in the correct position. As expected, the induced protein bound to the Ni column was the full-length one with the molecular weight of approximately 55.7 kDa.

There are several factors to consider when raising pAbs in laboratory animals. In fact, rabbits are commonly used for reasons of cost-effectiveness, ease of handling, and high amount of serum compared to mice. We used the young rabbits because immune function peaks at puberty and declines with age [12]. There were several reports of batch-to-batch variants when producing pAbs by immunizing animals, so two rabbits per antigen are recommended. In our study, two rabbits responded differently as the agglutination test exhibited more white precipitations with the antiserum from the first one (Fig. 4A). The number of injections and the amount of the ZmDREB2.7 fusion protein were tightly controlled. We used three booster doses that were double, triple, and four times the priming dose, respectively. In addition, the adjuvant was added to induce a high titer of antibodies without any side effects to the animal. A high quantity of anti-ZmDREB2.7 fusion protein serum was obtained from the raising pAbs experiment.

Conclusions

In conclusion, we successfully cloned the *ZmDREB2.7* gene into the pET-32a vector. The expression vector worked well in the *E. coli* Rosetta 1 that the thioredoxin-his₆-ZmDREB2.7 fusion protein was overexpressed. The optimized conditions for the production of the interested protein were five hours at 30°C using 0.05 mM of IPTG. The fusion protein was purified by IMAC column and used to raise pAbs in the rabbit. The obtained antiserum can specifically bind to the ZmDREB2.7 fusion protein.

ACKNOWLEDGEMENTS

The present research was supported by a grant from the Vietnam Ministry of Agriculture and Rural Development (MARD) named “Isolating genes related to drought

tolerance and constructing vectors for maize improvement”.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] S. Liu, X. Wang, H. Wang, H. Xin, X. Yang, J. Yan, F. Qin (2013), “Genome-wide analysis of *ZmDREB* genes and their association with natural variation in drought tolerance at seedling stage of *Zea mays* L”, *PLOS Genet.*, **9**(9), e1003790.
- [2] Y. Sakuma, K. Maruyama, Y. Osakabe, F. Qin, M. Seki, K. Shinozaki, K. Yamaguchi-Shinozaki (2006a), “Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought-responsive gene expression”, *Plant Cell*, **18**, pp.1292-1309.
- [3] Y. Sakuma, K. Maruyama, F. Qin, Y. Osakabe, K. Shinozaki, K. Yamaguchi-Shinozaki (2006b), “Dual function of an *Arabidopsis* transcription factor DREB2A in water-stress-responsive and heat-stress-responsive gene expression”, *Proc. Natl. Acad. Sci. USA*, **103**, pp.18822-18827.
- [4] J.Q. Chen, X.P. Meng, Y. Zhang, M. Xia, X.P. Wang (2008), “Over-expression of *OsDREB* genes lead to enhanced drought tolerance in rice”, *Biotechnol. Let.*, **30**(12), pp.2191-2198.
- [5] J. Chen, X. Xia, W. Yin (2009), “Expression profiling and functional characterization of a DREB2-type gene from *Populus euphratica*”, *Biochem. Biophys. Res. Commun.*, **378**(3), pp.483-487.
- [6] B. Pandey, P. Sharma, M. Saini, D.M. Pandey, I. Sharma (2014), “Isolation and characterization of dehydration-responsive element-binding factor 2 (DREB2) from Indian wheat (*Triticum aestivum* L.) cultivars, *Aust. J. Crop Sci.*, **8**(1), p.44.
- [7] F. Yesilirmak, Z. Sayers (2009), “Heterologous expression of plant genes”, *Int. J. Plant Genomics*, **2009**, ID 296482.
- [8] S. Sahdev, S.K. Khattar, K.S. Saini (2008), “Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies”, *Mol. Cell. Biochem.*, **307**(1-2), pp.249-264.
- [9] D.M. Francis, R. Page (2010), “Strategies to optimize protein expression in *E. coli*”, *Curr. Protoc Protein Sci.*, **61**(1), pp.5-24.
- [10] G.L. Rosano, E.A. Ceccarelli (2014), “Recombinant protein expression in *Escherichia coli*: advances and challenges”, *Front. Microbiol.*, **5**, p.172.
- [11] A.R. Gomes, S.M. Byregowda, B.M. Veeregowda, V. Balamurugan (2016), “An overview of heterologous expression host systems for the production of recombinant proteins”, *Adv. Anim. Vet. Sci.*, **4**(7), pp.346-356.
- [12] H.F. Stills (2012), “Polyclonal antibody production”, *The Laboratory Rabbit, Guinea Pig, Hamster, and Other Rodents*, pp.259-274.
- [13] J. Sambrook, D.W. Russell (2001), *Molecular cloning: a laboratory manual*, Cold Spring Harb Lab Press Cold Spring Harb NY, 999.
- [14] S. Mann, Y.P.P. Chen (2010), “Bacterial genomic G+C composition-eliciting environmental adaptation”, *Genomics*, **95**(1), pp.7-15.
- [15] R. Novy, D. Drott, K. Yeager, R. Mierendorf (2001), “Overcoming the codon bias of *E. coli* for enhanced protein expression”, *Innovations*, **12**, pp.1-3.
- [16] Y. Li (2009), “Carrier proteins for fusion expression of antimicrobial peptides in *Escherichia coli*”, *Biotechnol. Appl. Biochem.*, **54**(1), pp.1-9.
- [17] X. Xu, F. Jin, X. Yu, S. Ji, J. Wang, H. Cheng, W. Zhang (2007), “Expression and purification of a recombinant antibacterial peptide, cecropin, from *Escherichia coli*”, *Protein Expr. Purif.*, **53**(2), pp.293-301.

Identification of cassava bacterial blight-causing *Xanthomonas axonopodis* pv. *manihotis* based on *rpoD* and *gyrB* genes

Ngo Quang Huy¹, Mai Van Quan¹, Le Quang Man¹,
Duong Thi Nguyen², Trinh Xuan Hoat^{1*}

¹Plant Protection Research Institute, Vietnam Academy of Agricultural Sciences

²College of Agriculture and Forestry, Thai Nguyen University

Received 30 August 2018; accepted 30 November 2018

Abstract:

Cassava (*Manihot esculenta* Crantz) is one of the most important crops in Vietnam for providing food, starch sources, and raw materials for the production of bio-ethanol and other purposes. Cassava bacterial blight (CBB) disease, caused by *Xanthomonas axonopodis* pv. *manihotis* (*Xam*), is one of the most important factors affecting cassava production in Vietnam. A rapid and sensitive molecular tool is required to support the traditional method, primarily based on biochemical reactions, that was considered less sensitive and time-consuming. In the present study, in order to detect *Xam*, the Polymerase Chain Reaction (PCR) technique was applied with the two primer pairs, *rpoD*_17F/*rpoD*_1005R and *XgyrconpcrF1*/*Xgyrconpcr1*, to amplify partial sequences of *rpoD* and *gyrB* genes, respectively. The primers directed the amplification of about 900 bp DNA fragments from CBB-infected leaf and stem tissues. All the amplified DNA sequences were identical in each gene and to the *Xam* reference strains (accession numbers KP265376 and KP265378 for the *rpoD* gene, and KP265387 and KP265388 for the *gyrB* gene) from GenBank. The representative sequences of *rpoD* and *gyrB* genes were deposited in GenBank with accession numbers MF774491 and MF774490, respectively. DNA and phylogenetic analysis based on the sequences of *rpoD* and *gyrB* genes have confirmed that *Xam* is the causal agent of CBB in Dong Nai, Vietnam. This is the first confirmed identification of the causal pathogen of CBB disease in Vietnam using molecular tools, and this method is a reliable tool for the detection and identification of other plant bacterial pathogens.

Keywords: cassava bacterial blight, *gyrB*, PCR, *rpoD*, *Xanthomonas axonopodis* pv. *manihotis*.

Classification number: 3.1

Introduction

Cassava (*Manihot esculenta* Crantz) has high starch content, with approximately 30-60% dry matter, and has become one of the most important crops in the world that is both a food source as well as providing materials for a variety of industries. In Vietnam, cassava is a cash crop that provides food, in particular, and material for the production of bio-ethanol and other products. The most important characteristic of cassava is its drought tolerance and it is therefore widely grown in mountainous areas, in marginal soil types, and with non-irrigation systems with low fertiliser requirements. However, in recent years, cassava production in Vietnam has suffered from a number of diseases and insect

pests, especially Cassava pink mealybug (*Phenacoccus manihoti*), cassava witches' broom (CaWB), cassava anthracnose, cassava bacterial blight (CBB) diseases, and a new invasive disease called Sri Lankan cassava mosaic virus (SLCMV). On the basis of observations in the field, CBB is considered to be the third important disease of cassava in Vietnam, after CaWB and SLCMV. However, no study of it has been conducted in Vietnam until now. CBB disease caused by *Xam* is one of the important diseases of cassava in Africa and South America, affecting cassava plants from the planting period to harvesting, and reducing cassava production by up to 100% [1, 2].

Traditional methods for the detection and identification

*Corresponding author: Email: trinxuanhoatppri@gmail.com

of bacterial plant pathogens use selected media, analysis of bacterial colonies on agar plate, biochemical tests, pathogenicity tests, and the analysis of fatty acids [3, 4]; however, these methods are usually labour-intensive, time-consuming, and insufficient. The enzyme-linked immunosorbent assay (ELISA) and monoclonal antibodies have also been employed for the characterisation of bacterial pathogens [5]; however, they are no longer widely used because of their low specificity to the pathogens and because they cause cross-reactions.

In recent years, molecular approaches have become the best methods and are widely applied in the detection and identification of plant bacterial pathogens because they are highly specific, quick, and reliable [6, 7]. PCR and DNA analysis, including the Polymerase Chain Reaction (PCR), Amplified fragment length polymorphism (AFLP) [8], nested-PCR, dot-blot assays [9] and Repetitive elements based PCR (rep-PCR) [10], have been successfully used to detect and identify *Xam*.

To date, no information is available on the phenotypic or genetic variability of *Xam* populations in cassava-growing regions in Vietnam. In particular, there are no reports of the use of molecular techniques for detecting and identifying *Xam* that causes CBB disease in Vietnam. In this study, we focus on the application of a PCR technique using specific primer pairs for *rpoD* (RNA polymerase, sigma 70 (sigma D) factor) and *gyrase B* (*gyrB*) genes in combination with phylogenetic analysis to identify *Xam* from cassava plants with symptoms of CBB collected in Dong Nai province.

Materials and methods

Plant materials

Cassava stems and leaves from five plants showing typical symptoms of CBB were collected from CBB-infected fields in Dong Nai province. One healthy cassava plant was collected from the north of Vietnam as a negative control.

Isolation of bacteria and genomic DNA extraction

Isolation of *Xam* from CBB-infected leaves and stems was undertaken as described by [11] and [12], with some modifications as described below. Small pieces of leaf and stem tissue were excised at the border of the healthy and infected tissues. The tissues were then homogenised in 1 ml of sterilised distilled water and incubated for 25 minutes at room temperature. 100 µl of the resulting solution was streaked in LPG medium (5 g glucose, 5 g yeast extract, 5 g

peptone, and 15 g agar in a litre of distilled water). Freshly grown single colonies were used for bacterial cultures in 5 ml of liquid LPG medium at 30°C for 12 hrs and shaken at 150 rpm. 3 ml of the resulting bacterial cultures were subjected to genomic DNA isolation using a DNeasy® mini kit (QIAGEN), following the instructions of the manufacturer.

PCR assays

The two primer pairs were used for PCR amplification. The first primer pair corresponded to the *rpoD* gene, *rpoD*_17F (5'-ATCTGACCTACGCCGAAGTC-3') and *rpoD*_1005R (5'-CTGCTGCTCGGAGATGATCT-3') [13]; and the second primer pair was alternatively conserved PCR primers corresponding to the *gyrB* gene, *XgyrconpcrF1* (5'-AAGAGCGAGCTGTATCTGAAGGACGA-3') and *Xgyrconrpcr1* (5'-CGCGTCCTCGATGCGCACCTGCA-3') [14]. DNA was quantified by means of a UV-Vis Spectrophotometer Optima SP-3000*nano* (Indonesia) and diluted to 20 ng/µl. The PCR amplifications were conducted in 25 µl reaction mixtures containing 2.5 µl of 10×PCR buffer (including 15 mM MgCl₂), 0.75 µl of each primer (10 µM), 2 µl of dNTP mix (2.5 mM each), 0.125 µl of Taq polymerase, 1 µl of DNA (20 ng/µl), and up to 25 µl of H₂O. The PCR temperature conditions were 4 min at 94°C followed by 35 cycles of 94°C for 30 s, T_m (50°C for 45 s for the *XgyrconpcrF1*/*Xgyrconrpcr1* primer pair, and 60°C for 30 s for the *rpoD*_17F/*rpoD*_1005R primer pair), and 72°C for 1 min, with a final extension at 72°C in a Mastercycler Pro (Eppendorf, Germany). The PCR products were separated by electrophoresis in 1% agarose gel in a 1×TAE buffer containing 0.5 µg/ml ethidium bromide and photographed under UV light using a GelDoc-It® 310 Imaging System (United Kingdom).

Phylogenetic analysis

The target PCR fragments were cut out and purified using a QIAquick Gel Extraction kit (QIAGEN, Cat No./ID: 28704) and directly sequenced with both forward and reverse primers that were used in PCR assays with ABI3100 sequencer. The DNA sequences were subjected to the National Centre for Biotechnology Information BLAST search tool [15] to ascertain the closest match. The multiple DNA sequences were aligned with the clustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Phylogenetic trees were constructed using MEGA 7.0 [13] by means of the neighbour-joining method with default values and 1,000 bootstrap replications.

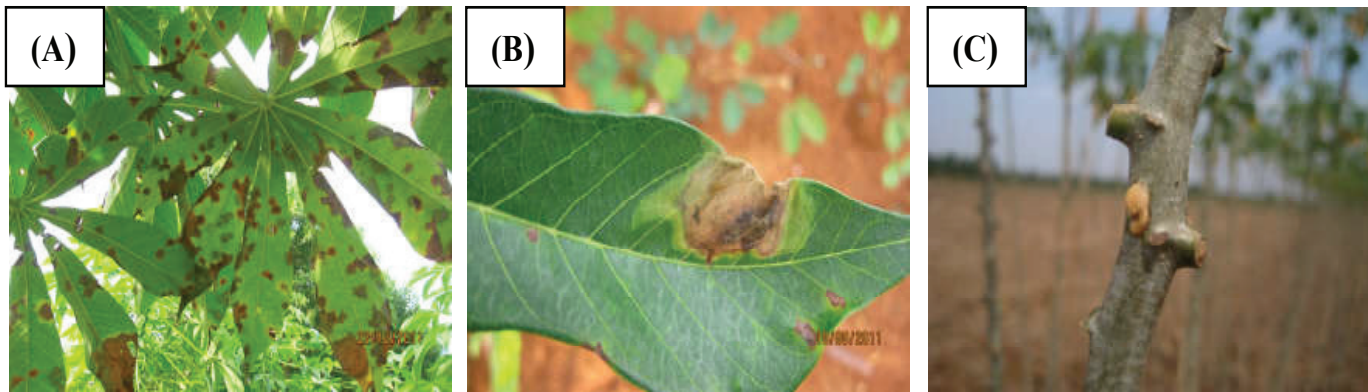


Fig. 1. Typical symptoms of CBB disease on cassava plants grown in Dong Nai province, Vietnam. (A, B) symptoms on leaves and (C) on stem.

Results and discussion

The symptoms of CBB disease

The typical symptoms of CBB disease in cassava caused by *Xam* include blighting, wilting, die-back of stems, formation of small angular leaf spots (Fig. 1A) that can become larger (Fig. 1B) and surrounded by yellow edge, vascular necrosis of the stem, and the production of exudates on the leaves and stem (Fig. 1C).

Detection and phylogenetic analysis of *Xam* based on partial *rpoD* gene sequences

PCR products with an expected length about 900 bp were amplified from two leaf samples and one stem sample by the *rpoD*_17F/*rpoD*_1005R primer pair. No amplicon was detected in the negative control (a healthy cassava plant) (Fig. 2). All three PCR products were purified with

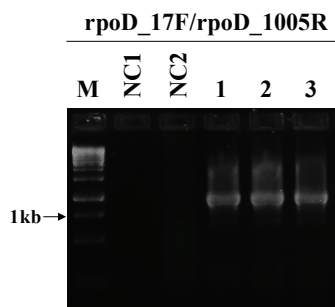


Fig. 2. PCR using primer pair *rpoD*_17F and *rpoD*_1005R for amplification of the *rpoD* gene of *Xam* isolated from the CBB-infected samples collected from Dong Nai province. M is the DM3100 DNA ladder (SMOBIO), NC1 is DNA extracted from a healthy cassava plant, NC2 is no DNA, 1-3 are two leaf isolates and one stem isolate from the CBB-infected leaf samples and stem sample, respectively, collected from Dong Nai province.

agarose gel and directly sequenced both directions using *rpoD*_17F and *rpoD*_1005 primers. Sequencing results showed that DNA sequences from the three isolates were identical. BLAST search results indicated that strains were 100% identical to strains of *Xam* (accession numbers: KP265372, FJ561609, and FJ561607). The CBB isolated from Dong Nai province that used *rpoD* as the target gene was designated CBBDNVN-*rpoD* and deposited in GenBank with the accession number MF774491.

The phylogenetic tree was constructed with nucleotide sequences of CBBDNVN-*rpoD* and 21 other *rpoD* sequences of *Xanthomonas* species from GenBank with the accession number shown in parentheses in the Fig. 3 and was analyzed by means of the neighbour-joining method with 1,000 bootstrap replications using MEGA7.0 [16]. The most ancestral *Xanthomonas* species within the phylogenetic tree were *X. axonopodis* pv. *manihotis* and *X. axonopodis* pv. *alfalfae*. The CBBDNVN-*rpoD* shares 100% homology with the *X. axonopodis* pv. *manihotis* and is therefore considered to be among the most ancestral species in this phylogenetic tree, confirming that it belongs to the pathovar *manihotis*. The remaining *Xanthomonas* species were separated into three major groups. Group I, supported by bootstrap values of 81%, comprised *X. translucens*, *X. oryzae*, *X. sacchari*, *X. theicola*, *X. albilineans*, and *X. hyacinthi*. Group II, supported by bootstrap values of 100%, comprised *X. vesicatoria*, *X. cassavae*, *X. dyei*, *X. vasicola*, *X. pisi*, and *X. bromi*. Group III, also supported by bootstrap values of 100%, comprised *X. melonis*, *X. fragariae*, *X. cynarae*, *X. cucurbitae*, *X. codiae*, and *X. alfalfae* subsp. *Citrumelonis* (Fig. 3).

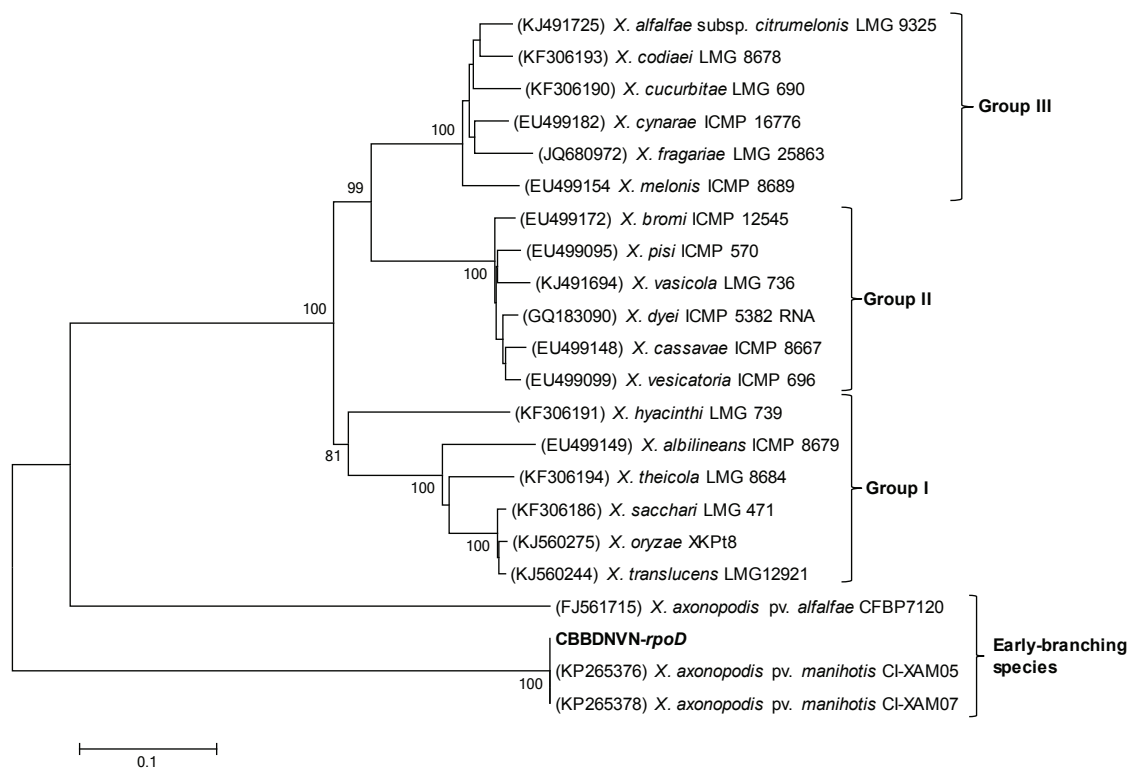


Fig. 3. Phylogenetic tree constructed by the neighbour-joining method, comparing the CBBDEVN-*rpoD* sequence with 21 other *rpoD* gene sequences of different *Xanthomonas* species from GenBank. Accession numbers are shown in parentheses. The numbers at the branches are confidence percentages obtained from 1,000 bootstrap replicates (only values above 80% are shown).

Detection and phylogenetic analysis of *Xam* based on partial *gyrB* gene sequences

Primer pair XgyrconpcrF1/Xgyrconpcr1 directed the amplification of about 900 bp of the target DNA sequence of the *gyrB* gene from two CBB-infected leaf samples and from one CBB-infected stem sample (Fig. 4).

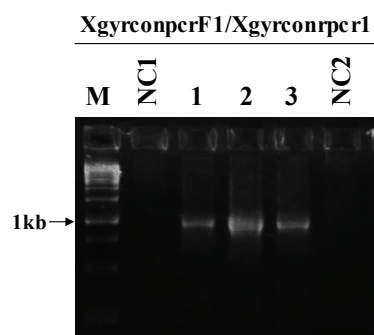


Fig. 4. PCR using primer pair XgyrconpcrF1/Xgyrconpcr1 for amplification of the *gyrB* gene from bacteria isolated from the CBB-infected samples collected from Dong Nai province. M is the DM3100 DNA ladder (SMOBIO), NC1 is DNA extracted from a healthy cassava plant, NC2 is no DNA, 1-3 are two leaf isolates and one stem isolate from the CBB-infected leaf samples and stem sample, respectively, collected from Dong Nai province.

The CBB isolated from Dong Nai province that used *gyrB* as the target gene was designated CBBDEVN-*gyrB*. This shares 100% homology with *X. axonopodis* pv. *manihotis* CI-XAM04 (accession number KP265387) and *X. axonopodis* pv. *manihotis* CI-XAM05 (accession number KP265288) which are the pathovars that cause CBB. They are early-branching species in the phylogram. The phylogenetic analysis confirms that CBBDEVN-*gyrB* (deposited in GenBank with the accession number MF774490) belongs to the pathovar *manihotis* and is the causal agent of CBB in Vietnam.

Difference in comparison with the phylogenetic tree based on the sequences of the *rpoD* gene, the remaining *Xanthomonas* species were separated into three major groups. Group I, supported by bootstrap values of 100%, comprised *X. hyacinthi*, *X. translucens*, *X. theicola*, *X. albilineans*, and *X. sacchari*. Group II, supported by bootstrap values of 99%, comprised *X. cassavae*, *X. codiae*, and *X. cucurbitae*. Group III comprised a large group and was divided into three clades IIIA, IIIB, and IIIC. Clade IIIA, supported by bootstrap values of 86%, comprised *X. vesicatoria* and *X. pisi*; clade IIIB comprised *X. cynarae*, *X. populi*, and *X. fragariae*; and clade IIIC, supported by bootstrap value of 100%, comprised *X. oryzae*, *X. vasicola*, *X. bromi*, *X. alfalfae*, *X. axonopodis*, and *X. melonis* (Fig. 5).

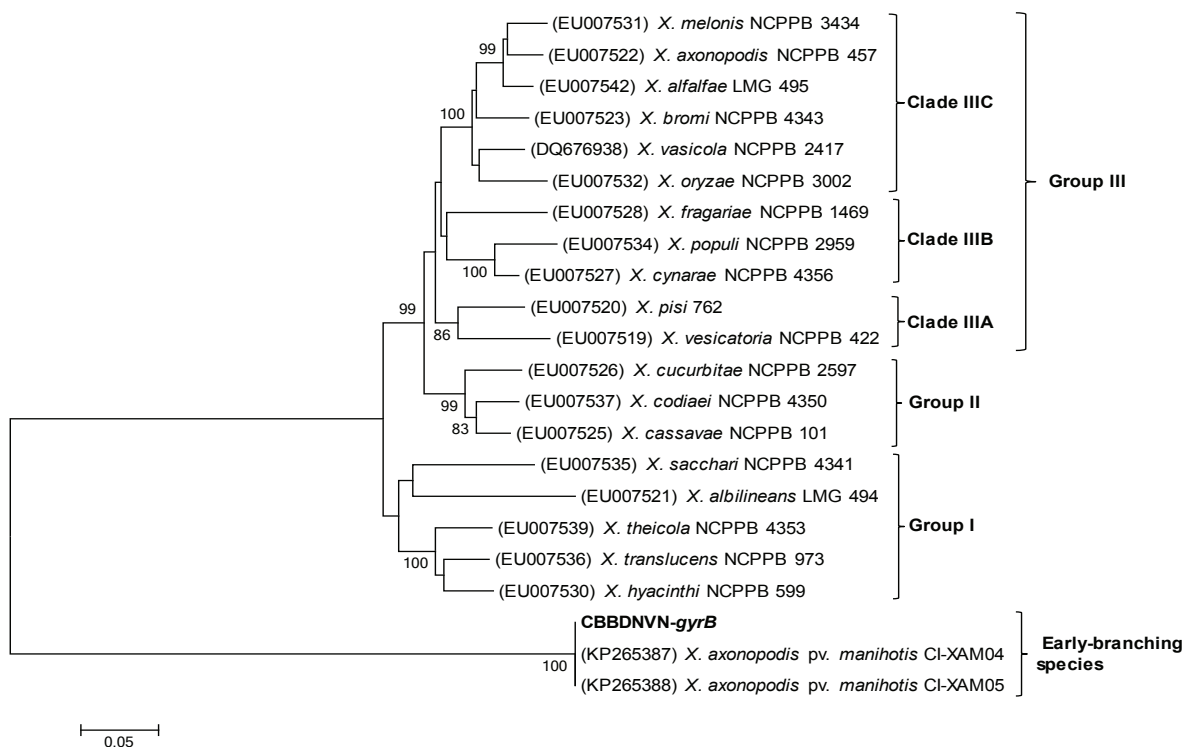


Fig. 5. Phylogenetic tree constructed by the neighbour-joining method, comparing the CBBDNVN-*gyrB* sequence with 21 other *gyrB* gene sequences of different *Xanthomonas* species from GenBank. Accession numbers are shown in parentheses. The numbers at the branches are confidence percentages obtained from 1,000 bootstrap replicates (only values above 80% are shown).

Xanthomonas comprises a large group of bacterial species with various physiological and pathological characteristics; they are causal agents of different diseases of many important crops. In recent years, molecular approaches using different target genes have been applied to distinguish between species within *Xanthomonas*. [11] developed a PCR method with specific primer pairs XV (5'-TTC-GGC-AACGGC-AGT-GAC-CAC-C-3') and XK (5'-TCA-ATC-GGA-GAT-TAC-CTGAGC-G-3') which proved to be a relatively sensitive method for detecting *Xam* in cassava stem and leaf lesions with the minimum number of cells (3×10^2 CFU/ml) could be detected. A specific, sensitive, quick nested-PCR technique using two primer pairs XK/XV and XV2 (5'-AACAGT-AAG-TCG-GTG-TCG-CC-3')/XK2 (5'-TCG-TTG-CCG-TGG-CTA-CC-3') successfully detected *Xam* from cultured cells, artificially inoculated seeds, and naturally infected seeds [9]. The phenotypic and genetic variability of *Xam* isolates was determined using DNA polymerase chain reaction-based fingerprinting (rep-PCR) [10].

Different genes have been used as targets for phylogenetic analysis, such as the 16S rRNA gene [17], the 16S-23S rRNA intergenic spacer region [18], *rpjB* and *atpD* genes [8], the *gyrB* gene [14, 19], and the *rpoB* gene (which encodes the β subunit of bacterial RNA polymerase) [20]. Previous

studies indicate that approaches based on the analysis of 16S rDNA sequences and the 16-23S intergenic region revealed very low diversity, with most of the species clustered into a single group [18, 21]; therefore, neither of these target genes would be reliable for the differentiation of the majority of *Xanthomonas* species. [14] constructed a phylogram using different *gyrB* sequences and indicate that the *gyrB* gene is a reliable target gene for rapid, accurate identification at the species level among *Xanthomonas* species. The *rpoB* gene, a highly conserved housekeeping gene with a single copy in the genome [22], has also been used as a target gene for the detection, characterisation, and discrimination of several bacterial species, including *Xanthomonas* [20].

The *rpoD* and *gyrB* genes have been used as molecular markers for the rapid detection and characterisation of *Xam* in Burkina Faso [23] and in the Ivory Coast [13], suggesting that these genes can be used for the study of a variety of bacterial pathogens.

In the present study, we applied a PCR technique using the two primer pairs, *rpoD*_17F/*rpoD*_1005R and *XgyrconprF1*/*XgyrconprC1*, and provided evidence that they are appropriate targets for the detection and phylogenetic analysis of *Xam* in the CBB-infected cassava plants. The method is quite specific and quick in detecting *Xam* from

bacterial cultures, suggesting that it is a good method for the detection and characterisation not only of the *Xanthomonas* genus but others too.

Conclusions

The causal agent of CBB disease in samples collected from Dong Nai province, Vietnam, is *Xanthomonas axonopodis* pv. *manihotis*. The PCR technique with primer pairs rpoD_17F/rpoD_1005R and XgyrconpcrF1/Xgyrconrper1, involving the amplification of the target *rpoD* and *gyrB* genes, is reliable method for the rapid detection and identification of the *Xam* that causes CBB disease.

ACKNOWLEDGEMENTS

This work was partially supported by a counterpart budget from the Ministry of Agriculture and Rural Development of Vietnam with the title “The project for development and dissemination of sustainable production system based on invasive pest management of cassava in Vietnam, Cambodia and Thailand”.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] A.A. Ogunjobi (2006), “Molecular variation in population structure of *Xanthomonas axonopodis* pv. *manihotis* in the south eastern Nigeria”, *Afr. J. Biotechnol.*, **5(20)**, pp.1868-1872.
- [2] V. Verdier, et al. (2004), “Recent progress in the characterization of molecular determinants in the *Xanthomonas axonopodis* pv. *manihotis* - cassava interaction”, *Plant Mol. Biol.*, **56(4)**, pp.573-584.
- [3] Y. Berthier, et al. (1993), “Characterization of *Xanthomonas campestris* pathovars by rRNA gene restriction patterns”, *Appl. Environ. Microbiol.*, **59(3)**, pp.851-859.
- [4] P. Yang, et al. (1993), “Application of fatty acid methyl esters for the taxonomic analysis of genus *Xanthomonas*”, *Syst. Appl. Microbiol.*, **16(1)**, pp.47-71.
- [5] B. Boher, et al. (1997), “Extracellular polysaccharides from *Xanthomonas axonopodis* pv. *manihotis* interact with cassava cell walls during pathogenesis”, *Mol. Plant Microb. Interact.*, **10(7)**, pp.803-811.
- [6] J.S. Hartung, et al. (1996), “Rapid and sensitive colorimetric detection of *Xanthomonas axonopodis* pv. *citri* by immunocapture and a nested-polymerase chain reaction assay”, *Phytopathol.*, **86(1)**, pp.95-101.
- [7] I.K. Toth, et al. (1998), “PCR-based detection of *Xanthomonas campestris* pv. *phaseoli* var. *fuscans* in plant material and its differentiation from *X. c.* pv. *Phaseoli*”, *J. Appl. Microbiol.*, **85(2)**, pp.327-336.
- [8] S. Restrepo, V. Verdier (1997), “Geographical differentiation of the population of *Xanthomonas axonopodis* pv. *manihotis* in Colombia”, *Appl. Environ. Microbiol.*, **63(11)**, pp.4427-4434.
- [9] S. Ojeda, V. Verdier (2009), “Detecting *Xanthomonas axonopodis* pv. *manihotis* in cassava true seeds by nested polymerase chain reaction assay”, *Can. J. Plant Pathol.*, **22(3)**, pp.21-24.
- [10] M.N. Chege, et al. (2017), “Phenotypic and genotypic diversity of *Xanthomonas axonopodis* pv. *manihotis* causing bacterial blight disease of cassava in Kenya”, *J. Appl. Biol. Biotechnol.*, **5(2)**, pp.38-44.
- [11] V. Verdier, et al. (1998), “Detection of the cassava bacterial blight pathogen, *Xanthomonas axonopodis* pv. *manihotis*, by polymerase chain reaction”, *Plant Dis.*, **82(1)**, pp.79-83.
- [12] A. Prokic, et al. (2012), “Detection and identification methods and new tests as developed and used in the framework of COST837 for bacteria pathogenic to stone fruits and nuts”, *J. Plant Pathol.*, **94(1)**, pp.127-133.
- [13] D. Kone, et al. (2015), “Confirmation of *Xanthomonas axonopodis* pv. *manihotis* causing CBB in Ivory Coast”, *Plant Dis.*, **99(10)**, p.1445.
- [14] N. Parkinson, et al. (2007), “Phylogenetic analysis of *Xanthomonas* species by comparison of partial gyrase *B* gene sequences”, *Int. J. Syst. Evol. Microbiol.*, **57(Pt 12)**, pp.2881-2887.
- [15] S.F. Altschul, et al. (1990), “Basic local alignment search tool”, *J. Mol. Biol.*, **215**, pp.403-410.
- [16] S. Kumar, et al. (2016), “MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets”, *Mol. Biol. Evol.*, **33(7)**, pp.1870-1874.
- [17] E.R. Moore, et al. (1997), “16S rRNA gene sequence analyses and inter- and intragenic relationships of *Xanthomonas* species and *Stenotrophomonas maltophilia*”, *FEMS Microbiol. Lett.*, **151(12)**, pp.145-153.
- [18] E.R. Goncalves, Y.B. Rosato (2002), “Phylogenetic analysis of *Xanthomonas* species based upon 16S-23S rDNA intergenic spacer sequences”, *Int. J. Syst. Evol. Microbiol.*, **52(Pt 2)**, pp.355-361.
- [19] N. Parkinson, et al. (2009), “Phylogenetic structure of *Xanthomonas* determined by comparison of *gyrB* sequences”, *Int. J. Syst. Evol. Microbiol.*, **59(Pt 2)**, pp.264-274.
- [20] M. Ferreira-Tonin, et al. (2012), “Phylogenetic analysis of *Xanthomonas* based on partial *rpoB* gene sequences and species differentiation by PCR-RFLP”, *Int. J. Syst. Evol. Microbiol.*, **62(Pt 6)**, pp.1419-1424.
- [21] L. Hauben, et al. (1997), “Comparison of 16S ribosomal DNA sequences of all *Xanthomonas* species”, *Int. J. Syst. Bacteriol.*, **47(2)**, pp.328-335.
- [22] D.J. Jin, C.A. Gross (1989), “Three *rpoBC* mutations that suppress the termination defects of rho mutants also affect the functions of nusA mutants”, *Mol. Gen. Genet.*, **216**, pp.269-275.
- [23] I. Wonni, et al. (2015), “Report of CBB caused by *Xanthomonas axonopodis* pv. *manihotis* in Burkina Faso”, *Plant Dis.*, **99(4)**, p.551.

Initial study of learning curves in robot-assisted radical prostatectomy

Le Chuyen Vu^{1*}, Vinh Hung Tran², Phuc Cam Hoang Nguyen², Van An Nguyen²,
Vu Phuong Do², Ngoc Chau Nguyen², Lenh Hung Do³

¹Medical University of Pham Ngoc Thach, Ho Chi Minh city

²Binh Dan Hospital, Ho Chi Minh city

³University of Medicine and Pharmacy, Ho Chi Minh city

Received 20 April 2018; accepted 13 February 2019

Abstract:

The application of robot-assisted laparoscopic techniques is new and generates numerous benefits for patients. Here, we summarise the experience of our first series through 52 cases of prostate cancer treated by robot-assisted radical prostatectomy (RARP) in the Department of Urology of Binh Dan Hospital, from December 2016 to September 2017, to study the learning curves of this procedure. In this clinical comparative study, 52 patients diagnosed with prostate cancer (clinical stage T1 to T3) received RARP with and without nerve sparing as well as standard pelvic lymphadenectomy. Patients were divided into 4 groups according to their surgeon (surgeons A, B, C, and D, with 22, 12, 10, and 8 patients, respectively) for comparison. Research variables were cancer stage, pre- and postoperative prostate-specific antigen (PSA) serum levels, Gleason scores, lymph node metastasis, estimated blood loss, surgery time, urinary incontinence, hospital stay, and complications. Mean age, PSA, and stage of cancer were statistically similar ($p>0.3$). Operative times were 194.55, 269.17, 236.00, and 306.88 min, respectively ($p<0.01$). Mean estimated blood losses were 363.64, 404.17, 322.22, and 253.75 ml, and were significantly different ($p<0.01$). Nine patients required blood transfusion. The lengths of hospital stay were 5.73, 12.92, 5.10, and 6.13 days, and were not similar among groups ($p<0.05$); however, drainage times and complication rates between groups ($p<0.01$) were statistically significant. The optimal learning curve for operative times was achieved after 20 cases. Our initial RARP results were relatively strong, suggesting that surgery could be safely performed with acceptable complications.

Keywords: learning curve, prostate cancer, radical prostatectomy, robot-assisted surgery.

Classification number: 3.2

Introduction

Radical prostatectomy is considered the gold standard in treating locally advanced prostate cancer. Currently, 3 main methods of surgery exist: open, classic endoscopic, and robot-assisted surgery. The application of robot-assisted surgery is a new step that engenders numerous benefits to patients. We studied the learning curves (LCs) of robot-assisted surgery through 52 cases of radical prostatectomy performed in the Department of Urology, Binh Dan Hospital, from December 2016 to September 2017.

Materials and methods

This longitudinal study focused on 52 patients with prostate cancer (clinical stage T1 to T3) who underwent robot-assisted radical prostatectomy (RARP) with and without preservation of neurovascular bundle, as well as standard pelvic lymphadenectomy. The study variables were cancer stage, pre- and postoperative prostate-specific antigen (PSA) levels, Gleason scores, lymph node metastasis, estimated blood loss during surgery, operative time, and postoperative urinary incontinence. We examined the improvement of these variables over an 11-month period, comparing the differences between the 4 console surgeons: surgeon A (22 patients), surgeon B (12 patients), surgeon C (10 patients), and surgeon D (8 patients).

Results

Age

Table 1. Age.

