

목차

통계학

표지잡음 존재시 분류율 향상을 위한 새로운 방법.....1

정혜림, 이석호

전자물리학

광촉매용  $C_3N_4$  시료의 광특성 분석.....11

이석주

생명공학

*Pichia pastoris* 에서 사람 락토펜리신으로부터 유래된 변형 락토펜리신의 세포 표면발현.....17

김대운, 이현환

**Characterization and identification of twelve bacterial strains isolated from fresh water and marine environments.....35**

Heeyoung Kang, Haneul Kim, Kiseong Joh



# 표지잡음 존재시 분류율 향상을 위한 새로운 방법<sup>1</sup>

정혜림, 이석호<sup>2</sup>

요약 : 표지잡음에 노출된 훈련자료를 통해 학습된 분류자는 정확한 분류규칙을 생성하지 못한다. 본 연구에서는 훈련자료에 내재된 표지잡음에 영향을 덜 받는 새로운 방법론을 제안한다. 이를 통해, 부정확하게 레이블링된 훈련자료를 통해서도 기존의 방법보다 오분류율을 획기적으로 줄일 수 있음을 보인다. 제안된 새로운 표지잡음 모형에 대한 효율적인 알고리즘을 제공하고, 컴퓨터 모의실험을 통해 기존의 방법으로는 훈련이 어려운 표지잡음 상황에서도 본 방법이 베이스 오분류율에 근접한 성능을 보임을 실증한다.

## 서론

본 연구에서는 두개의 그룹으로 분류하는 분류문제(binary classification)를 고려한다. 표지잡음(label-noise)은 통계모형을 학습하는 훈련자료(training data) 상의 그룹표지자(class label)가 뒤바뀐 경우를 의미한다.

이러한 표지잡음은 현실문제의 다양한 상황에서 발견된다. 의료서비스의 경우, 일반적으로 환자가 처음 병원에 방문시 임상과의 환자에 대한 문진 등 간접적인 방법으로 병의 유무를 진단한다. 초진시 병에 대한 확신이 없는 상태에서 정확한 진단키트를 사용하지 않는 것은 비용 및 시간 절감에 도움이 되기 때문이다. 하지만, 이는 임상과의 주관적 판단, 진료 당시 환자의 상태 및 부정확한 답변 등으로 정확한 진단이 이루어지기 어려울 수 있음을 내포한다. 이는 결과적으로 실제와는 다른 진단('병이 있음'을 '병이 없음'으로 진단하거나, 혹은 그 반대의 경우)이 이루어질 수 있음을 의미한다. 한편, IT 기업에서는 정확한 분류규칙을 얻기 위해 대규모의 훈련자료를 필요로 한다. 이를 손쉽게 취득하기 위해서

<sup>1</sup> 이 논문은 2018학년도 한국외국어대학교 학술연구비 지원에 의하여 이루어진 것임.

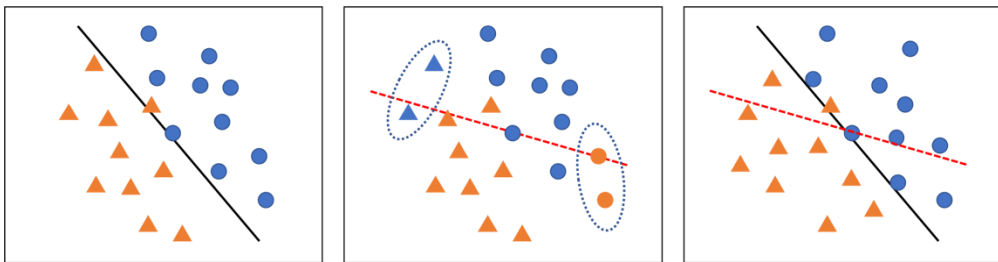
<sup>2</sup> 한국외국어대학교 자연과학대학 통계학과 교수

Amazon Mechanical Turks 와 같은 클라우드소싱을 통해 대규모자료를 저비용으로 획득한다. 클라우드소싱에 기여하는 참여자는 일정량의 보상을 받으며 자료에 대한 레이블링 작업을 대리한다. 이러한 방식은 기업에게 손쉽게 저렴한 비용으로 대규모 자료를 취득하게 하지만, 참여자의 비전문성 및 책임감결여로 인하여 정확한 레이블링을 기대하기 어렵다는 단점을 지닌다. 또다른 사례로 데이터보안의 경우, 악의적 침투자(malicious intruder)가 데이터베이스에 침투하여 고의로 자료의 레이블을 뒤집음으로써 데이터베이스 사용 주체가 올바른 의사결정을 내릴지 못하게 하는 사례가 빈번하다. 이렇듯, 표지잡음은 다양한 상황에서 훈련자료에 내재될 개연성이 높다. 표지잡음에 대한 상세한 조사는 Prenay and Verleysen (2014)를 참고하기 바란다.

대부분의 통계적, 기계학습적 분류모형은 올바른 학습을 위해 정확히 분류가 된 훈련자료를 필요로 한다. 표지잡음이 개입된 훈련자료를 통해 학습된 분류자(classifier)는 그 정확성을 담보할 수 없다. 즉, 새로운 자료(test data)에 대한 분류는 정확하지 않게 된다. 이를 개선하기 위해 몇가지 방법들이 제안되었다. 가장 손쉬운 방법으로써, 로지스틱회귀분류법 (logistic regression classification)에서 표지잡음에 대한 메커니즘을 확률모형에 고려한 방법이 흔히 사용된다. 이 방법의 성능은 표지잡음에 대한 확률모형의 정확도에 있다. 기존의 방법으로 흔히 사용하는 방법은 표지 전이확률(transition probability)을 모든 자료에 동일(random label-noise model)하게 가정하는 것이다. 이는 어떤 자료가 표지잡음인지 미리 알 수 없으므로 모든 훈련자료에 대해 동일한 개연성을 가정하는 방식이다. 이는 구현이 간단하면서도 표지잡음을 고려하지 않은 모형에 비해 향상된 분류율을 보인다. 구현이 간단하다는 이점에도 불구하고 이러한 단순한 전이확률은 현실에서 나타나는 표지잡음 메커니즘을 모형화하기에는 한계가 있다. 이를 개선하기 위해 Bootkrajang (2016)은 좀더 현실적인 표지잡음 메커니즘을 고려하기 위하여 분류경계(decision boundary)에서 가까운 자료에 대해 높은 전이확률을 모형화 하였다. 다양한 현실문제에서, 분류가 까다로운

자료에 대해 부정확한 표지가 생성됨을 고려한 것이다. 이를 위하여 Bootkrajang (2016)은 분류경계로부터 거리를 공변량으로 하는 모수적 전이확률모형(parametric label-noise model)을 제안하였다. 불분명한 자료는 대부분 분류경계에 놓이게 됨으로 이에 대해 높은 전이확률을 고려함으로써 좀더 현실성 있는 표지잡음 모형이 가능하다.

기존에 제안된 표지잡음 모형의 한계는 명확하다. 전술한 바와 같이, random label-noise model 은 단순한 구조로 인해 현실의 복잡한 표지잡음 구조를 포괄할 수 없다. Parametric label-noise model 은 훈련자료가 놓이는 다차원공간 상의 거리에 기반하고 있다. 일반적으로 차원수, 즉 설명변수의 수가 증가할수록 거리에 기반한 방법들은 급격한 성능저하가 수반된다. 이는 차원의 저주(curse of dimensionality)로 설명될 수 있다. 더욱이, 이 두가지 방법은 악의적 침입자에 의한 표지잡음을 설명하는데 한계를 보인다. 고의로 표지변환을 시도할 때, 가장 효과적으로 경향을 왜곡하는 경우는 주로 레버리지 효과가 큰 영향점(influence point)에 대해 표지변환을 시도하는 것인데, 이러한 영향점은 분류문제에서 주로 분류경계에 놓이기 때문이다. 이러한 메커니즘을 시각적으로 도시하면 아래와 같다.



표지잡음이 없는 경우의 분류자

표지잡음이 있는 경우의 분류자

검증자료에서의 성능

본 연구에서는 이러한 제약을 극복할 수 있는 방법을 제안한다. 이를 위해 각 자료에 대해 개별적인 전이확률을 고려한다. 로짓(logit) 변환된 전이확률에 대해

개별적인 절편(intercept)을 갖도록 추정한다. 이때 모수적접근을 하지 않으므로 이는 비모수적(nonparametric) 접근이라 할 수 있다. 개별절편의 개수는 자료개수에 비례하므로 추정을 위해 규제화기법(regularization)이 필요하다. 본 연구에서는 라소벌점(lasso penalization)을 통한 추정을 고려한다. 제안방법의 장점은 다음과 같다. 로짓변환 하에서 라소벌점화를 통한 추정은 명시적인 thresholding rule 을 제공한다. 이를 활용하면 좌표하강법(coordinate descent method)을 이용한 효율적인 알고리즘을 구현할 수 있다. 개별 절편추정치는 해당 자료가 표지잡음의 가능성이 높으면 0 이 아닌 값으로 추정되고, 낮으면 0 으로 추정된다. 따라서 절편추정치는 표지잡음에 대한 표지자(indicator)로 활용될 수 있다. 또한, 기존의 방법과 다르게, 악의적 표지잡음에 대해서도 베이즈분류자(Bayes classifier)에 근접한 분류자를 제공하여 향상된 분류율을 제공할 수 있으며, 이를 모의실험을 통해 실증한다.

## 방 법

$y_i \in \{0, 1\}$  ( $i = 1, 2, \dots, n$ ) 를 훈련자료의 표지변수(label)라 하자. 각 자료에 대한 설명변수  $x_i = (x_{i1}, \dots, x_{ip}) \in \mathbb{R}^F$  를  $p$  차원 공간상의 점으로 간주할 수 있다. 이에 대해 로지스틱모형을 가정하면 아래와 같은 확률모형을 고려할 수 있다.

$$\Pr(y|x; \Theta) = \prod_{i=1}^n \sigma(f_i)^{y_i} \sigma(-f_i)^{1-y_i}$$

여기서,  $f_i = f(x_i) = \beta_0 + x_i^T \beta$  이고  $\sigma(u) = 1/\{1 + \exp(-u)\}$  이다. 만일 표지잡음이 존재하지 않는다면, 위의 모형은 일반적인 로지스틱모형이 된다. 훈련자료의 표지  $y_i$  에 표지잡음이 있다고 가정하자. 표지잡음이 없는 실제표지자를  $z_i \in \{0, 1\}$  로 표기하자. 즉, 표지잡음이 있는 경우,  $(z_i, y_i) = (0, 1), (1, 0)$  이며 표지잡음이 없는 경우는  $(z_i, y_i) = (0, 0), (1, 1)$  이다.

대부분 현실문제에서 실제표지자는 알려져 있지 않다. 하지만, 이를 알고 있다고 가정한 경우, 위의 확률모형은 아래와 같이 표현할 수 있다.

$$\begin{aligned}\Pr(y|x; \Theta) &= \prod_{i=1}^n \sum_{z \in \{0,1\}} \Pr(z_i|x_i; \Theta_1) \Pr(y_i|z_i, x_i; \Theta_2) \\ &= \prod_{i=1}^n \left\{ (1 - \gamma_{i,1})^{y_i} \gamma_{i,1}^{1-y_i} \sigma(f_i) + \gamma_{i,0}^{y_i} (1 - \gamma_{i,0})^{1-y_i} \sigma(-f_i) \right\}\end{aligned}$$

위에서,  $\gamma_{i,0}$  은 실제표지자가 0 일때 1 로 표지잡음이 일어날 전이확률이고,  $\gamma_{i,1}$  은 실제가 1 일때 0 으로 표지잡음이 일어날 확률을 나타낸다. 위 모형을 로지스틱회귀분류모형에서의 일반적인 표지잡음모형이라 부른다.

표지잡음 모형에서 전이확률을  $\gamma_{i,0} = \gamma_{i,1} = \gamma$  로 동일하게 두면, 이를 대칭확률모형(symmetric random model)이라 부른다.  $\gamma_{i,0} = \gamma_0$ ,  $\gamma_{i,1} = \gamma_1$  로 두면, 이를 비대칭확률모형(asymmetric random model)이라 부른다. 이러한 확률모형은 모든 자료에 대한 전이확률이 동일할 때 성립한다. 이는 관측되지 않은 실제표지자  $z_i$  를 결측치로 간주할 때, EM 알고리즘을 이용하여 손쉽게 적합할 수 있다.

한편, 아래와 같이 전이확률을 각 자료가 결정경계로부터의 거리에 기반한 모수적 전이확률로 표현할 수 있다.

$$\log \frac{\gamma_{i,0}}{1 - \gamma_{i,0}} = a_0 + b_0 d_i, \quad \log \frac{\gamma_{i,1}}{1 - \gamma_{i,1}} = a_1 + b_1 d_i$$

여기서  $d_i$  는  $x_i$  가 결정경계인  $f(x) = c$  으로부터의 거리를 의미하고,  $b_0, b_1 \leq 0$  을 가정한다. 따라서 거리가 가까울수록 전이확률은 높게 나타난다 (Bootkrajang, 2016). 전이확률을 로짓변환을 통해 모형화하고 있으므로 이는 EM 알고리즘

하에서, 일반적인 로지스틱회귀법을 통해 손쉽게 절편과 기울기 모수를 추정할 수 있다.

본 연구에서는 전이확률을 개별절편을 가지도록 모형화한다. 따라서 전이확률은 다음과 같다.

$$\log \frac{\gamma_{i,0}}{1 - \gamma_{i,0}} = a_0 + b_{i,0}, \quad \log \frac{\gamma_{i,1}}{1 - \gamma_{i,1}} = a_1 + b_{i,1}$$

개별절편  $b_{i,0}$ ,  $b_{i,1}$  는 자료수가 증가하면 추정해야 할 모수의 개수가 비례하여 늘어난다. 따라서 추정가능하게 하기 위해 본 연구에서는 라소벌점(lasso penalization)을 통한 규제화기법을 사용한다. 라소추정에 의한 개별절편은 대부분 0 으로 추정되며, 전반적인 전이확률에 비해 높은 전이확률을 가질 것으로 예상되는 자료에 대해 해당 절편이 0 이 아닌 값으로 추정하게 된다.

본 개별절편모형(individual intercept model)을 위한 EM 알고리즘은 다음과 같다. E-step 은 실제표지자  $z_i$  에 대한 조건부기대값을 구한다. 이는 아래와 같이 명시적으로 얻을 수 있다.

$$\begin{aligned} \pi_i &= E[z_i | x, y; \Theta^o] = \Pr(z_i = 1 | x, y; \Theta^o) \\ &= \frac{\gamma_{i,1}^{1-y_i} (1 - \gamma_{i,1})^{y_i} \sigma(f_i)}{\gamma_{i,0}^{y_i} (1 - \gamma_{i,0})^{1-y_i} \sigma(-f_i) + \gamma_{i,1}^{1-y_i} (1 - \gamma_{i,1})^{y_i} \sigma(f_i)} \end{aligned}$$

여기서  $\Theta^o$  은 이전 단계에서 얻은 모수추정치를 의미하며, 위 식에 필요한 모수들은 이를 대입하여 계산한다. 따라서, 이를 이용하여 M-step 에 필요한 음의 로그가능도함수(negative loglikelihood function)은 아래와 같이 세가지 성분의 합으로 표현이 가능하다.



$$Q(\Theta; \Theta^o) = Q_1(\Theta_1; \Theta^o) + Q_2(\Theta_2; \Theta^o) + Q_3(\Theta_3; \Theta^o)$$

위에서 정의된 항들은 아래와 같다.

$$Q_1(\Theta_1; \Theta^o) = - \sum_{i=1}^n \{ \pi_i \log \sigma(f_i) + (1 - \pi_i) \log \sigma(-f_i) \}$$

$$Q_2(\Theta_2; \Theta^o) = - \sum_{i=1}^n (1 - \pi_i) \{ y_i \log \gamma_{i,0} + (1 - y_i) \log(1 - \gamma_{i,0}) \}$$

$$Q_3(\Theta_3; \Theta^o) = - \sum_{i=1}^n \pi_i \{ (1 - y_i) \log \gamma_{i,1} + y_i \log(1 - \gamma_{i,1}) \}$$

여기서  $\Theta_1 = \{\beta_0, \beta\}$ ,  $\Theta_2 = \{a_0, b_{1,0}, \dots, b_{n,0}\}$ ,  $\Theta_3 = \{a_1, b_{1,1}, \dots, b_{n,1}\}$  이며, 세 식은 세 묶음으로 주어진 모수들을 각기 개별적으로 추정이 가능하게 됨을 알 수 있다.  $\beta_0, \beta$ 의 추정은  $\Theta_1$ 을 통해 이루어지며, 이는 반응변수가  $\pi_i$ 인 로지스틱회귀를 이용해 구할 수 있다. 한편, 전이확률과 관련된 모수들은  $\Theta_2, \Theta_3$ 의 최소화 과정을 통해 구할 수 있다. 여기에 각기 라소별점  $\lambda_0 \sum_{i=1}^n |b_{i,0}|$ ,  $\lambda_1 \sum_{i=1}^n |b_{i,1}|$ 을 더하여 최소화한다. 이에 해당하는 개별절편의 추정식은 닫힌 해로 주어지며 그 구체적인 형태는 아래와 같이 주어짐을 확인할 수 있다.

$$\hat{b}_{i,0} = \begin{cases} 0 & \text{if } \lambda_0 \geq \frac{(1-\pi_i)e^{(1-y_i)a_0}}{1+e^{a_0}} \\ (2y_i - 1) \log \left( \frac{1-\pi_i-\lambda_0}{\lambda_0} \right) - a_0 & \text{if } \lambda_0 < \frac{(1-\pi_i)e^{(1-y_i)a_0}}{1+e^{a_0}} \end{cases}$$

$$\hat{b}_{i,1} = \begin{cases} 0 & \text{if } \lambda_1 \geq \frac{\pi_i e^{y_i a_1}}{1+e^{a_1}} \\ (2y_i - 1) \log \left( \frac{\lambda_1}{\pi_i - \lambda_1} \right) - a_1 & \text{if } \lambda_1 < \frac{\pi_i e^{y_i a_1}}{1+e^{a_1}} \end{cases}$$

위의 해는 **thresholding rule** 형태를 띄고 있음을 알 수 있다. 또한 이를 이용하면 좌표하강알고리즘을 통해 각 모수를 반복해법으로 손쉽게 구현할 수 있다. 한편,

$a_0, a_1$  은 뉴튼랩슨 방식으로 해를 구한다. 라소벌점의 벌점모수  $\lambda_0$  및  $\lambda_1$  은 AIC (Akaike's information criterion) 에 의거하여 결정한다.

## 실 험

제안방법의 성능을 확인하기 위한 모의실험을 수행하였다.  $i_0, i_1 = 1, \dots, n/2$  에 대하여 두 그룹의 자료를 아래와 같이 생성하였다.

$$x_{i_0}^0 \sim N_p((-c/2)\mathbf{1}_p, \Sigma_0), \quad x_{i_1}^1 \sim N_p((c/2)\mathbf{1}_p, \Sigma_1)$$

여기서  $c$  는 양의 상수로써, 다음과 같이 둔다.

$$c = 4z_\alpha / \sqrt{\mathbf{1}_p^T (\Sigma_0^{-1/2} + \Sigma_1^{-1/2})^2 \mathbf{1}_p}$$

여기서  $z_\alpha$  는 표준정규분포의 오른쪽 백분위수를 의미하여, 이 경우 베이즈 오분류율(Bayes error rate)은  $\alpha$  가 된다. 두 분포의 공분산행렬  $\Sigma$  의  $(i, j)$  원소는  $\Sigma_{ij} = \text{sign}(\rho)|\rho|^{|i-j|}$  로 둔다. 위의 분포로부터 훈련자료를 100, 200, 400, 800개를 생성하였다. 이에 대해 10%의 표지잡음을 인위적으로 생성하였다. 처음 설명변수에 대해 가장 레버리지가 높은 점들을 선별하여 표지잡음을 유발하였다. 훈련자료로부터 얻은 분류자의 성능을 확인하기 위해, 동일한 분포로부터 10000개의 검증자료(test data)를 표지잡음 없이 생성하였다. 이로부터 검증오분류율(test error rate)을 계산하였다. 다양한 상황에 대하여 실험을 하였으나, 공간상의 제약으로 인해, 변수 개수  $p=2$ , 독립공분산  $\rho=0$ , 표지잡음 10%의 경우에 대해 1000회의 모의실험을 통한 검증오분류율의 평균값 및 오차범위를 그림으로 정리하였다.