Age	All (*)	A	B	C	D
Min	49	55	49	52	62
Max	80	80	79	77	80
Mean	66.27	66.59	64.17	64.40	70.88
SD	8.96	8.12	8.58	7.06	5.82

All (*): means all patients.

*Corresponding author: Email: vulechuyen@gmail.com

No differences in age exists between All and groups A, B, and C. However, group D was older ($p < 0.05$; Table 1).

Preoperative PSA

Table 2. Preoperative PSA (ng/ml).

PSA	All	A	B	C	D
Min	4.5	7	18	9	5
Max	100	100	100	100	50
Mean	41.51	43.38	41.00	55.11	22.06
SD	28.81	35.51	24.77	38.07	13.53

Furthermore, no differences existed in PSA levels between All and groups A and B. Group C had higher levels and group D had lower levels ($p < 0.01$; Table 2).

Gleason scores and tumour stages

Table 3. Gleason scores.

Gleason	All	A	B	C	D
$\geq 3+4$	24	11	5	5	3
$4+3$	8	3	2	2	1
$\geq 4+4$	20	8	5	3	4
Stage					
1	8%	5.3%	9.1%	11.1%	14.3%
2	42%	50%	22.7%	38.9%	57.1%
3	50%	44.7%	68.2%	50%	28.6%

No differences existed in Gleason scores between the groups (Table 3).

Operative time (minutes)

Table 4. Operative time (OT).

OT	All	A	B	C	D
Min	105	120	105	150	225
Max	480	270	480	315	420
Mean	237.02	194.55	269.17	236.00	306.88
SD	69.14	46.72	109.15	60.91	56.06

The mean operative time was equal between All and groups B and C ($p > 0.05$), whereas it was shorter in group A and longer in group D ($p < 0.05$; Table 4).

Estimated blood loss (ml)

Table 5. Estimated blood loss (EBL).

EBL	All	A	B	C	D
Min	80	100	100	100	80
Max	1400	1000	1400	800	700
Mean	345.77	363.64	404.17	322.22	253.75
SD	205.45	209.96	272.01	223.76	201.35

Drainage time (days)

Table 6. Drainage time.

Drainage time	All	A	B	C	D
Min	1	1	3	1	3
Max	35	8	35	11	7
Mean	5.67	4.45	9.83	3.70	5.25
SD	4.05	2.15	9.42	2.87	1.49

Drainage time was different between groups ($p < 0.05$; Table 6).

Postoperative hospital stay (days)

Table 7. Hospital stay and complication rate.

Postoperative hospital stay	All	A	B	C	D
Min	2	2	4	3	5
Max	38	11	38	11	7
Mean	7.33	5.73	12.92	5.10	6.13
SD	4.33	2.45	9.70	2.56	0.99
Complications	13/52	4/22	6/12	2/10	1/8

Drainage time differed between groups ($p < 0.05$); however, it exhibited a close relationship with hospitalised time and complication rate in each group ($p < 0.01$; Table 7).

LCs

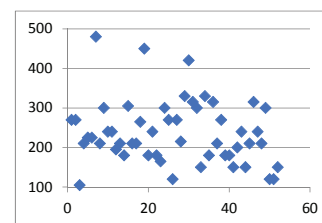


Chart 1. LC in time for All.

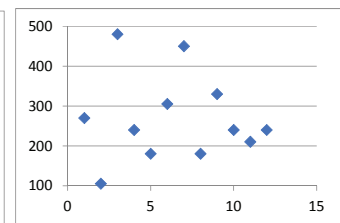


Chart 2. LC in time for group B.

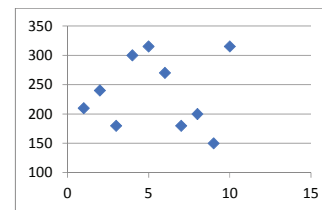


Chart 3. LC in time for group C.

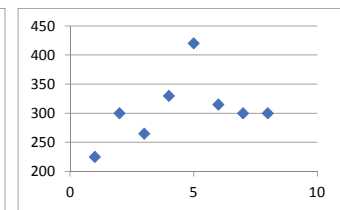


Chart 4. LC in time for group D.

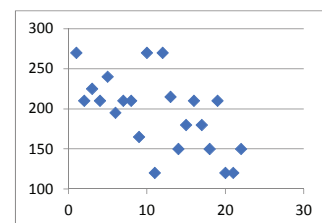


Chart 5. LC in time for group A.

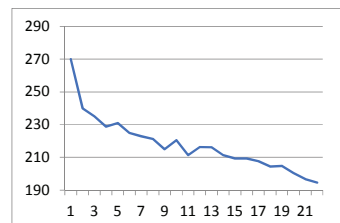


Chart 6. Group A: cumulative mean.

The LCs of All and group A were linear, whereas those of groups B, C, and D were nonlinear because of an insufficient number of cases. In conclusion, approximately 20 cases are required to achieve optimum efficiency in terms of LCs (Charts 1-6).

Postoperative recovery

Early complications in postoperative recovery included:

- Nine cases of abdominal fluid collection, in which 2 cases had to be redrained.
- Two cases of urethral catheters slipping, which necessitated reinsertion.
- One case of subcutaneous emphysema.
- One case of acute myocardial infarction.
- One case of intestinal occlusion, which necessitated emergency reoperation.

In addition, 16 out of the 52 patients suffered urinary incontinence from 3 weeks to 2 months postoperatively. PSA levels were measured every 3 months after surgery. In 49 cases, they dropped to a nadir of 0.01-0.4 ng/ml. In 3 cases of stage 3 prostate cancer, we performed salvage prostatectomy and PSA did not reach ideal levels (4.7, 11.3, and 14.7 ng/ml). Thus, we had to send these patients to the oncology department to begin androgen blockade therapy.

Discussion

The prostate gland is an organ of the genitourinary system; it is small and deep in the pelvis of men. As previously mentioned, radical prostatectomy is still considered the gold standard in treating local and locally advanced prostate cancer. It is used for prostate cancer in otherwise healthy patients and with a life expectancy over 10 years. Furthermore, it is the most common urology surgery at present [1, 2].

Open radical prostatectomy for cancer has been performed at Binh Dan hospital for a long time, although the number is not high. Since 2000, as endoscopic surgery has developed worldwide, laparoscopic surgery has developed quite strongly in our Urology Department. At the end of 2004, we began employing laparoscopic surgery for prostatectomy. During the period of 2004-2006, we performed laparoscopic radical prostatectomy for 23 cases of prostate cancer. Since then, we have performed over 200 cases using this type of surgery and have published numerous papers both nationally and internationally

regarding surgical and stitching techniques, patients' quality of life, and the value of the predictors of surgery [3-7].

However, this technique has increasingly had disadvantages exposed, such as a narrow surgical field, difficult operation, and long LC. When robotic surgery was born, these drawbacks were minimised, and both patients and urologists have increasingly advocated this type of surgery when the patient's choice of treatment is surgery rather than radiotherapy [8, 9].

The concept of LCs is a crucial topic in surgery and one of the less mentioned aspects. Abboudi, et al. [10] presented a noteworthy article that evaluated the concept of LCs in urological procedures. Specifically, the authors undertook a unified approach to systematically evaluate materials focusing on the LC of some urological procedures, including primarily radical prostatectomy and partial segmental kidney surgery.

Most studies have focused on prostatectomy but have poorly documented their methodological quality, including a series of surgeries that primarily limited the number of surgeons and selection of heterogeneous results to study the LC, focusing on short-term results [11-13].

By contrast, the literature on open surgery or open prostatectomy is of higher quality, including many large studies and the application of sophisticated statistical methods; however, robotic surgery is still preferable. With these limitations, we concluded that the duration of the LC for the operative time of robotic surgery is between 50 and 200 cases and the benefit of surgical margin is between 50 and 600 cases. Moreover, urine and erectile control were reported in 200 cases [10]. Thompson, et al. [14] evaluated the LC of an open surgeon with experience of more than 3,000 prostatectomy cases before beginning robot-assisted surgery. The study demonstrated that the effect of the robot overtook open surgery after 100 cases of sexual function scores and marginal rates of pT2 cancer, whereas approximately 150 cases was necessary to achieve urinary function. In addition, the efficiency of the robot continued to improve, with scores of sexual function increasing after 600-700 cases and urine continence increasing after 700-800 cases. Similarly, the negative margin was stable after 400-500 cases in pT2 and 200-300 in pT3-4. However, no evidence exists that further improvements can be achieved.

Davis, et al. studied the LCs of more than 71,000 prostatectomy cases in 300 hospitals in the United States, and between laparoscopic and open radical prostatectomy

revealed a longer operative time (4.4 vs. 3.3 hours), shorter hospital stay (2.2 vs. 3.2 days), and fewer complications (10.4% vs. 15.8%) for laparoscopic prostatectomy [15].

Other studies have shown improvements in the LCs of robotic surgeons when robotic surgery is part of formal residency/fellowships [16], graduate training [17], and/or regular surgical simulator practice [18]. A study of numerous surgeons who performed open prostatectomy and robots demonstrated that robotic surgery has fewer complications, shorter hospital stays, and reduced blood transfusion rates [19]. Moreover, the LC is enhanced in robotic arms, which have been shown to be superior in terms of team work as opposed to the surgeon's time undergoing surgery.

Good, et al. [20] studied match cases, comparing 531 open and 550 laparoscopic prostatectomies to investigate the LCs on blood loss, surgery time, and rate of complications (Clavien-Dindo grade III). The LC for the overall margin free rate was longer for the pT2 group than for the laparoscopy group but was shorter for urinary control. For apex margins, the long LC for the laparoscopy group and lower rates for the robot-assisted group ($p \leq 0.001$). They concluded that both methods have a long LC; however, we found significant benefits in lower margin rates and early improvements in urinary control over laparoscopy, particularly in the prostate apex.

In our series, it is too early to draw conclusions about our LC because of the experience of the console surgeons of only 8-22 cases. However, it can clearly be seen that the mean operative time decreased when the surgeon had more cases and was clearly identified in 20 cases (group A; Charts 5-6). Furthermore, we observed that the surgeon with the longest duration of surgery had the least blood loss (group D, 307 minutes and 237 ml; Table 4-5). One factor to consider was the correlation between the longest hospital stay and the highest incidence of complications (group B, 12.9 days, complication 6/12; Table 7). One of this study's weaknesses is that it did not investigate the margins of surgical specimens.

Conclusions

Although the number of patients was low and the follow-up time was short, our initial results for RARP suggest that this type of surgery can be performed safely with acceptable complications. The optimal LC for operative time was achieved after 20 cases.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] V.L. Chuyen, D.Q. Oanh, T.N.K. Linh (2010), "Mass screening of prostate cancer in Vietnam: status and our opinion", *Urologic Oncology*, **10(6)**, pp.673-676.
- [2] H.T. Dong, V.L. Chuyen (2014), "Epidemiology and prevention of prostate cancer in Vietnam", *Asian Pac. J. Cancer Prev.*, **15(22)**, pp.9747-9751.
- [3] V.V. Ty, V.L. Chuyen (2006), "Radical prostatectomy for prostate cancer, Initial experiences at Binh Dan Hospital", *Medical Ho Chi Minh city*, **10(1)**, pp.136-144.
- [4] V.L. Chuyen, V.T. Vu (2006), "Laparoscopic radical prostatectomy for prostate cancer, our experiences in first 23 cases at Binh Dan Hospital", *Medical Ho Chi Minh city*, **10(4)**, pp.173-182.
- [5] V.L. Chuyen, N.T. De (2012a), "Retroperitoneoscopic radical prostatectomy: factors influencing surgical results", *Medical Ho Chi Minh city*, **16(3)**, pp.88-92.
- [6] V.L. Chuyen, N.T. De (2012b), "Endoscopic extraperitoneal radical prostatectomy: surgical complications", *Medical Ho Chi Minh city*, **16(3)**, pp.93-97.
- [7] V.L. Chuyen (2013), "Survey the quality of life of the patients after having the endoscopic extraperitoneal radical prostatectomy", *Journal of Endoscopic Surgery*, **1**, pp.51-57.
- [8] J.A.Jr Smith (2016), "Radical prostatectomy which patients benefit most from surgery?", *J. Urol.*, **176(2)**, p.437.
- [9] V.R. Patel, A.S. Tully, R. Holmes, J. Lindsay (2005), "Robotic radical prostatectomy in the community setting the learning curve and beyond: initial 200 cases", *J. Urol.*, **174(1)**, p.26972.
- [10] H. Abboudi, M.S. Khan, K.A. Guru, et al. (2014), "Learning curves for urological procedures: a systematic review", *BJU Int.*, **114**, pp.617-629.
- [11] S.B. Farnham, T.M. Webster, S.D. Herrell, J.R.Jr Smith (2006), "Intraoperative blood loss and transfusion requirements for robotic assisted radical prostatectomy versus radical retropubic prostatectomy", *Urology*, **67(2)**, p.3603.
- [12] M. Froehner, V. Novotny, R. Koch, S. Leike, L. Twelker, M.P. Wirth (2013), "Perioperative complications after radical prostatectomy: open versus robot assisted laparoscopic approach", *Urol. Int.*, **90**, p.3125.
- [13] C.R. Ritch, C. You, A.T. May, S.D. Herrell, P.E. Clark, D.F. Penson, et al. (2014), "Biochemical recurrence free survival after robotic assisted laparoscopic vs open radical prostatectomy for intermediate and high-risk prostate cancer", *Urology*, **83(6)**, p.130915.
- [14] J.E. Thompson, S. Egger, M. Böhm, et al. (2014), "Superior quality of life and improved surgical margins are achievable with robotic radical prostatectomy after a long learning curve: a prospective single-surgeon study of 1552 consecutive cases", *Eur. Urol.*, **65**, pp.521-531.
- [15] J.W. Davis, U.S. Kreaden, J. Gabbert, R. Thomas (2014), "Learning curve assessment of robot-assisted radical prostatectomy

compared with open-surgery controls from the premier perspective database”, *J. Endourol.*, **28(5)**, pp.560-566.

[16] G.B. Di Pierro, P. Baumeister, P. Stucki, et al. (2011), “A prospective trial comparing consecutive series of open retropubic and robot-assisted laparoscopic radical prostatectomy in a centre with limited caseload”, *Eur. Urol.*, **59**, pp.1-6.

[17] T.J. Leroy, D.D. Theil, D.A. Duchene, et al. (2010), “Safety and perioperative outcomes during learning curve of robot-assisted laparoscopic prostatectomy: A multi-institutional study of fellowship-trained robotic surgeons versus experienced open radical prostatectomy surgeons incorporating robot assisted laparoscopic prostatectomy”, *J. Endourology*, **24**, pp.1665-1669.

[18] C.D. Lallas, J.W. Davis (2012), “Robotic surgery training

with commercially available simulation systems in 2011: a current review and practice pattern survey from the society of urologic robotic surgeons”, *J. Endourol.*, **26**, pp.283-293.

[19] F. Rozet, J. Jaffe, G. Braud, J. Harmon, X. Cathelineau, E. Barret, et al. (2007), “A direct comparison of robotic assisted versus pure laparoscopic radical prostatectomy: a single institution experience”, *J. Urol.*, **178(2)**, p.47882.

[20] D.W. Good, G.D. Stewart, A. Laird, J.U. Stolzenburg, D. Cahill, S.A. McNeill (2015), “A critical analysis of the learning curve and postlearning curve outcomes of two experience- and volume-matched surgeons for laparoscopic and robot-assisted radical prostatectomy”, *J. Endourology*, **29(8)**, pp.939-947.

Evaluation of graft function in renal retransplantation at Hue Central Hospital

Tam Vo¹, Ngoc Tuan Anh Dang^{1*}, Nhu Hiep Pham¹, Gia Khanh Pham²

¹Hue Central Hospital

²Vietnam Society of Organ Transplantation

Received 28 October 2018; accepted 18 February 2019

Abstract:

This study aims to assess the results of kidney retransplantations and the survival of second kidney allografts performed in our center. Methodology: of 550 kidney transplantations performed since July 2001, 26 were second kidney transplants. Many recipients were sensitized. All patients were treated with the same immunosuppressive regimen. Results: grafts were functionally effective in all patients with kidney retransplantations one month after the transplant. Graft survival was 100% after one year (17/17) and after five years (1/1). Graft survival was similar for second and primary kidney transplants performed during the same period of time. There were three cases of acute rejection, but all of these cases responded to the anti-rejection treatment. Conclusions: the rate of acute rejection was similar for both second and primary transplantations in our hospital, and the second graft outcome demonstrated effectiveness 12 months after transplantation.

Keywords: graft, renal retransplantation, survival.

Classification number: 3.2

Introduction

Since immunosuppressive drugs such as CsA, FK 506, or MMF were introduced, the short-term graft survival rate of kidney transplant increased from 10 to 20%, but the rate of chronic rejection remained high [1]. About 5 to 24% of graft dysfunction occurred within five years post-transplant, and 50 to 80% of rejection patients had to return to dialysis [2]. Graft failure has been one of the most common causes of end-stage renal disease, accounting for 25% of 30% of those on the waiting list [3].

A successful kidney transplant offers enhanced quality and duration of life and is more effective in terms of medical outcome and patient satisfaction relative to long-term dialysis [4]. However, retransplant recipients have a higher risk of immunity reactions relative to those receiving their first grafts [5-7]. The success of retransplantation has improved significantly due to pre-transplant screening and post-transplant treatment [5]. Recent reports have indicated that the success rate of kidney retransplantations relative to first kidney transplants is nearly identical one year after the transplant. Therefore, the evaluation of the outcome as well as an enhanced understanding of the characteristics of this group of patients is necessary for the optimal treatment of chronic renal failure.

Of 550 cases performed at Hue Central Hospital since 2001, 26 cases were retransplanted. We initiated a second kidney transplant from 2012; this number has increased every year. As we anticipated, the retransplant rate has increased rapidly. There were 10 patients with immunological risk (PRA>25%), including two cases with PRA of greater than 80%. This article seeks to evaluate the outcome of immunosuppressive therapy on retransplant patients and the renal function of patients at one month and one year post-transplantation.

Methods

Of 550 cases between 7/2001 and 4/2018, 26 were

*Corresponding author: Email: dangngoctuananh@yahoo.com

retransplant cases. The first one was performed in March 2012. Monitoring time was one month, three months, and 12 months post-transplantation. HLA antibodies were detected using micro ELISA or Luminex methods. Patients selected were either ABO-compatible or of the same blood type. Patients with high-HLA antibodies required at least three HLA matches. High-HLA mismatches were recommended for patients who did not have HLA class I antibodies or specific HLA class II antibodies. Crossmatches were performed through complement-dependent lymphocytotoxicity (CDC).

Living donors' details, such as age, gender, and serum creatinin were collected. Recipients' features such as gender, antibodies concentration, HLA mismatches, crossmatches, delay graft function, acute rejection episodes, creatinine intervals at one month, three months, and 12 months post-transplantation, characteristics of the first kidney transplant (graft survival time, interval between two transplants, cause of first graft loss and information on first immunosuppression regimen).

Return to dialysis or death with graft functions was referred to as graft loss. Acute rejections were defined by a combination of clinical criteria and the renal biopsy. Delayed graft functions were defined by requiring at least one dialysis session in the first seven days post-transplantation.

Immunosuppressive regimen

Despite the abundance of immunosuppressive drugs, all patients were treated with quadruple therapy consisting of induction therapy by antithymocyte globulin (ATG) horses or rabbits or monoclonal antibody rather than standard maintenance therapy. The dose of ATG rabbits and horses was 1-1.25 mg/kg/day for seven patients as opposed to 10 patients, respectively. Monoclonal antibody (Basiliximab) was used as the induction regimen for nine patients. Maintenance therapy with cyclosporin (two cases) or FK506 (24 cases) combined with MMF 2 g/day were used along with intravenous corticosteroid, followed by oral administration. The CsA dose was initiated at 7-8 mm/kg/day in two divided doses and adjusted by trough level to 150-200 ng/ml. Tarcrolimus dose was 0.1-0.15 mg/kg/day in two divided doses and adjusted by trough level 8-10 ng/ml. Prednisolone was administered at the rate of 0.4 mg/kg/day followed by 0.3 mg/kg/day and then 5 mg/day after three months. Acute rejections were treated with bolus methylprednisolon 500 mg/day x 3 days. ATG was used for treatment of steroid-resistant rejection. For acute mediated antibody, the plasmapheresis was indicated as $\pm$ IVIG [8, 9].

Results

Number of patients each year

Year	2012	2013	2014	2015	2016	2017	3/2018
Cases	1	0	3	3	7	9	3

Characteristics of donors

Gender	N	Average age	Average serum creatinin
Male	25	28.09 $\pm$ 8.05	88.59 $\pm$ 10.61 μ mol/l
Female	1		

100% were living donors.

Characteristics of recipients

Gender	N	Max	Min
Male	15	61	31
Female	6	55	30
Total	21	Average: 42.89 $\pm$ 9.59	

ABO blood type

Type (-donor-recipient)	N	Percentages
Same type	15	57.7%
ABO- compatability	11	42.3%
Total	26	100%

HLA matches

HLA match	N	Percentages
2/6	2	7.69 %
3/6	19	73.07 %
4/6	5	19.24%

There were some immunologically high-risk cases that required highly HLA-matched proportions and more extensive immunosuppression regimens.

HLA antibodies

HLA Antibodies	N	Percentages
<25%	16	61.53%
25-80%	8	30.77%
$\geq$ 80%	2	7.7%
Total	26	100%

Of two cases with $\text{PRA} \geq 80\%$ and one with acute mediated antibody required plasmapheresis, one had been desensitized pre-transplant with five plasmapheresis sessions pre-transplant.

Crossmatch

100% of retransplant cases were negative.

Timing of first graft survival

Timing of first graft survival (year)	N	Percentages
<1	1	3.8%
1-5	1	3.8%
5-10	11	42.4%
>10	13	50.0%

There were seven cases who had their first transplant at Hue Central Hospital and 16 in China or at other centers.

Serohepatitis

All patients in this group had been infected and had viral load below detectable levels before transplantation, with anti-HCV positive and HBsAg positive were nine and patients, respectively. All patients with hepatitis virus were placed on low-dose immunosuppression therapy if they had minimal antibodies levels and high HLA matches.

Virus	N	Percentages (%)
HBsAg +	1	4.8
Anti HCV +	9	34.62

Average blood creatinin at 1 month, 3 months, 12 months, 3 years, 5 years

The number of kidney retransplants which experienced normal functioning at one month, three months, twelve months, three years, and five years post-transplant were 26, 25, 17, 4, and 1, respectively.

Timing	1 month	3 months	1 year	3 years	5 years
Creatinin	89.19±36,16	87.96±13,9	94.83±20,16	88.75±10,5	85
Median	85	88.5	90	85.5	
Number	26	25	17	4	1

Acute rejection

There were three cases of acute rejection, which required three to five dialysis sessions; one case with acute mediated antibody needed plasmapheresis. Two cases were treated

with bolus solumedrol and ATG. All three cases responded positively to treatment, and kidney function returned to normal before discharge.

Dicussion

We initiated the second kidney transplant starting in March 2012. This number has increased each year because the patients who had previously been transplanted went up and because of immune tests such as the specific HLA antibody Luminex, the experience of the pathology of kidney transplant biopsies, immunosuppressive drugs such as ATG, IVIG... which are fully on the market, the use of plasmapheresis, and knowledge and experience of anti-rejection treatment with high immunologically high-risk recipients [8, 9].

The disadvantage of this article was that the number of patients remained minimal and the follow-up was relatively short. There was merely one five-year patient. Data analysis performed for 50,291 recipients from deceased brain donors was reported by UNOS from 1991 to 1997; Cecka demonstrated that the rate of patient survival five years post-transplant was 82% and that the graft function rate was 63% [1]; the research of Stephani's, et al. perceived the rate of second graft survival as only 2% lower than that among living donors or deceased brain donors. In contrast, in one study, Moss, et al. demonstrated that the short-term outcome of kidney retransplantation relative to first kidney transplantation was the same [4]. In our study, kidney retransplant function was successful in 100% of cases when evaluated at one month, three months, and 12 months. There were three cases that involved delayed renal function, but kidney function then returned to normal. Therefore, recipients from living donors, as well as new immunosuppressive agents, have been proven to produce effective results in short-term renal function. Particularly, graft survival rates interval at one year were not different between first transplants and retransplantations in this group.

Acute rejection is a formidable complication in kidney retransplant patients due to the immunological risk involved [2]. Among these 26 retransplant cases, in three cases, acute rejection accounted for 11.54%, equivalent to approximately 10% of transplants in 550 patients at Hue Central Hospital.

The improvement of graft survival reflects the effect of immunosuppressive drugs and the effective control of cross-reactions between donors and recipients [3]. Additionally, it would be fruitful to minimise cold ischemia time in renal transplantations from living donors and precisely evaluate

immunologically high-risk recipients for better prevention and treatment.

Conclusions

The acute rejection rates and graft function of kidney retransplantations in comparison with first kidney transplantations was not statistically significantly different from the interval at one year post-transplant at Hue Central Hospital.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] J.M. Cecka (1998), "The UNOS scientific renal transplant registry", *Clin. Transpl.*, pp.1-16.
- [2] Chethan Puttarajappa, Ron Shapiro, H.P Tan (2012), "Antibody mediated rejection in kidney transplantation: a review", *Journal of Transplantation*, 193724, Doi: 10.1155/2012/193724.
- [3] Christophe Legendre (2011), *La transplantation rénale*, 942pp, Médecine - Sciences Publications,.
- [4] A. Moss, J.S. Najarian, et al. (2000), "5,000 kidney transplants - a single center experience", *Clin. Transpl.*, pp.159-171.
- [5] R. Marcen, J. Pascual, et al. (2001), "Outcome of cadaveric renal transplant patients treated for 10 years with cyclosporine", *Transplantation*, **72**, pp.57-62.
- [6] Stephanie Coupel, Magali Giral-Classe, et al. (2003), "Ten year survival of second kidney transplants: impact of immunologic factors and renal function at 12 months", *Kidney International*, **64**, pp.674-680.
- [7] K.L. Womer, J.P. Vella, M.H. Sayegh (2000), "Chronic allograft dysfunction: mechanisms and new approaches to the therapy", *Semin. Nephrol.*, **20**, pp.126-147.
- [8] Monica Grafals, Enver Akalin (2009), "The highly sensitized renal transplant recipient", *Nephrology Rounds*.
- [9] Kwaku Marfo (2011), "Desensitization protocols and their outcome", *Clin. J. Am. Soc. Nephrol.*, **6**, pp.922-936.

Examination of the accuracy of the quantitative in-house kit for determining zinc concentration in seminal fluid

Thi Trang Nguyen*, Thi Minh Phuong Le, Thi Huyen Trang Do, Thi Quynh Dien Tran, Ngoc Thach Pham

Hanoi Medical University

Received 10 December 2018; accepted 20 February 2019

Abstract:

Zinc in seminal fluid originates primarily from the prostate gland. It is pivotal for male sexual function because it affects the quantity, quality and mobility of sperm. **Materials and methods:** the semen samples were obtained from 300 male partners of infertile couples who attended the Department of Biomedicine and Genetics at the Hanoi Medical University between the ages of 18 and 50 years; they were then analysed for routine seminal parameters. They were collected and analysed according to WHO 2010 guidelines. Seminal fluid was centrifuged at 1,500 rpm for 10 minutes, and floating fluid collected for zinc quantification using spectroscopy with 5-Br-PAPS was used as a color indicator. In a pH 8.6-buffer solution, in a buffered media, zinc reacts with specific complexing 5-Br-PAPS form a stable color compound. The optical density is directly proportional to the concentration of zinc in the semen. The results yielded a linear regression model of $y = 0.0666x + 1.2026$ with a correlation coefficient of $r = 0.9956$. The calibration function was $y = 0.9977x$ with $R^2 = 0.9995$. The repeatability was $SD = 0.004$, and the coefficient of variation was $CV\% = 0.27\% < 5\%$. In terms of intermediate precision, the standard deviation was $(SD) = 0.01$, and the coefficient of variation was $CV\% = 0.64\% < 5\%$. Trueness was $t_{ex} = 2.076 < t_t = 2.262$. Specificity and sensitivity were 100% at 64x dilution. Specificity and sensitivity were 100% and 99.05% respectively. A significant correlation was discovered between the two methods, with $r = 0.975$ and $p < 0.001$; the average difference between the two methods was 0.0002. **Conclusion:** successfully completed the kit for determining zinc concentration in semen by the colorimetric method.

Keywords: male infertility, seminal zinc, seminogram, spectrophotometric method.

Classification number: 3.2

Introduction

Infertility is “a disease of the reproductive system defined by the failure to achieve a clinical pregnancy after 12 months or more of regular unprotected sexual intercourse” (by WHO) [1, 2]. The incidence of infertility in the world is about 13 to 18%; the causes for men and women are similar [3]. Recently, aside from tests involving semen analysis, genetic testing and hormonal methods, biochemical tests involving fructose and zinc have become more common the diagnosis and treatment of infertility [4]. They have therefore been established as reliable indicators of human male fertility. Understanding the role of these indicators can help to enable a deeper understanding of the mechanisms that cause male infertility [5, 6].

Zinc is an essential factor for the male reproductive system. Zinc deficiency in the reproductive system causes hypogonadism and gonadal hypofunction [7, 8]. Zinc is crucial for maintaining normal functioning of the testes and prostate, affecting the quantity, quality and mobility of sperm [9, 10]. In Vietnam, tests for concentrations of zinc in seminal plasma were used in 2013 at the Hanoi Medical University Hospital, and these tests are increasingly popular in clinics. The necessity of using test kits to quantify zinc in semen is increasing. However, in Vietnam, no units nor establishments currently produce this test. Therefore, all tests must be imported from abroad with many intermediate costs, often increasing the cost of the test.

Materials and method

Objects

The study subjects were semen samples of male patients who visited for examination, counseling and testing semen at the Genetic Counseling Center and the Urology and Urology Clinic at Hanoi Medical University Hospital from June,

*Corresponding author: Email: trangnguyen@hmu.edu.vn

2017 to March, 2018. The procedure for collecting semen samples was conducted in accordance with the guidance of the World Health Organization in 2010. Participants were required to abstain from sex for 2 to 5 days. The semen was deposited into a sterile vial with no spermicide and analysed within 2 hours of sampling.

The formula used for a sample size for a descriptive study by S.K. Luanga and Lemeshow [11] is as follows:

$$n = Z_{1-\alpha/2}^2 \frac{p(1-p)}{\epsilon^2 p^2}$$

In this equation $Z_{1-\alpha/2}$: confidence factor (with 95% confidence, $Z_{1-\alpha/2}=1.96$); $\alpha=0.1$ (reliability); $\epsilon=0.10$; $p=95\%$ (reference process precision), n =number of experiments required, calculated by 21; a round of 30 was executed.

To calculate the sample size to determine sensitivity, specificity, and equivalence the following were used:

$Z_{1-\alpha/2}$: confidence factor (with 95% confidence, $Z_{1-\alpha/2}=1.96$).

According to Zahoor Ahmed and colleagues, in 2010, the percentage of men with low zinc concentrations in the azoospermia and oligospermia groups was $p=25\%$ [12]; ϵ is 0.2, and $n=1.96^2 \times 0.25 \times (1-0.25)$: $(0.2 \times 0.25)^2=288.12$, rounded to 300.

The sample size of 300 was employed to increase accuracy.

On the same sample of semen, zinc concentrations were measured using two methods: one involved the IVD kit (Zinc 5-Br-PAPS, Spinreact company, Spain), and one involved the improvement kit. The difference between the two kits was based on a Pearson correlation, T-test, and Bland-Altman plot.

The selection criteria required semen samples from male patients of reproductive ages between 18 and 50 years without acute illness who consented to participate in the study.

For the exclusion criteria, men with genital cancers, men with HIV, syphilis, and gonorrhea, men suffering from acute illness or mental illness, and men who did not consent to enroll in the study were excluded.

Method

The principle of the method involved direct colorimetric testing without deproteinisation of the sample. In a pH 8.6 buffer solution, zinc reacts with 5-Br-PAPS complexes and produces stable color. The optical density is directly

proportional to the concentration of zinc in the semen.

Measuring the concentration of zinc:

Materials:

Buffer A: Sodium bicarbonate (200 mmol/l), Sodium citrate (170 mmol/l), Dimethylglyoxime (4 mmol/l), Triton-X100 (1%), 5-Br-PAPS (0.08 mmol/l).

Buffer B: Salicyldoxime (2,9 mmol/l).

Working buffer, pH-8.6 (C):4A:1B.

Zinc standard (Merk).

Zinc color 5-Br-PAPS (Spinreac, Spain).

The procedure for zinc quantitative testing in semen:

The steps for the procedure of zinc quantitative testing in semen are as follows:

+ Step 1: the semen sample is centrifuged at 1,500 rpm for 10 minutes. This step is used to settle the sperm cells down to the bottom; only the top of the semen containing the zinc for testing is then gathered, as the sperm is not used in the test.

+ Step 2: 200 μ l of supernatant is added into 200 μ l TCA 370 μ mol/l, mixed thoroughly, and centrifuged at 10,000 rpm for 10 minutes. This step is used to remove the protein.

+ Step 3: 100 μ l of supernatant is into 2 ml of working buffer. This is then incubated at room temperature for 5 minutes.

+ Step 4: optical density (OD) is measured at 530 nm wavelength and 1 cm curvature. For control, 100 μ l of distilled water is substituted for semen. The color of the solution remained unchanged within 1 hour.

$[Zn] \mu\text{mol/l} = (\text{OD sample}/\text{OD blank}) \times C \text{ zinc standard } (\mu\text{mol/l})$.

The sensitivity and specificity are calculated according to the following formula:

The sensitivity (%) = $\frac{\text{True positive}}{\text{True positive} + \text{False negative}} \times 100\%$

The specificity (%) = $\frac{\text{True negative}}{\text{True negative} + \text{False positive}} \times 100\%$

Statistical analysis involved data processing using SPSS software version 20.0. The mean values were compared using student-t-tests. The analysis is to be meaningful when the coefficient was $p<0.05$.

Regarding ethical considerations, the research was approved by the ethical council of Hanoi Medical

University. The patient completely voluntarily participated in the study. All of the information from the database was kept under strict confidentiality. No names were recorded.

Results

Construction of linear regression equations and calibration function

Linear regression equations were used to assess whether the color intensity of the mixture is proportional to the concentration of zinc in the seminal plasma. The linear regression equation was used to calculate zinc concentrations in seminal plasma based on measured photometer densities.

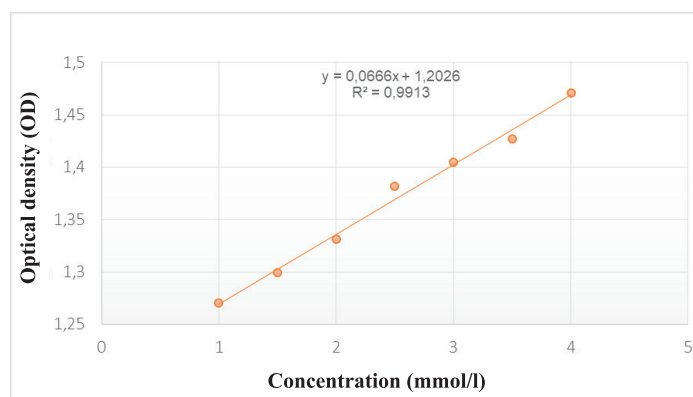


Fig. 1. Linear regression equations.

The regression equation of $y = 0.0666x + 1.2026$ was used, the correlation coefficient of $r=0.9956$ was employed (Fig. 1).

The calibration function is constructed with the standard zinc concentration threshold of 0; 0.5; 1; 1.5; 2 g/l the OD density was measured corresponding to each standard zinc concentration threshold based on the completed test procedure and the calibration curve was established as in Fig. 2.

The calculation function equation is $y = 0.9977x$; $R^2=0.9995$.

Determining the accuracy of the kit (Table 1)

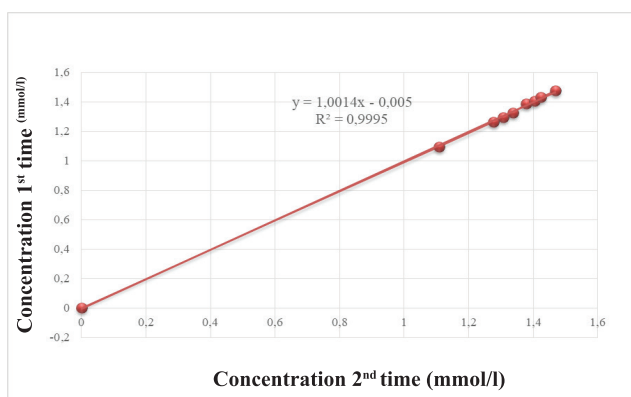


Fig. 2. Calibration curve.

Table 1. Repeatability, intermediate precision and trueness of the in-house kit.

Measurement	Zinc standard concentration (mmol/l)	Repeatability		Intermediate precision			Trueness	
		OD	The zinc concentration measured (mmol/l)	Zinc concentrations measured by different technicians (mmol/l)			OD	The zinc concentration measured (mmol/l)
				Technicians 1	Technicians 2	Technicians 3		
1	1.53	1.352	1.527	1.559	1.523	1.530	1.332	1.521
2	1.53	1.354	1.529	1.524	1.527	1.524	1.335	1.524
3	1.53	1.358	1.533	1.504	1.524	1.534	1.333	1.522
4	1.53	1.352	1.527	1.532	1.525	1.526	1.334	1.523
5	1.53	1.36	1.536	1.513	1.532	1.534	1.340	1.530
6	1.53	1.362	1.538	1.537	1.533	1.524	1.341	1.531
7	1.53	1.356	1.531	1.527	1.536	1.534	1.343	1.533
8	1.53	1.351	1.525	1.514	1.530	1.529	1.338	1.528
9	1.53	1.353	1.528	1.520	1.536	1.532	1.343	1.533
10	1.53	1.356	1.531	1.544	1.528	1.536	1.336	1.525
SD		0.004		0.01			$T_{ex}=2.076$	
CV%		0.27		0.64				

Repeatability: based on the results in the Table 1, a standard deviation (SD) of =0.004 and a coefficient of variation of CV%=0.27 were calculated. The variation coefficient of the in-house kit is within the allowable limits (CV%<5%).

Regarding the intermediate precision, based on the results obtained from the table above, standard deviation of (SD)=0.01 was obtained, so that the coefficient of variation was CV%=0.64%. The coefficient of variation lies within the CV%<5% limit.

Based on the above results, $t_{ex}=2.076$ was calculated. In addition, since $t_t=2.262$, $t_{ex}<t_t$ the testing standards were achieved.

Sensitivity and specificity (Tables 2, 3)

Table 2. Statistics of zinc quantitative results in three groups of patients.

Zinc concentration	The in - house kit	Commercial kit
Group 1 (Control group)		
Below normal concentration	43	43
Normal	57	57
Group 2 (Group with some abnormal seminal indexes)		
Below normal concentration	40	38
Normal	60	62
Group 3 (Azoospermia group)		
Below normal concentration	73	73
Normal	27	27

Table 3. Index of indicators to calculate sensitivity and specificity.

	True positive	False positive	True negative	False negative
Group 1	57	0	43	0
Group 2	60	0	38	2
Group 3	27	0	73	0
Total	208	0	90	2
Sensitivity	99.05%			
Specificity	100%			

Comment:

- True negative: below normal zinc concentrations when tested by both kits.
- False negative: below normal zinc concentrations when tested by the in-house kit above normal zinc concentrations when tested by the commercial kit.
- True positive: normal zinc concentrations when tested with both kits.
- False positive: normal zinc concentrations when tested by the in-house kit; below normal zinc concentrations when tested by the commercial kit.