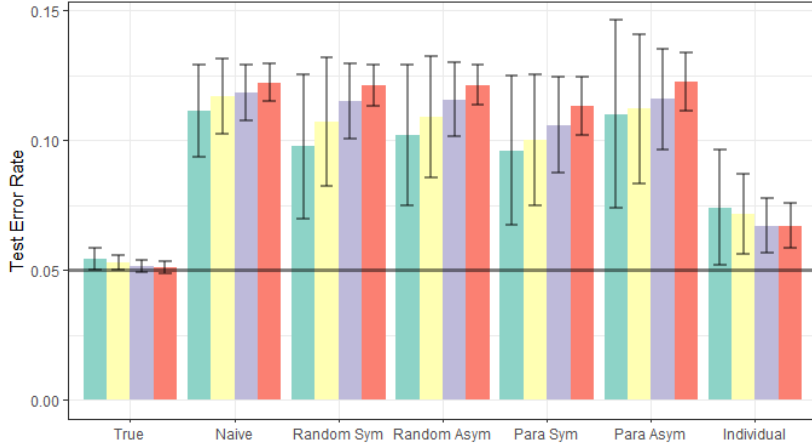


그림 1: 베이지오분류율 0.05인 경우

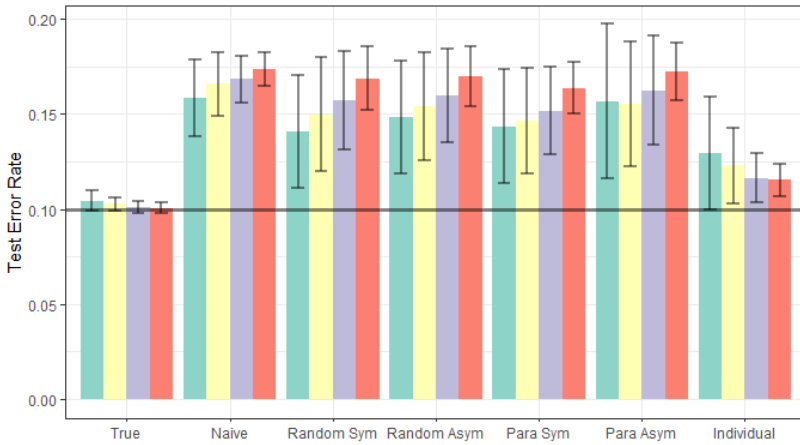


그림 2: 베이지오분류율 0.1인 경우

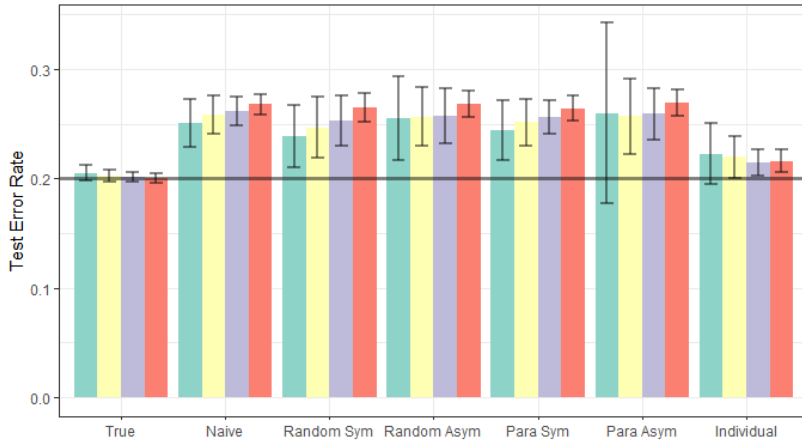


그림 3: 베이지오분류율 0.2인 경우

그림 1~3에서 4개의 막대는 왼쪽부터 훈련자료의 크기가 100, 200, 400, 800에 대응한다. 비교를 위해 실제표지자를 이용한 경우 (True), 표지잡음이 있는 표지자를 이용한 경우 (Naive), 대칭확률모형 (Random Sym), 비대칭확률모형 (Random Asym), 대칭모수모형 (Para Sym), 비대칭모수모형 (Para Asym), 그리고 제안모형 (Individual)을 함께 도시하였다. 베이지오분류율은 가로선으로 표기하였다. 그림에서 확인 가능하듯이 제안모형이 기존의 모형보다 매우 향상된 검증오분류율을 보여주고 있음을 확인할 수 있다.

### Acknowledgements

이 연구는 2018년 한국외국어대학교 학술연구비 지원에 의해 이루어진 것임.

### 참고문헌

- Bootkrajang, J. (2016) A generalized label noise model for classification in the presence of annotation errors. *Nuerocomputing*, **192**, 61-71.
- Prenay, B., Verleysen, M. (2014) Classification in presence of label noise: survey. *IEEE Transactions on Neural Networks and Learning Systems*, **25**, 845-869.

## 광촉매용 $C_3N_4$ 시료의 광특성 분석<sup>1</sup>

이석주<sup>2</sup>

요약 :  $g-C_3N_4/SnFe_2O_4$  composite의 광특성을 photoluminescence(PL)과 PL excitation(PLE) 방법을 통해 분석하였다. PLE spectrum은 광원의 power에 대해 보정되었으며, PL spectrum의 low energy tail 쪽에서 관측하였다. PL signal은 주로  $g-C_3N_4$ 에서의 흡수에 의한 발광으로 관측되지만  $SnFe_2O_4$  에서의 흡수 또한 PL spectrum에 영향이 있는 것으로 생각된다.

### 서론

Graphitic carbon nitride ( $g-C_3N_4$ )는 연료전지나 전해조에서의 촉매지지대<sup>[1]</sup> 혹은 벤젠/이산화탄소 활성화 반응<sup>[2]</sup>이나 furfuryl alcohol 환원반응의 촉매<sup>[3]</sup> 등 여러 응용에서 주목받고 있는 소재이다. Carbon nitride 물질은 carbon 을 대체할 수 있는 물질로 알려져 있고, 그 중 한 가지인  $\beta-C_3N_4$  는 공유 결합을 하여 carbon 결정인 다이아몬드에 버금가는 강도를 보이기도 한다.<sup>[4]</sup> Carbon 의 결합 상태에 따라 여러 구조가 만들어지며 그 중 이차원 구조가 graphite 이듯이,  $C_3N_4$  의 이차원 구조가 바로  $g-C_3N_4$  이다. Physical vapor deposition, chemical vapor deposition, 고온-고압 합성 등의 여러 방법으로  $g-C_3N_4$  이 만들어져 왔을 뿐만 아니라 nanoparticle 형태<sup>[5]</sup> 혹은 자성 특성을 가지도록 하는  $SnFe_2O_4$  와의 composite 형태로도<sup>[6]</sup> 연구되어왔다.

본 연구에서는 향상된 광촉매 특성을 가지도록 준비된  $g-C_3N_4/SnFe_2O_4$  composite 의 광특성을 photoluminescence(PL)과 PL excitation(PLE) 방법을 통해 분석해 보았다.

<sup>1</sup> 이 논문은 2018학년도 한국외국어대학교 학술연구비 지원에 의하여 이루어진 것임.

<sup>2</sup> 한국외국어대학교 자연과학대학 전자물리학과 교수

## 실 험

$g\text{-C}_3\text{N}_4$  는 melamine 을  $550^\circ\text{C}$ 로 4 시간 가열한 후 milling 하여 준비하였다.  $\text{SnFe}_2\text{O}_4$  는  $\text{SnCl}_2$  와  $\text{FeCl}_2\cdot 4\text{H}_2\text{O}$  를 precursor 로 이용한 solvothermal 방법을 통해 nanoparticle 형태로 준비하고,  $g\text{-C}_3\text{N}_4$  와 함께 역시 solvothermal 방법으로 섞어 power 를 만들었다.<sup>[6]</sup>

PL 과 PLE 의 광원은 Xe 1kW lamp 를 Gemini 180 double monochromator 로 분광하여 사용하였다. 0.5mm slit 을 통해 약 0.8nm ( $\sim 7\text{meV}$ )의 FWHM 을 가진 (0.83nm/10.5meV near 300nm 와 0.77nm/3.8meV near 500nm) 단색광을 만들었다. 광원의 파장에 따른 optical power 가 그림 1 에 보여지고 있으며, 약 45~160  $\mu\text{W}$  사이의 값을 가지고 있음을 알 수 있다. PL spectrum 은 1702 single spectrometer 와 Si photomultiplier tube 를 통해 얻었고, 모든 실험은 상온에서 진행되었다

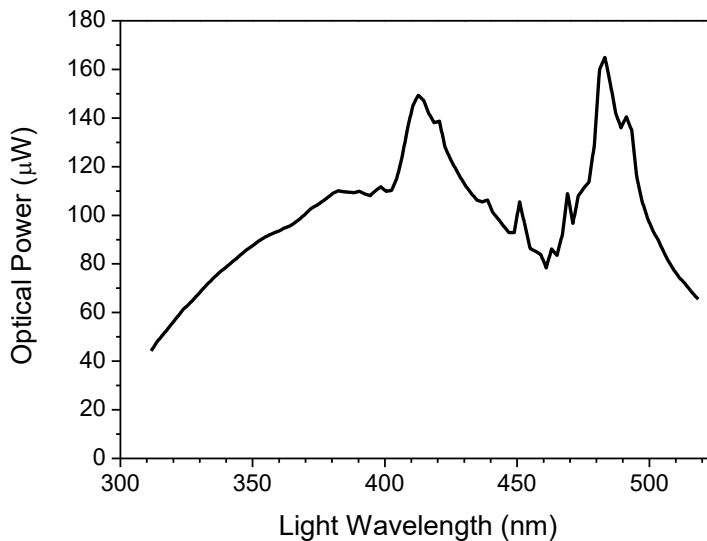


그림 1. PL/PLE 광원의 파장에 따른 optical power

## 결과 및 분석

SnFe<sub>2</sub>O<sub>4</sub>의 함유량이 다른 여러 시료에 대해 325nm 광원을 통해 얻은 PL spectrum 들이 그림 2에 보여지고 있다. 시료에 따라 달라지는 특성들은 다른 연구 결과<sup>[6]</sup>를 참조하도록 하고, 여기서는 가장 큰 세기를 보이는 시료에 대하여 PLE 결과를 살펴보도록 한다.