Sensitivity and specificity were calculated according to the following formula:

$$\text{The sensitivity (\%)} = \frac{\text{True positive}}{\text{True positive} + \text{False negative}} \times 100\%$$

$$\text{The specificity (\%)} = \frac{\text{True negative}}{\text{True negative} + \text{False positive}} \times 100\%$$

The sensitivity of the in-house kit is 99.05%, and the specificity of the self-mixing kit is 100%.

Assessing the capability of detecting the concentration of zinc between the in-house kit and the commercial standard kit (Fig. 3)

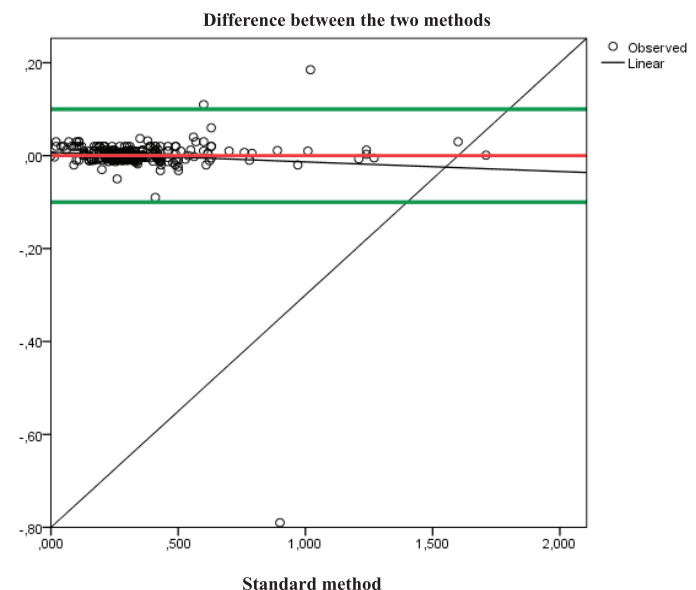


Fig. 3. Bland-Altman diagram assessing the capability of detecting zinc concentrations between the in-house kit and the commercial standard kit.

- Use semen samples of 300 patients. Each sample quantifies 2 times: using a commercial standard kit and in-house kit (the list of 300 patients tested is included in the appendix).

- The Pearson test results demonstrate a strong correlation between the concentrations of zinc measured by the two kits, with $r=0.975$ ($0.983-0.995$), $p<0.001$.

The in-house kit tends to yield higher results than the commercial standard kit. The difference between the two methods of detection is random (the value points are dispersed and follow no pattern), and the error deviation is not related to the zinc quantification results.

The mean difference between the two methods was minimal, or (0.0002) close to 0, with a standard deviation of 4.94%. Most cases were within ± 1.96 standard deviations.

Discussion

Concentrations of zinc in semen can be detected using the method involving 5-Br-PAPS. At suitable pH levels, zinc reacts with 5-Br-PAPS, creating a chelate with stable colors. The darkness of the mixture's color is proportional to the quantity of zinc in semen. Concentrations of zinc can be measured by the optical density of this mixture.

The researchers proceeded to measure the optical density two times in the sequence 1.0; 1.5; 2.0; 2.5; 3.0; 3.5; 4.0 mmol/l, obtaining the average. Ms-Excel was used to draw the linear line.

The linear regression equation is $y = 0.0666x + 1.2026$, zinc concentration is indicated by the independent variables, x , and optical density is the dependent variable, y .

The correlation coefficient of $r=0.9956$, and $0.995 < r < 1$ indicate a strong positive relationship between zinc concentrations and optical density.

• About construction of calibration function

When mixing chemicals, due to subjective or objective reasons, error factors may yield different results between the two tests. In this kit, to minimise these factors and to ensure result stability between the tests, each time chemicals were mixed, the researchers proceeded to build the calibration function.

The calibration function was $y = 0.9977x$, and the correlation relation coefficient was $R^2=0.9995$. Such as computational, quantitative results when using chemical batch after need with coefficient 1, meaning that no additional coefficients. No significant difference in the test results is apparent between the different batches of chemical tests.

The calibration function equation was $y = 0.9977x$, and the correlation coefficient was $R^2=0.9995$.

• Accuracy

In trials, particularly in quantitative testing, numerous error factors affect the test, producing inaccurate results. Therefore, to control these confounding factors, it is necessary to apply precision. Precision results only depend upon the random error factor that is not related to the actual results of the sample. When the precision of the standard deviation is lower, the variance is greater.

Precision is based on three parameters, including repeatability, intermediate precision, and reproducibility. This study only involved experiments that calculated repeatability and intermediate precision; since there is no reference laboratory, it is not possible to calculate the reproducibility.

In this study, the in-house kit has a repeatability with a CV% coefficient of 0.27%, so that the coefficient of variation does not exceed 5%, which indicates that it satisfies the requirements of the analysis.

The calculation of the intermediate yielded, the coefficient of variation of $CV\%=0.64\%$. This coefficient of variation is also valid for not more than 5%. Therefore, the process also meets the requirements of the analysis.

Therefore, when the effect of random error elements is the same, the concentration measured under different conditions has a tolerable range.

• About the trueness

The trueness of the method demonstrates that the degree of proximity between the result obtained and the actual value or accepted value is true (μ).

By experimentally testing the trueness, the result obtained is $t_{ex}=2.076$. Additionally, the table indicates that, the t_t value obtained is 2.262. This means that $t_{ex} < t_t$, and the concentration of zinc measured from this method exert the same effect as

the actual concentration of the sample. The process achieves the required accuracy of an analysis.

A good test kit should have high sensitivity and specificity, which means that the kit has low false positives and false negatives. These are two important criteria for evaluating the quality of an analysis kit.

The kit achieves a specificity of 100% and a sensitivity of 100% as well. It is therefore possible to use this kit to quantify zinc in semen with high reliability.

To reassert the accuracy of the in-house kit, this study continued to compare the results obtained by the kit with those of the standard kit.

Results demonstrated significant correlations between the two kits ($r=0.975$; $p<0.001$) and that, the mean difference between the two methods was 0.0002, equivalent to 4.94%. This difference was not statistically significant.

The chart illustrates that the difference is completely random and independent from the standard scale.

In the kit, proteins were transformed using TCA 370 mmol/l. TCA is a non-poisonous, common, and easy to purchase acid. Transforming proteins before mixing them with color indicators limits the possibility of proteins reacting to the color indicator and produces a mixture of inaccurate color. In this manner, the accuracy of obtained results is considerably improved.

Moreover, the in house test kit uses only simple, low-cost, common, and easy-to-buy chemicals and fewer, chemicals than commercial kits. The IVD kit used in this study is the Zinc 5-Br-PAPS Test (Spain), which is being used widely to measure zinc in semen in laboratories today. This kit also follows the principle of the colorimetric method, but indol is used as an expensive chemical that is difficult to purchase in Vietnam and requires the use of a color-rendering stopper. In addition, the kit must be imported from abroad, through many intermediate stages, resulting in a high cost of testing. This means that this kit is more suitable for use in Vietnam, where determination of zinc concentration in seminal fluid is in high demand; however, the average annual income remains only average.

With the achieved advancements, the researchers hope

their kit can soon be the subject of quantity production, replacing currently imported kits in health institutions.

However, this study has, only tested the kit on a laboratory scale. Assessment on an industrial scale is fundamental for the kit to become subject to quantity production.

Conclusions

Successfully complete the procedure to create the zinc quantification kit using the colorimetric method.

Calibration function: $y = 0.9977x$; correlation coefficients $R^2=0.9995$.

- Accuracy

- Repeatability: $SD=0.004$, coefficient of variation $CV\%=0.27\%<5\%$.

- Intermediate precision: $SD=0.01$, coefficient of variation $CV\%=0.64\%<5\%$.

- Trueness: $t_m=2.076 < t_c=2.262$.

- The sensitivity of the self-mixing kit is 99.05%, and its specificity is 100%.

- The results indicate a strong correlation between the two methods ($r=0.975$; $p<0.001$); the average difference between the two kits is 0.0002, equivalent to 4.94%. The difference has no statistical significance. Most cases yielded zinc concentration data within ± 1.96 standard deviations.

ACKNOWLEDGEMENTS

The authors would like to take this opportunity to extend our sincere thanks to Ministry of Health for providing financials support for the study. We also are grateful for the technical support of the Hanoi Medical University Hospital for the assay of the seminal zinc concentration.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] WHO (2010), *WHO Laboratory manual for the Examination and processing of human semen*, 5th edn., Geneva: WHO Press.
- [2] R.S. Swerdloff, C. Wang (2004), *The testis and male sexual function*, 22nd ed. Goldman L, Ausiello D, editors, Cecil Textbook of Medicine, Philadelphia (PA): Saunders Publisher, pp.1472-1483.

- [3] H. Doshi, O. Heana, T. Hemali (2008), "Zinc levels in seminal plasma and its relationship with seminal characteristics", *Journal of Obstetrics and Gynecology of India*, **58**, pp.152-155.
- [4] A. Agarwal, F.M. Bragais, E. Sabanegh (2008), "Assessing sperm function", *Urol. Clin. North Am.*, **35**(2), pp.157-171.
- [5] T.G. Cooper (2010), *WHO laboratory manual for the Examination and processing of human semen*, 5th edn., Bern: WHO Press, pp.1-157.
- [6] D.I. Lewis Jones, I.A. Aird, M.M. Biljan, et al. (1996), "Effects of sperm activity on zinc and fructose concentrations in seminal plasma", *Human Reproduction*, **11**(11), pp.2465-2467.
- [7] H.H. Sandstead, A.S. Prasad, A.R. Schulert, et al. (1967), "Human zinc deficiency, endocrine manifestations and response to treatment", *Am. J. Clin. Nutr.*, **20**(5), pp.422-442.
- [8] A.E. Omu, H. Dashti, S. Al-Othman (1998), "Treatment of asthenozoospermic with zinc sulphate: andrological, immunological and obstetric outcome", *Eur. J. Obstet. Gynecol. Reprod. Biol.*, **79**(2), pp.179-184.
- [9] S. Biswas, K.M. Ferguson, J. Stedronska, et al. (1978), "Fructose and hormone levels in semen: their correlations with sperm counts and motility", *Fertil. Steril.*, **2**(30), pp.200-204.
- [10] B.O.M. Saleh, N.K. Hussain, A.Y. Majid, et al. (2008), "Status of zinc and copper concentrations in seminal plasma of male infertility and their correlation with various sperm parameters", *The Iraqi Postgraduate Medical Journal*, **7**(1), pp.76-80.
- [11] S.K. Luanga, Lemeshow (1991), *Sample size determination in health studies - a practical manual*, 88pp.
- [12] Zahoor Ahmed, et al. (2010), "Seminal fructose in various classes of infertile patients", *Pak. J. Physiol.*, **6**(1), pp.36-38.

Field trial of beneficial effects of carotenoid-producing *Bacillus aquimaris* SH6 spores to whiteleg shrimps

Thi Tham Nguyen^{1,2}, Thi Huong Nguyen^{1,2}, Thi Ngoc Que Nguyen^{1,2},
Thi My Tran³, Thi Thu Huong Pham¹, Dinh Thang Nguyen¹, Thi Van Anh Nguyen^{1*}

¹Key Laboratory of Enzyme and Protein Technology, University of Science, Vietnam National University, Hanoi

²Faculty of Biology, University of Science, Vietnam National University, Hanoi

³ANABIO Research & Development JSC, Hanoi

Received 4 September 2018; accepted 21 November 2018

Abstract:

This field study was performed to demonstrate that *Bacillus aquimaris* SH6 spores administered as a feed supplement in adequate amounts confer beneficial effects to the whiteleg shrimp. Shrimps were administered either *B. aquimaris* SH6 spores at 1×10^6 CFU g⁻¹ pellet (SH6 group) or a mixture of SH6 and *B. subtilis* SH23 spores at 5×10^5 CFU g⁻¹ pellet for each strain (SH6-SH23 group). After a 28-day feeding period, the number of SH6 spores in the gut of whiteleg shrimp in the SH6 group and the SH6-SH23 group were 5.7×10^3 CFU gut⁻¹ and 7.7×10^3 CFU gut⁻¹, respectively. The total bacterial population of these two experimental groups increased about 10-75% compared to that of the control group. The astaxanthin level and red colour score were the highest in the SH6-SH23 group (1.66 µg.g⁻¹ shrimp; 28-29), followed by the SH6 spore group (1.42 µg.g⁻¹ shrimp; 27-28), and were the lowest in the control group (0.61 µg.g⁻¹ shrimp; 26-27). Nevertheless, the growth rate of shrimps was similar (7.8-8%) among the three groups. In conclusion, feed supplements containing SH6 and SH23 spores co-effectively improved live counts and diversity of microbiota in shrimp guts, as well as improved astaxanthin level and red colour of whiteleg shrimps.

Keywords: astaxanthin, *Bacillus aquimaris*, carotenoid, gut, spore.

Classification number: 3.4

Introduction

Probiotics play a fundamental role in health enhancement of the host. They have the ability to produce enzymes for food digestion and serve as a bio-competitor for inhibition of the growth of pathogenic microorganisms [1], increase in the immune response [2], improvement in water quality [3], and, subsequently, enhancement of the quantity and quality of shrimp aquaculture. Therefore, probiotics are considered the best choice for sustainable shrimp aquaculture development because of their safety characteristics [4].

In recent years, whiteleg shrimp (*Litopenaeus vannamei*) have become among the most valuable aqua-products based on their high nutrient and commercial value. This stems from their ability to provide a high level of astaxanthin [5], which is evident visually by their red body colour. Astaxanthin, a carotenoid pigment, is a natural antioxidant and is said to provide many human health benefits. However, fish, shrimps, and other crustaceans cannot produce astaxanthin themselves. Therefore, resources for carotenoids in general and for astaxanthin in particular attract great interest in the studies of environment and feed supplements. Natural algae in the shrimp's native habitat may be a source of astaxanthin. Regarding feed supplements, there are many reports on the screening non-pigmented effect of probiotics on shrimp aquaculture in both laboratory and field trials, mostly in *Bacillus* sp., such as *B. subtilis*, *B. licheniformis*, *B. coagulans* [6], and *Lactobacillus* sp., such as *L. acidophilus*, *L. plantarum* [4, 7]... However, the research on effects of pigmented *Bacillus* strains producing carotenoids for shrimps remains limited. Recently, we have published a paper reporting beneficial effects of *B. aquimaris* SH6

*Corresponding author: Email: vananhbiolab@gmail.com

spores on whiteleg shrimps at laboratory scale, including red colour, astaxanthin level, and growth rate [8]. However, we have not yet performed field trials to confirm these beneficial effects, and we have not yet determined whether SH6 spores can be combined with any other *Bacillus* probiotic strains for use as feed supplements for shrimps. Thus, in this study, we performed a trial in ponds of whiteleg shrimps at Quang Ninh province for feeds supplemented with either *B. aquimaris* SH6 spores or a mixture of the SH6 spores with non-pigmented *B. subtilis* SH23, in comparison to the conventional feed used. In addition, we also analysed the persistence of SH6 spores and their effect on microbiota in shrimp guts to explain a possible mechanism behind the effects.

Materials and methods

Preparation of spores

Orange-pigmented *B. aquimaris* SH6 (Accession No. KF443807) [8] and non-pigmented *B. subtilis* SH23 (Accession No. KP735610; a strain of shrimp intestinal

origin belonging to the GreenBio S1 product of ANABIO R&D JSC) are strains isolated from the whiteleg shrimp gut (*L. vannamei*). These strains have been isolated and identified based on the previous research project KLEPT.12.03 funded by the Ministry of Science and Technology for Key Laboratory of Enzyme and Protein technology. The two strains were used to produce spores in Difco Sporulation medium (DSM, Oxoid, England), following the methods described by Nicholson, et al. (1990) for the SH23 strain [9] and Ngo, et al. (2016) for the SH6 strain [8]. The spore shapes of SH6 and SH23 in the cultures were observed under a conventional microscope and shown in Fig. 1A. The sprayed powder contained condensed spores at concentrations of about 2×10^{10} CFU g⁻¹ pellet for SH6 and 1×10^{10} CFU g⁻¹ pellet for each strain (ratio 1:1) in the mixture (Fig. 1B).

Preparation of the feed pellets

Commercial feed pellets (Grobest Industrial Vietnam) were assigned as the 'control' feed. The commercial feed was further coated with either condensed SH6 spores

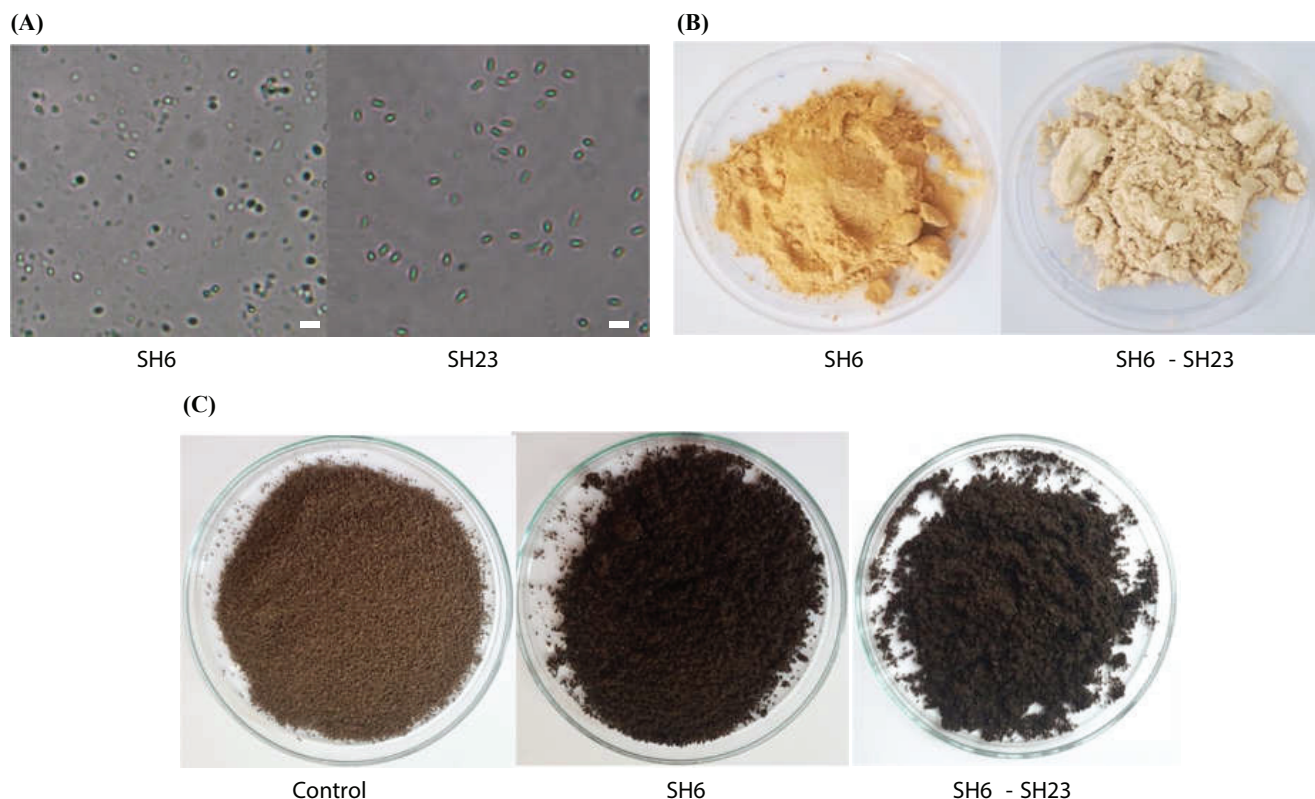


Fig. 1. Spore shapes and spore powders of *B. aquimaris* SH6 and *B. subtilis* SH23; non-coated and spore-coated feed pellets. Panel (A) Images of spore shapes of SH6 and SH23 under microscope. Scale bars are 2 μ m. Panel (B) The condensed spores of SH6 at 2×10^{10} CFU g⁻¹ pellet and mixture of SH6 and SH23 spores at ratio 1:1 of 1×10^{10} CFU g⁻¹ pellet for each strain. Panel (C) Control non-coated feed and feeds coated with SH6 spores at final concentration of 1×10^6 CFU g⁻¹ pellet or mixture of SH6 and SH23 spores 5×10^5 CFU g⁻¹ pellet.

to have the final concentration of 1×10^6 CFU g^{-1} pellet or a mixture of condensed SH6 and SH23 spores at ratio 1:1 to have the final concentration of each strain of 5×10^5 CFU g^{-1} pellet. These control and two supplemented feeds, as presented in Fig. 1C, were then coded randomly for blind field trial in ponds culturing whiteleg shrimps.

Experimental design for ponds, shrimps and feed supplements

Three ponds ($n=3$) with surface areas of about 0.8 ha were selected at Hai Dong village, Mong Cai city, Quang Ninh province, Vietnam. Twenty-five-day-old whiteleg shrimps (*L. vannamei*) which weighed about 2.05 ± 0.05 grams, were divided into three ponds and fed with different feed types: control group (control feed), SH6 group (*B. aquimaris* SH6 spores 1×10^6 CFU g^{-1} pellet), and SH6-SH23 group (each *B. aquimaris* SH6 and *B. subtilis* SH23 spores at 5×10^5 CFU g^{-1} pellet). Water in the ponds was changed every week and monitored regularly. Shrimps were maintained under the following conditions: temperature 26-28°C, pH = 7.5-8.5, DO ≥ 4 mg l^{-1} , 12 ppt salinity. These ponds were cultured at a density of (80 shrimps/ m^2). Shrimps were fed three times per day continuously for 28 days. After 28 days of the experiment, live counts of SH6 spores and major bacteria in shrimp gut, astaxanthin concentration in shrimp muscle, and growth rate of shrimps were evaluated.

Counts of SH6 and total bacteria in shrimp gut

On the 28th day, three shrimps in each group were sacrificed and their guts were collected for live counts of SH6 and total bacteria on MRS agar (for *Lactobacili* bacteria) and LB agar (for other bacteria). Each shrimp gut was prepared in 0.9% NaCl and then vigorously suspended by vortex until a homogenous suspension was obtained. Serial dilutions with 0.9% NaCl were made before plating on LB or MRS agar for live counts of total bacteria. Colony-forming units (CFU) were determined after a 24-hour incubation period for LB agar and after 48 hours for MRS agar plates. The single colony was collected based on their morphology and their species was identified using the 16S rRNA sequencing method. In brief, the extracted genomic DNA from each colony was used as template for PCR and partial 16S rRNA sequencing of a PCR-amplified 1,500 bp fragment using primers 27F: 5'-AGAGTTTGATCMTGGCTCAG-3' and 1527R: 5'-AAAGGAGGTGATCCAGCC-3', following the method described by Ngo, et al. (2016) [8]. Three colonies with the same morphology and colour were used to repeat the species identification three times. Then, the number of colonies with the same morphology and colour grown on the plates were taken from each group we counted.

Measurement of body colours and astaxanthin levels in shrimps

For comparing body colour, eight shrimps ($n=8$) in each

group were taken after 28 days, boiled to observe red colour level using the Roche index, SalmoFan™ standard colour [8]. For measuring astaxanthin concentration, eight shrimps ($n=8$) in each group were also taken at day 28. The astaxanthin level was extracted and identified spectrophotometrically at A_{480} , following the method described by Ngo, et al. and Khaneja, et al. [8, 10].

Weight gain measurement for shrimps

For determination of weight gain in shrimps, 100 shrimps at day 28 were taken from five positions of shrimp pond for each group in order to measure the average weight and compare it to that of day 0 following the equation: GR (growth rate) = $\frac{\text{final weight} - \text{initial weight}}{\text{days of feeding}} \times 100\%$. The survival

rate (%) of shrimps was calculated by the following formula: $SR = 100 \times (n_t/n_0)$, where SR is the survival rate, n_t is the number of shrimps at the time t , and n_0 is the number of shrimps at the initial.

Statistical analysis

Data of astaxanthin concentration, live counts, and growth rate among three treatment groups were compared using a Student's t-test, with the accepted significance level at $p < 0.01, 0.05, 0.001$. Statistical analyses were performed using the Analysis ToolPak in MICR Microsoft Excel Software (Microsoft, Redmond, WA). Experimental data were expressed as mean $\pm$ SD.

Results and discussion

Persistence of *B. aquimaris* SH6 spores in the shrimp gut and the cooperative effect of SH6 and SH23 to live counts and diversity of shrimp gut microbiota

For 28 days, three groups of shrimps were fed with pellets coated with or without either SH6 spores (1×10^6 CFU g^{-1} pellet) and a mixture of SH6 and SH23 spores (5×10^5 CFU g^{-1} pellet for each strain). The counts of *B. aquimaris* SH6 and other bacterial strains recovered in the shrimp intestines were determined by plating samples of excised GI-tracts onto selective agar (see Methods). As shown in Fig. 2A, our data revealed that SH6 colonies were present in the SH6 and SH6-SH23 groups; meanwhile, the SH6 colonies were not detectable in the control group. Surprisingly, the number of SH6 colonies in the SH6-SH23 groups (7.7×10^3 CFU gut $^{-1}$) was even higher than that in the SH6 group (5.7×10^3 CFU gut $^{-1}$) although the input dose of SH6 spores in the SH6-SH23 group was just half of that in the SH6 group. Our data in the field trial once again confirm that SH6 spores can colonize well into the intestinal epithelium of whiteleg shrimps and provide new information that feeding SH6 together with SH23 allows more effective residence of SH6 spores in the shrimp gut.

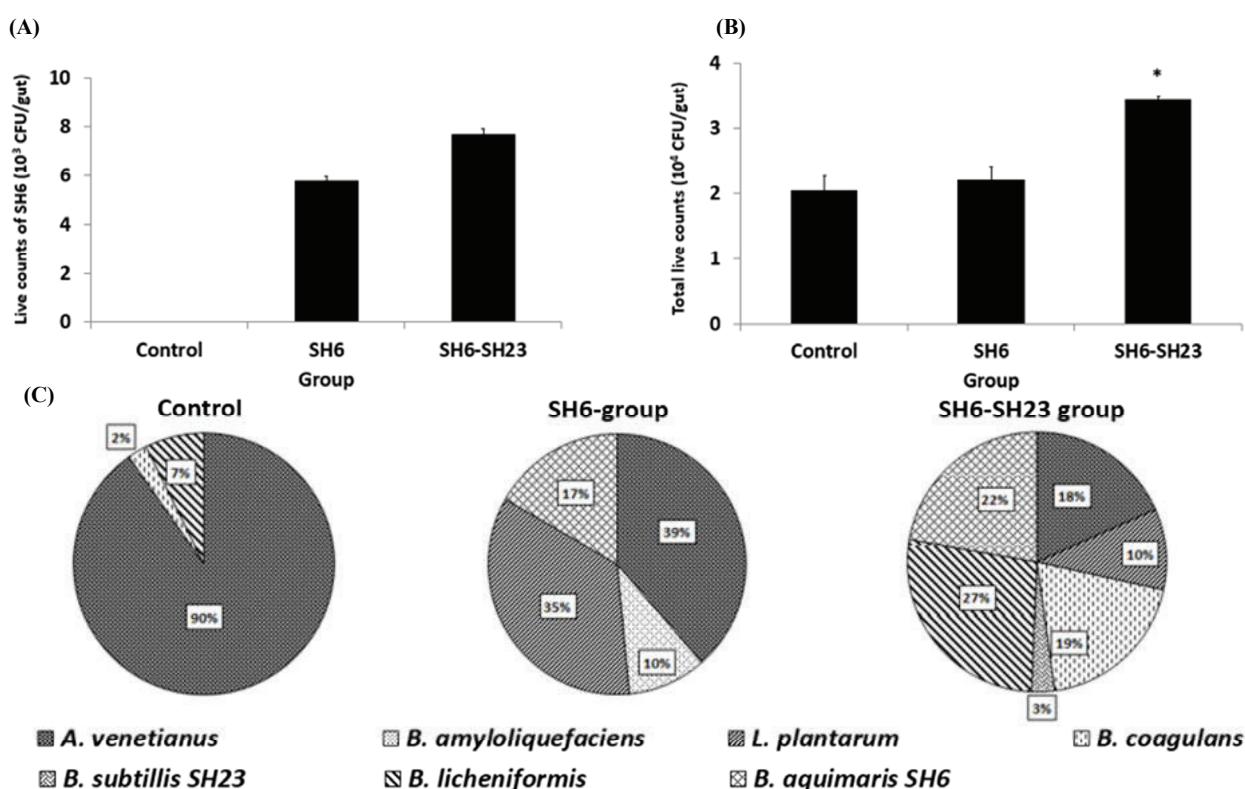


Fig. 2. Live counts of *B. aquimaris* SH6 and total bacteria in shrimp gut during the 28-day feeding. Experiment groups include SH6 spores (1×10^6 CFU g^{-1} pellet), SH6-SH23 spores (5×10^5 CFU g^{-1} pellet for each) and control (without supplements). Panel (A) Live counts of SH6 in shrimp gut. Panel (B) Live counts of total bacteria in shrimp gut. P values were generated by ANOVA using the Student's t-test for multiple comparisons to the control (* $p < 0.05$). Panel (C) Pie chart of useful bacterial population in shrimp gut.

Regarding the bacterial microbiota of shrimp gut among three groups, we found significant differences in both live counts and diversity of the morphology of colonies. In detail, the live counts in the control, SH6, and SH6-SH23 groups were 2.0×10^4 , 2.2×10^4 , and 3.5×10^4 CFU gut^{-1} in 28th day, respectively (Fig. 2B). However, no significant difference was noted between the SH6 group and the control group, whereas the 75% increase in live counts in the SH6-SH23 group compared to that in the control group ($p < 0.05$). Our data also revealed that the SH6 and SH6-SH23 groups exhibited more diversified useful bacterial population than the control group. As shown in Fig. 2C, there were six major useful species, including the *B. aquimaris* SH6 (22% population), and five other strains which closely related (>98% identity of 16S rRNA sequence) to the following species *B. licheniformis* (27% population), *B. coagulans* (19% population), *A. venetianus* (18% population), *L. plantarum* (10% population), and *B. subtilis* (3% population) in the shrimp gut of the SH6-SH23 group. These species have been either commonly used as probiotic bacteria or are naturally found in the intestines of many other aquatic species [11, 12]. For example, *B. subtilis*, *B. coagulans*, and *B. licheniformis* have been reported to play roles in

increasing the survival and net production of channel catfish and shrimps [13], increasing the survival of shrimps [14, 15], as well as decreasing luminous *Vibrio* densities [16]. *L. plantarum* is effective on the growth rate and microflora in the shrimp gut [17] while *A. venetianus* is naturally found useful in shrimp GI-tracts [11, 12]. In the group fed with SH6 spores only, we found only four major bacterial species, including *A. venetianus*, *L. plantarum*, *B. aquimaris* SH6, and *B. subtilis*, with respective population of 39, 35, 17, and 10% in the shrimp gut. When shrimps were fed with the control feed pellets, only three major species were found, including the naturally found species *A. venetianus* with 90% population, *B. licheniformis* with 7% population, and *B. coagulans* with 3% population. In conclusion, our data indicate that SH6 and SH23 strains cooperatively improve the live counts and diversity of bacterial microbiota in the whiteleg shrimp gut.

Colour and astaxanthin concentration in *L. vannamei* after 28 days of feeding

After the 28-day feeding, the body colour of shrimps in the different groups was measured, and the image of boiled shrimps is shown in Fig. 3A. Shrimps in the SH6-

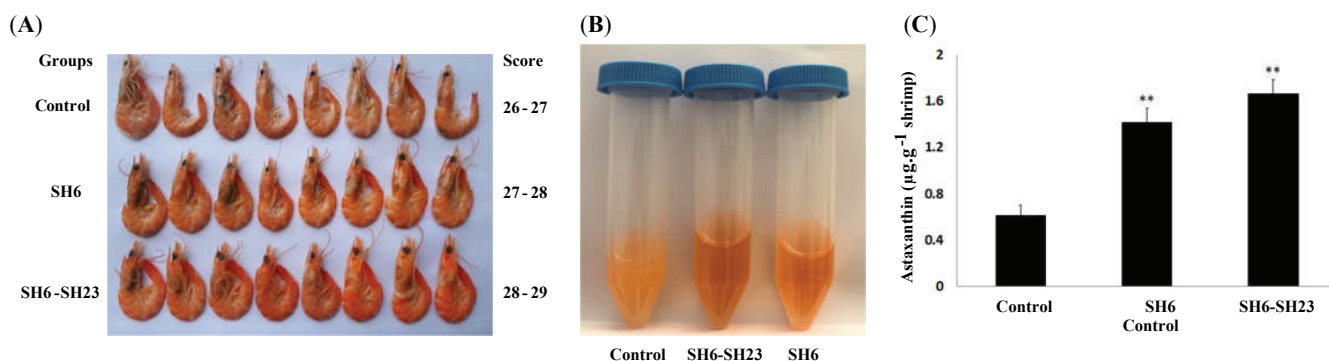


Fig. 3. Colour and astaxanthin level of shrimp after the 28-day feeding with and without SH6 spores. Experiment groups include SH6 (1×10^6 CFU g^{-1} pellet), SH6-SH23 (5×10^5 CFU g^{-1} pellet for each strain), and control (without supplements). Panel (A) Image of boiled shrimps and their colour variable scores indicating the levels of red pigmentation ($n=8$). Panel (B) Astaxanthin concentrations in extracts from shrimp muscle. P value was generated using Student's t-test for comparisons with the control (** $p < 0.001$). Panel (C) Image of astaxanthin extraction from shrimp muscle. Data are presented as arithmetic means and error bars are standard deviations.

SH23 group had the highest red scores of 28-29, followed by the SH6 group (red scores of 27-28), and the control group (red scores of 26-27). It has been reported that carotenoid-supplemented feed enhances the red or orange colour of shrimp when they are cooked [10, 18]. Moreover, carotenoids may provide enhanced immune function and disease resistance in intracellular oxygen supply for shrimp, allowing survival under the hypoxic conditions in the pond bottom [19]. Furthermore, carotenoid-producing bacteria have potential as a feed supplement for shrimp because of their nature-friendliness. In order to correlate the relationship between the shrimp colour and the presence of astaxanthin in shrimp tissue, we determined the astaxanthin levels in shrimp of each group and present the data in Fig. 3B and 3C. It is clear that astaxanthin level in the SH6-SH23 and SH6 groups were 2.72-fold and 2.3-fold higher, respectively, than in the control group ($p < 0.001$). Fig. 3C also confirms the significant different level of redness in extracts from shrimp muscle between the two experimental groups and the control group.

Our data suggest that SH6 combined with SH23 might have a stronger effect in producing a higher astaxanthin level in shrimp muscle, subsequently inducing a redder colour of shrimps. These data are also consistent with better colonization of SH6 spores and improvement of microbiota in the shrimp gut found in the SH6-SH23 group, in comparison to those found in the SH6 group. Taken together, our data in the field trial once again confirm that carotenoids produced by *B. aquimaris* SH6 were successfully converted to astaxanthin by whiteleg shrimps. These data also provide new information that the SH23 strain may play an effective role in improving the SH6 counts in the gut, thereby enhancing carotenoid production of SH6 and resulting in a higher astaxanthin level and redder colour in shrimps.

Growth performance

We then chose to further analyse the effects of the *B. aquimaris* SH6 and *B. subtilis* SH23 supplement on whiteleg shrimp growth performance. As shown in Fig. 4, the growth rate (%) of all three groups demonstrated positive values: $8.0 \pm 0.6\%$, $7.7 \pm 1.8\%$, and $7.8 \pm 1.1\%$ in the control, SH6, and SH6-SH23 groups, respectively. After the 28 days of experimental feeding, there was no statistically significant difference in growth rate between shrimps in the SH6 and SH6-SH23 groups in comparison to the control group. These data from the field trial are different from what we have observed at laboratory scale, where SH6 has demonstrated significant effect on the growth performance of shrimps [8]. We do not have a reasonable explanation for this field trial data but suppose that the size of shrimp samples we collected

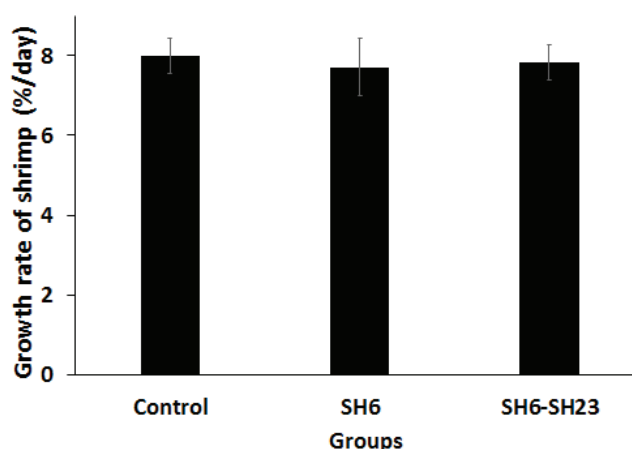


Fig. 4. The growth rate of shrimp after 4-week feeding with SH6 spores. Experiment groups include SH6 (fed with spores at 1×10^6 CFU g^{-1} pellet), SH6-SH23 (fed with SH6 spores at 5×10^5 CFU g^{-1} pellet and fed with SH23 spores at 5×10^5 CFU g^{-1} pellet), and control (without supplements). Data are presented as arithmetic means and error bars are standard deviations.

might be not representative for all shrimps in ponds. In the future, we need to collect more shrimps at various positions in the shrimp pond for weighing, perform field studies with a larger number of ponds, and include the control group of only *B. subtilis* SH23 in order to have better statistical data concerning growth rate.

Conclusions

This field study demonstrated that when *B. aquimaris* SH6 spores were administered as a single-strain feed supplement at a concentration of 1×10^6 CFU g⁻¹ pellet or as a double-strain probiotic together with *B. subtilis* SH23 at a concentration of 5×10^5 CFU g⁻¹ pellet, the following benefits were conferred to the whiteleg shrimps: 10-75% increased live counts and more diversity of bacterial population in the intestine, as well as significantly increased astaxanthin levels (2.3-2.7-fold) in the muscle of shrimps and higher red colour scores (27-29 vs. 26-27). Although more field studies should be performed in the future, the SH6 and SH23 spores exhibit to be potential co-effective strains for further development of spore-based mixture probiotics to improve astaxanthin level and red colour in whiteleg shrimps.

ACKNOWLEDGEMENTS

This study was funded by a grant from Third World Academy of Science (TWAS) to A.T.V. Nguyen under project number 16-549 RG/BIO/AS_G - FR3240293311, and a grant from Vietnam National University_Hanoi, Key Laboratory of Enzyme and Protein Technology (KLEPT) to D.T. Nguyen under project number KLEPT.18.01. The authors would like to express sincere thanks to TWAS and KLEPT for funding the study.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] A.C. Sánchez-Ortiz, et al. (2015), "Isolation and characterization of potential probiotic bacteria from pustulose ark (*Anadara tuberculosa*) suitable for shrimp farming", *Lat. Am. J. Aquat. Res.*, **43**, pp.123-136.
- [2] C.N. Zheng and W. Wang (2017), "Effects of *Lactobacillus pentosus* on the growth performance, digestive enzyme and disease resistance of white shrimp, *Litopenaeus vannamei*", *Aquaculture Res.*, **48**, pp.1-11.
- [3] F. Vargas-Albores, M.A. Porchas-Cornejo, F. Vargas-Albores, M. Martínez-Porchas, E. Villalpando-Canchola, T. Gollas-Galván, L.R. Martínez-Córdova (2017), "Bacterial biota of shrimp intestine is significantly modified by the use of a probiotic mixture: a high throughput sequencing approach", *Helgol. Mar. Res.*, **71**(1), pp.1-10.
- [4] F.J. Gatesoupe (1991), "The effect of three strains of lactic bacteria on the production rate of rotifers, *Brachionus plicatilis*, and their dietary value for larval turbot, *Scophthalmus maximus*", *Aquaculture*, **96**, pp.335-342.
- [5] H. Zokaieifar, N. Babaei, C.R. Saad, M.S. Kamarudin, K. Sijam, J.L. Balcazar (2014), "Administration of *Bacillus subtilis* strains in the rearing water enhances the water quality, growth performance, immune response, and resistance against *Vibrio harveyi* infection in juvenile white shrimp, *Litopenaeus vannamei*", *Fish Shellfish Immunol.*, **36**(1), pp.68-74.
- [6] H.A. Hong, J.M. Huang, R. Khaneja, L.V. Hiep, M.C. Urdaci, S.M. Cutting (2008), "The safety of *Bacillus subtilis* and *Bacillus indicus* as food probiotics", *J. Appl. Microbiol.*, **105**, pp.510-520.
- [7] H. Jyonouchi, L. Zang, M. Gross, Y. Tomita (1994), "Immunomodulating actions of carotenoids: enhancement of in vivo and in vitro antibody production to T-dependent antigens", *Nutr. Cancer*, **21**(1), pp.47-58.
- [8] H.T. Ngo, T.T.N. Nguyen, Q.M. Nguyen, A.V. Tran, H.T.V. Do, H.A. Nguyen, et al. (2016), "Screening of pigmented *Bacillus aquimaris* SH6 from the intestinal tracts of shrimp to develop a novel feed supplement for shrimp", *J. Appl. Microbiol.*, **121**(5), pp.1357-1372.
- [9] W.L. Nicholson and P. Setlow (1990), "Sporulation, germination, and outgrowth", *Molecular Biological Methods for Bacillus*, pp.391-450, John Wiley and Sons.
- [10] R. Khaneja, L. Perez-Fons, S. Fakhry, L. Baccigalupi, S. Steiger, E. To, et al. (2010), "Carotenoids found in *Bacillus*", *J. Appl. Microbiol.*, **108**, pp.1889-1902.
- [11] K. Supamattaya, S. Kirirantnikom, M. Boonyaratpalin, L. Borowitzka (2005), "Effect of a *Dunaliella* extract on growth performance, health condition, immune response and disease resistance in black tiger shrimp (*Penaeus monodon*)", *Aquaculture*, **248**(1-4), pp.207-216.
- [12] L. Verschuere, G. Rombaut, P. Sorgeloos, W. Verstraete (2000), "Probiotic Bacteria as Biological Control Agents in Aquaculture", *Microbiol. Mol. Biol. Rev.*, **64**(4), pp.655-671.
- [13] M.D. Ibrahim (2015), "Evolution of probiotics in aquatic world: Potential effects, the current status in Egypt and recent prospectives", *J. Adv. Res.*, **6**(6), pp.765-791.
- [14] M. Li, B. Wang, M. Zhang, M. Rantalainen, S. Wang, H. Zhou, et al. (2008), "Symbiotic gut microbes modulate human metabolic phenotypes", *Proc. Natl. Acad. Sci. USA*, **105**(6), pp.2117-2122.
- [15] M.S. Kim, E. Min, J.H. Kim, J.K. Koo, J.C. Kang (2015), "Growth performance and immunological and antioxidant status of Chinese shrimp, *Fenneropenaeus chinensis* reared in bio-floc culture system using probiotics", *Fish and Shellfish Immunol.*, **47**(1), pp.141-146.
- [16] T. Mohammadian, M. Alisahi, M.R. Tabandeh, M. Ghorbanpoor, D. Gharibi (2017), "Effect of *Lactobacillus plantarum* and *Lactobacillus delbrueckii* subsp. *bulgaricus* on growth performance, gut microbial flora and digestive enzymes activities in *Tor grypus*", *Iran. J. Fish. Sci.*, **16**(1), pp.296-317.
- [17] D.J.M. Moriarty (1999), "Control of luminous *Vibrio* species in panacoid aquaculture ponds", *Aquaculture*, **164**, pp.351-358.
- [18] S. Chaiyapechara, W. Rungrasamee, I. Suriyachay, Y. Kuncharin, A. Klanchui, N. Karoonuthaisiri, et al. (2012), "Bacterial Community Associated with the Intestinal Tract of *P. monodon* in Commercial Farms", *Microb. Ecol.*, **63**(4), pp.938-953.
- [19] S. Ziaei-Nejad, M.H. Rezaei, G.A. Takami, D.L. Lovett, A.R. Mirvaghefi, M. Shakouri (2006), "The effect of *Bacillus* spp. bacteria used as probiotics on digestive enzyme activity, survival and growth in the Indian white shrimp *Fenneropenaeus indicus*", *Aquaculture*, **252**(2-4), pp.516-524.

Studies of common antibiotic resistance-associated genes of *Acinetobacter baumannii*

Le Phuong Trinh^{1*}, Van Son Ta^{2*}, Thi Ngoc Han Luu², Mai Phuong Nguyen¹,
Thi Huong Dinh³, Van Hung Nguyen³, Thi Ha Doan³, Bao Yen Pham¹

¹Key Laboratory of Enzyme and Protein Technology, University of Science, Vietnam National University, Hanoi

²Faculty of Biology, University of Science, Vietnam National University, Hanoi

³National Lung Hospital

Received 11 October 2018; accepted 10 December 2018

Abstract:

This study, the presence of 4 common beta-lactamase encoding genes including OXA-23, OXA-51, VIM and IMP was determined simultaneously by multiplex PCR, either directly with sputum samples or with purified DNA, giving results within 3-4 hours instead of 2 days compared to the culture method. Within 58 samples positive with *Acinetobacter baumannii* collected from the hospital, OXA-51 fragment was dominant (81%), whereas, VIM fragment was only identified in one sample (DNA and sputum samples) (1.7%). The antibiograms indicated for 19 samples were used to find correlation between genotype and resistance phenotype.

Keywords: *Acinetobacter baumannii*, antibiotic resistance-associated genes, beta-lactamase, multiplex PCR, susceptibility.

Classification number: 3.5

Introduction

Pneumonia is a form of acute respiratory infection that affects the lungs. When an individual has pneumonia, the alveoli are filled with pus and fluid, which make breathing painful and limits oxygen intake. Pneumonia is the single largest infectious cause of death in children worldwide. Pneumonia killed 920,136 children under the age of five around the world in 2015, accounting for 16% of all deaths of children under five years old [1]. Pneumonia is caused by a number of infectious agents, including viruses, bacteria, and fungi. Antibiotics can be effective for many of the bacteria that cause pneumonia. Unfortunately, antibiotic resistance is growing amongst the bacteria that cause pneumonia, often arising from the overuse and misuse of antibiotics [2]. The most dangerous antibiotic-resistant bacteria that causes pneumonia is *Acinetobacter baumannii*.

A. baumannii is a pleomorphic aerobic gram-negative bacillus belonging to the *Moraxellaceae* family of the *Gammaproteobacteria* class. It can survive on surfaces for a long time. *A. baumannii* has several resistance mechanisms to many classes of antimicrobials, including beta-lactamase, multidrug efflux pumps, aminoglycoside-modifying enzymes, permeability defects, and alteration of target sites. The beta-lactam-resistant mechanism plays the most important role in multidrug resistance in *A. baumannii*. It involves the hydrolysis of beta-lactam-based drugs by beta-lactamases. These enzymes are divided into four classes: A, B, C, D, and their antibiotic-resistant abilities are not equivalent. Classes B and D are more concerning because beta-lactamases in class B are mostly strong penicillinase, and, for genes in class D, some OXAs can hydrolyse wide-range cephalosporins and carbapenem [3]. Studying genes of this class is essential because this can help doctors reduce the time of diagnosis and treatment for patients suffering from *A. baumannii* pneumonia.

*Corresponding author: Email: violetph2009@gmail.com

Recently, there have been several antimicrobial-susceptibility testing methods available for *A. baumannii*, such as cultured-based methods (disk diffusion, E-test [4]), genotypic methods (polymerases chain reaction, DNA hybridisation), and single-cell imaging method [5]. These all have the same goal: to provide a reliable prediction of the sensitivity of *A. baumannii* to respond therapeutically to a particular antibiotic treatment. Cultured-based methods are the gold standard and are routinely used in hospitals; however, they require considerable time for analysis. Single-cell imaging is a novel form of testing prototypes, with very short time for analysis; however, it has been applied only in research, not in practice. Genotyping methods can be an effective alternative to these two methods, with shorter test result turn-around time.

In Vietnam, most studies have focused only on the level of resistance of *A. baumannii* to antimicrobials often used in hospital. In the study of V.Q. Nguyen (2013) on 32 strains of *A. baumannii* isolated from 98 patients suffering from pneumonia at Hue Central Hospital, the results showed that this bacteria resistant to all of the 17 antibiotics tested [6]. T.H.P. Ngo, et al. (2013) generated a similar result: 100% of the *A. baumannii* specimens were resistant to most of the common antibiotics, except Colistin [7]. H.H. Tran, et al. (2017) investigated the presence of beta-lactamases in classes B and D, including *NDM-1*, *OXA-23*, *OXA-51*, *IMP*, and *VIM*, in *A. baumannii* from patients at three major hospitals in Hanoi [8]. However, this study does not mention the correlation between gene sequences and antibiogram to support diagnosis. These studies of the antibiotic resistance of *A. baumannii* are only in the initial phase and are not systematic. Therefore, our study focuses on four beta-lactamase-encoding genes, *OXA-23* and *OXA-51* in class D and *VIM* and *IMP* in class B, that may affect the susceptibility of *A. baumannii* to antibiotics.

Materials and methods

Materials

Specimens: a total 58 samples, including purified DNA (52/58) and sputum (6/58), identified positive for *A. baumannii* were collected by the National TB Reference Lab at the National Lung Hospital. DNA samples were extracted from cultured *A. baumannii*. Sputum was treated with NaLC-NaOH and then heated at 95°C for 10 minutes before being stored at -20°C.

Primers: the list of primers used for the detection of the beta-lactamase-encoding genes of *A. baumannii* are listed in Table 1.

Table 1. Primer list for *OXA-23*, *OXA-51*, *IMP*, and *VIM* fragments.

Beta-lactamase gene	Product size (bp)	Sequence (5'-3')
OXA-23	501	F- GATCGGATTGGAGAACCAGA R- ATTTCTGACCGCATTTCCAT [9]
OXA-51	353	F- TAATGCTTTGATCGGCCTTG R- TGGATTGCACTTCATCTTGG [9]
VIM	390	F- GATGGTGTGGTTCGCATA R- CGAATGCGCAGCACCAG [10]
IMP	188	F- GGAATAGAGTGGCTTAAYTCTC R- CCAAACYACTASGTTATCT [10]

Methods

Multiplex PCR for identifying the four target genes simultaneously: identification of the four beta-lactamase-encoding genes' co-presence was achieved by means of multiplex PCR. The multiplex PCR was designed to include 1U Dream Taq polymerase (Thermo Scientific), 1X Green Taq Buffer (Thermo Scientific), dNTPs, four primer pairs at concentrations of 0.25-0.5 μ M, DNA and sputum samples, and sterile deionised water. The reaction was conducted in a volume of 12.5 μ l and occurred according to the following thermal cycles: 94°C, 5 minutes; 30 cycles of 94°C, 30 seconds, 50°C, 45 seconds, 72°C, 50 seconds, and finally, 72°C, 5 minutes. Four pairs of primers used in the amplified reaction mixture: the 501 bp fragment corresponding to the *blaOXA-23* gene, the 353 bp gene coded for the *blaOXA-51* gene, the 390 bp gene encoding the *blaVIM* gene, and the 188 bp fragment within the *blaIMP* gene.

Analysis of PCR products by means of agarose electrophoresis: in this study, 5 μ l of multiplex PCR product was analysed using 3% agarose gel electrophoresis (3 g/100 ml) (Clever Scientific, England) with 1X RedSafe DNA staining solution (Intron Biotechnology, Korea), which replaced the traditional Ethidium bromide stain. Following gel electrophoresis, the gel was observed under UV light (GelDoc, BIO-RAD).

Results and discussion

Optimisation of multiplex PCR for detecting the four beta-lactamase coding genes, *OXA-51*, *OXA-23*, *VIM*, and *IMP*: primer specificity was initially confirmed through the presence of single bright bands at calculated molecular weights for each target gene in a single PCR. To establish the multiplex PCR, which simultaneously amplified the four fragments in one amplification reaction, the annealing temperature (50°C, 52°C) (Fig. 1A), primer concentration (Fig. 1B), and DMSO concentrations (0-10%, 2% interval) were examined. As a result, the optimal condition was conducted at 50°C annealing temperature, 0.2 μ M of *OXA-23* and *OXA-51* primer, and 0.3 μ M for *VIM*, 0.4 μ M for *IMP*,

and DMSO concentration of 8%. The results showed four specific bands corresponding to the four target genes of *A. baumannii* and no extra band. Moreover, when performing PCR with DNA and sputum samples, the quality of PCR products was comparable; therefore, sputum could be used directly for the multiplex PCR, reducing testing time.

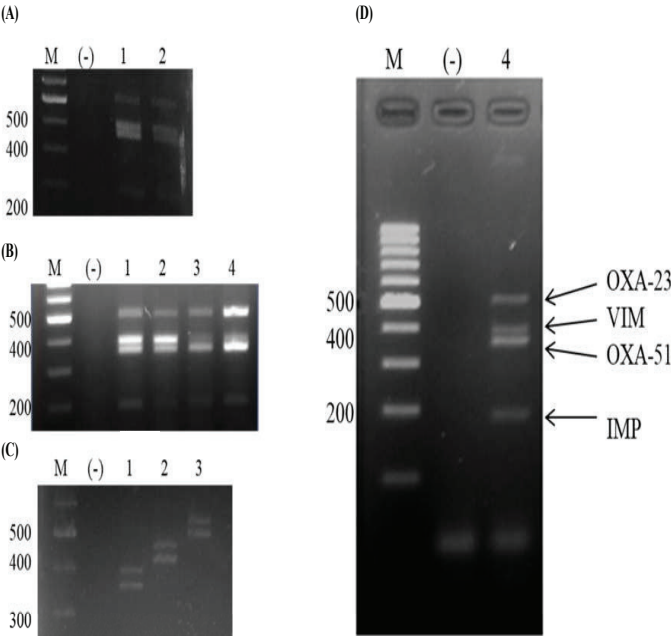


Fig. 1. Optimisation of multiplex PCR for the four targets. (A) Annealing temperature optimisation at 500C (1) and 520C (2) was examined. (B) Primer optimisation with different concentrations of *OXA-23*, *OXA-51*, *VIM*, and *IMP*, 0.2; 0.2; 0.3; 0.4 μ M (1); 0.2; 0.1; 0.1; 0.3 μ M (2); 0.3; 0.1; 0.1; 0.3 μ M (3); 0.3; 0.3; 0.2; 0.4 μ M (4). (C) Electrophoresis running for 50 minutes (1), 45 minutes (2), 40 minutes (3). (D) Final multiplex PCR. (M): marker; (-): negative.

In order to resolve the 353-bp and 390-bp bands of *OXA-51* and *VIM*, respectively, three different running times (40-50 minutes, 5-minute interval) were tested and the final electrophoresis condition was at 100 volts for 45 minutes.

Prevalence of the four targeted resistance-associated genes

Screening of the 58 samples (52 DNA and 6 sputum samples) using the established multiplex PCR showed that the presence of *OXA-51* was the highest (47/58, 81%), following by *OXA-23*: (63.8%, 37/58), and *IMP* (8.6%, 5/58); *VIM* (1.7%), the lowest, was found in only one sample (Fig. 2). Such prevalence is in line with the findings of other studies [9, 10]. However, compared to others, the sample size in our study was relatively small; therefore, the study should be expanded to obtain more statistically significant data.

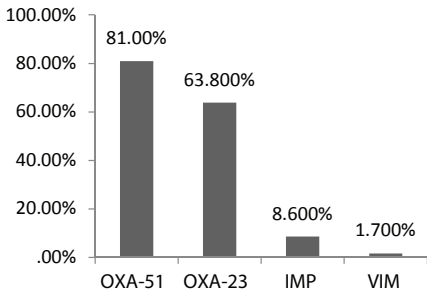


Fig. 2. Percentages of the four genes, *OXA-51*, *OXA-23*, *IMP*, and *VIM* in the 58 samples.

The consistently lower percentages of *VIM* and *IMP* are explained by the ineffective amplification of these fragments. Our established multiplex PCR also showed the unequal intensity of the four bands in one sample carrying all four genes, in which the *VIM* band was the faintest (Fig. 1D). In future studies, RT-PCR may be a method for increasing the sensitivity of detection and also allowing investigation of the copy number of the gene.

Correlation between genomic profiling and antibiotic susceptibility

The results of the antibiotic sensitivity test of eight β -lactam antibiotics, as indicated in 19/58 samples, revealed that three antibiotics, CRO, FEP (a cephalosporin), and

Table 2. Genomic profiling and antibiotic susceptibility.

No	PCR		Culture-based method							
	<i>OXA-23</i>	<i>OXA-51</i>	<i>IMP</i>	<i>DOR</i>	<i>MEM</i>	<i>CAZ</i>	<i>CRO</i>	<i>FEP</i>	<i>TCC</i>	<i>TZP</i>
30	-	+								
31	-	-								
36	+	+								
37	+	+								
38	-	+								
39	+	+								
41	-	+								
45	+	+								
48	+	+								
49	+	+								
50	+	+								
51	+	+								
52	+	+								
53	+	+								
54	+	+								
55	+	+								
56	+	+								
57	+	+								
58	+	+								
Σ	15	18	15	16	16	17	17	17	17	15

: No information
 : Sensitive
 : Intermediate
 : Resistant

IMP: imipenem; *DOR*: doripenem; *MEM*: meropenem; *CAZ*: ceftazidime; *CRO*: ceftriaxone; *FEP*: cefepime; *TCC*: tic + A. clavulanic; *TZP*: pipe + tazobactam.

TCC were the most resisted (17/19), while IMP and TZP were less resisted (15/19); however, the difference between the two groups was insignificant (Table 2). An examination of the molecular characteristics of these 19 samples showed the presence of *OXA-51* in 18 samples and *OXA-23* in 15 samples, whereas no samples carried *VIM* and *IMP*.

The sample designated 31, lacking both *OXA-51* and *OXA-23*, was resistant to only TCC and intermediately to CRO, while being sensitive to all the other six antibiotics. The remaining 14 samples containing both genes were resistant to all seven or eight antibiotics tested. The results confirms the association of *OXA-51* and *OXA-23* with β -lactam resistance, which had been addressed in previous studies. The results also promote the importance of examining these resistant genes for supporting drug prescription and treatment in addition to the conventional antibiotic sensitivity test.

Conclusions

The established multiplex PCR was successful in simultaneously detecting four beta-lactamase-encoding genes: *OXA-23*, *OXA-51*, *VIM*, and *IMP*. In a total of 58 samples, the *OXA-51* gene predominated, with more than 80% carrying the target. *VIM* was identified in only a single sample, approximately 1.7%. The presence of *OXA-51* and *OXA-23* is likely linked to the β -lactam resistance of *A. baumannii*, a finding that could be used for prescribing effective antibacterial drugs.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] WHO (2016), <http://www.who.int/news-room/fact-sheets/detail/pneumonia>, received on 7 November 2016.
- [2] American Thoracic Society (2015), *Top 20 Pneumonia Facts*.
- [3] M.F. Lin, C.Y. Lan (2014), "Antimicrobial resistance in *Acinetobacter baumannii*: from bench to bedside", *World J. Clin. Cases*, **2**(12), pp.787-814.
- [4] Biomerieux (2012), <http://www.biomerieux-usa.com/clinical/etest>.
- [5] Ö. Baltekin, A. Boucharin, E. Tano, et al. (2017), "Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging", *Proc. Natl. Acad. Sci. USA*, **114**(34), pp.9170-9175.
- [6] V.Q. Nguyen (2013), "Percentage of drug-resistant bacteria in patients caused by mechanical pneumonia in the recovery room after surgery A in Hue Central Hospital", *Journal of Practical Medicine*, **899**, pp.36-39 (in Vietnamese).
- [7] T.H.P. Ngo, Q.H. Nguyen, H.N. Cao, et al. (2013), "Antibiotic resistance of *Acinetobacter baumannii* detected at Pasteur institute, Ho Chi Minh city", *Journal of Science*, Ho Chi Minh city University of Pedagogy, **47**, pp.112-118 (in Vietnamese).
- [8] H.H. Tran, D.N. Tran, M. Matsui, et al. (2017), "Emergence of New Delhi metallo-beta-lactamase 1 and other carbapenemase-producing *Acinetobacter calcoaceticus-baumannii* complex among patients in hospitals in Hanoi, Vietnam", *Eur. J. Clin. Microbiol. Infect. Dis.*, **36**(2), pp.219-225.
- [9] H. Vahaboglu, F. Budak, M. Kasap, et al. (2006), "High prevalence of OXA-51-type class D beta-lactamases among ceftazidime-resistant clinical isolates of *Acinetobacter* spp: co-existence with OXA-58 in multiple centres", *J. Antimicrob. Chemother.*, **58**(3), pp.537-542.
- [10] M.D. Aksoy, S. Çavuşlu, and H.M. Tuğrul (2015), "Investigation of metallo beta lactamases and oxacilinases in carbapenem resistant *Acinetobacter baumannii* strains isolated from inpatients", *Balkan Med. J.*, **32**(1), pp.79-83.

Optimal culture conditions for mycelial growth and fruiting body formation of Ling Zhi mushroom *Ganoderma lucidum* strain GA3

Bich Thuy Thi Nguyen¹, Nghien Xuan Ngo¹, Ve Van Le^{1*}, Luyen Thi Nguyen¹, Ry Kana¹, Huy Duc Nguyen²

¹Faculty of Biotechnology, Vietnam National University of Agriculture

²Faculty of Agronomy, Vietnam National University of Agriculture

Received 17 November 2018; accepted 12 February 2019

Abstract:

The objective in this study is to investigate optimal culture conditions for mycelial growth and fruiting body formation of the Ling Zhi mushroom, *Ganoderma lucidum* strain GA3. The results of the study show that the optimal media and temperature for the mycelial growth are potato, glucose, and agar (PGA) supplemented with rice bran, and 25-30°C, respectively. Strain GA3 is able to grow in a wide pH range, between 4 and 12. The most favourable substrate mixture for the formation and development of the fruiting body is 87% sawdust + 4% corn powder + 8% rice bran + 1% calcium carbonate (CaCO₃).

Keywords: fruiting bodies, Ling Zhi mushroom, media, mycelium.

Classification number: 3.5

Introduction

Ganoderma lucidum (Fr.) Karst (Polyporaceae), known as the Ling Zhi mushroom, belonging to the family Polyporaceae (or Ganodermaceae) of the order Aphyllophorales, has been recognised as one of the most highly valued medicinal mushrooms in East Asian countries for more than 2,000 years. As with other medical mushrooms, Ling Zhi is well-known for containing various chemical substances, with approximately 119 different triterpenes and several types of polysaccharides [1]. The basidiocarp, mycelia, and spores of *Ganoderma lucidum* (*G. lucidum*) are widely used in the treatment and prevention of many diseases, such as hepatitis, hypertension, hypercholesterolemia, and gastric cancer [2, 3].

Due to its bioactive components, irregular distribution in the wild, and the increasing demand for it, the Ling Zhi mushroom is artificially cultivated on various substrates for mycelial biomass and fruiting body production [4, 5]. Grain, sawdust, wood logs, and cork residues have been used as basal substrates for the artificial cultivation of *G. lucidum* [6-9]. A combination of beech sawdust supplemented with 2.5% malt extract and 10% wheat bran has been found to be the best substrate mixture for the cultivation of *G. lucidum* [10]. According to Jandaik, et al. (2013) [11], *G. lucidum* cultivated on paddy straw supplemented with wheat bran exhibited the maximum yield (82.5 g) and biological efficiency (27.5%). As previously reported by Boh, et al. (2007) [8] and Zhou, et al. (2012) [12], the biological efficiency of *G. lucidum* is strictly involved in the environmental factors such as temperature, humidity, oxygen, light, and carbon dioxide. In Vietnam, several studies have focused on the classification and distribution of the family *Ganodermataceae* [13]. Forty-three species belonging to the genus *Ganoderma* sourced from highland regions have been identified [14]. Of these, five species have

*Corresponding author: Email: vanvecnshk53@gmail.com.

been successfully cultivated: *G. lucidum*, *G. applanatum*, *G. australe*, *G. colossum*, and *G. subresinosum*. The search for *G. lucidum* strains that can possibly enhance the mushroom's disease resistance, yield, and medical value plays an essential role in its cultivation. However, at the time of writing, to our knowledge, only a few studies had selected *G. lucidum* strains that could potentially produce high yields for commercial cultivation and adapt to a broad range of climatic conditions in Vietnam. In the course of a previous investigation into strains from our mushroom resource bank with such potential, strain GA3 was found to be able to adapt better to the climatic conditions in Vietnam than were other strains. In order to achieve a high biological yield and reduce the time required to cultivate *G. lucidum*, identifying the optimal media, and chemical, physical, and biological factors is considered as among the most crucial strategies. To this end, the present study sets out to determine the optimal culture conditions for mycelial growth and fruiting body formation for strain GA3.

Materials and methods

Mushroom strain

The *G. lucidum* strain GA3 used in this study was collected in Japan. Pure mycelial cultures were isolated from internal tissue following the protocol described by Jonathan and Fasidi (2003) [15]. The culture was maintained on a PGA medium in complete darkness and stored in a refrigerator at 5-7°C for further study.

Effect of different media on mycelial growth

Four different kinds of culture media - Raper, PGA, PGA supplemented with rice bran extract, and PGA supplemented with fresh oyster mushroom extract- were used to ascertain the optimal media for promoting the vegetative growth of strain GA3. To prepare the PGA, PGA supplemented with rice bran extract, and PGA supplemented with fresh oyster mushroom extract media, after peeling, potatoes were cut into small pieces, and then boiled in 500 ml distilled water for 30 minutes. Twenty grams of rice bran (PGA supplemented with rice bran extract) and 25 g of fresh oyster mushroom (PGA supplemented with fresh oyster mushroom extract) were extracted using 250 ml of warm and hot water, respectively. The crude extract obtained was filtered by means of a steel mesh. Thereafter, these two liquids were mixed thoroughly. Twenty grams of glucose and 20 g of agar were dissolved and added to the medium. The final volume of the media was increased to one litre by adding water. The media were sterilised by autoclaving at it 121°C for 60 minutes. The composition of the culture media is shown in Table 1.

Table 1. Composition of various culture media for mycelial growth.

Composition of media (g/l)	Media			
	Raper	PGA	PGA supplemented with rice bran extract	PGA supplemented with fresh oyster mushroom extract
Glucose	20	20	20	20
Yeast extract	2	-	-	-
Peptone	2	-	-	-
Potatoes	-	250	250	250
KH ₂ PO ₄	0.46	-	-	-
K ₂ HPO ₄	1	-	-	-
MgSO ₄ ·7H ₂ O	0.5	-	-	-
Rice bran	-	-	20	-
Fresh oyster mushroom	-	-	-	25
Agar	20	20	20	20
pH	7	7	7	7

Effect of temperature on mycelial growth

Following the media experiment, strain GA3 was inoculated on PGA supplemented with rice bran at pH 7 and incubated in darkness at four different temperatures (20±1°C; 25±1°C, 30±1°C and 35±1°C).

Effect of different initial pH levels on mycelial growth

The growth of *G. lucidum* strain GA3 on PGA supplemented with rice bran at 30±1°C and different pH levels was tested between pH 3.0 and 12.0 in increments of 1.0 pH units. pH levels were initially adjusted by using 1M sodium hydroxide (NaOH) or hydrochloric acid (HCl).

Effect of substrate mixtures on fruiting body formation

To investigate the most favourable substrate mixtures for fruiting body formation, *G. lucidum* was cultivated on rubber (*Hevea brasiliensis*) wood sawdust as the basal substrate with different types of supplements added, as indicated in Table 2.

Table 2. Composition of substrate mixtures for fruiting body formation.

Composition (%)	Treatment				
	1	2	3	4	5
Sawdust	87	87	87	87	87
Corn powder	4	4	4	4	4
Rice bran	8	6	4	2	0
Wheat bran	-	2	4	6	8
CaCO ₃	1	1	1	1	1

Data collection

Important characteristics of mycelial morphology such as texture (cottony, floccose), density (high, moderate, low), and colour (off-white, white) were recorded by means of visual observation. Diameter growth (mm) was measured at 5, 7, and 9 days after inoculation. The mycelial growth rate was calculated as follows: $V = D/T$, where V is the mycelial growth rate (mm/day), D is as the diameter growth (mm), and T is the duration of mycelial growth (days).

The period of surface colonisation (days) was defined as the time required for the mycelium to grow throughout the media and establish total colonisation on the bag surface. The period of primordia formation (days) was defined as the time required for the formation of primordia from the time of inoculation. The length of stalk (cm) and width of fruiting body (cm) were measured. Biological efficiency was measured as the ratio of the mass of dry fruiting body (g) per dry mass of substrate (g) and expressed as a percentage.

Statistical analysis

The data obtained were statistically analysed using GraphPad Prism (version 7.0, GraphPad Software Inc., San Diego, CA). Each treatment was replicated three times. Differences among the means of groups were assessed using two-way ANOVA followed by Tukey multiple range test, with $p < 0.05$. Values in the same column in a table with different letters were considered to differ significantly.

Results and discussion

Effect of media on mycelial growth of strain GA3

Nutrients, temperature, light, and pH are known to be significant factors that noticeably influence on the mycelial growth of mushrooms. To investigate the optimal media for rapid mycelial growth of strain GA3, diameter growth, and mycelial morphology (texture, density, and pigmentation) were recorded in the four different culture media. As shown in Fig. 1 and Table 3, strain GA3 was able to grow in all four kinds of media. Flocculence and whiteness were identified as the main mycelial morphology and pigmentation, respectively. Maximal mycelial growth was observed on PGA supplemented with rice bran extract, with an average mycelial growth rate of 9.29 ± 0.11 mm/day. PGA was determined to be unsuitable media for mycelial growth of GA3. PGA supplemented with rice bran exhibited high mycelial density. In contrast, the mycelial density of GA3 was found to be moderate on Raper and PGA supplemented with oyster mushroom fresh but low on PGA media. These results suggest that PGA supplemented with rice bran may be considered the optimal media for

mycelial growth of strain GA3 and this was therefore selected for further optimisation. Jayasinghe, et al. (2008) [16] report that Hamada (dextrose, ebirose, hyponex yeast-extract), Glucose Peptone (glucose, malt-extract, peptone, yeast-extract), Yeast-Malt extract (dextrose, malt-extract, peptone, yeast-extract), Mushroom Complete (malt-extract, peptone, yeast-extract, K_2HPO_4 , $MgSO_4$, KH_2PO_4) and Lilly (asparagine, maltose, $MgSO_4$, KH_2PO_4) are suitable media for the growth of *G. lucidum*. As reported by Badalyan, et al. (2015) [17], the mycelial morphology of *G. lucidum* is white and felt/cottony, with denser aerial mycelium in the centre during the initial incubation period stage of growth.

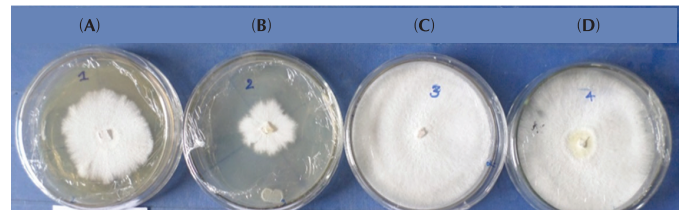


Fig. 1. Mycelial growth on different media, incubated in darkness for 9 days at 25°C, pH 7. (A) Raper; (B) PGA; (C) PGA supplemented with rice bran; (D) PGA supplemented with oyster mushroom fresh.

Table 3. The influence of different culture media on the mycelial growth performance of strain GA3 at 25°C, pH 7.

Media	Diameter growth (mm) after days			Mycelial characteristics		
	5	7	9	Density	Texture	Pigmentation
Raper	20.39±0.43 ^a	35.33±0.63 ^a	46.16±1.42 ^a	Moderate	Floccose	White
PGA	14.50±0.48 ^b	23.16±0.17 ^b	29.94±1.13 ^b	Low	Floccose	White
PGA + rice bran	24.28±0.62 ^c	60.5±0.82 ^c	83.68±1.06 ^c	High	Floccose	White
PGA + oyster mushroom fresh	21.28±0.22 ^d	57.50±0.35 ^d	71.39±0.11 ^d	Moderate	Floccose	White

Effect of temperature level on mycelial growth of strain GA3

As with the media, temperature is one of the most significant physical factors affecting both the growth of mycelium and fruiting body formation. According to Jayasinghe, et al. (2008) [16], favourable mycelial growth of *G. lucidum* was recorded at 25-30°C. However, it is worth noting that a range of temperatures, between 30 and 35°C, was found suitable for mycelial growth of *G. lucidum* [18, 19]. To ascertain the optimal temperature for favourable mycelial growth of strain GA3, the mycelial growth pattern was recorded at four different temperatures, 20±1°C, 25±1°C, 30±1°C, and 35±1°C for 5, 7, and 9 days. The results indicate that the incubation temperature has a significant influence on the growth of strain GA3. Optimal mycelial growth was observed at 30°C, followed by 25°C and 20°C, as indicated in Fig. 2 and Table 4.

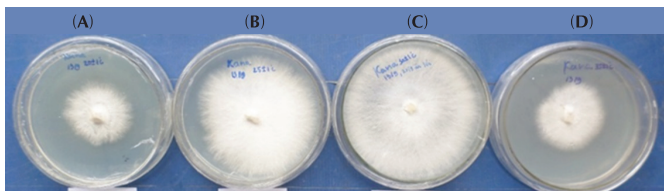


Fig. 2. Mycelium growth of strain GA3 grown in darkness on PGA medium supplemented with rice bran for 7 days at pH 7 after inoculation at various temperatures. (A) $20\pm 1^\circ\text{C}$; (B) $25\pm 1^\circ\text{C}$; (C) $30\pm 1^\circ\text{C}$; (D) $35\pm 1^\circ\text{C}$.

Table 4. The influence of different temperatures on mycelial growth performance. Strain GA3 was grown on PGA medium supplemented with rice bran at pH 7.

Temperature ($^\circ\text{C}$)	Diameter growth (mm) after days			Mycelial density		
	5	7	9	Density	Texture	Pigmentation
20 ± 1	27.89 ± 0.15^a	43.22 ± 0.20^a	76.22 ± 0.37^a	Moderate	Floccose	White
25 ± 1	39.83 ± 0.33^b	66.61 ± 0.20^b	85.89 ± 0.31^b	High	Floccose	White
30 ± 1	48.94 ± 0.20^c	74.94 ± 0.48^c	90.00 ± 0.15^c	Moderate	Floccose	White
35 ± 1	21.05 ± 0.28^d	39.50 ± 0.29^d	54.39 ± 0.42^d	Moderate	Floccose	White

Effect of different initial pH levels on mycelial growth of strain GA3

One of the most important chemical factors, pH can affect cell membrane function, the uptake of various nutrients, cell morphology and structure, the solubility of salts, the ionic state of substrates, enzyme activity, and product biosynthesis [20]. Rai (2003) [18] has reported that *G. lucidum* prefers an acidic pH for vegetative growth. In addition, a pH range from 4.0 to 6.5 was found to be the optimal initial pH for the growth of *G. lucidum*, as previously described by Veena and Pandey (2006) [19]. According to Kapoor and Sharma (2014) [21], *G. lucidum* can grow in a broad range of pH values, from 3.0 to 11.0, though the highest mycelial growth rate was observed at the pH 5.0 level. As indicated by Jayasinghe, et al. (2008) [16], the optimal pH for mycelial growth varies widely and is strongly related to the genotype of strain. Remarkably, strain GA3 is capable growing in a wide pH range, from 4 to 12, as shown in Fig. 3 and Fig. 4.

Effect of substrate mixtures on fruiting body formation of strain GA3

As previously reported, the yield and biological efficiency of *G. lucidum* relate not only to the kind of sawdust but also the supplements used [22]. Sawdust is used as the basal substrate in mixtures for cultivating *G.*

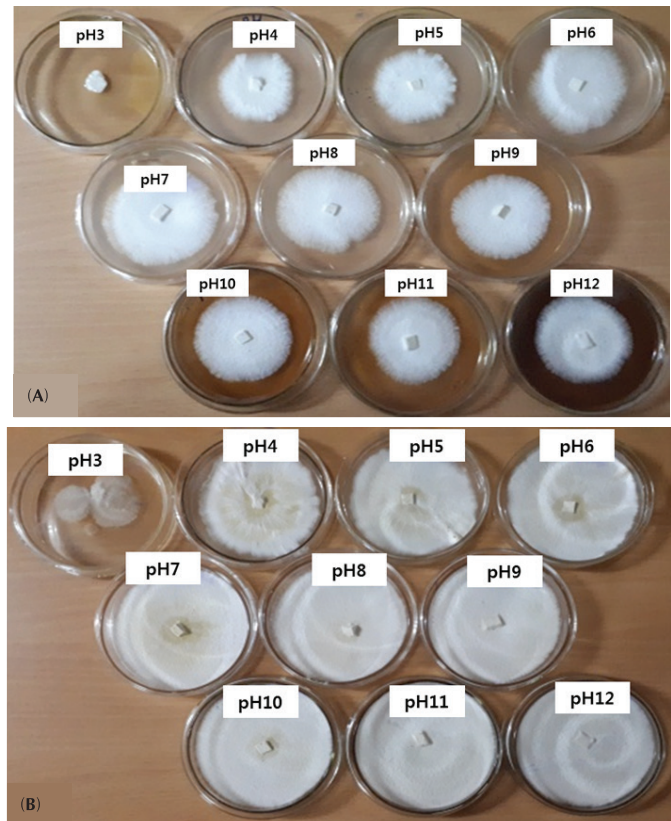


Fig. 3. Mycelium growth of strain GA3 grown in darkness on PGA medium supplemented with rice bran for 3 days (A) and 7 days (B) after inoculation at 30°C .

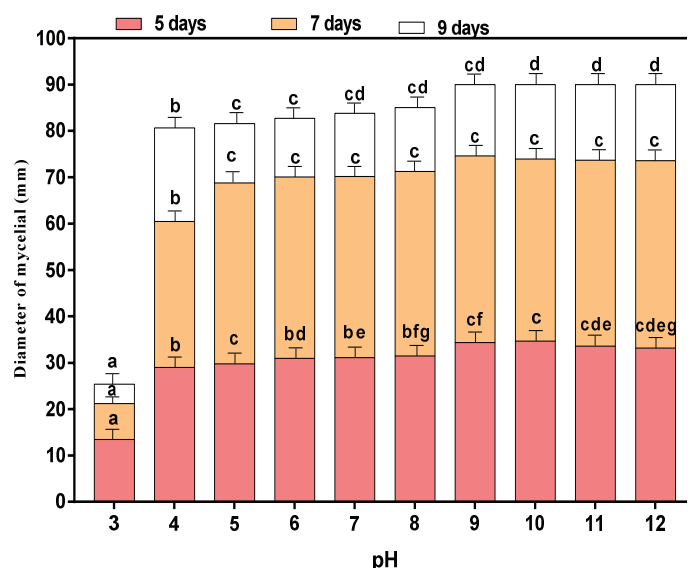


Fig. 4. Effect of initial pH on mycelial growth of strain GA3.

lucidum [23]. Compared to poplar and beech sawdust, oak sawdust was observed to support the cultivation of *G. lucidum* and produced the highest biological efficiency [22]. The effect of various kinds of substrate mixtures on fruiting body formation of *G. lucidum* was investigated in this study.