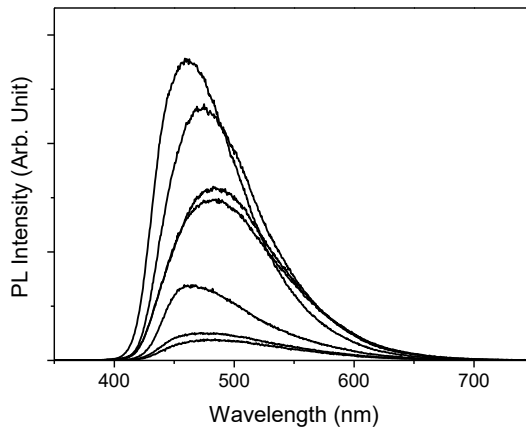


그림 2. 여러 g-C<sub>3</sub>N<sub>4</sub>/SnFe<sub>2</sub>O<sub>4</sub> composite 시료들의 PL spectrum

PL spectrum의 low energy 쪽 shoulder인 525nm의 빛을 관측하며 광원의 파장을 움직인 PLE spectrum이 오른쪽 그림 3에 보여지고 있다. 520nm보다 긴 쪽의 파장에서는 광원의 꼬리 부분이 관측되어 크기가 증가하는 모습을 보이고 있고, 여러 다른 feature들은 주로 광원 크기에 의한 것으로, 광원 크기에 대해 보정한 PLE spectrum을 그림 4에서 볼 수 있다.

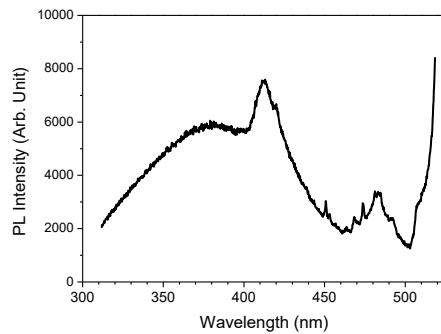


그림 3. 광원 power 보정 전의 PLE spectrum

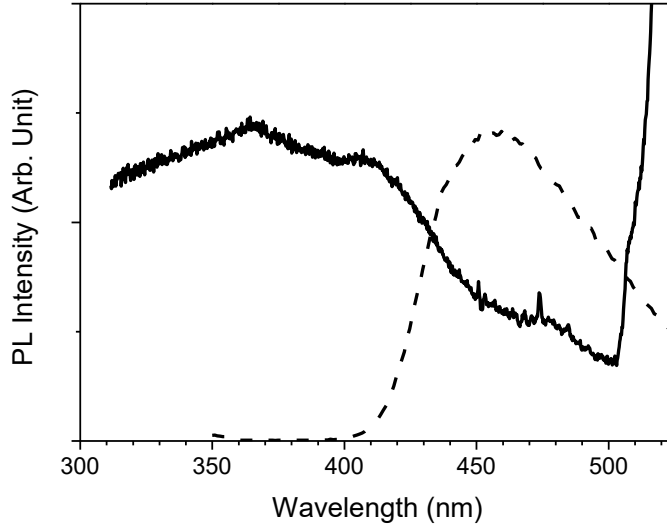


그림 4. PL spectrum (dashed) 과 525nm 에서 측정된 PLE spectrum (실선)

그림 4 의 보정된 PLE spectrum 에서는 기대대로 광원에서 유래하는 뾰족한 feature 들이 사라져 있다. 474nm 의 peak 이 남아 있는 것은 아마도 광원의 power 측정시의 문제일 것으로 생각되며, 이후에 좀 더 정확히 알아보아야 할 것이다. 실제 시료에서 보여주는 feature 들은 360nm, 410nm 근처의 broad 한 peak 과 480nm 근처의 shoulder 들로, 흡수 스펙트럼들과 비교하여 보면 앞의 peak 들은 g-C<sub>3</sub>N<sub>4</sub> 의 band edge transition 이며 뒤의 shoulder 는 SnFe<sub>2</sub>O<sub>4</sub> 의 absorption 이 관측되는 부분과 일치함을 알 수 있다. 순수한 SnFe<sub>2</sub>O<sub>4</sub> 시료에서는 PL spectrum 이 관측되지 않으므로<sup>6</sup> 위 결과는 SnFe<sub>2</sub>O<sub>4</sub> 에서 g-C<sub>3</sub>N<sub>4</sub> 로의 energy transfer mechanism 이 존재한다는 가능성을 보여준다. 물론 순수한 g-C<sub>3</sub>N<sub>4</sub> 시료도 420nm 에서 620nm 근처까지 서서히 줄어드는 absorption tail 을 가지고 있으므로 여러 시료를 비교한 후에야 정확한 결론을 내릴 수 있을 것이다.



## 결론

g-C<sub>3</sub>N<sub>4</sub>/SnFe<sub>2</sub>O<sub>4</sub> composite 시료에 대해 PLE spectrum을 측정한 결과, g-C<sub>3</sub>N<sub>4</sub>의 흡수에 의한 luminescence가 주 여기 원인이기는 하나, SnFe<sub>2</sub>O<sub>4</sub> 물질에서의 흡수도 emission에 영향을 주는 것으로 생각된다. 이는 광촉매의 역할에 제한점이 될 수 있으므로 앞으로 더 연구가 필요하다.

## Acknowledgements

이 연구는 2018년 한국외국어대학교 학술연구비 지원에 의해 이루어진 것임.

## 참고문헌

- [1] N. Mansor, T.S. Miller, I. Dedigama, A.B. Jorge, J. Jia, V. Brázdová, C. Mattevi, C. Gibbs, D. Hodgson, P.R. Shearing, C.A. Howard, F. Corà, M. Shaffer, D.J.L. Brett, P.F. McMillan, *Electroch. Acta* **222**, 44 (2016)
- [2] X. Chen, L. Zhang, B. Zhang, X. Guo, X. Mu, *Sci. Rep.* **6**, 28558 (2016)
- [3] A. Thomas, A. Fischer, F. Goettmann, M. Antonietti, J.-O. Müller, R. Schlögl, J.M. Carlsson, *J. Mater. Chem.* **18**, 4893 (2008)
- [4] A.L. Liu, M.L. Cohen, *Science* **245**, 841 (1989)
- [5] Q. Guo, Y. Xie, X. Wang, S. Lv, T. Hou, X. Liu, *Chem. Phys. Lett.* **380**, 84 (2003)
- [6] Y. Jia, S.J. Rhee, C. Liu, *Mat. Lett.* **188**, 338 (2017)



## *Pichia pastoris* 에서 사람 락토페리신으로부터 유래된 변형 락토페리신의 세포 표면발현<sup>1</sup>

김대운, 이현환<sup>2</sup>

**ABSTRACT** : 사람 락토페린(human lactoferrin, hLf)은 주로 초유에서 대량으로 분비되는 항균, 항 바이러스 및 항 곰팡이 활성을 가진 단백질이다. 본 연구에서는 hLf로부터 유래된 항균 펩타이드 Lactoferricin (hLfcin)의 변형체인 HLSA를 methylotrophic yeast인 *Pichia pastoris*의 세포 표면에 발현하고자 하였다. 사용된 HLSA는 hLfcin 아미노산 서열에서 20번부터 31번 아미노산 서열 중 Gln24, Asn26, Met27을 Lys24, Ala26, Ala27으로 치환한 것으로 항균활성이 높게 나타나도록 본 연구실에서 특정 amino acid residue를 변형시킨 것이다. HLSA의 N-terminal에 Ssp intein(Ssp DnaB intein)을, C-terminal에 Mxe intein(Mxe gyrA intein)을 융합시켜 발현되도록 하였고, *Saccharomyces cerevisiae*의 glycosylphosphatidylinositol (GPI)-anchoring protein을 사용하여 Ssp(CBD)-HLSA-Mxe 융합단백질을 *Pichia pastoris* SMD1168의 세포 표면에 위치하도록 발현시켰다. 발현된 융합단백질은 FITC-labeled immunofluorescence와 Flow cytometric analysis에 의해 세포표면에 위치함이 밝혀졌다. 또한 Ssp(CBD)-HLSA-Mxe intein 융합단백질이 pH와 DTT에 의해 CBD affinity column 상에서 절단되어 HLSA가 분리되는 것을 Western blot에 의해 확인하였다.

## 서론

인체의 초유에서 다량으로 분비(7~8g/L)되는 항균, 항 바이러스, 항 진균 단백질인 사람 락토페린 (Human lactoferrin, hLf)은 innate immunity 를 부여하는 매우 중요한 단백질이다(1, 2, 3). hLf 의 항균 작용은 주로 N-lobe 의 N-terminal domain 에 있는 1-47 아미노산으로 이루어진 짧은 peptide 가 담당하는 것으로 밝혀졌다. 이 짧은 peptide 를 사람 락토페리신(Human lactoferricin, hLfcin)이라 하며 Cys20 과 Cys37 사이의 disulfide bond 에 의해 loop structure 가 형성되는데, 이를 구성하는 18 residues 가 항균 활성 능력을 가지는 것으로 알려져 있다 (4, 5, 6, 7,

<sup>1</sup> 이 논문은 2018학년도 한국외국어대학교 학술연구비 지원에 의하여 이루어진 것임.

<sup>2</sup> 한국외국어대학교 자연과학대학 생명공학과 교수

8). 18 residues 중 특이적으로 높은 항균활성을 가지는 부분인 20-31 residues 는 native molecule 보다 높은 항균 활성을 나타낸다. 본 연구의 대상인 HLSA 는 hLfcin 20-31 에서 높은 항균 활성을 나타내는 것으로 알려진 특정 아미노산 서열을 기초로 하여 Gln24, Asn26, Met27 을 Lys24, Ala26, Ala27 로 치환한 것이다(9).

HLSA 는 양전하를 갖기 때문에 음전하를 띠는 bacteria membrane 의 Lipopolysaccharide(LPS)를 인식하여 binding 함으로써 pore 를 형성하고 osmotic shock 를 유도하거나 내부에 침투하여 intracellular molecules 에 binding 하여 bacteria 의 성장을 억제시킨다(10, 11, 12). 그러나, 유전자 조작 기술을 이용하여 미생물로부터 antimicrobial peptide 를 생산할 경우 숙주 미생물의 생장이 억제되거나 사멸하는 어려움이 존재한다(13, 14). 본 연구에서는 숙주 미생물의 성장 저해 및 사멸의 문제를 해결하기 위하여 cytotoxic protein 인 HLSA 를 protein splicing segment 인 intein 과 융합시켜 발현 시키고자 하였다.