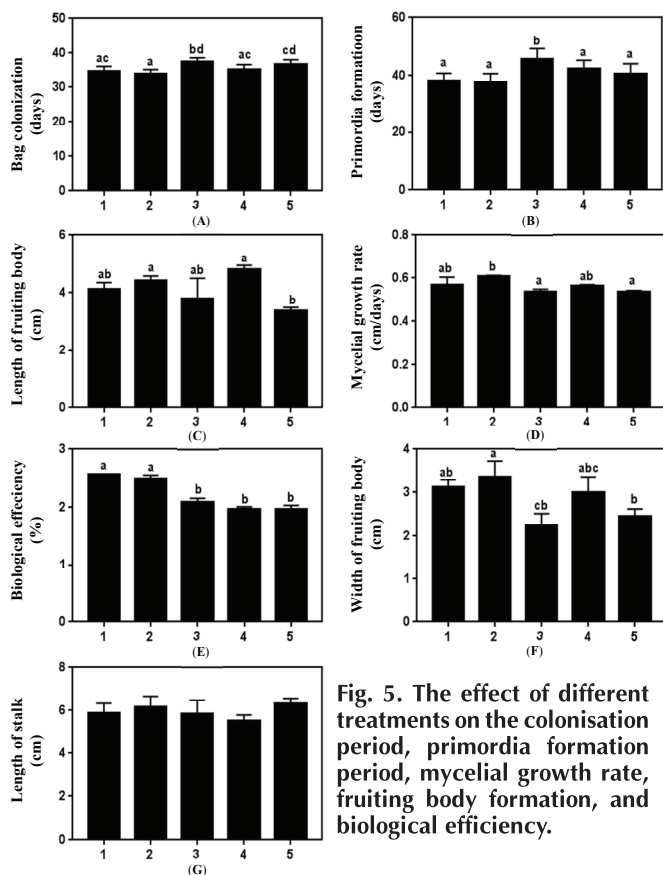


Fig. 5. The effect of different treatments on the colonisation period, primordia formation period, mycelial growth rate, fruiting body formation, and biological efficiency.

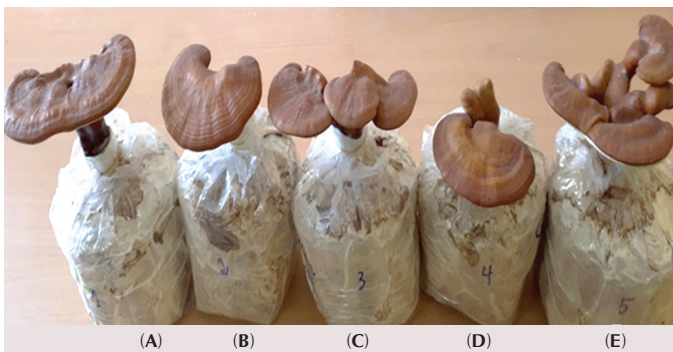


Fig. 6. Fruiting body of *Ganoderma lucidum* strain GA3 cultivated using different treatments. (A) treatment 1; (B) treatment 2; (C) treatment 3; (D) treatment 4; (E) treatment 5.

In this experiment, five treatments were used in order to ascertain the most effective treatment for the growth of mycelial and for fruiting body formation. To cultivate strain GA3, corn powder, rice bran, and wheat bran were used as major nutrients to supplement the substrate. The period required for surface colonisation, primordia formation, the length of the fruiting body, the mycelial growth rate, biological efficiency, the width of fruiting body, and the length of the stalk were monitored and are shown in Figs.

5A, 5B, 5C, 5D, 5E, 5F, and 5G, respectively. The results presented in Figs. 5A, 5B indicate that strain GA3 is able to grow to form primordia in all five treatments. The pilei of strain GA3 were found to be kidney-shaped (Fig. 6). Compared with other treatments, treatment 2 (87% sawdust + 4% corn powder + 6% rice bran + 2% wheat bran + 1% CaCO_3) and treatment 1 (87% sawdust + 4% corn powder + 8% rice bran + 1% CaCO_3) reduced the time required for surface colonisation, bag colonization, and primordia formation and resulted in a significantly higher growth rate (Figs. 5A, 5B, and 5D). As expected, the greatest fresh mass (17.16 g) and biological efficiency (2.56%) were obtained with treatment 1, followed by treatment 2 (16.66 g and 2.49%) (Fig. 5E). Rice bran and corn powder are known to have high vitamin content, especially vitamin B2. This is may be due to the presence of high percentage of rice bran in the substrate mixtures of treatments 1 and 2. In contrast, treatments 4 and 5 showed a lower biological efficiency (13.13 g and 1.96%, and 13.20 g and 1.97%), respectively. Therefore, treatment 1 (87% sawdust + 4% corn powder + 8% rice bran + 1% CaCO_3) is considered the most suitable substrate combination for cultivating strain GA3.

Conclusions

The optimal conditions for mycelial growth of strain GA3 were observed at 25-30°C on PGA media supplemented rice bran. Strain GA3 grew in a wide pH range, from 4 to 12. Of the five treatments used for cultivating of *G. lucidum*, treatment 1 (87% sawdust + 4% corn powder + 8% rice bran + 1% CaCO_3) was the most suitable substrate mixture for improving biological efficiency.

ACKNOWLEDGEMENTS

This study was partially funded by Vietnam National University of Agriculture as a “research working group project”.

The authors declare that there are no conflicts of interest regarding the publication of this article.

REFERENCES

- [1] S. Wachtel-Galor, J. Yuen, J.A. Buswell, I.F.F. Benzie (2011), “*Ganoderma lucidum* (Lingzhi or Reishi): a medicinal mushroom”, *Herbal Medicine: Biomolecular and Clinical Aspects* (2nd edition), CRC Press/Taylor & Francis.
- [2] E.I. Erkel (2009), “The effect of different substrate mediums on yield of *Ganoderma lucidum* (Fr.) Karst”, *J. Food Agric. Environ.*, **77**(3), pp.841-844.
- [3] T. Mizuno, W-Y. Wang, J. Zhang, H. Kawagishi, T. Nishitoba, and J.X. Li (1995), “Reishi, *Ganoderma lucidum* and *Ganoderma tsugae*: bioactive substance and medicinal effects”, *Food and Reviews*

International, **11(1)**, pp.151-166.

[4] N. Ninluam, W. Potiprasert, U. Romreun, E. Bangyeekhun (2016), "Cultivation of Lingzhi mushroom, *Ganoderma lucidum*, by using sugarcane bagasse", *Veridian E-Journal, Science and Technology Silpakorn University*, **3(6)**, pp.390-397.

[5] S.T. Chang, J.A. Buswell (2008), "Safety, quality control and regulatory aspects relating to mushroom nutraceuticals", *Proc. 6th Intl. Conf. Mushroom Biology and Mushroom Products*, pp.188-195.

[6] S.T. Chang, J.A. Buswell (1999), "*Ganoderma lucidum* (Curt.: Fr.) P. Karst. (Aphyllophoromycetideae): a mushrooming medicinal mushroom", *Int. J. Med. Mushrooms*, **1(2)**, pp.139-146.

[7] S.P. Wasser, P. Coates, M. Blackman, G. Cragg, M. Levine, J. White (2005), *Encyclopedia of Dietary Supplements*, Marcel Dekker, pp.680-690.

[8] B. Boh, M. Berovic, J. Zhang, L. Zhi-Bin (2007), "*Ganoderma lucidum* and its pharmaceutically active compounds", *Biotechnol. Annu. Rev.*, **13**, pp.265-301.

[9] H. Riu, G. Roig, J. Sancho (1997), "Production of carpophores of *Lentinus edodes* and *Ganoderma lucidum* grown on cork residues", *Microbiologia SEM*, **13**, pp.185-192.

[10] M. Azizi, M. Tavana, M. Farsi and F. Oroojalian (2012), "Yield performance of Lingzhi or Reishi medicinal mushroom, *Ganoderma lucidum* (W. Curt.: Fr.) P. Karst. (Higher Basidiomycetes), using different waste materials as substrates", *Int. J. Med. Mushrooms*, **14(5)**, pp.521-527.

[11] S. Jandaik, R. Singh, M. Sharma (2013), "Comparative growth characteristics and yield attributes of Lingzhi or Reishi medicinal mushroom, *Ganoderma lucidum* (Higher Basidiomycetes) on different substrates in India", *Int. J. Med. Mushrooms*, **15(5)**, pp.497-503.

[12] X.W. Zhou, K.Q. Su and Y.M. Zhang (2012), "Applied modern biotechnology for cultivation of *Ganoderma* and development of their products", *Appl. Microbiol. Biotechnol.*, **93(3)**, pp.941-963.

[13] D.T. Ha, L.T. Loan, T.M. Hung, L.V.N. Han, N.M. Khoi, L.V. Dung, B.S. Min and N.D. Nguyen (2015), "An improved HPLC-DAD

method for quantitative comparisons of triterpenes in *Ganoderma lucidum* and its five related species originating from Vietnam", *Molecules*, **20(1)**, pp.1059-1077.

[14] P.D.N. Nguyen, H.T. Do, B.D. Le (2013), "Characteristics of ecological factors and their distribution of Ganodermataceae Donk. in highlands of Vietnam", *J. Biol.*, **35**, pp.198-205.

[15] S.G. Jonathan, I.O. Fasidi (2003), "Studies on *Psathyrella atroumbonatai* (Pegler), a Nigerian edible fungus", *Food Chem.*, **81(4)**, pp.481-484.

[16] C. Jayasinghe, A. Imtiaz, H. Hur, G.W. Lee, T.S. Lee, U.Y. Lee (2008), "Favorable culture conditions for mycelial growth of Korean wild strains in *Ganoderma lucidum*", *Korean J. Mycol.*, **36(1)**, pp.28-33.

[17] S.M. Badalyan, A.V. Shnyreva, M. Lotti, and A. Zambonelli (2015), "Genetic resources and mycelial characteristics of several medicinal polypore mushrooms (Polyporales, Basidiomycetes)", *Int. J. Med. Mushrooms*, **17(4)**, pp.371-384.

[18] R.D. Rai (2003), "Successful cultivation of the medicinal mushroom Reishi (*Ganoderma lucidum*) in India", *Mushroom Res.*, **12**, pp.87-91.

[19] S.S. Veena and M. Pandey (2006), "Evaluation of the locally available substrates for the cultivation of indigenous *Ganoderma* isolates", *J. Mycol. Pl. Pathol.*, **36(3)**, pp.434-438.

[20] V. Elisashvili (2012), "Submerged cultivation of medicinal mushrooms: bioprocesses and products (Review)", *Int. J. Med. Mushrooms*, **14(3)**, pp.211-239.

[21] P. Kapoor and B.M. Sharma (2014), "Studies on different growth parameters of *Ganoderma lucidum*", *International Journal of Science, Environment and Technology*, **3(4)**, pp.1515-1524.

[22] O.K. Gurung, U. Budathoki, and G. Parajuli (2012), "Effect of different substrates on the production of *Ganoderma lucidum* (Curt.: Fr.) Karst", *Our Nature*, **10(1)**, pp.191-198.

[23] E. Erkel (2009), "Yield performance of *Ganoderma lucidum* (Fr.) Karst cultivation on hosts containing different protein and carbohydrate source", *Afr. J. Agric. Res.*, **4(11)**, pp.1331-1333.

Characterization of newly isolated thermotolerant yeasts and evaluation of their potential for use in *Cayratia trifolia* wine production

Doan Thi Kieu Tien^{1,2}, Huynh Xuan Phong¹, Mamoru Yamada³,
Ha Thanh Toan¹, Ngo Thi Phuong Dung^{1*}

¹Biotechnology Research and Development Institute, Can Tho University, Vietnam

²Faculty of Food Technology and Biotechnology, Can Tho University of Technology, Vietnam

³Faculty of Agriculture, Yamaguchi University, Japan

Received 9 July 2018; accepted 19 October 2018

Abstract:

Thermotolerant ethanologenic yeasts have attracted the interest of many scientists due to the current challenges caused by increasing global temperature, the benefits associated with processing at high temperatures, and the potential to reduce cooling costs. The objectives of this study are to characterize the selected thermotolerant yeasts and to evaluate their use in *Cayratia trifolia* fermentation at high temperatures. A total of 151 yeast strains isolated from 53 samples of *Cayratia trifolia* were studied for their morphology, physiology, biochemistry, and their phylogenetic relationship. Based on the results of tests for thermotolerance ability (37-45°C) and ethanol tolerance capacity (9-12% v/v), 57 of the 151 yeast isolates were selected to be tested for use in wine fermentation from three-leaf cayratia at 37°C. Thirty isolates that were found to have high fermentation ability and that produced an ethanol concentration of between 6.0 and 9.9% (v/v) were selected for identification using amplified 26S rDNA sequences. The yeasts were identified as follows: *Candida glabrata* (BL2.1, CT1.1, CT1.3, CT2.3, HG2.1), *Candida tropicalis* (KG1.1, KG3.2, CM3.3, HG3.3, TG1.1, TG3.1), *Candida nivariensis* (DT1.2, CM3.2, ST2.1, BT1.2), *Pichia kudriavzevii* (KG2.1, KG5.1, AG2.1, AG2.3, AG4.2, DT3.2, LA1.3, CM4.4, BT2.1, BT3.3, TV4.2, CT4.2, VL1.1), *Clavispora lusitaniae* (TG4.2), and *Saccharomyces cerevisiae* (HG1.3). The phylogenetic tree constructed using MEGA 6 with bootstrap analysis performed by repeating the data 1,000 times revealed that the selected yeast strains were closely related. The newly isolated strain of *S. cerevisiae* HG1.3 producing the highest ethanol concentration of 9.9% (v/v) in *Cayratia trifolia* wine fermentation at 37°C was selected for further study.

Keywords: *Cayratia trifolia*, ethanol fermentation, ethanol tolerance, *Saccharomyces cerevisiae*, thermotolerant yeast.

Classification number: 3.5

Introduction

Cayratia trifolia (L.) Domin is a rich source of biologically active compounds with antioxidant properties that can reduce tumor growth [1, 2]. It is used as a medicinal ingredient and in alcoholic wines. Currently, fermentation products are being researched for quality, yield and scale, for their application in industrial production to meet consumer demand. Wine, which is an indispensable drink that contributes greatly to supporting human health, is made from a variety of ingredients other than grapes. Temperature is a factor that significantly affects the fermentation capacity of yeast. In summer, the temperature in the South of Vietnam increases dramatically, particularly with global warming [3]. Thus, the use of thermotolerant yeast strains is essential for dealing with climate change. Furthermore, high temperature fermentation has several advantages, such as a reduction in the cost of cooling fermentation vats, higher saccharification yields, continuous removal of ethanol, and decreased risk of bacterial contamination [4-7]. Therefore, the use of thermotolerant yeast strains in ethanol production contributes to lowering manufacturing expenses.

The aims of this study are to isolate thermotolerant yeasts and evaluate their fermentation capacity for the production of three-leaf cayratia (*Cayratia trifolia* L.) wine.

Materials and methods

Culture and materials

Fifty-three samples of *Cayratia trifolia* were collected from 13 provinces in the Mekong Delta region. This was carried out in three phases:

I: the *C. trifolia* berries were collected from the four provinces of Kien Giang, An Giang, Dong Thap, and Long An.

*Corresponding author: ntpdung@ctu.edu.vn

II: then, berries was collected from the four provinces of Can Tho, Hau Giang, Vinh Long, and Tien Giang.

III: finally, berries was collected from the five provinces (Ca Mau, Bac Lieu, Soc Trang, Ben Tre, and Tra Vinh).

The berries were brought fresh to the Laboratory of Food Microbiology at the Biotechnology Research and Development Institute, Can Tho University and were processed immediately.

The microbiological medium used was YPD broth (g/l, D-glucose 20, peptone 5, yeast extract 5) with 20 g/l of agar added to make a YPD agar medium.

Research method

Isolation of yeast strains:

Five grams of each *Cayratia trifolia* sample was added to 100 ml of YPD broth and incubated at 30°C, 150 rpm for 24-48 hours. Yeast colonies were selected, streaked on YPD agar, and incubated at 30°C. Purified yeast cultures were stored in YPD agar slants at 4°C.

Examination of morphological, physiological, and biochemical characteristics:

Morphological characteristics: the shapes and dimension of colonies and cells were observed under a microscope and recorded.

Glucose, sucrose and maltose fermentation ability: after 24 hours' incubation, yeast suspensions were inoculated into Durham tubes containing a 2% (w/v) sucrose or maltose solution and incubated at 30°C. The accumulated CO₂ in the inner Durham tubes was measured after 48 hours.

Urea anabolism: yeast isolates were inoculated into tubes containing 3 ml of Stuart's Urea broth and the change in the color of the medium was recorded after incubating at 30°C for 48 hours.

Gelatin liquefaction: yeast isolates were inoculated into tubes containing 3 ml of gelatin medium and then incubated at 30°C for 48 hours. The tubes were immediately cooled and the gelatin liquefaction recorded.

Testing the thermo- and ethanol-tolerant capacity of yeast isolates:

Thermo-tolerance: yeast isolates were streaked onto YPD agar and then incubated at 30, 35, 37, 39, 41, 43, 45 and 47°C for 48 hours. The formations of the colonies that appeared on the medium were recorded.

Ethanol tolerance: yeast isolates were streaked onto YPD agar supplemented with 0, 3, 6, 9, 12 and 15% (v/v) of ethanol and then incubated at 37°C for 48 hours. The formations of the colonies that appeared on the medium

were recorded.

Screening for the ethanol fermentation capacity of yeast isolates:

This test was carried in Durham tubes containing a 2% (w/v) glucose solution and three-leaf cayratia juice (pH 4 and 22°Brix) incubated at 30°C. The accumulated CO₂ in the inner Durham tubes was measured at 6-hour intervals for 48 hours.

Testing ethanol fermentation from three-leaf cayratia juice:

The selected yeast isolates were inoculated into YPD broth and incubated for 48 hours. Then, 1 ml of yeast cell suspension (10⁸ cells/ml) was inoculated into 99 ml of three-leaf cayratia juice (pH=4 and 22°Brix) and incubated at 37°C. The pH, °Brix and ethanol concentration were determined.

Identification of selected yeast isolates:

The DNA of selected yeast isolates was extracted and used for nucleotide sequencing. The divergent D1/D2 (500 bp) domain of the LSU rRNA gene was amplified with the specific primers NL-1 (5'-GCATATCAATAAGCGGAGGAAAAG) and NL-4 (5'-GGTCCGTGTTTCAAGACGG) [8]. Nucleotide sequences were aligned and compared with the database on the National Center for Biotechnology Information website. The identification was conducted at the Department of Biological Chemistry, Faculty of Agriculture, Yamaguchi University, Japan.

Analytical method and statistical analysis:

The pH was measured with a digital pH meter (Sartorius PB-20). The total dissolved solids of the saccharified liquid (°Brix) was measured using a manual refractometer (FG102/112, Euromex-Holland). The alcohol content was determined using the distillation method [9]. The experimental data were statistically analyzed using Statgraphics Centurion XV software from Manugistics Inc., USA.

Results and discussion

Morphological, physiological, and biochemical characteristics of yeast isolates

One hundred and fifty-one yeast strains were isolated and purified from 53 *C. trifolia* berry samples. The yeast strains were cultured on YPD agar medium for 36 hours at 30°C and were investigated for their colony and cell morphology. Based on cell morphology and physiological and biochemical characteristics, the 151 yeast isolates were divided into 7 groups (Table 1).

Table 1. Summary of yeast cell shape.

Group	Cell conformation	Name of yeast isolate*			No. of isolate
		Group I	Group II	Group III	
1	Small spherical		CT3.2, CT4.5, HG1.1, HG1.3, HG4.3, VL3.3, TG1.1, TG3.1	CM1.1, CM1.2, ST1.3	11
2	Large spherical		CT3.3, CT4.1, HG4.4, VL1.2, VL3.1		5
3	Small oval	KG1.1, KG1.3, KG2.1, AG1.3, AG2.1, AG4.1, DT1.3, DT3.1, LA2.1	CT1.1, CT1.2, CT1.3, CT2.3, CT3.1, CT4.2, HG1.2, HG2.1, VL4.2, TG2.2, TG4.3, TG4.4	CM1.3, CM2.1, CM4.1, CM4.3, BL1.1, BL2.1, BL2.3, BL3.1, BL4.3, ST2.1, ST2.3, ST3.1, ST4.3, BT1.2, BT3.1, TV2.1, TV2.2, TV3.2	39
4	Large oval	KG1.2, KG2.2, KG2.3, KG3.1, KG4.2, KG5.1, KG5.2, AG1.1, AG2.4, AG3.2, DT1.1, DT1.2, DT2.1, DT2.3, DT3.2, LA1.1, LA1.2, LA1.3, LA3.1, LA3.2, LA3.3	CT2.1, CT2.2, CT4.4, HG3.1, HG3.3, VL2.2, VL4.4, TG1.2, TG2.3, TG4.2	CM2.2, CM3.1, CM4.2, BL1.2, BL2.2, ST1.2, BT1.3, BT2.1, BT3.2, BT3.3, TV1.2, TV4.4	43
5	Short ellipse	DT4.2, DT4.3, LA4.1	CT4.3, HG3.2, HG4.1, VL2.1, VL4.3, TG2.1, TG4.1	BL4.2, ST1.1, ST2.2, BT4.2	14
6	Elongated ellipse	KG3.2, AG1.2, AG2.3, AG3.1, AG4.2, DT2.2, LA2.2, LA2.3, LA3.4	HG2.2, HG4.2, HG4.5, VL1.1, VL1.3, VL3.2, VL4.1, TG3.2	CM4.4, CM3.2, CM3.3, TV4.2, BL3.2, ST3.3, BT1.1, TV2.3, TV3.1, TV4.1	27
7	Apiculate ellipse	KG4.1, AG2.2, DT4.1, LA4.2		BL4.1, ST3.2, ST4.1, ST4.2, BT2.2, BT4.1, TV1.1, TV4.3	12
Total					151

*Notes: Group I: yeast isolates from three-leaf cayratia were collected from Kien Giang, An Giang, Dong Thap, and Long An; Group II: yeast isolates from three-leaf cayratia were collected from Can Tho, Hau Giang, Vinh Long, and Tien Giang; Group III: yeast isolates from three-leaf cayratia were collected from Ca Mau, Bac Lieu, Soc Trang, Ben Tre, and Tra Vinh.

Colony morphology of yeast isolates: the colonies of yeast isolates measured 1-4 mm in diameter and 0.1 mm in height. Some colonies had smooth surfaces while others had rough surfaces. The margins of colonies were also diverse and included entire, undulate, serrated, filiform and lobate. The colonies of yeast were creamy white or white in color.

Cell morphology of yeast isolates: cell shape of yeast isolates were diverse but can be categorized into 4 main forms: spherical, ovoid, elliptical and cylindrical. There were also differences in the dimensions of yeast isolates but generally cell length was approximately 3-11 µm and cell width was approximately 2-5 µm.

Budding and endospore formation: yeast isolates in group 1, 2, 3, 4, 5 and 6 grew by multilateral budding, while isolates in group 7 grew by bipolar budding. All the yeast isolates had the ability to sporulate when nutritionally deficient except those in group 7. Although the endospore dimensions of yeast

isolates were not homogeneous, each cell had four ascospores. Yeast tends to form four ascospores after meiosis in their sexual reproduction [10].

Glucose, sucrose and maltose fermentation ability: of 151 yeast isolates, 138 were capable of using glucose and 101 of using sucrose as a carbon source for fermentation after 24 hours. Most strains in groups 1, 2, 4, 5 and 6 were capable of fermenting sucrose while none of the strains in group 3 could ferment this sugar. Of 151 yeast isolates, 104 were able to ferment maltose. The sugar fermentation capacity of the yeast strains was assessed by measuring the amount of CO₂ generated during fermentation [11]. Thus, testing the ability to consume sugar was one of criteria for classification the yeast and was also used to select appropriate yeast strains for fermenting different substrates.

Urea assimilation: of 151 yeast isolates, 26 were able to use urea as a source of nitrogen. None of the yeast isolates in groups 1 and 2 were capable of urea resolution. Yeasts belonging to Ascogenous species were able to resolve urea, while those of the Basidiomycetous species had this capability [12].

Gelatin liquefaction: of the 151 yeast isolates, 32 had the capacity to liquify gelatin using gelatinase. This capacity of yeasts was also often associated with protease activity, but only some yeast species were capable of producing protease [11].

The ethanol- and thermo-tolerant capacities of the yeast isolates

Thermotolerant ability: all yeast isolates could grow well in the temperature range 30-35°C. Ten of the 151 yeast strains showed high heat resistance by growing at 45°C. However, the number of yeast colonies generally decreased when the incubation temperature was increased. Among 48 yeast isolates with a high fermentation capacity, 10 isolates (BT2.1, TG2.3, VL1.1, HG4.3, LA1.1, DT3.2, AG4.2, AG3.1, AG2.3, AG2.1) could grow at temperature of 45°C and 38 isolates were able to grow at 43°C after 48 hours of incubation. Based on the results of the thermotolerant screening test, 141 yeast isolates that could grow at 37-45°C were selected for further testing of their ethanol tolerant ability.

Ethanol tolerant ability: when the ethanol concentration in the culture medium was increased, the number of yeast colonies that developed in the medium gradually decreased. This can be explained for causing affect to the yeast growth. Of the 141 isolates, 27 could tolerate an ethanol concentration of up to 12% (v/v), and 64 could tolerate a 9% (v/v) ethanol concentration after 48 hours of incubation.

Screening of ethanol fermentation ability of yeast isolates

The results reveal that 57 out of 64 yeast isolates were able to ferment tubes containing 2% (w/v) glucose solution and three-leaf cayratia juice after 48 hours. Yeast strains including KG2.2, KG3.1, DT1.2, CM3.2 and BT1.2 showed highest fermentation abilities which created maximum CO₂ amount

in Durham tubes within a 6-hour fermentation. The two yeast isolates KG4.1 and AG3.2 had no fermentation capacity. A total of 57 isolated strains that could grow at 37-45°C and tolerate 9-12% (v/v) ethanol were evaluated for their ability to ferment three-leaf cayratia at 37°C.

Ethanol fermentation by selected yeast isolates at high temperatures

The ethanol fermentation ability of 30 out of the 57 selected yeast isolates is presented in Table 2. These isolated yeast strains showed the best fermentation activity and an ethanol content of at least 6.0% (v/v). The highest ethanol concentration was produced by strain HG1.3, which reached at 9.9% (v/v). Isolates HG1.3, CM3.2 and AG2.1 produced the highest ethanol concentration in each group at 9.9, 8.95 and 8.0% (v/v), respectively. The obtained ethanol concentrations of these novel thermotolerant yeasts were better than many thermotolerant

yeasts isolated from drainage samples containing hot spring water. Isolates collected from hot spring water could generate maximum ethanol concentrations of approximately 7.0-7.2% (w/v) at 30°C with a nutritional substrate containing 15% (w/v) glucose [13].

At 37°C, there was a clear difference in ethanol concentration produced by 39 tested yeast isolates. The growth of yeast cells also went up when the temperature was increased to a level within the tolerance threshold of the yeast, but the amount of ethanol produced was reduced. Enzymes which control microbial activity and fermentation are sensitive to high temperatures which can denature their tertiary structure and deactivate them [14]. The five yeast isolates BT3.3, BT2.1, HG2.1, HG3.3, VL1.1 and TG4.2 showed lower fermentation ability, whereby ethanol concentrations reached only around 6.0% (v/v).

Table 2. Ethanol producing capacity of 30 selected yeast isolates at 37°C.

No.	Isolate	Thermo- tolerance (°C)	Ethanol-tolerance % (v/v)	CO ₂ in Durham tube (24 h)	Ethanol (% v/v)
Group I: Isolates of yeast from three-leaf cayratia collected from the four provinces of Kien Giang, An Giang, Dong Thap, and Long An					
1	KG1.1	41	12	30 (18 h)	7.12 ^{bc*}
2	KG2.1	41	12	19	6.25 ^{def}
3	KG3.2	43	9	8	6.08 ^{efg}
4	KG5.1	43	12	14	6.49 ^{cde}
5	AG2.1	45	12	30 (18 h)	8.00 ^a
6	AG2.3	45	12	18	6.23 ^{def}
7	AG4.2	45	12	20.67	6.16 ^{defg}
8	DT1.2	43	12	30 (6 h)	6.19 ^{defg}
9	DT3.2	45	12	29.33	6.68 ^{bcd}
10	LA1.3	43	12	17.33	6.84 ^{bc}
Group II: Isolates of yeast from three-leaf cayratia collected from the four provinces of Can Tho, Hau Giang, Vinh Long, and Tien Giang					
11	CT1.1	43	9	30	8.05 ^b
12	CT1.3	42	9	30 (18 h)	6.4 ^c
13	CT2.3	43	12	30 (18 h)	6.4 ^c
14	CT4.2	41	9	30	8.2 ^b
15	HG1.3	43	9	30	9.9 ^a
16	HG2.1	41	12	25.7	6.1 ^d
17	HG3.3	41	12	30 (18 h)	6 ^{de}
18	VL1.1	45	12	24	6 ^{de}
19	TG1.1	43	9	30 (18 h)	6.4 ^c
20	TG3.1	41	9	30	6.4 ^c
21	TG4.2	41	12	23.3	6 ^{de}
Group III: Isolates of yeast from three-leaf cayratia collected from the five provinces of Ca Mau, Bac Lieu, Soc Trang, Ben Tre, and Tra Vinh					
22	CM3.2	43	12	30 (6 h)	8.95 ^a
23	CM3.3	41	9	30	7.01 ^b
24	CM4.4	39	9	19.7	6.56 ^{cd}
25	BL2.1	43	9	30 (6 h)	6.61 ^{cd}
26	ST2.1	39	9	30 (6 h)	6.48 ^{cd}
27	BT1.2	39	9	30 (6 h)	6.79 ^{bc}
28	BT2.1	45	9	30 (18 h)	6.09 ^e
29	BT3.3	41	12	30 (18 h)	6.09 ^e
30	TV4.2	39	12	30	6.32 ^{de}

*Note: values in the table were the average values of triplication. The average values in a group with the same letter were not significantly different at the 95% confidence level.

Generally, wine fermentation of *C. trifolia* berries using thermotolerant yeast showed the same trend, whereby ethanol concentrations decreased when temperatures were increased. In this study, ethanol concentration were lower than that produced in the optimal temperature. At high temperatures, the accumulation of intracellular ethanol in yeast cells was increased, which stalled yeast growth. As a result, the fermentation ability of the yeast was affected and lower ethanol concentrations were generated [15].

Identification of selected yeast isolates

The results of aligning the 26S rDNA sequences of 30 selected yeast strains with the GenBank database (NCBI) along with an analysis of their morphology, physiology, and biochemistry indicated that all strains belonged to one of the four genera *Saccharomyces*, *Candida*, *Clavispora* and *Pichia*. The results of identification of 30 selected yeast isolates are presented in Table 3. *S. cerevisiae* was popularly utilized for alcoholic fermenting in industrial manufacturing. *S. cerevisiae* was able to yield an ethanol concentration of between 7.4 and 7.7% (w/v) fermenting molasses at room temperature. This species is also likely to grow at high temperatures ranging

from 40 to 44°C [16, 17]. Thus, it was decided to use the thermotolerant yeast *S. cerevisiae* HG1.3 to make wine from the fresh berries of *C. trifolia*.

The genetic relation of selected thermotolerant yeasts was determined by constructing a phylogenetic tree based on the 26S rDNA gene using MEGA 6 software (Neighbor-Joining). The phylogenetic tree for 30 selected yeast strains is shown in Fig. 1.

Table 3. The identification results of 30 selected yeast isolates.

No.	Genera	Species	Name of yeast isolate	No. of isolate
1	<i>Candida</i>	<i>Candida tropicalis</i>	KG1.1, KG3.2, CM3.3, HG3.3, TG1.1	5
		<i>Candida nivariensis</i>	DT1.2, ST2.1, BT1.2	3
		<i>Candida glabrata</i>	BL2.1, CT1.1, CT1.3, CT2.3,	4
2	<i>Pichia</i>	<i>Pichia kudriavzevii</i>	KG5.1, AG2.3, AG4.2, CM4.4, BT2.1, BT3.3, CT4.2, VL1.1	8
3	<i>Clavispora</i>	<i>Clavispora lusitanae</i>	TG4.2	1
4	<i>Saccharomyces</i>	<i>Saccharomyces cerevisiae</i>	HG1.3, CM3.2, AG2.1, TV4.2, DT3.2, LA1.3, KG2.1, TG3.1, HG2.1	9
Total				30

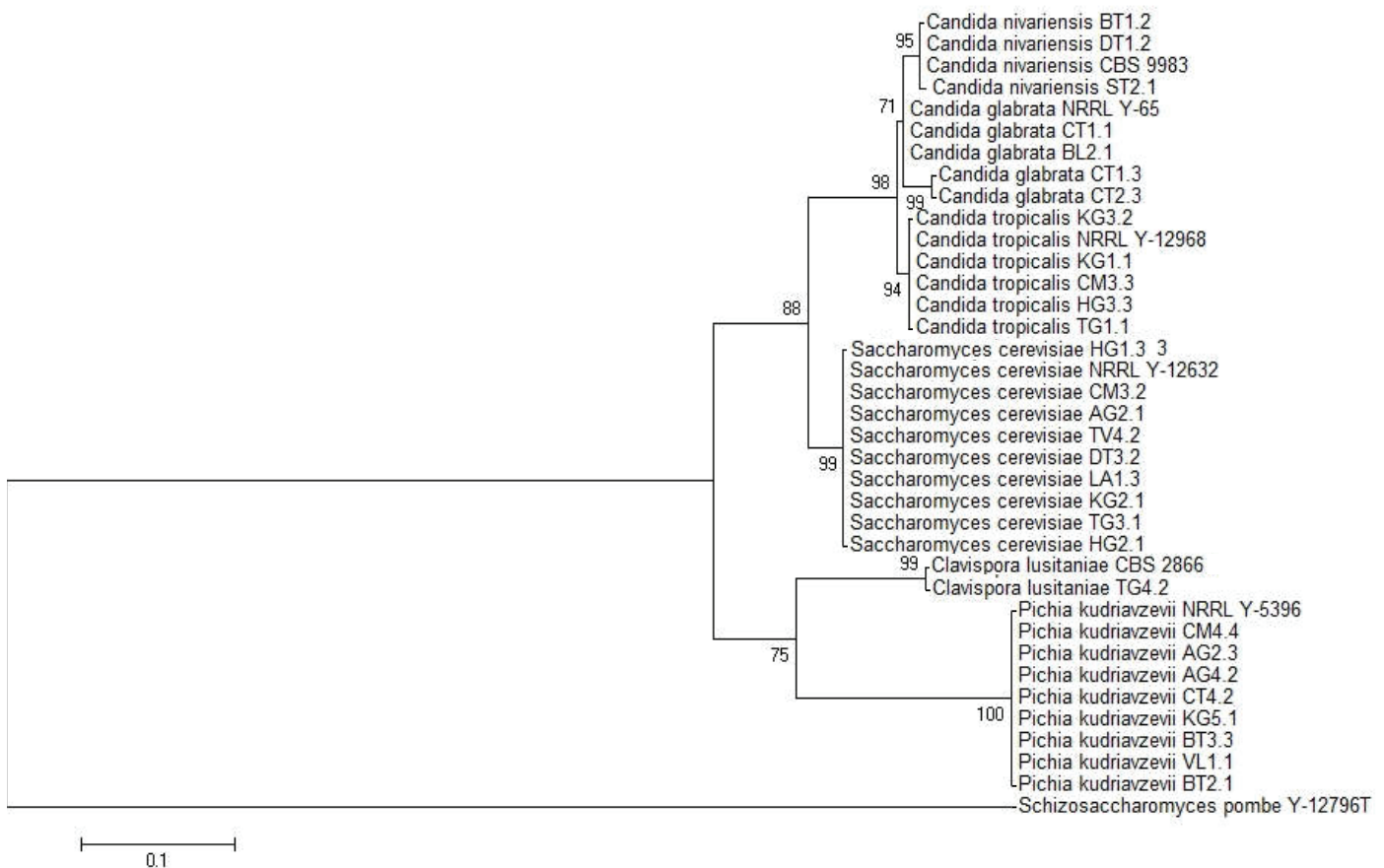


Fig. 1. Phylogenetic tree of 30 selected yeast strains.

The phylogenetic tree showed the genetic relation of the selected thermotolerant yeasts. It indicated that *Saccharomyces cerevisiae* HG1.3, CM3.2, AG2.1, TV4.2, DT3.2, LA1.3, KG2.1, TG3.1, and HG2.1 are the most closely related strains because of their high reliability (with 100% Bootstrap) and that the first distinct branch is *Candida nivariensis*.

Conclusions

The diversity of yeast isolates purified from *C. trifolia* berry samples was examined, and a number of ethanol- and thermo-tolerant ethanologenic yeasts were found. The feasibility of fermentation products from *C. trifolia* by the selected yeast isolates at high temperature was confirmed. This study indicated the promising applications of such isolates for the controlled *C. trifolia* wine fermentation at high temperature.

ACKNOWLEDGEMENTS

This research was jointly sponsored by the Ministry of Science and Technology of Vietnam (Contract Nr. 09/2014/HD-NDT), the Advanced Program in Biotechnology, Can Tho University, and the New Core-to-Core Program (2014-2019).

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] A.K. Gupta, M. Shamar (2007), "Review on Indian medical plant", *Council of Medical Research*, **7**, pp.879-882.
- [2] P.C. Perumal, S. Sowmya, P. Pratibha, B. Vidya, P. Anusooriya, Starlin, S. Ravi, and V.K. Gopalakrishnan (2015), "Isolation, structural characterization and in silico drug-like properties prediction of a natural compound from the ethanolic extract of *Cayratia trifolia* (L.)", *Pharmacognosy Reviews*, **7(1)**, pp.121-125.
- [3] H.X. Phong, N.T.C. Giang, S. Nitiyon, M. Yamada, P. Thanonkeo, and N.T.P. Dung (2016), "Ethanol production from molasses at high temperature by thermotolerant yeasts isolated from cocoa", *Can Tho University Journal of Science*, **3**, pp.32-37.
- [4] I.M. Banat, P. Nigam, D. Singh, R. Marchant, A.P. McHale (1998), "Review: ethanol production at elevated temperatures and alcohol concentrations: part I - yeasts in general", *World Journal of Microbiology and Biotechnology*, **14**, pp.809-821.
- [5] M. Roehr (2001), *The biotechnology of ethanol: classical and future applications*, 1st edition, Wiley-VHC, Weinheim.
- [6] S. Limtong, C. Sringiew, W. Yongmanitchai (2007), "Production of fuel ethanol at high temperature from sugar cane juice by a newly isolated *Kluyveromyces marxianus*", *Bioresource Technology*, **98**, pp.3367-3374.
- [7] B.A. Abdel-Banat, H. Hoshida, A. Ano, S. Nonklang, R. Akada (2010), "High-temperature fermentation: how can processes for ethanol production at high temperatures become superior to the traditional process using mesophilic yeast", *Applied Microbiology and Biotechnology*, **85**, pp.861-867.
- [8] K. O'Donnell (1993), "Fusarium and its near relatives", *The Fungal holomorph: mitotic, meiotic and pleomorphic speciation in Fungal systematics*, Reynolds D.R. & Taylor J.W. (Eds), pp.225-233, CAB International, Wallingford.
- [9] W.Z., Bruce, C.F. Kenneth, H.G. Barry, S.N. Fred (1995), *Wine analysis and production*, pp.447-449, Chapman and Hall, New York.
- [10] A.M. Neiman (2005), "Ascospore formation in the yeast *Saccharomyces cerevisiae*", *Microbiology and Molecular Biology Reviews*, **69(4)**, pp.565-584.
- [11] P.C. Kurtzman, J.W. Fell, T. Boekhout, V. Robert (2011), "Methods for isolation, phenotypic characterization and maintenance of yeast", *The Yeast, a taxonomic study*, **1**, 5th ed., Elsevier, Amsterdam.
- [12] J.A. Barnett, R.W. Payne, D. Yarrow (1983), *Yeast: characteristics and identification*, Cambridge University Press, Cambridge.
- [13] R. Ueno, N. Urano, S. Kimura (2002), "Effect of temperature and cell density on fermentation by a thermotolerant aquatic yeast strain isolated from a hot spring environment", *Fisheries Science*, **68**, pp.571-578.
- [14] M. Phisalaphong, N. Srirattana, W. Tanthapanichakoon (2006), "Mathematical modeling to investigate temperature effect on kinetic parameters of ethanol fermentation", *J. Biochem. Eng.*, **28**, pp.36-43.
- [15] T. D'Amore, C.J. Panchal, I. Russell, G.G. Stewart (1990), "A study of ethanol tolerance in yeast", *Critical Reviews in Biotechnology*, **9(4)**, pp.287-304.
- [16] W.R. Abdel-Fattah, M. Fadil, P. Nigam, I.M. Banat (2000), "Isolation of thermotolerant ethanologenic yeasts and use of selected strains in industrial scale fermentation in an Egyptian distillery", *Biotechnology and Bioengineering*, **68(5)**, pp.531-535.
- [17] K.N. Sree, M. Sridhar, K. Suresh, I.M. Banat, L. Venkateswar-Rao (2000), "Isolation of thermotolerant, osmotolerant, flocculating *Saccharomyces cerevisiae* for ethanol production", *Bioresource Technology*, **72(1)**, pp.43-46.

Remote sensing applications for analysing the impacts of land cover changes on the upper part of the Dong Nai river basin

Hung Pham^{1,2*}, Van Trung Le², Le Phu Vo²

¹Department of Natural Resources and Environment, Lam Dong province

²Ho Chi Minh city University of Technology - Vietnam National University, Ho Chi Minh city

Received 10 October 2018; accepted 3 January 2019

Abstract:

In recent years, activities related to socio-economic development have led to land cover (LC) changes in the upper part of the Dong Nai river basin. The use of remote sensing applications to analyse the impacts of these changes plays an important role in the managing the sustainability of the river basin. This paper introduces a solution for analysing the impacts of LC changes on the water balance in the upstream catchment of the Dong Nai river in Lam Dong province. Landsat images were used for mapping and monitoring major changes over the last 20 years. Rainfall and water discharge data was collected from the local hydrometeorological stations to identify the impacts of the LC changes on the runoff in the catchment area. The results show that the forest area was reduced by more than 223,576 ha (23%). The main changes were an increase in the agricultural area from 18.2 to 31.3% and in water bodies from 0.9 to 2.2%. The latter was due to hydropower development projects in the catchment area. The LC changes caused by the changes in the hydrological conditions of the river basin have had a significant impact on water resources. The identification of the main LC changes in the catchment area could be useful for establishing a policy to protect the headwater forests and mitigate against future impacts.

Keywords: hydrological conditions, land cover change, Landsat images, remote sensing, upper part of Dong Nai river basin.

Classification number: 4.1

Introduction

Land cover (LC) is the physical material on the Earth's surface, and LC maps play an important role in Earth system studies and ecosystem management [1]. Land cover changes can be related to natural processes, such as flooding and erosion, and anthropogenic activities, including urbanization and agriculture. Annually updated LC information is valuable for formulating socio-economic development policies and as data for environmental management applications, such as vulnerability and risk assessment [2]. Characterising and mapping LC is essential for multiple purposes, including planning and managing natural resources (e.g. land or water resource development, flora and fauna conservation), modelling environmental variables, and understanding the spatial distribution of habitats. Remote sensing and digital image processing enable observation, mapping, monitoring, and assessment of LC to be conducted at a range of spatial and temporal scales [3].

Remote sensing provides comprehensive thematic maps based on an image classification for visual or computer-aided analysis to assess past LC changes [4]. The choice of classification algorithm depends on many factors, including ease of use, speed, scalability, the interpretability of the classifier, the kind of data, the statistical distribution of classes and target accuracy. Unsupervised classification is typically used when limiting the knowledge and availability of the LC types [5, 6].

Clustering algorithms, including k-mean and ISODATA, run iteratively until convergence of an optimal set of clusters is achieved. Post-classification refinement techniques, such as merging and splitting clusters, are necessary before labeling because automatically produced

*Corresponding author: Email: hungmtk25@yahoo.com

clusters do not necessarily correspond with LC types [7, 8]. Parametric supervised classifiers are typically used when expert knowledge and the availability of the LC types are sufficient. However, supervised classification with algorithms, such as maximum likelihood, minimum distance, and discriminant analysis, is difficult to perform with multi-temporal data containing many spectral features and multi-modal distributions [9]. Other approaches involve various classifiers used in parallel or in succession, which can be either supervised or unsupervised [10]. Nonparametric classifiers, such as k-nearest neighbours (kNN), decision trees (DT), neural networks (NN), support vector machines (SVM), random forests (RF), and hierarchical classification based on multi-source and multi-temporal data and geo-knowledge (HC-MMK), impose boundaries of arbitrary geometries and provide higher flexibility although they involve computationally intense iterative processes [11]. Nonparametric classifiers that focus on decision rules of class boundaries are more suitable when the statistics and distribution of LC types are unknown [12].

Change image production uses post-classification change detection technique through cross-tabulation [13]. The success of this technique depends on the reliability of the maps created using image classification. Large-scale changes such as the construction of new hydroelectric reservoirs or major urban development might be mapped reasonably easily, whereas for evolutionary changes, such as erosion, colonization and degradation, the boundaries may be indistinct and the class-labels uncertain [14].

Land use and land cover (LULC) changes alter the hydrological system and can have potentially significant

effects on water resources [15]. In addition, the impact of LULC change on watershed hydrology are interlinked with climate change impacts [16].

According to the People's Committee of Lam Dong Province [17], the forest area of Lam Dong was 513,529 ha in 2014 and accounted for 52.5% of the provincial area. This report indicates that the forest area was reduced around 8% in 10 years. However, in the upper part of Dong Nai (UPDN), most of which belongs to Lam Dong province, historical LC change has yet to be examined in detail. In order to analyse past LC changes that impact upon the flow regime in the river basin in 10-year intervals (1994, 2004, and 2014), a changes detection technique was applied using a supervised maximum likelihood classification (MLC) algorithm. The impact of LC change on the flow regime in the UPDN river basin was assessed largely using hydrometeorological data collected along with Landsat images. The objectives of this study are to create LC maps and to observe LC changes over a 20-year period (1994-2014). In order to achieve these objectives, investigations were conducted into the effects of past LC changes and the effects of these changes on water discharges in the downstream part of the river basin. Specifically, the impact of headwater forest change and hydropower development on the flow regime in the UPDN was assessed.

Study area

The study area is located in the UPDN river basin (Fig. 1), which covers an area of 972,460 ha and belongs to the provinces of Lam Dong, Dak Nong, and Dong Nai. The upstream catchment area of the Dong Nai River has a

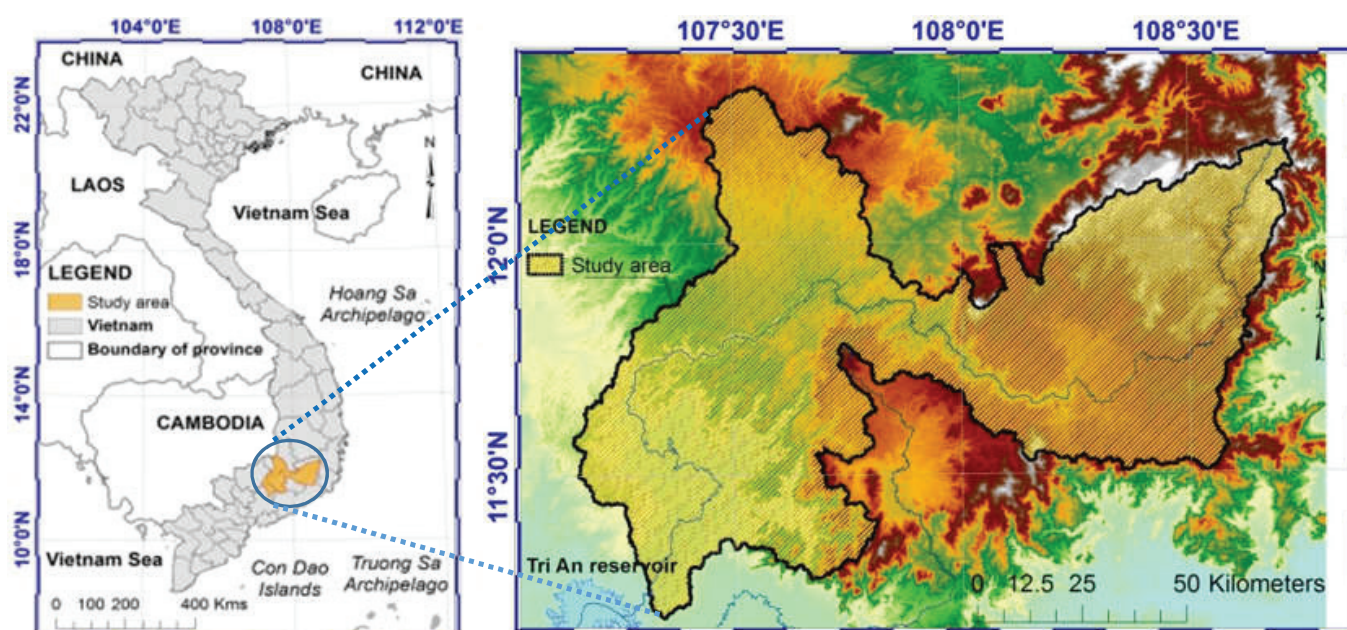


Fig. 1. Location of the upper part of Dong Nai river basin.

tropical wet climate with two seasons: the rainy season from May to November and the dry season from December to April. Over the past 33 years from 1981-2014, the average annual temperature was 22°C, annual precipitation was 2,500 mm, and annual humidity was 83% [18]. Forest cover is found mainly at high elevations in the West and North. The agricultural areas are characterised by small fields generally in close proximity to rivers.

Materials and methods

Landsat data

Image data from Landsat-5 TM (1994, 2004) and Landsat-8 OLI/TIRS (2014) covering the study area was downloaded from the United States Geological Survey (USGS) website (<http://earthexplorer.usgs.gov>), as summarized in Table 1. The criteria for the selection were that cloudless images be available and that the data be collected at a ground measurement station (Ta Lai gauge).

Table 1. Characteristics of Landsat images.

Year	Image	Landsat_Scene_ID	Resolution	Date_Acquired
1994	Landsat-5 TM	LT51240521994007BKT00	30x30 m	1994-01-07
2004	Landsat-5 TM	LT51240522004355BKT01	30x30 m	2004-12-20
2014	Landsat-8 OLI/TIRS	LC81240522014030LGN01	30x30 m	2014-01-30

Geometric correction

The original sub-scenes of Landsat images comprised of a significant amount of bands data, which was combined into one image (6 bands) by function layer stacking using ENVI 4.5 software. For this study, geometric correction was carried out using a ground control point from the available maps (Topographic maps of Lam Dong province in 2010, scale 1:100,000) to geocode the 2014 image. This image was then used to register the images from 2004 and 1994. The geometric correction was done by calculating the root mean square error (RMSE) between the two images, which was less than 0.2 pixels. Corrected geometric images were then cut (subset) into the UPDN river basin.

Training sample data

Training sample data was used to create an LC map with seven main classes, which are listed in Table 2.

Table 2. Land cover classes of the study area.

Type of LC	Description
(1) Water bodies	Natural (Lakes, Rivers, etc.) or man-made water bodies (e.g. Reservoirs)
Forest (2) <i>Broadleaf evergreen forest</i> (3) <i>Mixed forest</i> (4) <i>Coniferous forest</i>	All forests: evergreen broadleaf forest, coniferous forest (pine), mixed forest (bamboo and broadleaf forests, pine, and broadleaf forest, etc.)
(5) Built-up residential areas (6) Seasonal agricultural land (7) Perennial agricultural land	Residential areas, roads and built-up Rice fields, soybean, potato Rubber, coffee, tea, etc.

The training sample data was created based on the GIS data, the land use map of the area (provincial land use planning maps for the period 2010-2020), and the vector data for polygons of training sample data, so-called region of interest (ROI) is used in classification method of MLC. In addition, Google Earth images were deployed to support the selection of LC types for the training sample polygons by integrating Arc Google Tool with ArcGIS 10.1.

The result of the LC classification was evaluated based on ground truth data collected at test sites. The error matrix was used to indicate the quality of LC classifications in 1994, 2004, and 2014. Three natural forest classes (broadleaf evergreen forest, mixed forest, coniferous forest) were combined came under the definition of forest for the purposes of assessing LC changes. This meant that seven classes were categorized into five main classes: water bodies, forest areas, built-up residential areas, seasonal agricultural land, and perennial agricultural land [19, 20].

Land cover classification

The maximum likelihood pixel-based classification method is the most commonly used technique for Landsat images [21]. This study used the MLC method for Landsat 5 TM and Landsat-8 OLI/TIRS. The accuracy assessment is reflected by overall accuracy and Kappa coefficient in which overall accuracy included user's accuracy and producer's accuracy.

The thematic map used to analyse LC change trends in the UPDN river basin is shown in Fig. 2. The LC map was created using images from (A) Landsat-5 TM 1994, (B) Landsat-5 TM 2004 and (C) Landsat-8 OLI/TIRS 2014. The area for each type of LC in the river basin and the cover percentages in are summarised in Tables 3-5.

Results and discussion

Image classification: supervised classification was carried out using MLC, and the same training data was used

for each image. This proved an efficient solution for the visualisation of LC in the basin. The results indicate that the average forest cover decreased from 72.68% of the river basin area in 1994 to 49.97% in 2014. This finding can assist managers in undertaking further analysis regarding forest cover change trends with the aim of achieve sustainable development in the UPDN river basin.

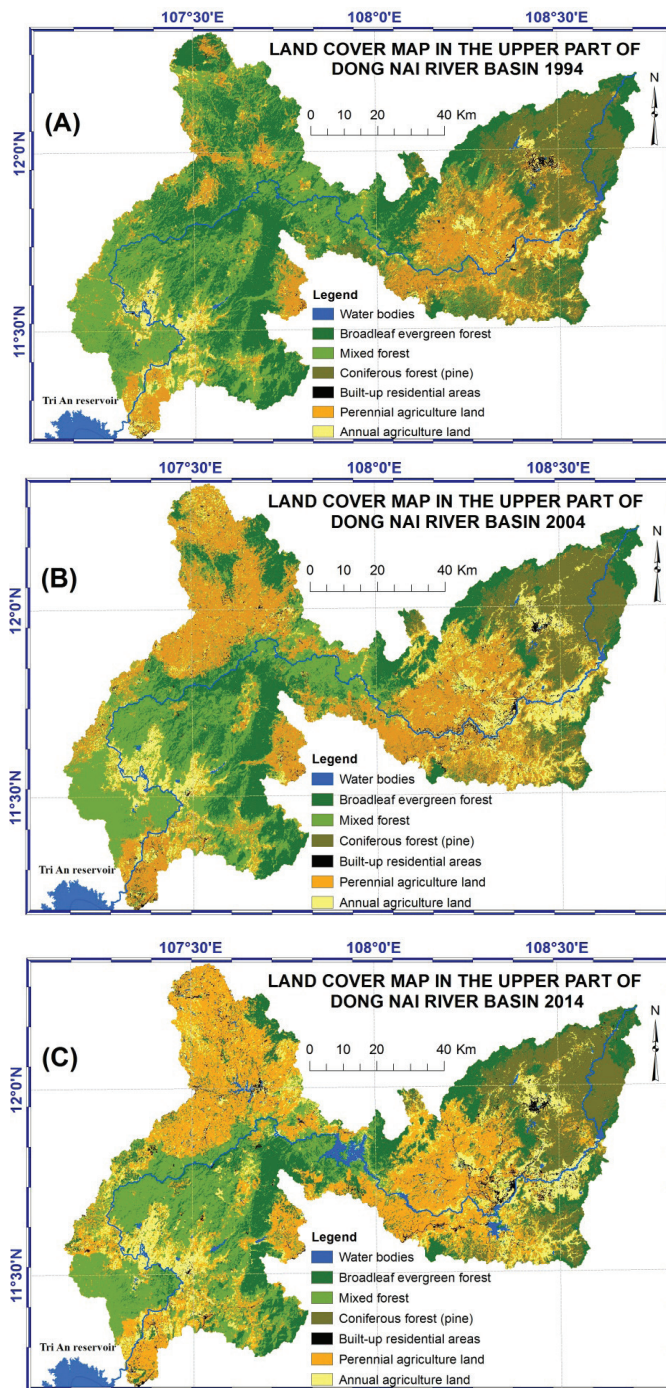


Fig. 2. Land cover map created using different images: (A) Landsat-5 TM 1994, (B) Landsat-5 TM 2004, (C) Landsat-8 OLI/TIRS 2014.

Table 3. Area of land cover and cover percentage (1994).

Types of Land cover	Area (ha)	Percentage (%)
Water bodies	8,505	0.87
Forest areas	706,803	72.68
- Broadleaf evergreen	283,257	29.13
- Mixed forest	283,616	29.16
- Coniferous forest	139,930	14.39
Built-up residential	7,922	0.81
Seasonal agricultural	177,033	18.20
Perennial agricultural	72,197	7.42
Total	972,460	100.00

Table 4. Area of land cover and cover percentage (2004).

Types of Land cover	Area (ha)	Percentage (%)
Water bodies	8,557	0.88
Forest areas	520,359	53.51
- Broadleaf evergreen	188,318	19.37
- Mixed forest	219,435	22.56
- Coniferous forest	112,606	11.58
Built-up residential	19,305	1.99
Seasonal agricultural	292,927	30.12
Perennial agricultural	132,312	13.61
Total	972,460	100.00

Table 5. Area of land cover and cover percentage (2014).

Types of Land cover	Area (ha)	Percentage (%)
Water bodies	21,590	2.22
Forest areas	485,908	49.97
- Broadleaf evergreen	178,720	18.38
- Mixed forest	194,050	19.95
- Coniferous forest	113,138	11.63
Built-up residential	24,274	2.50
Seasonal agricultural	304,231	31.28
Perennial agricultural	136,457	14.03
Total	972,460	100.00

Classification accuracy assessment: an assessment of the quality of LC classifications in 1994, 2004 and 2014 indicated that all seven classifications have very good overall accuracy (77.7-87%). In all cases, the Kappa coefficient had a high value (0.74-0.85). The user's accuracy and producer's accuracy for the LC maps are shown in Table 6. Therefore, the thematic map was used to analyse LC change trends and their impacts on the regime flow in the UPDN river basin.

The results show that the highest accuracy was for water bodies and the lowest accuracy was for broadleaf evergreen forest (Prod. = 57.02, 67.28, and 74.40% for 1994, 2004, and 2014, respectively).

Table 6. Summary of classification accuracy for the land cover map in 1994, 2004 and 2014.

Class name	1994		2004		2014	
	User (%)	Prod. (%)	User (%)	Prod. (%)	User (%)	Prod. (%)
Water bodies	98.26	98.64	98.10	99.95	97.88	98.90
Broadleaf evergreen forest	90.38	57.02	94.53	67.28	86.85	74.40
Mixed forest	65.54	87.19	74.53	89.89	71.01	87.74
Coniferous forest	91.42	92.52	96.47	97.93	94.65	95.62
Built-up residential areas	70.51	81.55	92.94	93.89	91.95	78.04
Seasonal agricultural land	89.69	79.35	92.84	76.73	90.20	80.90
Perennial agricultural land	69.03	90.51	80.82	92.62	77.77	84.94
Overall accuracy (OA)	77.7%		87.0%		84.3%	
Kappa	0.74		0.85		0.81	

This result can be explained by the fact that the river basin is partially covered by areas with high-density coffee trees, causing confusion between broadleaf forest and perennial agricultural land. Furthermore, the user's accuracy for the mixed forest class was also low (user = 65.5, 74.53, and 71.01%, for 1994, 2004, and 2014, respectively). This can be attributed to the fact that the Landsat images were taken in the dry season, when spectral signatures of mixed forest pixels are most similar to measured perennial plant spectra. Moreover, the accurate classification was a good match with the land use planning maps of Lam Dong province for the periods of 2000-2010 and 2010-2020 [22, 23].

Detection change: to analyse LC change, three natural forest classes (broadleaf evergreen forest, mixed forest, and

coniferous forest) were grouped under the forest definition and thematic maps containing five main LC classes were created. The LC map for 1994 was overlaid onto the LC map for 2014 in order to identify the regions where major changes had occurred in the five LC classes between 1994 and 2014.

The results show that the for 1994, 2004, and 2014 the forest area occupied 706,803 ha (72.68%), 520,359 ha (53.51%), and 485,908 ha (49.97%), respectively. This means that the area of forest coverage changed significantly over the 20 years from 1994 to 2014. This result is consistent with trends reported by the UN (2005) for the period 1990-2000, during which tropical forests in South-East Asia were reduced from 53.9% in 1990 to 48.6% in 2000 [24]. However, the forest area did not change much between 2004 and 2014, only dropping from 53.51% (2004) to 49.97% (2014), as shown in Tables 3-5.

In contrast, there was a significant increase of seasonal agricultural land and perennial agricultural land in the 10 years from 1994 to 2004. This indicates that the demand for agricultural land increased due to local socio-economic development. The area of seasonal agricultural land was 177,033 ha (18.20%) in 1994, 292,927 ha (30.12%) in 2004, and 304,231 ha (31.28%) in 2014, whereas perennial agricultural land accounted for 72,197 ha (7.42%) in 1994, 132,312 ha (13.61%) in 2004, and 136,457 ha (14.03%) in 2014.

The area of water bodies fluctuated over the study period measuring 8,505 ha (0.87%) in 1994, 8,557 ha (0.88%) in 2004, and 21,590 ha (2.22%) in 2014. This fluctuation can be explained by many reasons including climate conditions (change in annual rainfall), water use and land use change. The increase in the area covered by water bodies in the period 2004-2014 also reflects the recent construction of the large hydropower plants Dai Ninh (300 MW), Da Dang 2 (34 MW), Dong Nai 3 (180 MW), Dong Nai 4 (340 MW), Dong Nai 2 (70 MW) and Dong Nai 5 (150 MW), which came into operation in 2008, 2009, 2010, 2012, 2013, and 2014, respectively [18, 25].

Residential coverage was 7,922 ha (0.81%) in 1994, 19,305 ha (1.99%) in 2004, and 24,274 ha (2.50%) in 2014. This reflects the low levels of urbanisation and population growth in the basin.

Table 7 summarises the results of the changes in area for each LC class during the period from 1994 to 2014.

Table 7. Cross-tabulation of land cover classes between 1994 and 2014 (area in ha).

	1994						Class Total
	Water	Forest	Built-up	Perennial Agri.	Seasonal Agri.	Row Total	
Water	5,122	11,484	147	5,210	2,769	24,732	24,733
Forest	895	448,091	414	25,976	7,877	483,253	483,309
Built-up	200	11,202	2,435	14,092	5,794	33,722	33,743
Perennial Agri.	464	168,920	1,431	91,231	16,821	278,875	278,885
Seasonal Agri.	1,821	67,188	3,493	40,462	38,920	151,885	151,912
Class Total	8,502	706,885	7,919	176,972	72,182	-	-
Class Changes	3,382	259,168	5,488	85,830	33,282	-	-
Image Difference	16,231	-223,576	25,824	101,913	79,730	-	-

Notes: The 'Class Total' row shows the total number of pixels in each initial state class. The 'Class Total' column shows the total number of pixels in each final state class. The 'Row Total' column is a class-by-class summation of all final state pixels that fell into the selected initial state classes. The 'Class Changes' row shows the total number of initial state pixels that changed classes. The 'Image Difference' row is the difference between the total number of equivalently classed pixels in the two images, computed by subtracting the initial state class total from the final state class total.

Overall, the results show that the area of the forest cover decreased by 223,576 ha from an average cover of 72.68% of the natural area in 1994 to 49.97% in 2014. The agricultural land area and water surface (bodies) area also increased in the same period due to the construction of the hydropower reservoirs.

Figure 3 shows major changes from forest to other land classes in the river basin.

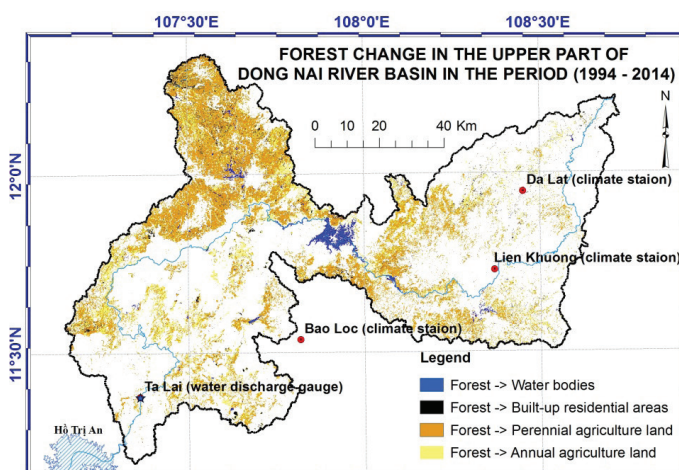


Fig. 3. Changes of forest into other land classes during the period from 1994 to 2014.

Land cover change impacts: in order to analyse the impacts of LC changes on the water balance in the upstream catchment area of the Dong Nai river, the difference between the area of each LC type must be assessed. Table 7 shows that the area of forest in the UPDN river basin was reduced by 223,576 ha (22.99%) over the 20-year period 1994-2014 due to the conversion of forests into built-up, perennial agricultural, and seasonal agricultural land. The changes for these LC types can be explained by a decrease in the level of evapotranspiration in the river basin. The increase in the area of water bodies caused by the recent development of hydropower projects had an impact on evapotranspiration and the annual water balance of the catchment in the dry season due to an increase in water consumption caused by irrigation practices. In this study, the impact of LC changes on hydrology can be analysed on water discharges in the river basin that affects the downstream part of the Dong Nai river to serve the local socio-economic development. In order to identify the impact of LC change on water discharges in the river basin, rainfall data was collected from three weather stations (Da Lat, Lien Khuong, Bao Loc) and discharge data was collected from Ta Lai gauge, as shown in Fig. 3. The hydrometeorological data was collected along with the Landsat images in 1994, 2004, 2014. The yearly rainfall and yearly discharge total for Ta Lai gauge is shown in Fig. 4.

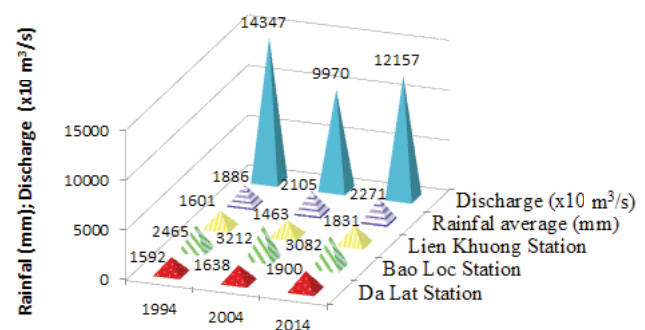


Fig. 4. Yearly rainfall at three weather stations and yearly discharge total at Ta Lai gauge.

The distribution of the mean monthly rainfall for the three climate stations and the monthly discharge are shown in Figs. 5, 6. Obviously, the average rainfall for the three meteorological stations did not change significantly, but the total runoff at the downstream part of the river basin changed dramatically in 2014. The flow in the dry season of year 2014 is higher than it was in 1994. At the same time, water discharges in the river basin for the 2014 rainy season were lower than those in 1994. This can be explained by the hydropower operations in the river basin. For example, water transportation for the Dai Ninh hydropower (300 MW) plant reduced the total water discharge in the lower river.

This is evidence that land use change influenced the flow regime in the river. These findings provide local managers with information on natural resources and environmental management practices to protect headwater forests.

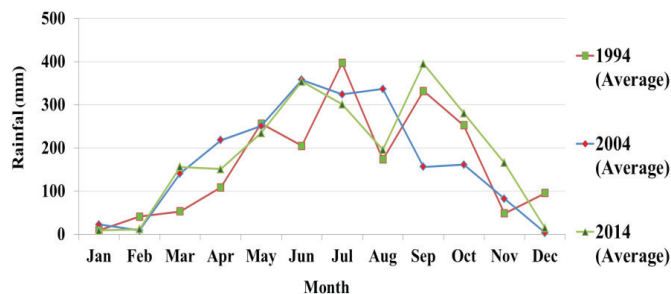


Fig. 5. Monthly rainfall average from three weather stations in 1994, 2004, and 2014.

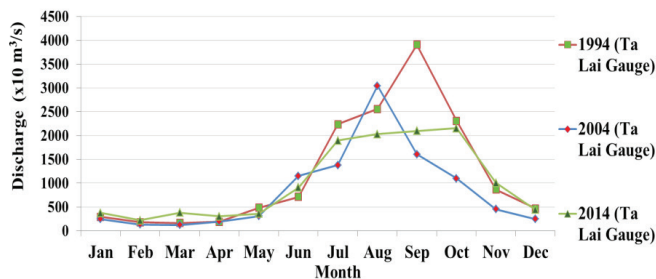


Fig. 6. Monthly discharge observed at Ta Lai gauge in 1994, 2004, and 2014.

Conclusions

The study results show that using Landsat images with algorithm of maximum likelihood supervised classification (MLC) together with generally available data is a comprehensive approach for analysing the impacts of LC changes on the UPDN river basin. The analysis of these results shows that forest area was reduced by more than 223,576 ha (23%) over the 20 years from 1994 to 2014. The agricultural area increased from 18.2% to 31.3% and water bodies also increased from 0.9% to 2.2% due to hydropower development projects in the catchment area. These results indicate that the LC changes were caused by changes in the hydrological conditions of the river basin, which have a significant impact on water resources.

The average rainfall at the three meteorological stations did not change significantly but the total runoff at the downstream part of the river basin changed dramatically in 2014. Land cover change and cascade hydroelectric reservoirs are the major causes of erratic river flow regimes. These changes have had a negative effect on the water quality of the Dong Nai river.

The findings are useful for informing management practices in the watershed area. An analysis of future LC changes and their impacts on the UPDN river basin based on high resolution images would be helpful for the creation of the suitable solutions to the sustainable watershed management.

ACKNOWLEDGEMENTS

The authors would like to thank Ho Chi Minh city University of Technology, Vietnam National University, Ho Chi Minh city, and the Department of Natural Resources and Environment of Lam Dong province for supporting this study.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] P. Gong, J. Wang, L. Yu, Y. Zhao, Y. Zhao, L. Liang, Z. Niu, X. Huang, H. Fu, S. Liu, Congcong Li, X. Li, W. Fu, C. Liu, Y. Xu, X. Wang, Q. Cheng, L. Hu, W. Yao, H. Zhang, P. Zhu, Z. Zhao, H. Zhang, Y. Zheng, L. Ji, Y. Zhang, H. Chen, A. Yan, J. Guo, L. Yu, L. Wang, X. Liu, T. Shi, M. Zhu, Y. Chen, G. Yang, P. Tang, B. Xu, C. Giri, N. Clinton, Z. Zhu, J. Chen, and J. Chen (2013), "Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data", *International Journal of Remote Sensing*, **34**(7), pp.2607-2654.
- [2] C. Gomez, J.C. White, and M.A. Wulder (2016), "Optical remotely sensed time series data for land cover classification: a review", *ISPRS Journal of Photogrammetry and Remote Sensing*, **116**, pp.55-72.
- [3] J. Rogon and D.M. Chen (2004), "Remote sensing technology for mapping and monitoring land cover and land use change", *Progress in Planning*, **61**, pp.301-325.
- [4] G.M. Foody (2002), "Status of land cover classification accuracy assessment", *Remote Sensing of Environment*, **80**, pp.185-201.
- [5] Y. Chen and P. Gong (2013), "Clustering based on eigenspace transformation - CBEST for efficient classification", *ISPRS Journal of Photogrammetry and Remote Sensing*, **83**, pp.64-80.
- [6] H.D. Eva, A.S. Belward, E.E.D. Miranda, C.M.D. Bella, V.R. Gond, O.T.H. Er, S.J. Ones, M. Sgrenzaroli, and S. Fritz (2004), "A land cover map of South America", *Global Change Biology*, **10**, pp.731-744.
- [7] M.A. Wulder, J.C. White, M. Granny, H.J. Hall, J.E. Luther, A. Beaudoin, D.G. Goodenough, and J.A. Dechka (2008), "Monitoring Canada's forest, part 1: completion of the EOSD land cover project", *Can. J. Remote Sensing*, **34**(6), pp.549-562.
- [8] T.R. Loveland, B.G. Reed, J.F. Brown, D.O. Ohlen, Z. Zhu, L. Yang, and J.W. Merchant (2000), "Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data", *International Journal of Remote Sensing*, **21**, pp.1303-1330.

- [9] H. Glanz, L. Carvalho, D. Sulle-Menashe, and M.A. Friedl (2014), "A parametric model for classifying land cover an evaluating training data based on multi-temporal remote sensing data", *Journal of Photogrammetry and Remote Sensing*, **97**, pp.219-228.
- [10] J. Chen, A. Liao, X. Cao, L. Chen, X. Chen, C. He, G. Han, S. Peng, M. Lu, W. Zhang, X. Tong, and J. Mills (2014), "Global land cover mapping at 30 m resolution: a POK-based operational approach", *Journal of Photogrammetry and Remote Sensing*, **103**, pp.7-27.
- [11] G. Lei, A. Li, J. Bian, Z. Zhang, H. Jin, X. Nan, W. Zhao, J. Wang, X. Cao, J. Tan, Q. Liu, H. Yu, G. Yang, and W. Feng (2016), "Land cover mapping in Southwestern China using the HC-MMK approach", *Remote Sensing*, **8(4)**, pp.1-22.
- [12] G.M. Foody and A. Mathur (2006), "The use of small training sets containing mixed pixels for accurate hard image classification: training on mixed spectral responses for classification by a SVM", *Remote Sensing of Environment*, **103**, pp.179-189.
- [13] A. Shalaby and R. Tateishi (2007), "Remote sensing and GIS for mapping and monitoring land cover and land-use changes in the Northwestern coastal zone of Egypt", *Applied Geography*, **27**, pp.28-41.
- [14] G.M. Foody and D.S. Boyd (1999), "Detection of partial land cover change associated with the migration of inner-class transitional zones", *International Journal of Remote Sensing*, **20**, pp.2723-2740.
- [15] P.D. Wagner, S. Kumar, and K. Schneider (2013), "An assessment of land use change impacts on the water resources of the Mula and Mutha rivers catchment upstream of Pune, India", *Hydrology and Earth System Sciences*, **16**, pp.2233-2246.
- [16] L.M. Mango, A.M. Melesse, M.E. McClain, D. Gann, and S.G. Setegn (2011), "Land use and climate change impacts on the hydrology of the upper Mara river basin, Kenya: results of a modeling study to support better resource management", *Hydrology and Earth System Sciences*, **15**, pp.2245-2258.
- [17] People's Committee of Lam Dong province (2015), *Decision No. 299/QĐ-UBND dated 28/1/2015 on reporting the results of forest investigation of Lam Dong province in 2014*, Da Lat, Lam Dong, Vietnam.
- [18] Statistical Bureau of Lam Dong province (2015), *Statistical Yearbook of Lam Dong province*, Da Lat, Lam Dong, Vietnam.
- [19] Y. Zhao, P. Gong, L. Yu, L. Hu, X. Li, C. Li, H. Zhang, Y. Zheng, J. Wang, Y. Zhao, Q. Cheng, C. Liu, S. Liu, and X. Wang (2014), "Towards a common validation sample set for global land-cover mapping", *International Journal of Remote Sensing*, **35**, pp.4795-4814.
- [20] V.D. Sy, M. Herold, F. Achard, R. Beuchle, J.G.P.W. Clevers, E. Lindquist, and L. Verchot (2015), "Land use pattern carbon loss following deforestation in South America", *Environmental Research Letters*, **10**, pp.1-15.
- [21] D. Phiri and J. Morgenroth (2017), "Developments in landsat land cover classification methods: a review", *Remote Sensing*, **9(967)**, pp.1-25.
- [22] Government of Vietnam (2003a), *Land use planning map the period of 2000-2010 of Lam Dong province*, Ha Noi, Vietnam.
- [23] Government of Vietnam (2013b), *Land use planning map the period of 2010-2020 of Lam Dong province*, Ha Noi, Vietnam.
- [24] UN (2005), *World and regional trends*, Millennium Indicators Databases.
- [25] Department of Natural Resources and Environment of Lam Dong province (2013), *Report 456/BC-STNMT dated 11/28/2013 on hydropower development projects in Lam Dong province*, Da Lat, Lam Dong, Vietnam.

The characteristics of seismic activity in the ladder zone of Da river hydro-electric plants

Dinh Trong Cao^{1*}, Van Dung Le¹, Anh Tuan Thai¹, Thanh Hai Dang¹, Dinh Trieu Cao²

¹Institute of Geophysics, VAST

²Institute of Applied Geophysics, VUSTA

Received 4 June 2018; accepted 16 October 2018

Abstract:

In this paper, the authors describe a study examining the characteristics of seismic activity in the ladder zone of Da river hydro-electric plants. The study results show that:

- The occurrence of natural earthquakes in the ladder zone of Da river hydro-electric plants is the strongest in Vietnamese territory. An earthquake magnitude of $M=6.7$ has been observed in the Tuan Giao and Lai Chau areas, with a b coefficient value (Gutenberg-Richter law) of 0.935.
- Natural earthquakes with magnitudes in the range of 6.0-6.9 seem to occur in 13 seismic sources, these being: Muong Te, Nam Nhe, Muong Nhe, Lai Chau - Dien Bien, Son La, Tuan Giao, Phong Tho, Muong La - Bac Yen, Song Da (Da River), Nam Tong, Mu Cang Chai, Lao Cai - Ninh Binh (Phan Sipan), and Mai Chau. Of these, an apparent risk of strong earthquakes, with a magnitude of $M=6.9$, appears in Muong Te, Lai Chau - Dien Bien, Tuan Giao, Yen Chau, and Mai Chau areas.
- Triggered earthquakes with a magnitude $M>4.0$ have occurred in the region of the Hoa Binh and Son La hydro-electric plant reservoirs, which have a maximum depth of over 100 meters, and a geological structure comprising mainly limestone.
- No likelihood of triggered earthquakes in the region of the Hoa Binh hydro-electric plant reservoir; however, in the Son La hydro-electric reservoir region, a triggered earthquake is expected in the future with a strong magnitude of $M=4.3$, or possibly even $M=5.0$.

Keywords: ladder zone of Da river hydro-electric plants, natural earthquake, triggered earthquake.