Methylotrophic yeast 인 *Pichia pastoris* 는 외래유전자를 숙주세포의 chromosomal DNA 에 삽입시킬 수 있어 발현의 안정성이 높다. 또한 alcohol oxidase 1(AOX1) promoter 를 사용하여 외래 유전자를 대량 발현시키거나, 발현된 단백질을 세포 외부로 분비시키기에 유용한 숙주로 알려져 있다(13, 15, 16, 17). *S. cerevisiae* 의 GPI anchoring motif 를 이용한 surface display system 은 다양한 연구가 진행되었으나 HLSA 를 *P. pastoris* 세포 표면에 발현시킨 연구는 전무하다. 본 연구에서는 산업적으로 널리 사용되는 *P. pastoris* 에서 HLSA 를 발현시키기 위해 intein-based system 을 도입했으며, *S. cerevisiae* 의 GPI-anchored protein 을 사용하여 *P. pastoris* 세포 표면에 발현시키고자 하였다.

## 재료 및 방법

### 1. 균주 및 plasmids

본 연구에서 사용된 균주와 plasmid는 Table 1과 같으며, 사용된 primer는 table 2와 같다.

Table 1. Bacterial strains and plasmids.

Strains/Plasmids	Genotype	References
<b>Strains</b>		
<i>P. pastoris</i> SMD 1168	<i>his4mut+pep4</i>	Invitrogen, Co.
<i>E. coli</i> Top10F'	F <sup>'</sup> { <i>Iacq</i> , Tn10(TetR)} <i>mcrAΔ(mrr-hsdRMS-mcrBC)Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara leu) 7697 galU galK rpsL(StrR) endA1 nupG</i>	ATCC13032
THLSA	SMD1168 integrated by pTWIN 1-HLSA TIP630	This study
<b>Plasmids</b>		
pGEM-T	<i>Amp<sup>r</sup></i> , T/A cloning vector	Promega, Co.
pBHA-HLSA	<i>Amp<sup>r</sup></i> , sub-cloning vector with synthesized HLSA gene	Bioneer, Co
pBHA-TWIN 1	<i>Amp<sup>r</sup></i> , sub-cloning vector with synthesized Ssp(CBD) & Mxe intein	Bioneer, Co
pTWIN 1-HLSA	pBHA-TWIN 1 containing HLSA	This study
pPIC9K	<i>Amp<sup>r</sup></i> , <i>Kan<sup>r</sup></i> , <i>his+</i>	Invitrogen, CO
pPIC9K-700k	<i>Amp<sup>r</sup></i> , <i>his+</i> containing 630bp <i>TIP</i> gene	This study
pPIC9K-THLSA	pPIC9K containing Ssp(CBD) & Mxe intein, HLSA	This study

Table 2. Primer sequence.

Primer	Sequence ( 5'→3')
a-factor(F)	ACTACTATTGCCAGCATTGCTG
3'AOX (R)	TCCTACAGTCTTACGGTAAACG
pTWIN check(F)	CCAGAGATTGAAAAGTTGTC
pTWIN check(R)	AAGAGCAACTAGTGCATC
SapI-HLSA(F)	GCTCTTCCAACAAATGTTTCCAATGGAAG
HLSA-SapI(R)	GCTCTTCCGCACCTCACCTTACGTGC
TIP630-NotI(R)	GCGGCCGCTTATAACAATAAAGCAGCTGCAC
HLSA(R)	CGACGGAAAATGCAGACCTT
HLSA check(F)	GGTGTTCGATTTGACCGTGCCAGG

\* F : Forward, R : Reverse.

## 2. 균주의 배양

*E. coli* 는 LB(Luria Bertani)배지 (0.5% yeast extract, 1% tryptone, chloride in 1 liter of distilled water)를 사용하여 37. C, 230 rpm으로 배양하였다. 형질전환체의 선별 등 필요시 ampicillin (50 $\mu$ g/ml)을 첨가하여 배양하였다.

Methylotrophic yeast *P. pastoris*는 YPD(1% yeast extract, 2% peptone, 2% dextrose in 1 liter of DW)를 사용하여 30. C, 230rpm에서 배양하였다. 형질전환체의 선별을 위해서는 MD 한천 배지 (4 $\times$ 10<sup>-5</sup>% biotin, 1.34% yeast nitrogen base (YNB), 2% dextrose, 1.5% agar in 1 liter of DW)를 사용하였다. 재조합 유전자 발현을 위해서 BMGY(1% yeast extract, 4 $\times$ 10<sup>-5</sup>% biotin, 2% peptone, 100 mM potassium phosphate buffer pH 6.0, 1.34% YNB, 1% glycerol in 1 liter of DW)와 BMMY(1% yeast extract, 4 $\times$ 10<sup>-5</sup>%

biotin, 2% peptone, 100 mM potassium phosphate buffer, pH 6.0, 1.34% YNB, 1% methanol in 1 liter of DW)를 사용하였고, 30. C, 230 rpm에서 배양하였다.

### 3. Codon optimization

HLSA와 Ssp DnaB intein (Ssp intein), Mxe gyrA intein (Mxe intein)을 *P. pastoris*에서 발현 시키기 위하여 유전자 합성(Bioneer Co.)시에 codon usage에 맞게 optimization시켰다. 이때 Ssp intein은 CBD를 포함하고 Mxe intein은 CBD를 포함하지 않게 design하였다. 이를 Ssp(CBD)-Mxe라 표시하였다

### 4. HLSA 세포 표면 발현을 위한 재조합 균주 개발

HLSA와 Ssp (CBD)-Mxe intein의 융합 단백질을 세포 표면에 발현시키기 위하여 pPIC9K-THLSA 재조합 플라스미드를 제작하였다. 유전자 합성을 통해 얻은 pBHA-HLSA를 주형으로 SapI-HLSA(F) 와 HLSA-SapI(B) primer를 이용하여 HLSA 유전자를 증폭시켜 이를 pGEM-T vector에 ligation한 후, *E. coli* TOP10F'으로 형질전환 하였다. 형질전환체로부터 플라스미드를 추출하여 제한효소 SapI 으로 절단하여 SapI-HLSA-SapI fragment를 획득하였다. 이 fragment를 SapI로 절단한 pBHA-TWIN 1에 삽입하여 pTWIN 1-HLSA를 제작하였다. pTWIN 1-HLSA를 EcoRI으로 절단하여 얻은 Ssp(CBD)-HLSA-Mxe fragment를 동일한 효소로 절단한 pPIC9K-700k로 삽입하였다. 이 재조합 플라스미드를 *E. coli*에 형질전환 하고 이로부터 얻은 플라스미드를 pPIC9K-THLSA라 명명하였다. 이 플라스미드를 제한효소 SalI으로 절단하여 선형으로 만든 후 이를 *P. pastoris* SMD1168에 electroporation에 의해 형질전환 하였다. 형질전환은 Jo 등 (1)의 방법을 약간 변형하여 사용하였다. *P. pastoris* SMD1168의 chromosomal DNA에 integration된

transformant를 찾기 위해 MD 한천 배지에 도말 한 다음 나타나는 colony를 선별하였다. 선별된 colony들로부터 chromosomal DNA를 추출한 후, 프라이머 HLSA check(F)와 TIP630-NotI (R)을 사용하여 PCR을 수행하여 integration을 확인하였다. 선별된 재조합 균주를 *P. pastoris* THLSA라 명명하였다

### 5. 유전자 발현 및 확인

재조합 균주 *P. pastoris* THLSA로부터 Ssp(CBD)-HLSA-Mxe의 발현을 유도하기 위하여 BMGY 배지에서 30. C, 230 rpm 조건으로 하룻밤 배양하였다. 이를 새로운 BMMY media로 OD<sub>600</sub>=1.0으로 조정하여 30. C, 230 rpm 에서 96시간 배양하였다. 단백질 발현 유도를 위해 1% (v/v) 메탄올을 24시간 마다 첨가하였다.

Ssp-HLSA-Mxe의 발현을 확인하기 위해서 Western blotting을 수행하였다. 배양이 끝난 BMMY 배양액을 4°C, 3,500 ×g에서 10분간 원심분리 하여 cell을 harvest 하였다. Harvest 한 cell을 1.2 mm glass beads와 buffer A(20 mM Tris/HCl, pH 7.5, 200 mM NaCl, and 5 mM MgCl<sub>2</sub>)를 넣어 Bead beater로 lysis한 후, 원심 분리하여 상층액을 회수하였다. 이 후 sample buffer(2% SDS, 0.1% bromophenol blue, 10% glycerol, 100 mM dithiothreitol, 6 M urea, 50 mM Tris-Cl, pH 6.8)를 넣고 10분간 boiling하여 SDS-PAGE를 수행하였다. SDS-PAGE가 끝난 gel을 polyvinylidene difluoride (PVDF) membrane으로 electrophoretic transfer하고 primary antibody로 mouse anti-CBD antibody (NEB, Inc.), secondary antibody로 goat anti-mouse IgG AP antibody (Enzo, Inc.)를 사용하여 발현을 확인하였다. 기질로는 4-nitrobluetetrazolium chloride (NBT)와 5-bromo-4-chloro-3-indolyl phosphate (BCIP) (Thermo scientific Inc.)를 사용하였다.



## 6. Immunofluorescence 와 Flow cytometric analysis

Ssp(CBD)-HLSA-Mxe 융합단백질의 표면발현을 확인하기 위하여 FITC-labeled Immunofluorescence 와 Flow cytometric analysis를 진행하였다 (1). Methanol 첨가에 의한 발현 유도가 끝난 cell을 원심분리(4. C, 6,000 xg, 10 min)하고, PBS로 3번 wash 한 후 300 $\mu$ l의 PBS (containing 1mg BSA/ml, mouse anti-CBD antibody(1:300)) 첨가하여 1시간 반응시켰다. 이를 다시 원심분리하고 PBS로 3번 wash 한 다음 PBS 300  $\mu$ l(containing 1mg BSL/ml, anti-mouse antibody(1:300), FITC conjugated)를 처리하여 30분 반응시켰다. 반응액을 원심분리 후 PBS로 3번 wash 한 다음 FITC-labeled cell을 confocal laser scanning microscope를 사용하여 분석하였다. 또한 융합 단백질의 표면발현은 Facscalibur flow cytometer에 의해 분석하였으며 이때 negative control 로서 *P. pastoris* SMD1168을 사용하였다.