Classification numbers: 4.2, 4.3

Introduction

The triggering of earthquakes by the filling of artificial water reservoirs has been recognised for over six decades, the first instance of this being pointed out by Carder (1945) at Lake Mead in the United States of America. To date, over 90 sites globally have been identified where earthquakes have been triggered by such activity. Destructive earthquakes exceeding magnitude 6 occurred at Hsinfengkiang, China in 1962; Kariba, Zambia-Zimbabwe border in 1963; Kremasta, Greece in 1966; and Koyna, India in 1967. The Koyna earthquake of $M=6.3$, that occurred on 10 December 1967, is so far the largest earthquake to be scientifically recognised as ‘triggered’. There is argument about whether the Sichuan, China, $M=7.9$ earthquake of 12 May 2008, which claimed over 80,000 lives, was triggered by the filling of the nearby Zipingpu reservoir [1]. In Vietnam, the study of reservoir triggered seismicity was initiated after the occurrence of earthquakes in the vicinity of Hoa Binh hydro-electric reservoir in 1989. The reservoir was impounded in May 1988. When the water level reached 80 m in December 1988, some weak earthquakes were detected in the vicinity of the dam. Earthquake activity continued, with some earthquakes of $M=2.4$ being detected during the period January-April, 1989. On 14 April 1989, seven earthquakes were recorded, including two events of $M=3.8$ and 3.7 with a hypocentral depth of 5 km. On 23 May 1989, the largest earthquake, measuring $M=4.9$, occurred [2, 3]. Another earthquake, of $M=4.7$, occurred at Song Tranh 2, Vietnam on 15 November 2012, and was assumed to be triggered by reservoir impoundment, as this region was otherwise regarded as largely aseismic [4-7].

Currently, five hydro-electric plants are operating in the area of the Da river main stream, namely: Hoa Binh, Son La, Lai Chau, Huoi Quang, and Ban Chat - the highest capacity hydro-electric plants in Vietnam (Table 1). Meanwhile, the Da river valley is assessed as an area of high seismic activity [8-11]; a representative earthquake is that of Tuan Giao, which reached $M=6.7$, and which occurred in June 1983. This earthquake was recorded by the system of seismostations and is known as the strongest recorded earthquake in Vietnam of the end of

*Corresponding author: Email: cdtrieu@gmail.com

20th and beginning of the 21st centuries.

In order to contribute to warnings given to the authorities on the safety of operations on the Da river ladder hydro-electric plant system, the authors of this paper have carried out the following:

1. A review of the characteristic manifestations of seismic activity (natural earthquakes), and
2. An assessment of the hazard of triggered earthquakes

around hydro-electric reservoirs.

In this paper, a comprehensive study is described which estimates b coefficient values and the densities of released energy, as well as the earthquake generating sources based on fault segmentation. It also predicts the maximum earthquake magnitude of triggered earthquakes based on an Artificial Neural Network algorithm. The main aim of this paper is to give an overview of the characteristics of seismic activity in the ladder zone of Da river hydro-electric plants.

Table 1. Basic parameters of hydro-electric reservoirs on the Da river and its two main branches.

No	Content	Unit	Lai Chau hydro-electric plant	Son La hydro-electric plant	Hoa Binh Hydro-electric plant	Nam Chien hydro-electric plant	Ban Chat hydro-electric plant	Huoi Quang hydro-electric plant
1	River name		Da river	Da river	Da river	Chien spring	Nam Mu spring	Nam Mu spring
2	Coordinate		102°59'33"E	103°59'42" E	105°19'26" E	104°16'23"E	103°50'0"E	103°52'25"E
			22°08'20"N	21°29'47" N	20°48'03"N	21°39'44" N	21°51'26"N	21°41'05"N
3	Valley area	km ²	26.000	43.760	51.700	332	1.929	282
4	Average water rise level	m	295	215	117	945	475	370
5	Total capacity	10 ⁹ m ³	1,215	9,260	4,871	0,154	2,13	0,18
6	Type of dams		Gravity concrete	Gravity concrete	Soil-Rock mixtures	Vault/dome Gravity concrete	Roll-Rammer concrete	Gravity concrete
7	Elevation to the top of dam	m	303	228,1	123	952,5	48	374
8	Max. elevation of dam	m	137	137	128	139	132	104
9	Top length of dam	m	611	1,000	734	312,5	424,45	267
10	Capacity of installed machine	MW	1,200	2.400	1,920	210	220	520
11	Number of engines	engine	3	6	8	2	2	2

Characteristics of seismic activity (natural earthquake occurrence) in the ladder zone of Da river hydro-electric plants

Earthquake manifestation

A catalogue of earthquakes in the studied areas was established up to the end of 2010, based on all data sources [11], as follows: 1. The results of an investigation of Paleo-earthquakes; 2. The catalogue of historical earthquakes and surveyed earthquakes; 3. The catalogue of earthquakes of the Institute of Geophysics; 4. The catalogue of earthquakes updated from ISC publications; 5. The catalogue of earthquakes updated from publications by NOAA; 6. The catalogue of earthquakes updated from NEIC publications.

In total, 326 earthquakes have been recorded in the ladder zone of Da river hydro-electric plants (Fig. 1); of these, two are Paleo-earthquakes; 36 occurred before 1976 (these are mainly earthquakes determined according to historical data and survey data with magnitudes ≥ 4.0); and 288 occurred between 1976 and 2010 (with maximum magnitudes ≥ 3.0). Greater attention is paid to the strong earthquakes, with magnitudes of $M=6.7-6.8$, which have been found to have

occurred within the study area [8, 9]. These earthquakes are:

1. The Tuan Giao Paleo-earthquake (with a definite age of around 420 years ago).
2. The Binh Lu Paleo-earthquake (with a definite age of around 480 years ago).
3. The Tuan Giao earthquake, which occurred at 14:18 (Hanoi time) on 24 June, 1983 in a mountain area about 11 km north of Tuan Giao town. The magnitude of this earthquake was determined to be $M=6.7\pm 0.2$. Maximum seismic intensity in the epicentral area was $I_0=8-9$ (MSK64 scale). The source parameters have been determined as follows: the length $L=33.1$ km; the width $W=21.2$ km; the source area $S=21.2\times 33.1=686.2$ km²; and the focal depth $20-3=17$ km.

Some manifestation rules of seismic activities

Rule of earthquake recurrence (Gutenberg-Richter law):
The Gutenberg-Richter law of earthquake recurrence for the studied areas is presented in Fig. 2, with the b value thus given as 0.935, for which the function has a form of $\text{Log}[N(M)/T]=-0.935M+3.923$. Data for the earthquakes used in computing the function of distribution is the catalogue of earthquakes for the studied area from 1976 to 2010. According to Cao Dinh Trieu [9], earthquake data recorded during this period is sufficient the magnitude of completeness (M_0) of earthquake catalogue in the northwest region has a value of $M_0=3.5$. If we assume that the repeated cycle of the Tuan Giao and Binh Lu earthquakes is 420 and 480 years, respectively, then it is likely these two Paleo-earthquakes will also have had a similar magnitude to the Tuan Giao earthquake which occurred in 1983 ($M=6.7$), with

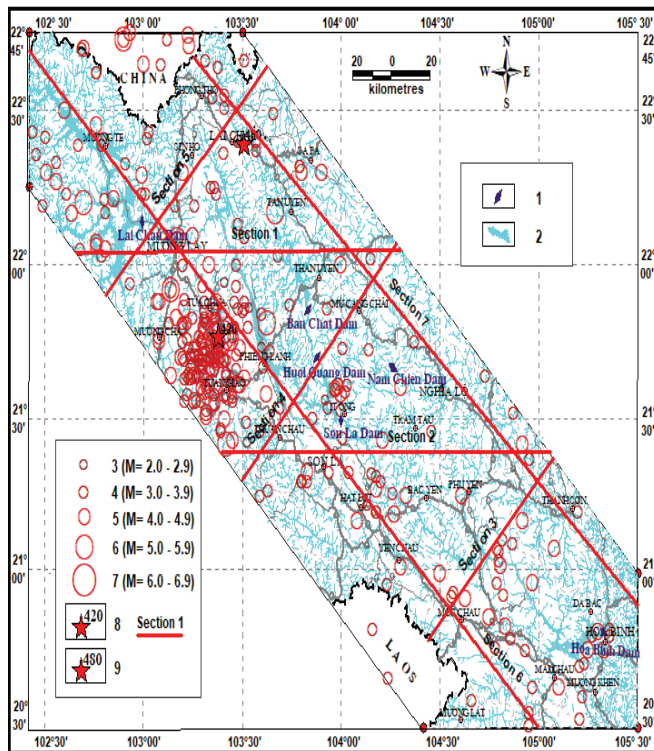


Fig. 1. Map of earthquake epicentres in the ladder valley of Da river hydro-electric plants and analysis of lines of energy generated. 1. Hydro-electric dam; 2. Valley of hydro-electric plant; 3-7. Earthquake epicentres and magnitudes; 8. Tuan Giao Paleo-earthquake and its definite age; 9. Phong Tho Paleo-earthquake and its definite age.

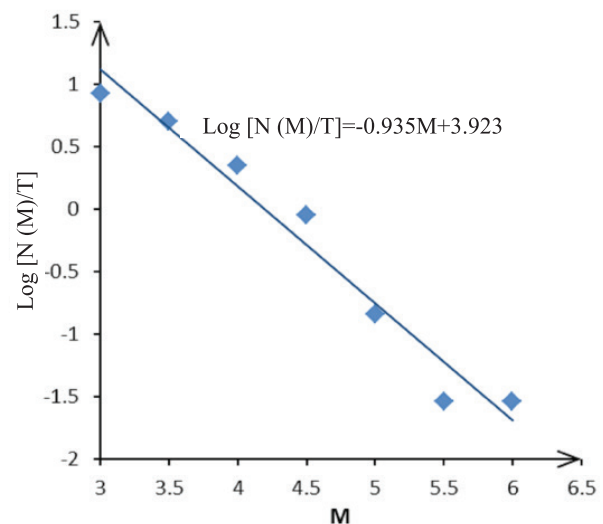


Fig. 2. Graph of the Gutenberg-Richter law for Da river ladder valley hydro-electric plants.

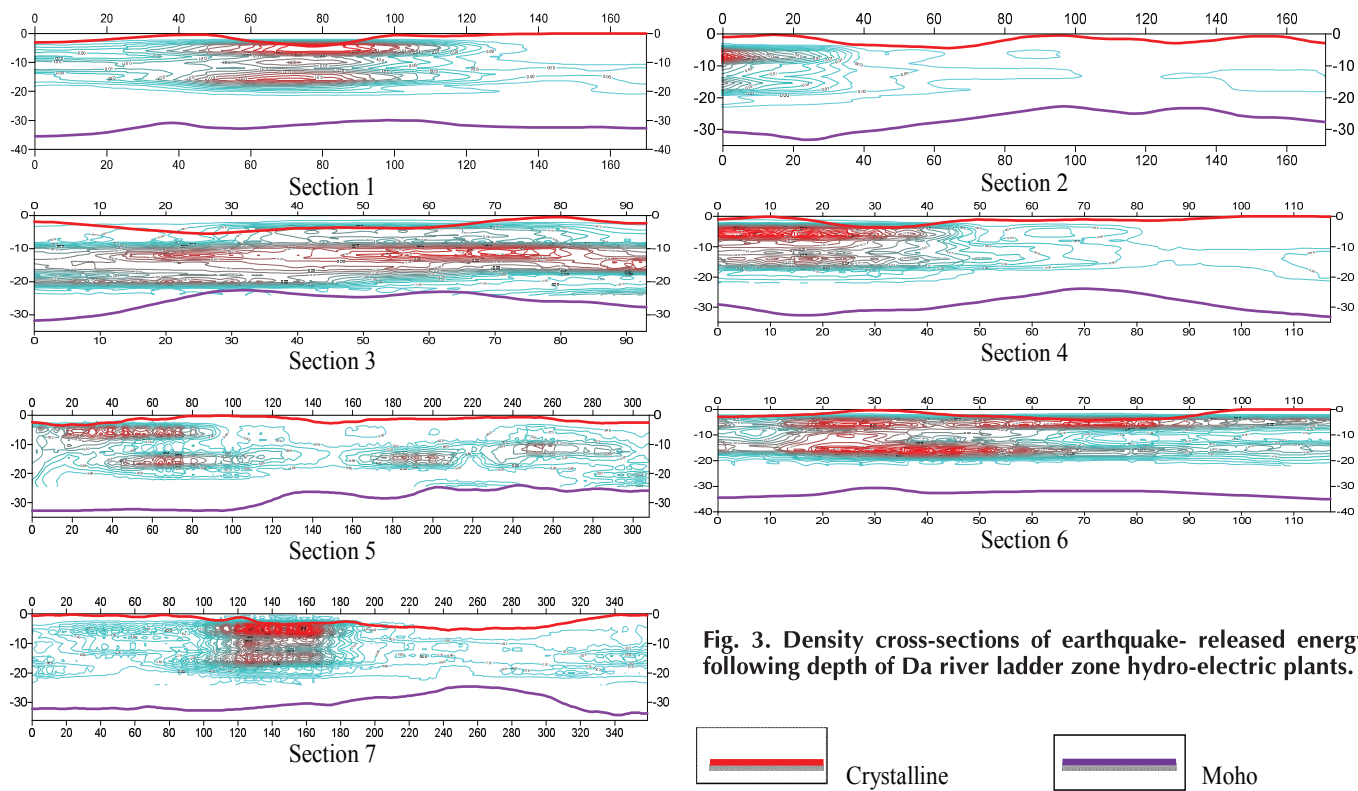
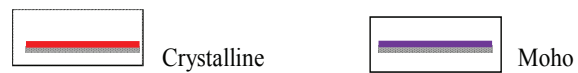


Fig. 3. Density cross-sections of earthquake- released energy following depth of Da river ladder zone hydro-electric plants.



a b value of 0.935. The determined repeated cycle of Tuan Giao will therefore be 418 years, which is approximately the age of the Tuan Giao Paleo-earthquake [8, 9].

Earthquake generating layer: the earthquake generating layer of the ladder zone of Da river hydro-electric plants is established based on determination of the density of energy released through seismic activity [11]:

$$\text{Log} E = 1.5M + 4.2 \quad (E \text{ is expressed in Joules}) \quad (1)$$

A computation of released energy was carried out for seven sections (Fig. 1) with varying hypocentral depths of 2, 4, and 6 km, respectively (the sections being 2 km apart) up to a depth of 24 km (Fig. 3). Based on the results for the above-mentioned computed sections, upper and lower boundaries of earthquake-generating layers, coinciding with zero levels of released energy, were established. This allowed for the building of a map of the upper and lower boundaries and the thickness of earthquake-generating layers in the study area, as presented in Fig. 4. The results show that the thickness of earthquake-generating layers varies in the range of 9 km to 18.5 km; this means that earthquake could be observed should have a minimum hypocentral depth of around 2-3 km, and a maximum hypocentral depth of around 20-22 km.

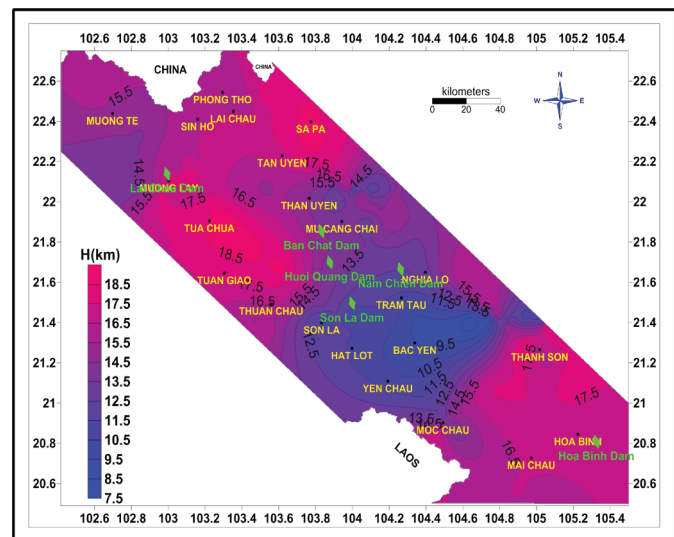


Fig. 4. Thickness of earthquake-generating layers in ladder zone of Da river hydro-electric plants.

Earthquake-generating sources

Earthquake-generating sources in this study are the line-sources [8, 11, 12]. Procedures for terminating sources and source characteristics have been applied in the following order:

1. The determination and classification of faults based on geological and geophysical data;

2. Recognition of active faults on the basis of identification signs.

3. Active faults capable of generating earthquakes with a magnitude $M \geq M_0$ ($M_0=3.5$ for Northwest zone) being determined as earthquake-generating sources.

4. Earthquake-generating segmentation based on gravitational and magnetic fields, the structure of the Earth's crust, and signs of new tectonics and geomorphology.

5. Width of sources based on the formulas: $W \text{ (km)} = H \cdot \tan \alpha + w$ and $\log W \text{ (km)} = 0.25M_{\max} - 0.35$; where: $W \text{ (km)}$ is the width of the seismic source and $M_{\max}$ is the earthquake magnitude, H : the fault depth, α : the dip angle of the fault, w : the zone of dynamic faulting.

Based on the described methodology, there are 13 earthquake seismic-generating sources in the ladder zone of Da river hydro-electric plants which have been discovered by the authors (Fig. 5). These are: Muong Te, Nam Nhe, Muong Nhe, Lai Chau - Dien Bien, Son La, Tuan Giao, Phong Tho, Muong La - Bac Yen, Song Da (Da river), Nam Tong, Mu Cang Chai, Lao Cai - Ninh Binh (Fansipan), and Mai Chau. Numerous segments of these sources are connected to hydro-electric plant reservoirs, giving rise to a high risk of triggered earthquakes that warrants increased attention (Fig. 6).

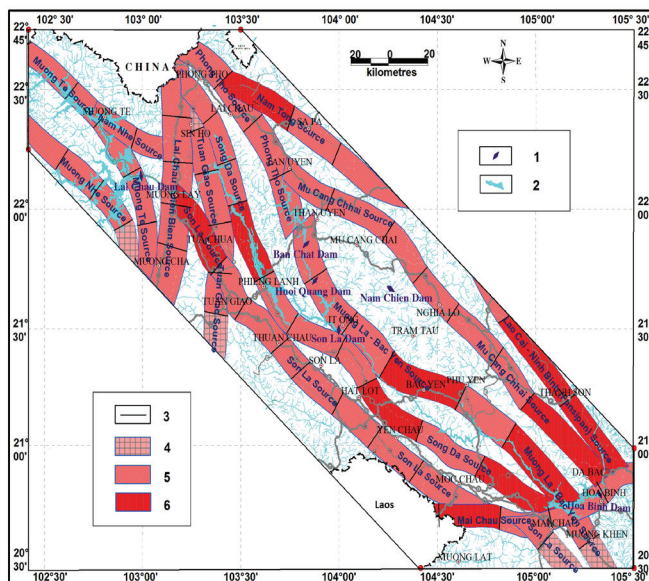


Fig. 5. Earthquake-generating sources and predictions of $M_{\max}$ following faulting segmentation of ladder zone of Da river hydro-electric plants. 1. Hydro-electric dam; 2. Reservoir area; 3. Source segmentations; 4. Source segment for earthquake prediction with magnitude of $M_{\max}=5.5-5.9$; 5. Earthquake-forecasting segments with $M_{\max}=6.0-6.4$; and 6. Earthquake-forecasting segments with $M_{\max}=6.5-6.9$.

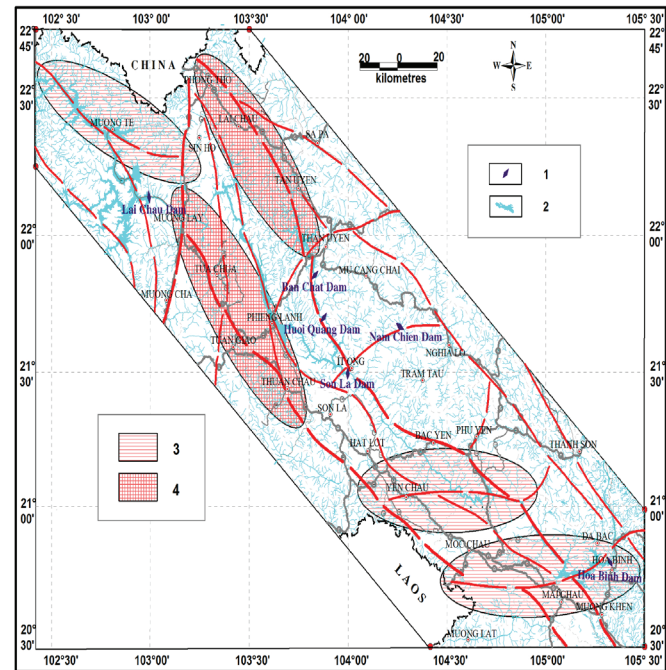


Fig. 6. Areas of high hazard of strong earthquake occurrence ($M_{\max}=6.5-6.9$) in the ladder zone of Da river hydro-electric plants: 1. Hydro-electric dam; 2. Reservoir area; 3-4. Area of strong earthquake risk, of $M_{\max}=6.5-6.9$ (3. Predicted, 4. Occurred).

Hazard of natural earthquake occurrence in the ladder zone of Da river hydro-electric plants

High risk from earthquakes with a maximum magnitude ($M_{\max}$) may be determined in the ladder zone of Da river hydro-electric plants based on faulting segmentation

'Maximum earthquake' is the largest earthquake that can be caused by faulting in a given fixed time or period of time. With the aim of assessing the largest earthquake possible in the area under study, the empirical formula and criteria for fault segmentation established by Cao Dinh Trieu [12] are used:

$$\log L \text{ (km)} = 0.6M_{\max} - 2.5 \quad (2)$$

where $L \text{ (km)}$ is the fault length and $M_{\max}$ is the maximum earthquake magnitude.

The obtained results for natural earthquake occurrence show that the largest earthquake likely along the sources in the ladder zone of Da river hydro-electric plants would have a maximum magnitude of $M=6.9$ (Fig. 5).

Determination of maximum earthquake magnitude ($M_{\max}$) based on an Artificial Neural Network algorithm

The results for prediction of earthquake $M_{\max}$ based on an

Artificial Neural Network algorithm indicates [13, 14] there is a risk of a strong earthquake of $M_{\max}=6.0-6.9$ occurring at the Son La generating source. For the remaining sources, this hazard is in the range of $M_{\max}=5.0-5.9$ (Fig. 7).

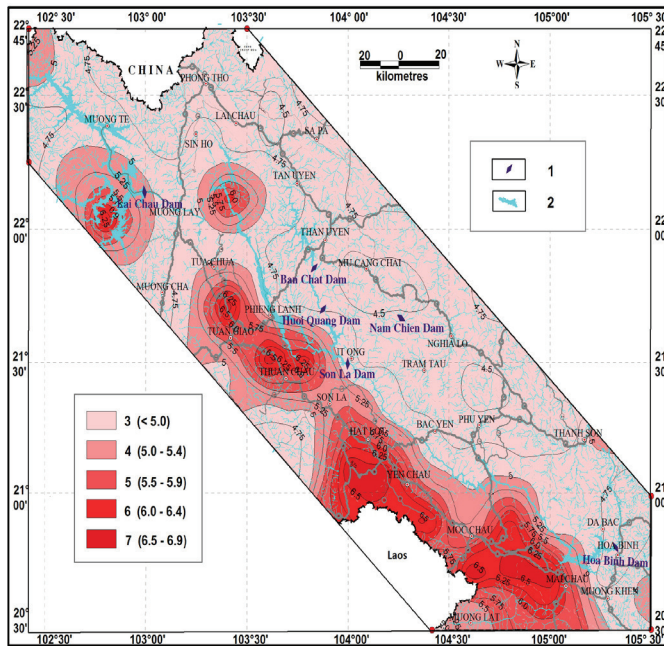


Fig. 7. $M_{\max}$ of the ladder zone of Da river hydro-electric plants determined on the basis of an Artificial Neural Network algorithm. 1. Hydro-electric dam; 2. Reservoir area; 3. $M_{\max}<5.0$; 4. $M_{\max}=5.0-5.4$; 5. $M_{\max}=5.5-5.9$; 6. $M_{\max}=6.0-6.4$; 7. $M_{\max}=6.5-6.9$.

Localization of strong earthquake occurrence based on dependent rule between hypocentral distance and magnitude, and knots of the active faults

The dependent rule between hypocentral distance (δ_m) and magnitude (M) was considered by using 120 earthquakes of $M \geq 4.5$, which have occurred in northern Vietnamese territory. Distances between earthquakes of the same magnitude (M) within a variation of five units ($M=4.5-4.9$; $5.0-5.4$; $5.5-5.9$; $6.0-6.4$, and $6.5-6.9$) were measured, selected, and counted, and their distribution diagrams established. An average distance and error of determination of distance value was then established (Table 2). Next, the function of the average distance distribution between hypocenters of identical level, according to magnitude (M), was constructed.

The empirical function expressing the dependence between hypocentral distance δ_m and magnitude (M) has the form:

$$\lg \delta_m = 0.6M_s - 1.976 \quad (3)$$

Table 2. Distribution of average earthquake hypocentral distance according to magnitude (M).

No	Strength level (M)	Average hypocentral distance (km)	Error (km)
1	4.5-4.9	7.5	± 3
2	5.0-5.4	15	± 6
3	5.5-5.9	30	± 10
4	6.0-6.4	60	± 14
5	6.5-6.9	105	± 18

Given the location of Paleo-earthquakes and Tuan Giao's earthquake of $M=6.7$ (1983), and the relationship between strength of earthquake and hypocentral distance, fault junction and strong earthquake location, five areas are determined to carry a particular risk of strong earthquake occurrence: Muong Te, Lai Chau, Tuan Giao, Yen Chau, and Mai Chau (Table 3). Of these, strong earthquakes have occurred in the areas of Lai Chau and Tuan Giao, and earthquakes of magnitude $M=4.5-4.9$ have occurred in Muong Te, Yen Chau, and Mai Chau, which have not reached the predicted maximum magnitude ($M_{\max}=6.5-6.9$). Hence, the likelihood of strong earthquakes in these areas in the future is very high.

Table 3. Areas which have high risk of earthquake occurrence of magnitude (M) 6.5-6.9.

No	Name of area	$M_{\max}$ observed	$M_{\max}$ forecast
1	Muong Te	4.5-4.9	6.5-6.9
2	Lai Chau	6.5-6.9	6.5-6.9
3	Tuan Giao	6.5-6.9	6.5-6.9
4	Yen Chau	4.5-4.9	6.5-6.9
5	Mai Chau	4.5-4.9	6.5-6.9

Manifestation of triggered earthquake activity in the ladder zone of Da river hydro-electric plants

Earthquake occurrence related to human activity is known as "triggered" activity. A number of human activities can lead to the occurrence of triggered earthquakes, amongst which can be listed: large explosions in the ground, the impoundment of reservoirs (high irrigation dams, hydro-electric dams, etc.), the pumping of liquid into deep rock layers, the discharging of water from water-bearing

formations on the ground or close to the ground surface, and pit exploitation. Triggered earthquakes occurred at the Hoa Binh and Son La hydro-electric plants following reservoir impoundment [15, 16].

Occurrence of triggered earthquake at Hoa Binh hydro-electric reservoir in the period 1989-1991

In May 1988, water impoundment at Hoa Binh hydro-electric reservoir began. Initially, the water level grew gradually, but this accelerated at the onset of the rainy season in July and August of that year. In December 1988, when the water level had reached an elevation of over 80 m, a number of weak earthquakes ($M < 2.0$) were observed/recorded near the dam area. From January to April 1989, a number of earthquakes with a magnitude of $M = 2.4 - 2.8$ were recorded near the dam. On one particular day, 14 April 1989, there were seven earthquakes, of which two were $M = 3.8$ and 3.7 , with hypocentres of around 5 km. After a quiet few weeks, on 23 May an earthquake of a magnitude of 4.9 ($M = 4.9$) occurred, which caused ground-shaking in the Hoa Binh region of an intensity of 7 degrees (in the MSK-64 scale). After this event, weaker earthquakes continuously occurred. On 27 May, another earthquake occurred, with a magnitude of $M = 4.0$, which made the ground shake to a degree of 5-6 in Hoa Binh town. Since then, there has been frequent earthquake activity; whilst relatively weak, it is picked up by seismographs [16] (Fig. 1).

The earthquake of $M = 4.9$ which occurred on 23 May 1989 in Hoa Binh has been specified as a triggered earthquake linked to the water reservoir. It occurred when the impoundment of the water reservoir had been complete for a certain time (it had been eight months since the date of full water impoundment). Its hypocentral depth was shallow, at around 5 km; prior to the occurrence of the main shock, a fore-shock appeared, and the magnitude of the after-shock matched that of the main shock. A gradual decline of seismic activity has been observed as well as high activity when the water levels are suddenly increased. The strongest measurement of the magnitude of the Hoa Binh triggered earthquake was $M_{\max} = 4.9$.

The Hoa Binh triggered earthquake had a Gutenberg-Richter reading as follows [15]:

$$\log \frac{N(M)}{T} = 4.688 - 1.2474M \quad (4)$$

Triggered earthquake of Son La hydro-electric reservoir

The study of the triggered earthquake of the Son La hydro-electric reservoir was conducted through access to the

independent governmental Science and Technology project code DTDL.2009T/09 for the period 2009-2011 [16]. A temporary system of earthquake observation stations (six in total) was installed at the beginning of 2009, and operated until the end of 2012, to monitor earthquake activity around the water reservoir at the initial impoundment. During this period, 400 earthquakes, of a magnitude of $M = 1.0 - 3.0$, and hypocentral depths of less than 10 km, were observed. These earthquakes were mainly concentrated in the submerged areas of the dam within the Muong La - Bac Yen and Da river (Song Da) faults. The largest earthquake in the bed of the Son La hydro-electric reservoir, measuring $M = 4.3$, was recorded on 19 July 2014.

The Gutenberg-Richter law as applied to the triggered earthquakes of the Son La hydro-electric reservoir is as follows (Fig. 8) [15]:

1. For the group of triggered earthquakes from Huoi Quang lowlands to Son La lowlands:

$$\log \frac{N(M)}{T} = 3.6603 - 1.2008M \quad (5)$$

2. For the group of earthquakes from the Pa Uon bridge to Son La dam:

$$\log \frac{N(M)}{T} = 2.9917 - 1.0716M \quad (6)$$

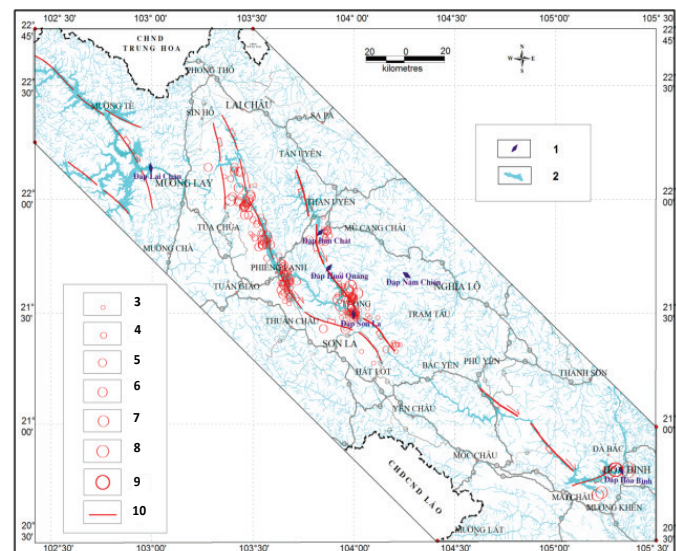


Fig. 8. Distribution of the epicenters triggered earthquakes of Hoa Binh and Son La hydro-electric reservoirs. 1. Hydro-electric dam; 2. Reservoir areas; 3-9. Epicenters and corresponding magnitude ($M = 1.0 - 1.4$; $M = 1.5 - 1.9$; $M = 2.0 - 2.4$; $M = 2.5 - 2.9$; $M = 3.0 - 3.4$; $M = 3.5 - 3.9$; $M = 4.0 - 4.4$; $M = 4.5 - 4.9$); 10. Source segment capable of generating triggered earthquake.

Geological features of Hoa Binh and Son La hydro-electric reservoirs

The triggered earthquakes of Hoa Binh and Son La hydro-electric reservoirs were mainly concentrated in the deep fault zones interconnecting with the reservoirs. The hard limestones of Dong Giao formation ($T_2a\ dg$) were broken down and became water-logged due to compression under high water columns, creating a change in stress pressures around the onset of the triggered earthquakes (Photo 1). In the below cross-section of Dong Giao limestone formation, the structure consists mainly of a strip/ribbon/belt form, of smooth and tight grain; glaze form or creak grain form of light colour (white, yellowish, pink, greenish, brown white); and also black bitumen limestone and silica limestone of bright colour. In the middle of the cross-section, mainly limestone of blocked form is evident [15].



Photo 1. Limestone of Dong Giao formation covering the reservoir bed of Hoa Binh and Son La hydro-electric reservoirs.

Prediction for the maximum triggered earthquake

For prediction of the maximum magnitude of a triggered earthquake ($M_{\max,te}$), the relationship between $M_{\max}$ and b values for natural and triggered earthquake distribution [15] were used:

$$b_{ne} \times M_{\max,ne} = b_{te} \times M_{\max,te} \quad (7)$$

where: b_{ne} is the b coefficient value of natural earthquakes; b_{te} is the b coefficient value of triggered earthquakes; $M_{\max,ne}$ is the maximum natural earthquake to have occurred

in the determined source; and $M_{\max,te}$ is the maximum triggered earthquake predicted to occur in the source if this is interconnected with the water reservoir and conditions for a triggered earthquake being generated are satisfied.

The procedure for determining sources and maximum triggered earthquake was conducted as follows [17, 18]:

Step 1 - Determination of triggered earthquake sources:

Determination of fault segments interconnecting with hydro-electric reservoirs following reservoir impoundment was carried out based on the generating sources and the maximum magnitude of natural earthquakes to have occurred. The fault segments, where earthquakes occurred following reservoir impoundment, are here considered as sources of triggered earthquakes.

Step 2 - Forecasting maximum value of triggered earthquakes based on formula (7):

- Determination of the b_{ne} value of the Gutenberg-Richter distribution function for natural earthquakes in the research areas of the hydro-electric reservoirs and their adjacent areas was conducted. The results show: a value of $b_{ne}=0.6041$ [15] for Song Tranh 2 (Tranh river 2); and $b_{ne}=0.935$ for the ladder zone Da river hydro-electric plants (including Hoa Binh, Son La, Lai Chau, Ban Chat, and Huoi Quang hydro-electric plants - Fig. 4) [15].

- a value of $b_{te}=1.2474$ for Hoa Binh triggered earthquakes (4); and for Song Tranh 2 hydro-electric reservoir, $b_{te}=0.8317$ [8]; while in the Son La reservoir two source segments are available (Fig. 8). For the formulas (5) and (6): from the lowlands of Huoi Quang to the lowlands of Son La, the value is $b_{te}=1.2008$; and from Pa Uon bridge to Son La dam, it is $b_{te}=1.0716$.

- The estimation of maximum magnitude for natural earthquakes ($M_{\max,ne}$) for the triggered earthquake-generating sources based on segmentation of faults shows that [12]: in Hoa Binh, $M_{\max,ne}=6.5$; from the lowlands of Huoi Quang to the lowlands of Son La, it is $M_{\max,ne}=6.5$; and from Pa Uon Bridge to Son La dam, it is $M_{\max,ne}=6.7$.

- The predicted maximum magnitude for triggered earthquakes ($M_{\max,te}$) for the Hoa Binh reservoir is $M_{\max,te}=4.9$ (Fig. 9); from the lowlands of Huoi Quang to the lowlands of Son La, it is $M_{\max,kt}=4.9$; and from Pa Uon Bridge to Son La dam, it is $M_{\max,kt}=5.0$ (Fig. 10).

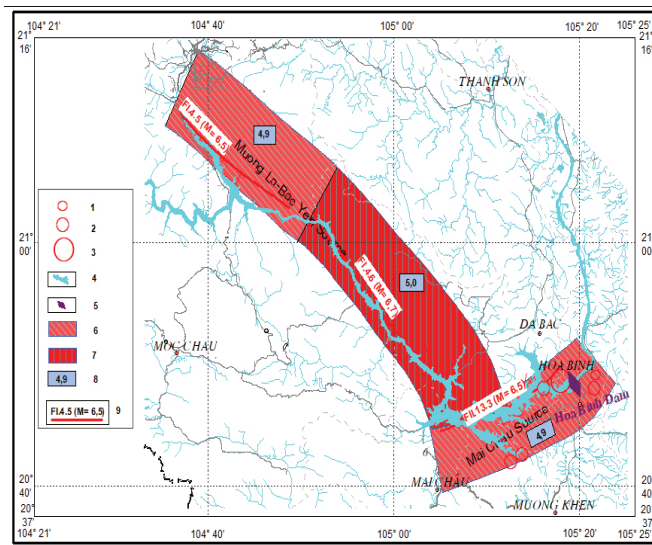


Fig. 9. Source area and maximum value of triggered earthquake for Hoa Binh hydro-electric reservoir. 1. Epicentres of triggered earthquakes of magnitude $M \leq 3.0$; 2. $M=3.7-3.9$; 3. $M=4.0-4.9$; 4. Reservoir area; 5. Hoa Binh hydro-electric dam; 6. Source segment forecasting of maximum value of natural earthquake, $M_{\max}=6.5$; 7. Source segment forecasting of maximum value of natural earthquake occurrence, $M_{\max}=6.7$; 8. Maximum value of triggered earthquake risk in the determined source segment; 9. Symbols of source, source segment, and maximum value of natural earthquake (FI.4 is Muong La - Bac Yen source, 5 is the fifth segment top-down; FII.13 is Mai Chau source, 3 is the third segment).

Conclusions

Based on a comprehensive assessment of the characteristics of seismic activity in the ladder zone of the Da river hydro-electric plants, the following can be asserted:

1. The largest natural earthquakes to have occurred in the ladder zone of Da river hydro-electric plants are the Lai Chau Paleo-earthquake (480 years ago), Tuan Giao Paleo-earthquake (420 years ago), and Tuan Giao earthquake, which occurred in 1983. The magnitudes of these two Paleo-earthquakes are equivalent to the Tuan Giao earthquake of 1983, i.e. $M=6.7 \pm 0.2$.

2. The b value of natural earthquakes in the ladder zone of Da river hydro-electric plants is 0.935. The measurement reflects the level of natural earthquake activity in the research area and also the measurement needed for computing and delimiting natural and triggered earthquake activity around the reservoirs following full water impoundment.

3. 13 earthquake-generating sources in the ladder zone of Da river hydro-electric plants were discovered: Muong Te, Nam Nhe, Muong Nhe, Lai Chau - Dien Bien, Son La, Tuan Giao, Phong Tho, Muong La - Bac Yen, Song Da (Da river),

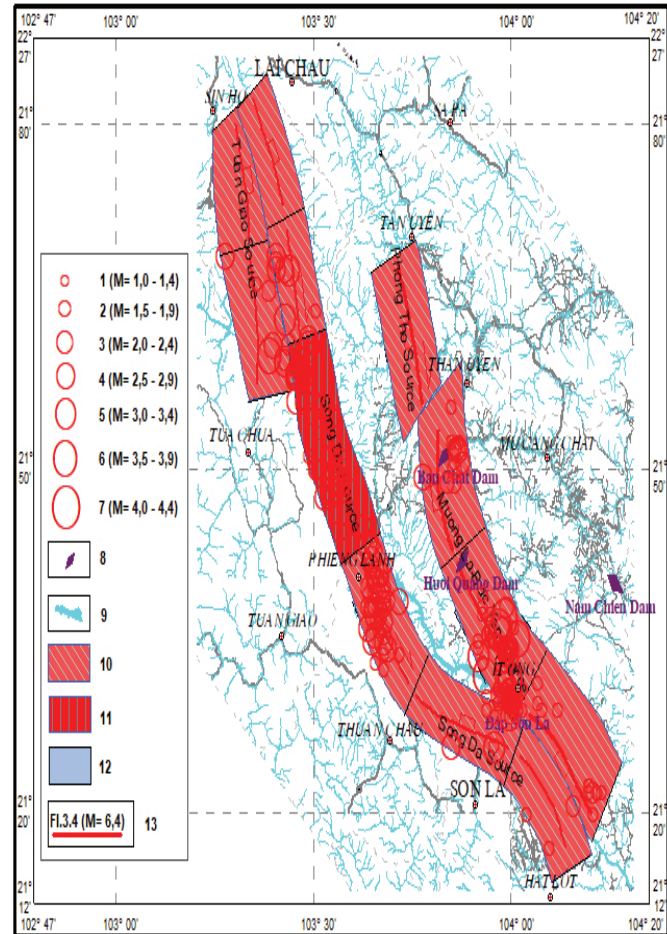


Fig. 10. Source area and maximum value of triggered earthquake for Son La hydro-electric reservoir. 1. Epicentre of triggered earthquake with values of a magnitude of $M=1.0-1.4$; 2. $M=1.5-1.9$; 3. $M=2.0-2.4$; 4. $M=2.4-2.9$; 5. $M=3.0-3.4$; 6. $M=3.4-3.9$; 7. $M=4.0-4.4$; 8. Hydro-electric location; 9. Reservoir areas; 10. Source segment forecasting maximum value of natural earthquake, $M_{\max}=6.0-6.4$; 11. Source segment forecasting maximum value of natural earthquake, $M_{\max}=6.5-6.9$; 12. Maximum value of triggered earthquake risk in the determined source segment; 13. Symbols of source, source segment, and maximum value of natural earthquake (FI.3 is Phong Tho source; 4 I is the fourth segment top-down; FI.4 is Muong La - Bac Yen source; 1,2,3 are the segments of 1,2,3 top-down; FII.4 is Tuan Giao source and FII.5 is Da river source).

Nam Tong, Mu Cang Chai, Lao Cai - Ninh Binh (Fansipan), and Mai Chau. The largest earthquake which might occur within these earthquake-generating sources would be of a magnitude not exceeding 6.9.

4. Large natural earthquakes with an $M_{\max}$ of up to 6.9 are predicted to occur in the following areas: Muong Te, Lai Chau, Tuan Giao, Yen Chau, and Mai Chau. Earthquakes of a maximum magnitude of $M=6.7$ have been observed in the Lai Chau and Tuan Giao areas; this maximum value has

to date not been observed in the areas of Muong Te, Yen Chau, or Mai Chau. If the distance regulation, degree of magnitude, and junction points of the earthquake-generating faults are correct, then there is a notable risk of imminent earthquake occurrence in the Muong Te, Yen Chau, and Mai Chau areas.

5. Triggered earthquakes of medium and above-medium magnitudes ($M \geq 4.0$) have occurred at the Hoa Binh and Son La hydro-electric reservoirs, where the maximum depth to the bottom is over 100 m, and the geological structure of the reservoir bed consists of broken down limestone. When reservoir water accumulates, endomose phenomena in the breakdown tectonical zones caused by water column pressure, change the pressure stress in the direction of previous triggered earthquakes.

6. The largest observed triggered earthquake at the Hoa Binh reservoir was of a magnitude of $M=4.9$. This maximum value is also predicted for triggered earthquakes. This means that earthquake activity in the Hoa Binh reservoir reached the maximum value. No likelihood of triggered earthquake occurrence at the Hoa Binh hydro-electric plant is predicted, since no earthquake has been observed at the Hoa Binh reservoir since 1996. At the Son La reservoir, triggered earthquakes may continue to occur in the future, with a magnitude exceeding 4.3, and possibly reaching a maximum of 5.0.

ACKNOWLEDGEMENTS

The research team would like to convey their sincere thanks to the Ministry of Science and Technology for their support in providing funds for the implementing of the research theme of Governmental Science and Technology Project coding- DTDLCN.27/15.

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] H.K. Gupta (2011), *Encyclopedia of Solid Earth Geophysics*, Springer, Netherlands.
- [2] Nguyen Dinh Xuyen, Nguyen Thanh Tung, et al. (1996), "Database for mitigative solution of risk caused by earthquake in Vietnam", *The summary report of independent theme of National research programme code KT-DL92-07 archive of the Institute of Geophysics* (in Vietnamese).
- [3] Nguyen Ngoc Thuy (2008), "Study of modern tectonic and earthquake in order to assessment stable operation of the Hoa Binh hydro-electric", *The summary report of independent theme of National research programme code DTDL.2005/19G archive of the Institute of Geophysics* (in Vietnamese).
- [4] Cao Dinh Trieu, C.D. Trong, L.V. Dung, T.A. Tuan, and D.Q. Van, and H.V. Long (2014), "Triggered earthquake study in Thanh River no. 2 (Vietnam), hydro-electric reservoir", *Journal Geological Society of India*, **84**, pp.319-325.
- [5] Gahalaut Kalpna, Thai Anh Tuan, and N. Purnachandra Rao (2016), "Rapid and delayed earthquake triggering by the Song Tranh 2 reservoir, Vietnam", *Bull. Seismol. Soc. Am.*, **106**(5), pp.2389-2394, doi:10.1785/0120160106.
- [6] Nguyen Van Giang, J. Wiszniowski, B. Plesiewicz, G. Lizurek, D.Q. Van, and L.Q. Khoi (2015), "Some characteristics of seismic activity in the Song Tranh 2 Reservoir, Quang Nam, Vietnam by local seismic network data", *Earth Sciences*, **4**, pp.101-111, doi:10.11648/j.earth.20150403.13.
- [7] Thai Anh Tuan, N. Purnachandra Rao, Kalpna Gahalaut, Cao Dinh Trong, Le Van Dung, Cao Chien, K. Mallika (2017), "Evidence that earthquakes have been triggered by reservoir in the Song Tranh 2 region, Vietnam", *Journal of Seismology*, **21**(5), pp.1131-1143.
- [8] Cao Dinh Trieu (2002), "Some characteristic features of seismotectonic conditions and seismic regime of the Tuan Giao earthquake area", *Journal of Geology, Series B*, **19-20**(9-10), pp.54-68 (in Vietnamese).
- [9] Cao Dinh Trieu (2010), *Earthquake hazards in Vietnam*, Publishing House for Science and Technology, Hanoi, 304pp.
- [10] Cao Dinh Trieu, Ngo Gia Thang, Mai Xuan Bach, Pham Nam Hung, Bui Anh Nam (2010a), "Bac Yen and Mai Son earthquakes occurred in 26 November 2009", *Journal of Geology, Series A*, **320**(9-10), Hanoi, pp.241-252 (in Vietnamese).
- [11] Cao Dinh Trieu, Le Van Dung, Thai Anh Tuan (2010b), "Risk of earthquake in the Northwest region and its adjacent", *Journal of Geology, Series A*, **320**(9-10), pp.253-262 (in Vietnamese).
- [12] Cao Dinh Trieu, Nguyen Duc Vinh (2012), "Fault segmentation in assessment of maximum earthquake in Vietnam", *Journal of Geology*, **331-332**(8-9), pp.59-68 (in Vietnamese).
- [13] Cao Dinh Trong, Cao Dinh Trieu, Nguyen Duc Vinh (2011), "Application of the Neuron network in prediction magnitude of earthquake in northwest region", *Journal of the Earth Sciences*, **33**(2), pp.51-163 (in Vietnamese).
- [14] M.T. Hagan, H.B. Demuth and M. Beale (1996), *Neural Network Design*, PWS Publishing Company, Boston, MA.
- [15] Cao Dinh Trong, Nguyen Anh Duong, Thai Anh Tuan, Cao Dinh Trieu (2016), "Characteristics of triggered earthquake in ladder Da river hydro-electric plant", *Journal of Geology, Series A*, **361-362**(11-12), pp.80-92 (in Vietnamese).
- [16] Le Tu Son (theme director), et al. (2012), "Study of triggered earthquake forecasting in Son La hydro-electric reservoir", *The summary report of independent theme of National research programme code- DTDL.2009T/09 archive of the Institute of Geophysics*, 271pp (in Vietnamese).
- [17] D.W. Simpson and T.N. Narasimhan (1990), "In homogeneities in rock Properties and their influence on reservoir triggered seismicity", *Gerlands Beitrage zur Geophysik*, **99**, pp.205-219.
- [18] H.K. Gupta, B.K. Rastogi, H. Narain (1972), "Common features of the reservoir associated seismic activities", *Bulletin of the Seismological Society of America*, **62**, pp.481-492.

River information management and early flood release in response to climate change in Vietnam

Tran Dinh Hoa*, Do Hoai Nam

Vietnam Academy for Water Resources

Received 22 October 2018; accepted 11 January 2019

Abstract:

Increasing flood risks in a changing climate tend to put greater pressure on water-related infrastructure, existing operations, and management practices. This paper introduces preliminary research results on river information management and flood-risk reduction based on an early flood-release approach that has the goals of better reservoir operation, adapting to climate change, and ensuring dam safety in Vietnam. Early flood release is performed using inflow prediction information derived from a medium-range global numerical weather-prediction model. The results show that peak discharge and inundation areas are remarkably reduced, and are useful for improving the safety of dams and flood-risk management in downstream areas.

Keywords: early flood release, flood detection, numerical weather prediction, river information.

Classification number: 5.2

Introduction

The trend of increasingly heavy rain in a changing climate will directly affect the management and development of river basins in the future; in particular, it will have a strong impact on the safety of water-related infrastructure such as embankment dams. Most existing dams were designed based on frequency analyses of historical rainfall patterns and extreme events, but this excluded consideration of climate change impacts [1]. Even so, the design criteria do not take into account recent changes in frequency and severity, as described in the assessment reports of the Intergovernmental Panel on Climate Change [2]. For example, in the August 2002, flood water levels in Dresden, Germany, exceeded historical flood levels of the last few centuries [3].

Vietnam is heavily influenced by the tropical monsoon climate that comprises a distinct wet and a dry season. It is hence considered to be a water-abundant country. However, water distribution varies extremely between the wet and dry seasons. Approximately 80% of annual runoff occurs in the wet season. Results from recent studies show that changes in extreme rainfall events seem to be more crucial than changes in the average climate conditions [4]. Short-term precipitation intensities (e.g., the highest precipitation amount in a three-day period and the total precipitation when precipitation is greater than the 95th percentile of precipitation on very wet days) representing risks of flooding (in terms of frequency and scale) are expected to increase in most parts of the country in the near future (2015-2039), with the highest increases to occur in the northeast region and Ho Chi Minh city vicinities. Increasing flood risks tend to put more pressure on water-related infrastructure, existing operations, and management practices. This paper introduces preliminary research results on river information management and flood-risk reduction that are based on an early flood-release approach for ensuring better reservoir operation in adapting to climate change and ensuring

* Corresponding author: Email: tranhoa08@gmail.com

dam safety in terms of a research cooperation agreement between the Vietnam Academy for Water Resources and the Foundation of River Basin Integrated Communications of Japan (FRICS).

Methodology and data

River information management

The objectives of river information management are to provide accurate, reliable, and quick collection, processing, and dissemination of the necessary information to assist with decision-making processes. For example, the amount of rainfall forecast by numerical weather-prediction (NWP) models monitored by weather satellites or radar, and observed at rain stations and as water levels in rivers. All of this information is particularly important to increase the efficiency of river-related infrastructure and to enhance river administration services through the dissemination of information to the public.

Basically, river information management includes three processes: (i) data collection, (ii) data processing, and (iii) data provision.

Data collection: there are a number of data types currently available for the river information management. These consist of point data (e.g., rainfall amounts, water levels, and discharge), area data (e.g., rainfall amounts estimated by weather satellites and radar), and image data (e.g., images recorded by closed-circuit television cameras).

Data processing: the data collected from different sources are processed and analysed to check for any missing measurements or irregular values and to remove any noise from the data.

Data provision: information is to be provided to the public via the internet and cellular phones. This includes information on water levels, rainfall amounts, flood forecasts and warnings, and dam-related parameters (notice of release of water from the reservoir, reservoir water storage, etc.).

Conceptual framework for flood-risk reduction

In light of the distinct water distribution in the wet and dry seasons, thousands of reservoirs have been built across Vietnam. They show the capabilities for flood control; however, not all floods have been entirely avoided because flood control storages are no longer able to accommodate the increasing inflow of flood runoff induced by intensified short-term precipitation. As a result, new reservoir operation rules for flood control have emerged as a vital tool in attempts to reduce flood risk.

The concept of early flood release applied to those

reservoirs with controlled gates is illustrated in Fig. 1. Theoretically, the flood control storage would increase when the reservoir starts releasing water downstream before any incoming flood, while the efficiency of peak discharge cutoff depends on the forecast horizon, the longer lead time, and the greater efficiency of flood-risk reduction. Implementing flood-risk reduction based on the early-release operation approach to increase the volume of flood control storage is considered an appropriate solution to cope with climate change and to ensure dam safety. This approach is widely applied in such developed countries as Japan, the USA, and those in the Europe.

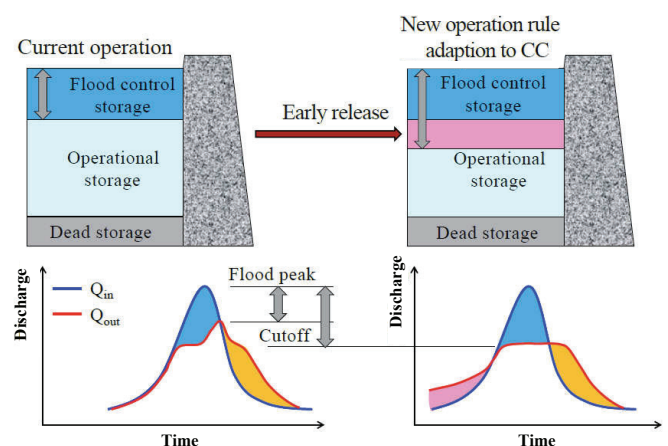


Fig. 1. Concept of reservoir operation for adapting to climate change (CC).

Incoming flow forecast

River flow forecast methods in general vary, depending on rainfall input. Conventional flood predictions based on real-time or near real-time observations of rainfall in the river basins and other hydrologic parameters provide a relatively short lead time because they are dependent on the runoff response of the river basins considered. The lead times for the forecasts are quite short for small and steeply sloped river basins. However, the forecasts show high accuracy.

NWP-based flood forecasts: benefiting from increased computational power, high-resolution NWP models are available to the public and offer a better forecast of rainfall and the forecast horizon. These advanced features allow the generating of short-term forecasts of inflow into the reservoirs. The NWP-based forecasts tend to promote higher efficiency of flood-risk reduction because of the forecast horizon of NWP models, which are either a few days or up to 10 days for the short and medium forecast ranges, respectively. However, there are inherent uncertainties in such forecasts: the longer the forecast horizon, the greater

the uncertainty of the forecast.

In this paper, the medium-range rainfall prediction by a global NWP model operated at the Japan Meteorological Agency (JMA) is used to drive short-term flood forecasts. The NWP model has spatial resolution of 0.5 degrees and 60 vertical layers. In terms of forecast range, the model provides a quantitative estimation of the accumulated rainfall every 6 h. This is issued four times per day at 00:00, 06:00, 12:00, and 18:00 coordinated universal time (UTC) for the lead time of 84 h, and every 12 h (00:00 and 12:00 UTC) for lead time up to 132 h.

Rainfall-runoff inundation model: the Rainfall-runoff inundation (RRI) model, a two-dimensional (2D) model developed by the International Center for Water Hazard and Risk Management (ICARM), Japan, is introduced in this paper. The detailed model structure is documented on the ICARM website (http://www.icharm.pwri.go.jp/research/rri/rri_top.html). In short, the RRI model is capable of simulating rainfall-runoff and flood inundation simultaneously. The RRI model treats spatial hydrological processes and inundation analyses on a grid-cell basis. The 2D diffusive wave model is applied to simulate overland flow on the slope grid cells; while channel flow is routed using the 1D diffusive wave model. Flow exchange between the river channel and slope is calculated using overflowing formula, a function of water-level and levee-height conditions.

Case studies

The Huong and Vu Gia-Thu Bon river basins are large river systems in Central Vietnam. River tributaries begin in the mountains and run through narrow floodplains along the coastline and finally empty into the East Sea of Vietnam. Given the effects of the topography and climate pattern, the Huong and Vu Gia-Thu Bon river basins have more rainfall than other river basins in the region. Thus, these river basins have a higher risk of flooding, especially with large floods. Recent statistics show that large floods are becoming more extreme and more frequent. Most of the large floods of the last 50 years occurred during the 1995-2010 period. This statistic implies that, in a changing climate, there are significant increases in the frequency and severity of floods, which result in the exposure of water infrastructure to high flood levels, especially at the highly vulnerable earth dams in the region.

Results and discussion

In this paper, flood detection and forecast using NWP model outputs are examined and used to perform early

flood release and inundation mapping. In order to reduce inherent uncertainties in the forecasts, it is suggested that a cascading process of the forecasts is implemented. First, flood detection is performed using the medium-range rainfall prediction of a global NWP model. Second, for as long as a flood is detected, detailed flood forecasts are realised using rainfall prediction by the NWP model and observations by weather radar and ground stations across the river basin. Third, optimal reservoir operation is applied and potential downstream inundation areas are analysed by the RRI model. More importantly, the flood forecasts are regularly renewed on a daily basis to improve their reliability for the data assimilation technique that improves the estimation of the model's initial state for the runs that follow.

Flood detection

As a pioneer among such centres, JMA offers the most advanced NWP model outputs and is continually working to improve its products. Figs. 2, 3 show the forecasts by the JMA NWP model for 24-h precipitation accumulation (issued at 7:00 AM) on September 28th and 29th, 2009, in the course of the influence of typhoon Ketsana on the Huong and Vu Gia-Thu Bon river basins. It is interesting to observe that the model is capable of capturing extreme rainfall, approximately 400-500 mm/day, near the centre of the cyclone. This indicates a very high risk of large-scale flooding in the region.

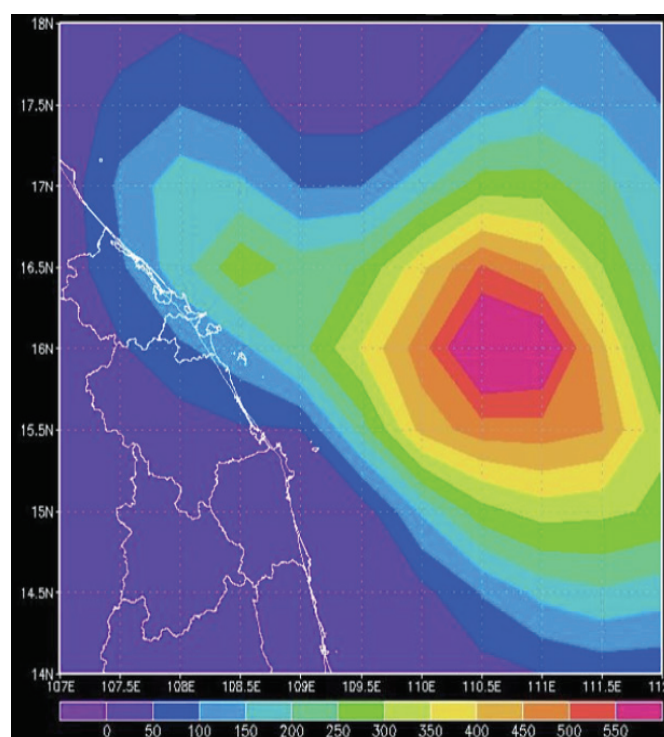


Fig. 2. 24-h precipitation accumulation (mm) forecast issued at 7:00 AM on September 28th, 2009.

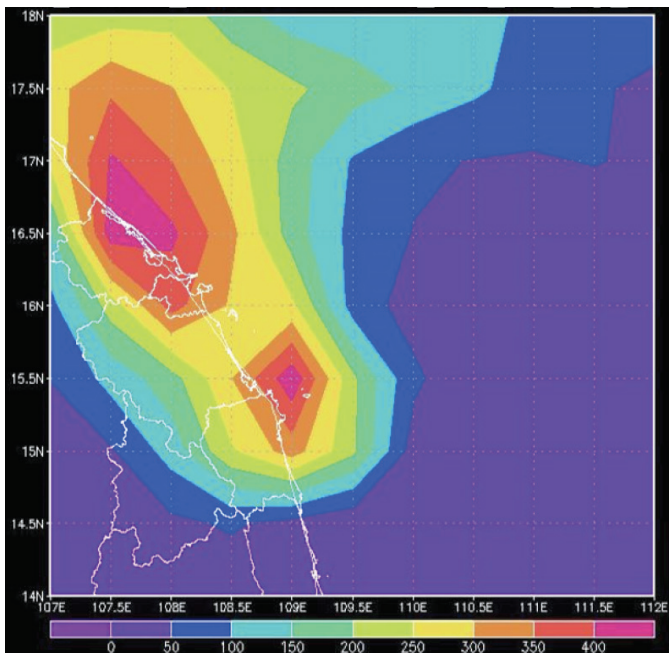


Fig. 3. 24-h precipitation accumulation (mm) forecast issued at 7:00 AM on September 29th, 2009.

Early flood release and inundation mapping

Once potential floods have been identified, appropriate reservoir operation rules can be considered [5]. Early flood release was first applied at A Vuong hydropower dam in the Vu Gia-Thu Bon river basin during the arrival of the typhoon Ketsana [6]. Based on the inflow prediction information, advance reservoir release was performed two days before the peak discharge occurred, as illustrated in Fig. 4. The results show that, compared to the actual situation as applied with the existing operation rule, peak discharge is considerably reduced when the new operation rule is applied. The peak discharge reduction rate is approximately 40%. The reduction rate is, however, likely to increase further as the forecast lead time is extended.

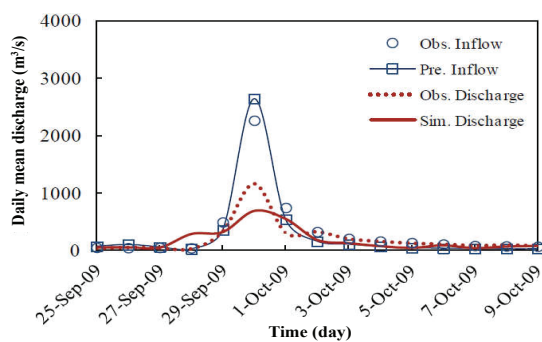


Fig. 4. Application of early flood release at A Vuong hydropower dam during typhoon Ketsana, September 28th - October 2nd, 2009 [7].

A similar procedure was then performed by FRICS at Huong Dien hydropower dam in the Huong river basin, and further analyses of downstream flood inundation were conducted using the RRI model [7]. Figs. 5 and 6 depict the water level and flow information at Huong Dien and Phu Oc stations, respectively. It can easily be seen that the peak discharge was cut significantly, especially when the optimized release was applied. The peak discharge reduction rate is up to 50%.

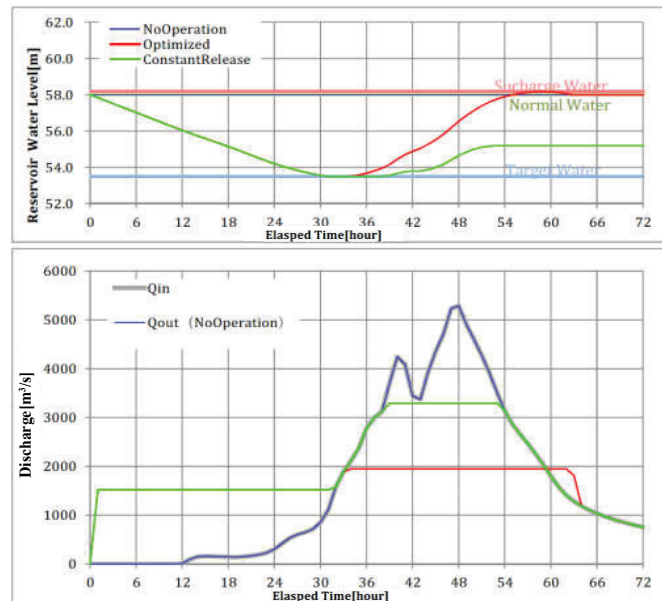


Fig. 5. Reservoir water level and discharge at Huong Dien hydropower dam with the implementation of advance release, two days before the peak discharge [7].

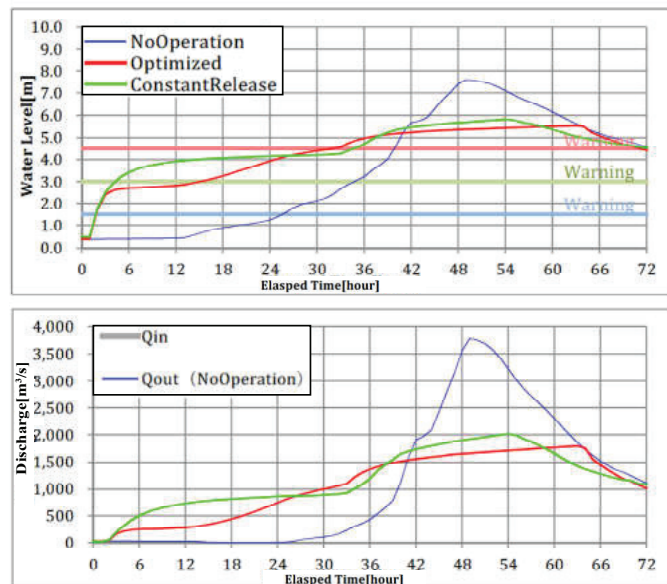


Fig. 6. Water level and discharge at Phu Oc station with (i) normal (no) operation, (ii) optimized release, and (iii) constant release [7].

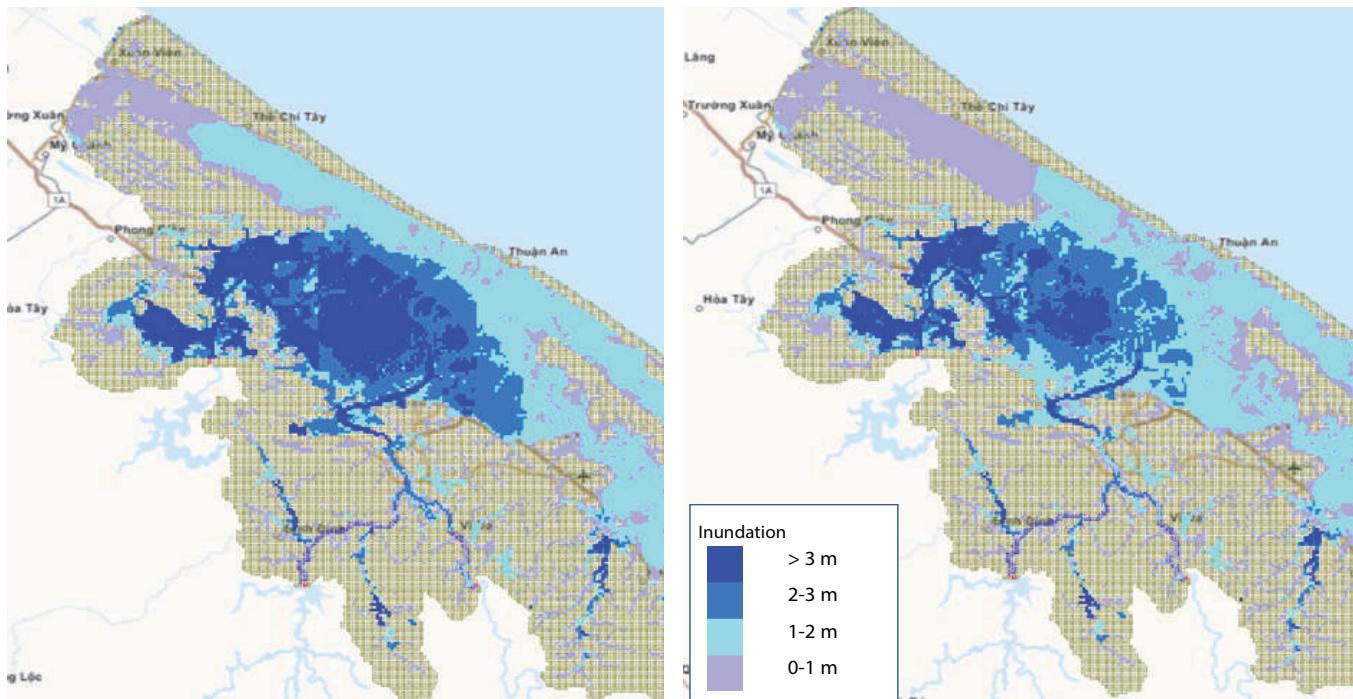


Fig. 7. Inundation risks in downstream areas without (upper) and with advance release (lower) during typhoon Ketsana, Sep. 28th - Oct. 2nd, 2009.

As a result, inundation risks for the downstream areas were analysed and are presented in Fig. 7. The figure depicts the remarkable reduction in the size of the inundation areas and in flooding severity when the advance release was performed using NWP-based inflow prediction. Such an inundation forecast is very useful for the real time operation of reservoirs in the river basins and for the implementation of flood risk-reduction measures in the downstream areas.

Conclusions and remarks

A new model of river information management and a flood risk-reduction approach in response to climate change has been introduced to promote informed decision-making about flood-risk reduction. The new model is crucially important for improving the safety of dams and flood risk management in downstream areas in that it allows the making of more informed decisions and the controlling of the timing of the storing and discharging of water from the dams.

ACKNOWLEDGEMENTS

This work belongs to the Research Cooperation Agreement between the Vietnam Academy for Water Resources (VAWR) and the Foundation of River & Basin Integrated Communications, Japan (FRICS).

This work was financially supported by Vietnam's National Foundation for Science and Technology Development (NAFOSTED) for a basic research project

(Code: 105.08-2014.23).

The authors declare that there is no conflict of interest regarding the publication of this article.

REFERENCES

- [1] WTO (2009), *Guidelines on analysis of extreme in a changing climate in support of informed decisions for adaptation*.
- [2] Intergovernmental Panel on Climate Change (2007), *The physical science basis*, Cambridge University Press, Cambridge, UK and New York, USA.
- [3] U. Ulbrich, et al. (2003), "The central European floods of August 2002: part 1 - rainfall periods and flood development", *Weather*, **58**(10), pp.371-377.
- [4] Ministry of Agriculture and Rural Development (2015), *Near future changes in extreme rainfall over vietnam projected by CMIP5 high-resolution climate models*, Vietnam-Japan Workshop on Estuaries, Coasts and Rivers Hoi An, Vietnam.
- [5] T. Kojiri, S. Ikebuchi, H. Yamada (1989), "Basinwide flood control system by combining prediction and reservoir operation", *Stochastic Hydrology and Hydraulics*, **3**(1), pp.31-49.
- [6] D.H. Nam, K. Udo, A. Mano (2012), "Inflow forecast using downscaled rainfall from global NWP for real-time flood control", *Journal of Japan Society of Civil Engineers, series B1 (Hydraulic Engineering)*, **68**(4), pp.181-186.
- [7] A. Terakawa, et al. (2014), "A development of the prototype of flood management information system for the Huong river basin, Vietnam", *Kasen Jouhou Shinpojuumu Kouen Shuu* (Lectures Collection of Symposium on River and Basin Integrated Communication).

Vietnam Journal of Science, Technology and Engineering (VJSTE): Manuscript Template

General Instructions on using this template and submitting a manuscript to VJSTE: Thank you for preparing a manuscript for submission to VJSTE. Using this template, or following the guidelines below, will help us in processing your paper. Our goal is to be able to identify each section of your manuscript so that we can accurately record the title, authors, abstract, etc. and also be able to enrich it by including reference links and an accurate layout.

You can submit your paper at <http://vietnamscience.vn>

So that we can identify the parts of your paper, and even if you do not use our template, please begin each section with the specific key words listed below. Please do not use paragraph breaks in the title, author list, or abstract.

Title

Authors

Affiliations

Abstract

One Sentence Summary

Main Text

Acknowledgements

References

Several of these headings are optional, for example, not all papers will include a one-sentence summary, figures or tables.

Please use the .docx format (all version after Word 2007).

More specific formatting instructions are provided in the actual template, which follows.

Title: How to Format a Science Paper (replace with your real title)

Authors: The author list should be one single paragraph (no breaks). Authors should be listed by first name or initial followed by last name and separated by commas. Use superscript numbers to link affiliations, and symbols *†‡ for author notes. For example, X. Jones¹*, P. Smith^{1,2}

Affiliations:

¹Affiliations should be preceded by superscript numbers corresponding to the author list, and each affiliation should end with a period.

²Each affiliation should be a separate paragraph.

Correspondence to: include the email addresses or telephone number of the corresponding author(s). Please use the asterisk () symbol for the corresponding author information.

Abstract: The abstract should be about 100-150 words, and organized in this structure: An opening sentence that sets the question that you address and is comprehensible to the general reader, background content specific to this study, results, and a concluding sentence. It should be one paragraph only.

One Sentence Summary: A brief summary of the main result of your paper, without excessive jargon and less than 150 characters.

Main Text: In general, this should include a brief (1-2 paragraph) introduction, followed by a statement of the specific scope of the study, followed by results and then interpretations. Please avoid statements of future work or claims of priority, and avoid repeating the conclusions at the end.

All of the Figures and Tables should be cited in order (which should be cited as, for example, figure S1, and table S1). You can include page breaks if you would like to place the figures within the text.

Subheads: These can be included in Research Articles or Reviews and should be brief.

Equations can be included. We do not recommend using the native Word 2007, 2008, 2010 or 2011 equation editor. This can in some cases produce less reliable MathML, the online markup language we use, which may result in display errors. Instead, use the legacy equation editor in word (Insert menu; select insert object; select word equation) or use Mathtype (recommended). If you enter equations in simple LaTeX, check that they will convert accurately (Word 2007 and higher can convert simple LaTeX equations).

Fig. #: (Begin each figure caption with a label, “**Fig. 1.**” for example, as a new paragraph.).

Fig. 1. The figure caption should begin with an overall descriptive statement of the figure followed by additional text. They should be immediately after each figure. Figure parts are indicated with capital letters (**A**). If you prefer, you can place both the actual figures and captions logically through the text near where they are cited rather than at the end of the file (but not both). If a paragraph in the main text begins with the name of a figure, write out “Figure” in full (e.g., <para>“Figure 1 shows....”).

Fig. 2. You can place graphics above each equation as part of this file.

Table #: (Begin each table caption with a label “**Table 1.**”, etc.) as a new paragraph.

Table 1. Start this caption with a short description of your table. Format tables using the Word Table commands and structures. Do not create tables using spaces or tabs characters.

Acknowledgments: The acknowledgments should include a statement about where the data reported in the paper are presented, archived, or available. If in an archive, include the accession number or a placeholder for it. Please also include relevant funding information such as grant numbers and funding agencies. You can also include a statement of author contributions here. The acknowledgements can be multiple paragraphs.

References:

References should be cited in bracket with a number [1]. Multiple reference citations are separated by commas [2, 3] or if a series, dashes [4-6].

References are listed in numerical order, and in the same order in which they are cited in text. The reference list should include all and only those references you have cited in the text. The reference list appears at the end of the paper.

Please include the above heading “References”.

Please do not mix in references with explanatory notes.

The reference consist of: Family name(s) and initials of all authors (year of publication), “Titles of journal articles” (optional), *Title of journal, book or other publication*, **volume(number)**, page numbers.

Example of a reference list:

[1] T. Binzoni, T.S. Leung, D. Boggett, D.T. Delpy (2002), “A new near infrared laser-Doppler flowmeter for deep tissue perfusion monitoring”, *MAGMA*, **14**, pp.74-75.

[2] T. Binzoni, T.S. Leung, D. Boggett, D.T. Delpy (2003), “Non-invasive laser Doppler perfusion measurements of large tissue volumes and human skeletal muscle blood RMS velocity”, *Phys. Med. Biol.*, **48(1)**, pp.2527-2549.

EDITORIAL COUNCIL**STANDING MEMBERS:**

Bao Chau Ngo (*University Avenue Chicago, USA*)
 Gia Khanh Pham (*Chairman of Vietnam Organ Transplant, Vietnam*)
 Thuc Tran (*Vietnam Panel on Climate Change, Vietnam*)
 Van Hieu Nguyen (*Vietnam Academy of Science and Technology, Vietnam*)

MEMBERS:

Akos Redey (*University of Pannonia, Hungary*)
 Chi Buu Bui (*Vietnam Academy of Agricultural Sciences, Vietnam*)
 Clark Gedney (*Bio Media Center for Instructional Design, Biological Sciences, Purdue University, USA*)
 Dang Xuan Tran (*Hiroshima University, Japan*)
 Dindo Campilan (*International Center for Tropical Agriculture (CIAT) - Asia Regional Office, Philippines*)
 Dinh Duc Nguyen (*University of Engineering and Technology (UET) - Vietnam National University, Hanoi, Vietnam*)
 Edgar Andreas Bohner (*VTT Technical Research Centre of Finland, Finland*)
 Franz Nestmann (*Institute for Water and River Basin Management, Karlsruhe Institute of Technology, Germany*)
 Gopalakrishnakone P (*National University Health System, Singapore*)
 Henry Nguyen (*University of Missouri, Columbia, USA*)
 Hung Viet Pham (*University of Science - Vietnam National University, Hanoi, Vietnam*)
 John Nieber (*University of Minnesota, USA*)
 Pierre Darriulat (*European Organization for Nuclear Research; Vietnam Academy of Science and Technology, Vietnam*)
 Stefan Norra (*Institute of Applied Geosciences, Karlsruhe Institute of Technology, Germany*)
 Tran Binh Le (*University of Science and Technology of Hanoi, Vietnam*)
 Tri Nang Tran (*University of Minnesota, USA*)
 Trong Trinh Phan (*Institute of Geological Sciences - Vietnam Academy of Science and Technology, Vietnam*)
 Truong Duc Pham (*University of Birmingham, UK*)
 Tu Bao Ho (*Japan Advanced Institute of Science and Technology, Japan*)
 Van Suong Hoa (*Concordia University, Canada*)
 Van Thang Nguyen (*Vietnam Institute of Meteorology, Hydrology and Climate change, Ministry of Natural Resources and Environment, Vietnam*)
 Van Toi Vo (*Tufts University, Massachusetts, USA; Vietnam National University in Ho Chi Minh City, Vietnam*)

ASSISTANT EDITORS

Dang Khanh Tran (*Agricultural Genetics Institute*)
 Huu Song Le (*108 Hospital*)
 Tien Dung Le (*Ho Chi Minh city*)
 Van Tuan Tran (*Vietnam National University, Hanoi*)
 Xuan Thanh Bui (*Ho Chi Minh city University of Technology*)

EDITORIAL OFFICE

113 Tran Duy Hung Str., Cau Giay Dist., Hanoi
 Tel: (84.24) 39436793; Fax: (84.24) 39436794
 Email: khcnvn@most.gov.vn
 Website: vietnamscience.vn

PUBLICATION LICENCE

No.1153/GP-BTTTT 26th July 2011

No.2528/GP-BTTTT 26th December 2012

No.592/GP-BTTTT 28th December 2016