## 7. HLSA의 분리 및 확인

발현된 Ssp(CBD)-HLSA-Mxe로부터 HLSA만을 분리하기 위하여 IMPACT (Intein Mediated Purification an Affinity Chitin-binding Tag) system, NEB, Inc.)을 이용하였다. 원심 분리로 회수된 균을 buffer A로 resuspension하고 bead beater를 이용하여 cell lysis를 수행하였다. 원심 분리 후 상층액을 chitin resin(NEB, Inc.)이 packing된 column에 flow시켜 Ssp(CBD)-HLSA-Mxe를 resin에 binding시켰다. Binding을 마 친 후 resin의 20배 volume의 buffer B1 (500 mM NaCL, 20 mM HEPES, pH 8.5, 1 mM EDTA)으로 non-binding protein들을 제거하고 buffer B3 (500 mM Nacl, 20 mM HEPES, pH8.5, 1 mM EDTA, 40 mM DTT)를 넣어 40 시간 incubation함으로써 Ssp-HLSA와 Mxe intein의 cleavage를 유도하였다. 다시 resin의 20배 volume의

buffer B1으로 cleavage된 fragment를 제거하고 buffer B2 (500 mM NaCl, 20 mM Na-HEPES, pH 6.5, 1 mM EDTA)를 soaking하여 Ssp와 HLSA의 cleavage를 유도하였다. Cleavage reaction이 끝난 후 buffer B1을 사용하여 HLSA를 회수하였다.

## 결과 및 고찰

### 1. HLSA 세포 표면 발현을 위한 재조합 균주 제조

Ssp(CBD)-HLSA-Mxe 융합단백질을 *P. pastoris* 에서 발현시키기 위해 유전자 합성을 통하여 codon usage에 맞게 optimization 시켰다. 또한 HLSA를 세포 표면에 발현 시키기 위해 *S. cerevisiae* 의 *TIP* 유전자를 사용하여 재조합 플라스미드 pPIC9K-THLSA를 제작하였다(Figure 1). pPIC9K에는  $\alpha$ -factor signal sequence와 메탄올에 의해 발현이 유도되는 *AOX1* 프로모터를 가진다. HLSA의 세포표면 발현을 위해서 HLSA를 *TIP* 유전자와 연결하여 *AOX1* 프로모터 하에 발현되도록 하였다. 완성된 재조합 플라스미드는 HLSA check(F)와 TIP630-NotI(R) primer를 이용한 PCR을 통하여 올바른 발현 방향을 확인 하였다. 그 결과 1.6kb 의 DNA가 증폭된 것을 확인하였으며 (Figure 3), sequencing한 결과 염기서열이 정확히 일치하였다 (data not shown).

완성된 재조합 플라스미드 pPIC9K-THLSA를 SalI을 이용하여 선형으로 만든 후, electroporation을 통해 *P. pastoris* SMD1168에 형질전환 시킨 후 genomic DNA에 삽입 시켰다(Figure 2). 그 후 MD 한천 배지에 나타난 colony들을 선별한 다음, genomic DNA를 추출하여 PCR을 통해 chromosomal integration을 확인하였다. Ssp intein의 중간 부분에서부터 *TIP* 유전자 사이를 PCR로 증폭시킴으로써

삽입여부를 확인하였으며, 이 때 primer는  $\alpha$ -factor (F)와 HLSA(R)를 사용하였다. 그 결과 1.1kb 의 DNA 가 증폭되었으며 sequencing을 통하여 최종적으로 염기서열이 일치하는 것을 확인하였다(Figure 3). 이 재조합 균주를 THLSA로 명명하였다.

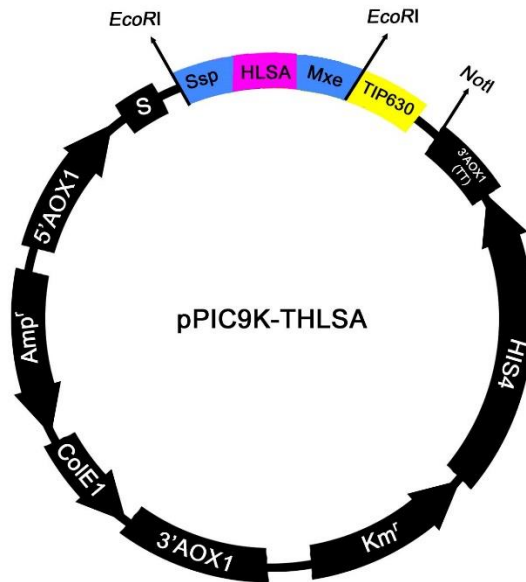


Figure 1. Structure of recombinant plasmid pPIC9K-THLSA. HLSA gene is controlled by the *AOX1* promoter for the expression of HLSA. For the location at the surface of the cell, TIP anchoring motif is used (see details in textbook)

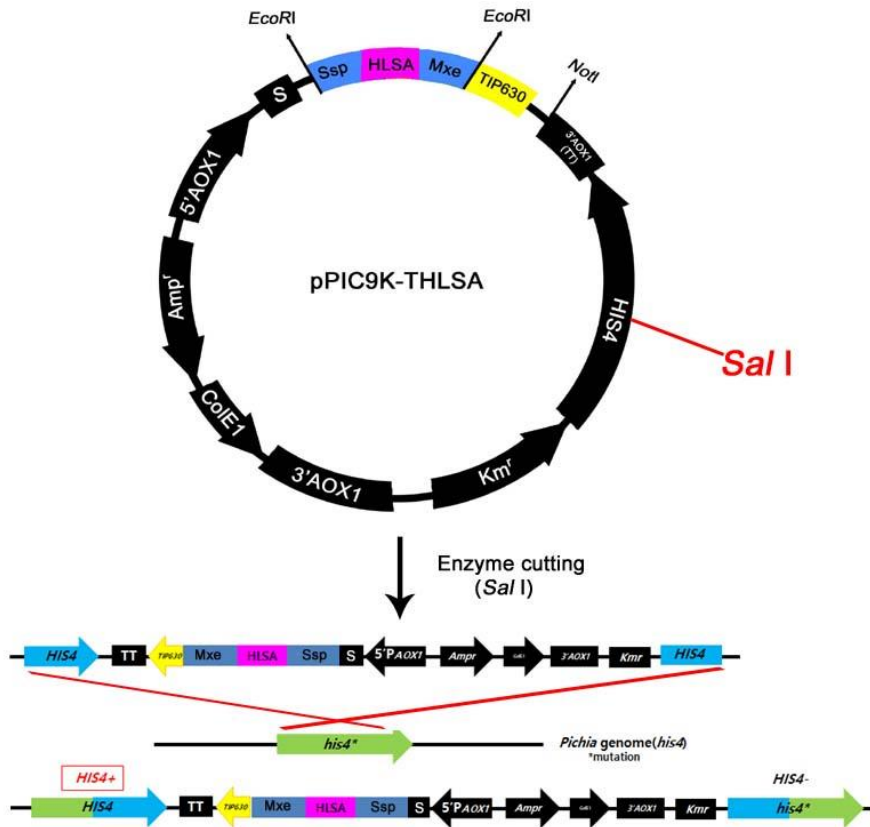


Figure 2. Schematic diagram of integration strategy into the chromosome of *Pichia pastoris* SMD1168.

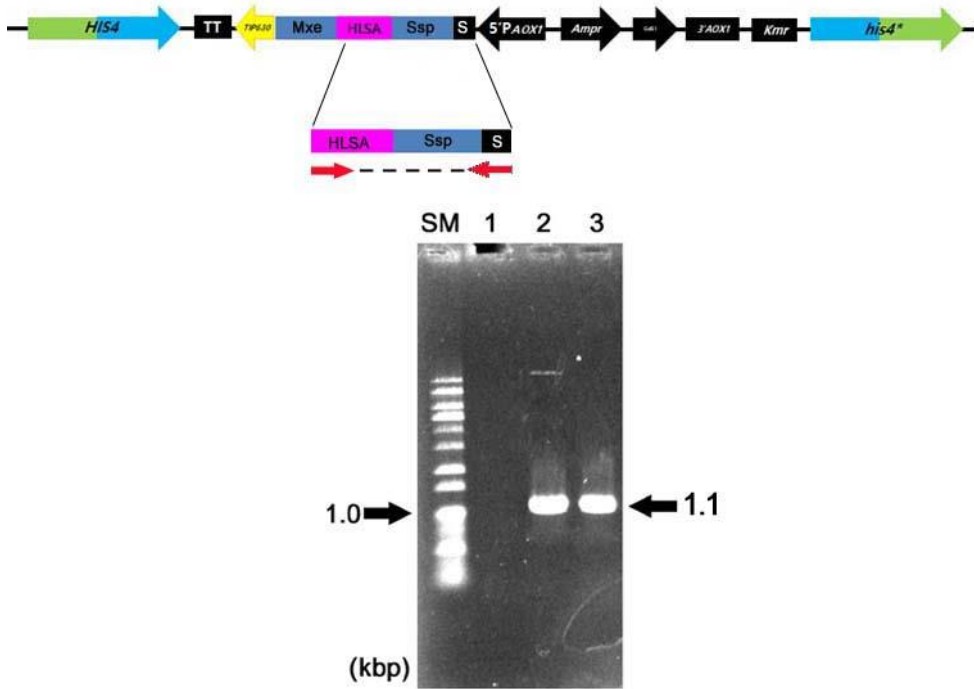


Figure 3. Confirmation of integration of Ssp(CBD)-HLSA-Mxe by PCR. SM: DNA size marker, Lane 1: Negative control (*P. pastoris* SMD1168), Lane 2: Positive control (pPIC9K-THLSA), Lane 3: THLSA. Arrows(red) indicate primers used for the PCR.

## 2. 융합 단백질 발현 확인

발현된 Ssp(CBD)-HLSA-Mxe intein 융합단백질을 확인하기 위해 Western blotting을 실시하였다. 재조합 균주 THLSA를 배양하여 methanol로 발현을 유도한 후 세포를 파쇄한 후 SDS-PAGE와 Western blotting에 의해 확인한 결과 약 72 kDa의 단백질을 확인 하였다 (Figure 4). 이는 Ssp(CBD) intein (26 kDa), Mxe intein (23 kDa), TIP (22 kDa), 그리고 HLSA (1.5 kDa)를 가진 Ssp(CBD)-HLSA-Mxe intein의

분자량과 일치하였다. 따라서 HLSA를 포함하는 융합단백질이 재조합 균주 THLSA로부터 발현됨을 알 수 있다.

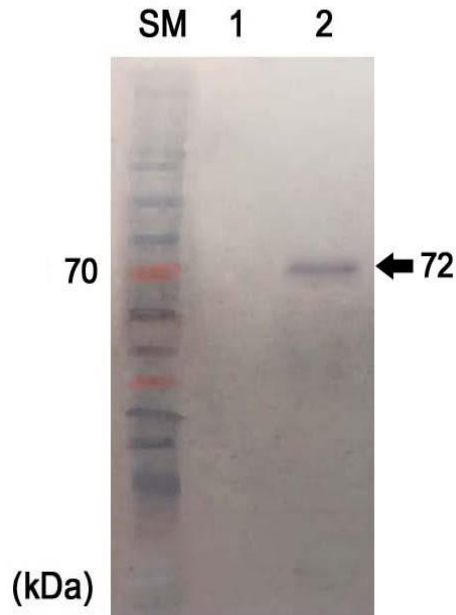


Figure 4. Identification of fused protein Ssp(CBD)-HLSA-Mxe intein expressed from recombinant *P. pastoris* THLSA by Western blotting. SM : protein size marker, Lane 1: Negative control (*P. pastoris* SMD1168), Lane 2 : THLSA. Arrow indicate fused protein containing HLSA.

### 3. 융합단백질 Ssp(CBD)-HLSA-Mxe의 세포표면 발현 확인

발현된 Ssp(CBD)-HLSA-Mxe의 위치를 확인하기 위하여 immunofluorescence 및 confocal microscopy로 분석 하였다. Mouse anti-CBD와 FITC-conjugated anti-mouse antibody를 사용하여 반응한 후 confocal microscope를 사용하여 재조합 균주 THLSA의 세포표면에서 형광을 관찰하였다. 그러나 음성 세포인 *P. pastoris*

SMD1168의 표면에는 관찰되지 않았다(Figure 5A). 발현된 융합단백질이 세포 표면에 위치하는 것은 Flow cytometric analysis에 의해서도 증명되었다. Figure 5B에서 보는 바와 같이 융합 단백질 Ssp(CBD)-HLSA-Mxe가 발현된 재조합 균주 *P. pastoris* THLSA만이 CBD antibody에 반응하며 negative control인 *P. pastoris* SMD1168에는 반응을 보이지 않았다. 이는 융합단백질이 *P. pastoris*의 세포 표면에 발현되었음을 나타내는 것이다.

#### 4. Ssp-HLSA-Mxe complex로부터 HLSA의 분리 및 정제

발현된 Fused HLSA의 Ssp intein에는 CBD가 존재한다. 본 연구에서는 column내에 packing되어 있는 chitin resin과 CBD의 affinity를 이용하여 분리 및 정제를 진행하였다. Cell lysis 후 lysate를 column에 흘려 융합단백질 Ssp(CBD)-HLSA-Mxe intein를 chitin resin에 binding시켜주었다. Mxe intein과 Ssp intein의 cleavage를 위하여 buffer B3와 buffer B2를 각각 순서대로 처리하였다. 반응이 끝난 후 HLSA를 회수 하였다. 한편 Ssp(CBD)-HLSA로부터 HLSA가 절단되었음을 확인하기 위하여 chitin resin으로부터 Ssp(CBD)도 회수 하였다. Cleavage의 여부를 확인하기 위해 Western blotting을 수행한 결과 회수된 HLSA fraction으로부터는 HLSA를 확인할 수 없었다. 이는 HLSA가 아미노산 11개로 구성된 아주 작은 분자량을 가지기 때문에 SDS-PAGE에서 확인이 힘들기 때문인것으로 사료된다. 그러나 절단 반응이 끝난 후 chitin resin으로부터 회수된 Ssp(CBD)가 분자량 26 kDa에서 확인되었다. 이는 Ssp(CBD)-HLSA로부터 HLSA가 절단되었음을 간접적으로 나타내는 것으로 사료된다.

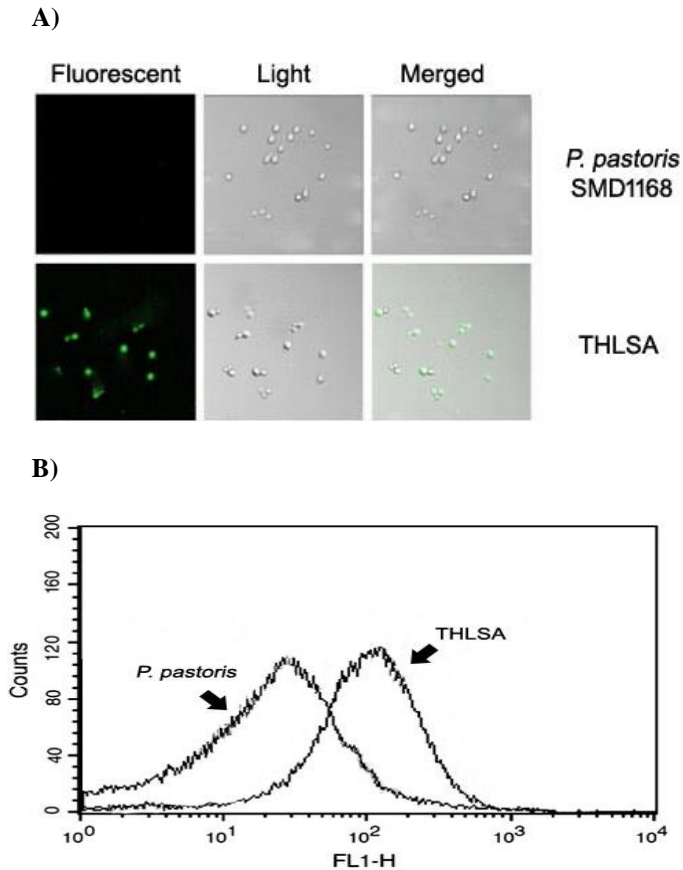


Figure 5. Immunofluorescence and localization of the fused Ssp(CBD)-HLSA-Mxe intein in recombinant *P. pastoris* THLSA. A) Fluorescence and confocal laser scanning microscopy. Cells were labeled by FITC-conjugated antibody. Fluorescence microscope (left column), light microscope (middle column) and merged image (right column) are shown. B) Flow cytometric analysis. *P. pastoris* SMD1168 harboring pPIC9K was used as control



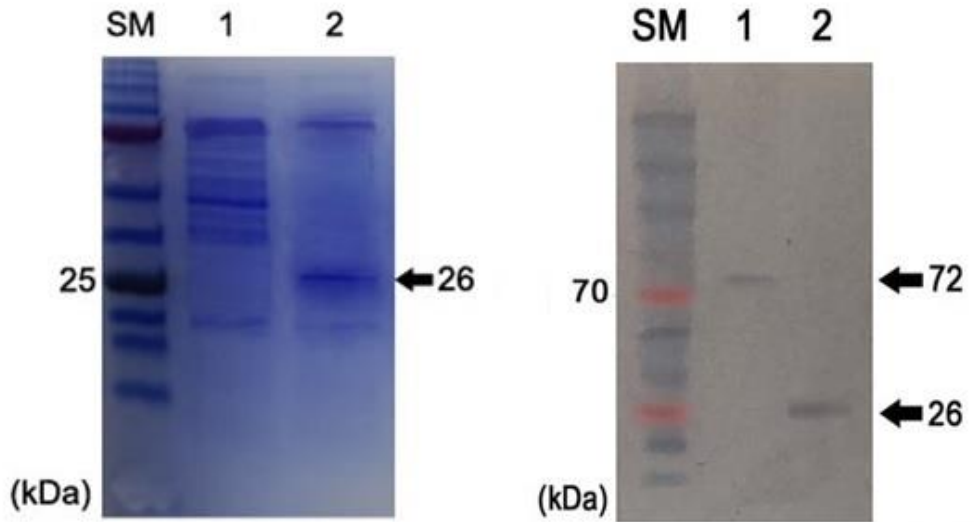


Figure 6. Confirmation of cleavage of HLSA from Ssp(CBD)-HLSA-Mxe intein by SDS-PAGE and Western blotting. SM : protein size marker, Lane 1 : Cell lysate of THLSA, Lane 2: Isolated Ssp(CBD) intein after cleavage from chitin resin (see details in text).

### Acknowledgements

이 연구는 2017년 한국외국어대학교 학술연구비 지원에 의해 이루어진 것임.

## 참고문헌

1. Jo J, Im E, Kim S, Lee H. Surface display of human lactoferrin using a glycosylphosphatidylinositol-anchored protein of *Saccharomyces cerevisiae* in *Pichia pastoris*. *Biotechnol Lett*. 2011;33(6):1113-1120.
2. 원수진, 조재형, 김승환, 권혁진, 이현환. *Pichia pastoris* 에서 사람 락토페린 N-lobe 의 발현과 항균활성. *미생물학회지*. 2015;51(3):271-279.
3. Bellamy W, Takase M, Yamauchi K, Wakabayashi H, Kawase K, Tomita M. Identification of the bactericidal domain of lactoferrin. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology*. 1992;1121(1-2):130-136.
4. Haversen L, Kondori N, Baltzer L, et al. Structure-microbicidal activity relationship of synthetic fragments derived from the antibacterial alpha-helix of human lactoferrin. *Antimicrob Agents Chemother*. 2010;54(1):418-425.
5. Tomita M, Wakabayashi H, Shin K, Yamauchi K, Yaeshima T, Iwatsuki K. Twenty-five years of research on bovine lactoferrin applications. *Biochimie*. 2009;91(1):52-57.
6. Turchany JM, Aley SB, Gillin FD. Giardicidal activity of lactoferrin and N-terminal peptides. *Infect Immun*. 1995;63(11):4550-4552.
7. Wakabayashi H, Yamauchi K, Takase M. Lactoferrin research, technology and applications. *Int Dairy J*. 2006;16(11):1241-1251.

8. Yamauchi K, Tomita M, Giehl TJ, Ellison RT,3rd. Antibacterial activity of lactoferrin and a pepsin-derived lactoferrin peptide fragment. *Infect Immun*. 1993;61(2):719-728.
9. Daly R, Hearn MT. Expression of heterologous proteins in *Pichia pastoris*: A useful experimental tool in protein engineering and production. *Journal of molecular recognition*. 2005;18(2):119-138.
10. Chen YQ, Zhang SQ, Li BC, et al. Expression of a cytotoxic cationic antibacterial peptide in *Escherichia coli* using two fusion partners. *Protein Expr Purif*. 2008;57(2):303-311.
11. Lueking A, Holz C, Gotthold C, Lehrach H, Cahill D. A system for dual protein expression in *Pichia pastoris* and *Escherichia coli*. *Protein Expr Purif*. 2000;20(3):372-378.
12. Ellison RT,3rd, Giehl TJ. Killing of gram-negative bacteria by lactoferrin and lysozyme. *J Clin Invest*. 1991;88(4):1080-1091.
13. Seung-Hawn, Kim., Lee HH. Chromosomal integration of the vitreoscilla hemoglobin gene (vgb) enhances cell growth and expression of human lactoferrin displayed on the surface of *Pichia pastoris*. *GRADUATE SCHOOL HANKUK UNIVERSITY OF FOREIGN STUDIES*. 2015:10-21.
14. 김성집. *Pichia pastoris* 에서 glycosylphosphatidylinositol(GPI)-anchored protein 을 이용한 cyclodextrin glycosyltransferase(CGTase) 세포표면발현 균주 개발. *한국외국어대학교 대학원*. :7-9.

15. 송가은. *Escherichia coli* 내에서 항균활성을 갖는 재조합 사람 락토페리신의 발현 및 활성. *한국외국어대학교 대학원*. 2012:8-10.
16. Lonnerdal B. Nutritional roles of lactoferrin. *Curr Opin Clin Nutr Metab Care*. 2009;12(3):293-297.
17. Tursi A, Elisei W, Brandimarte G, Giorgetti GM, Modeo ME, Aiello F. Effect of lactoferrin supplementation on the effectiveness and tolerability of a 7-day quadruple therapy after failure of a first attempt to cure helicobacter pylori infection. *Med Sci Monit*. 2007;13(4)

## Characterization and identification of twelve bacterial strains isolated from fresh water and marine environments<sup>1</sup>

Heeyoung Kang, Haneul Kim and Kiseong Joh<sup>2</sup>

**ABSTRACT:** Twelve aquatic bacteria were isolated from fresh water and marine environments in Korea. These isolates were mainly assigned to the phylum *Proteobacteria* and class *Flavobacteria*. Phylogenetic analysis based on 16S rRNA gene sequence revealed that strains HME8514, HME8540, HME8620, HME8633, HMF4573, HMF4667, HME8202, HME8262, HME8336, HMF2457, HMF4132 and HMF6043 were most closely related to *Acinetobacter kookii* (98.6 % sequence similarity), *Flavobacterium aquaticum* (98.8 %), *Massilia terrae* (99.9 %), *Lysobacter olei* (98.7 %), *Brevundimonas variabilis* (100 %), *Brevundimonas staleyii* (98.8 %), *Maribacter lutimaris* (100 %), *Litoreibacter ponti* (98.9 %), *Maribacter litorisediminis* (99.7 %), *Maricaulis maris* (99.0 %), *Bacillus algicola* (100 %) and *Maricaulis maris* (99.3 %). In this study, basic informations including Gram reaction, colony and cell morphology, biochemical characteristics, isolation source and strain IDs were described.

### Introduction

Water covers more than three-quarters of the Earth's surface, comprising the fresh water and marine ecosystems. Fresh water environments have extremely low concentrations of salt, and include some wetlands, streams, rivers, ponds and lakes. In contrast, marine regions have higher concentrations of salt, and include some wetlands, salt marshes, mangroves, lagoons and bays. Bacteria and other microorganisms have been reported to live in aquatic ecosystems and play foundational roles, capturing the sun's energy through photosynthesis, and decomposition of organic matter, releasing nutrients stored in organic tissues (Shklar-McAuley *et al.*, 1986). The studies of aquatic bacterial communities have been shown to be characterized by *Proteobacteria*, *Actinobacteria*, *Bacteroidetes*, *Cyanobacteria*, *Verrucomicrobia* and *Planctomycetes* (Logue *et al.*, 2008; Gajigan *et al.*, 2018). The *Proteobacteria* are particularly recognized Gram-negative bacteria of agricultural, industrial, and medical significance (Kersters *et al.*, 2006). The *Cytophaga-Flavobacteria*, belonging to

---

<sup>1</sup> This work was supported by Hankuk University of Foreign Studies Research Fund in 2018 and the project on survey of indigenous species of Korea of the National Institute of Biological Resources (NIBR) under the Ministry of Environment (MOE).

<sup>2</sup> Department of Bioscience and Biotechnology, Hankuk University of Foreign Studies

the phylum *Bacteroidetes*, are chemoorganotrophic and especially proficient in degrading various biopolymers (Kirchman *et al.*, 2002). We have isolated about 160 bacterial strains from the domestic aquatic environments in Korea for 6 years, and reported their characteristics. Here, we described the phylogenetic, morphological and biochemical characteristics of twelve isolated bacterial strains.

## Materials and Methods

Bacterial isolates were obtained from fresh water and marine environment sources. The fresh water samples were collected from reservoir of Hankuk University of Foreign Studies campus, Gyeong-an stream and wetland located in Yongin and Gwangju city. Marine environment samples such as seawater, lagoon and sand of seashore were collected from Buan, Gangneung, Boseong and Jeju. The collected samples were allowed to grow on R2A and marine agars by standard dilution-plating method. These plates were cultured at 25°C, 30°C and 37°C for 2-3days, routinely. The pure cultures were obtained by repetitive streaking and preserved at -80 °C in distilled water supplemented with 20 % (v/v) glycerol. Colony morphology was observed on agar plates after cells were cultivated to their stationary phase. Cellular morphology and cell size were examined using either transmission electron or scanning electron microscopy. Gram reaction was carried out using a Gram-staining kit according to the standard procedures. API 20NE strips (bioMérieux) were inoculated according to the manufacturer's instructions. Oxidase activity determined using 1 % *N,N,N,N*-tetramethyl-*p*-phenylenediamine dihydrochloride (Sigma). Bacterial identification was performed by amplification and sequencing of 16S rRNA genes. Amplification of 16S rRNA gene was carried out as described Woodman *et al* (2008). The 16S rRNA PCR product was sequenced using an Automatic Sequencer 3730XL (Applied Biosystems) with universal primers (27F and 1492R; Lane, 1991). The sequences of twelve strains and related bacterial species were obtained from EzBioCloud server and edited using the BioEdit program (Yoon *et al.*, 2018; Hall, 1999). Multiple alignment was performed using the CLUSTAL\_W program (Thompson *et al.*, 1994). The phylogenetic tree was constructed by using neighbor-joining method within MEGA7 software package with bootstrap resampling method with 1000 replicates (Saitou & Nei, 1987; Kumar *et al.*, 2016; Felsenstein, 1985).

## Results and Discussion

The six bacterial strains, isolated from fresh water environments, were distributed into five families of the phyla *Proteobacteria* and *Bacteroidetes*: Strains HMF4573 and HMF4667 were most closely related to *Brevundimonas variabilis* (100 %, 16S rRNA gene sequence similarity) and *Brevundimonas staleyii* (98.8 %) within family *Caulobacteraceae* of the class *Alphaproteobacteria*. Strains HME8620, HME8514, HME8633 and HME8540 were most closely related to *Massilia niabensis* (98.3 %, belonging to the family *Oxalobacteraceae* of within *Betaproteobacteria*), *Acinetobacter kookii* (98.6 %, belonging to the family *Moraxellaceae* within *Gammaproteobacteria*), *Lysobacter olei* (98.7 %, belonging to the family *Lysobacteraceae* within *Gammaproteobacteria*) and *Flavobacterium aquaticum* (98.8 %, belonging to the family *Flavobacteriaceae* within *Bacteroidetes*), respectively.

Marine bacterial isolates were assigned the phyla *Bacteroidetes*, *Firmicutes* and *proteobacteria*. Strains HME8202 and HME8336 were related to *Maribacter lutimaris* (100 %) and *Maribacter litorisediminis* (99.7 %) within family *Flavobacteria*, respectively. Strain HME8262 was a member of species *Litoreibacter ponti* (98.9 %, belonging to the family *Rhodobacteraceae* within *Alphaproteobacteria*) and strain HMF4132 was a members of species *Bacillus algicola* (100 %, belonging to the family *Bacillaceae* within *Bacilli*), respectively. Strains HMF2457 and HMF6043 were related to *Maricaulis maris* (99.0 % and 99.3 %, belonging to the family *Hyphomonadaceae* within *Alphaproteobacteria*). The phylogenetic analysis results of twelve isolates present in Figure 1 and Table1. The isolated strains were mainly Gram-stain-negative, rod-shaped and oxidase produced, but strain HMF4132 was Gram-stain-positive, and oxidase activity of strain HME8514 was negative. The detailed morphological (Figure 2) and biochemical characteristics of twelve strains are also shown in the species description section.

### Description of *Flavobacterium aquaticum* HME8540

Cells are Gram-stain-negative, non-flagellated and rod-shaped. Colonies are circular, convex, entire and yellow-colored after incubation at 30 °C on R2A agar for 2days. Oxidase is produced. With API 20NE strips, nitrate reduction, indole production, glucose

fermentation, aesculin hydrolysis, arginine dihydrolase, urease, gelatinase and  $\beta$ -galactosidase (PNPG test) activities and assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetyl-glucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid are negative. Strain HME8540 (=NIBRBAC000002219) was isolated from a fresh water stream located on Yongin, Republic of Korea. The 16S rRNA gene sequence of strain HME8540 has been deposited in the GenBank database under accession no. KC157036.

#### **Description of *Brevundimonas variabilis* HMF4573**

Cells are Gram-stain-negative, non-flagellated and rod-shaped. Colonies are circular, convex, entire and orange-colored after incubation at 25 °C on R2A agar for 3days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis and  $\beta$ -galactosidase (PNPG test) activity and assimilation of D-glucose, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid and phenylacetic acid are positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase urease and gelatinase activities and assimilation of L-arabinose, D-mannose, D-mannitol, *N*-acetyl-glucosamine and trisodium citrate are negative. Strain HMF4573 (=NIBRBAC000498442) was isolated from fresh water wetland, located on Gwangju, Republic of Korea. The 16S rRNA gene sequence of strain HMF4573 has been deposited in the GenBank database under accession no. KY047394.

#### **Description of *Brevundimonas staleyi* HMF4667**

Cells are Gram-stain-negative, flagellated and rod-shaped. Colonies are circular, convex, entire and yellow colored after incubation at 25 °C on R2A for 3days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis and assimilation of D-glucose, D-maltose, adipic acid and malic acid are positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase, urease, gelatinase and  $\beta$ -galactosidase (PNPG test) activities and assimilation of L-arabinose, D-mannose, D-mannitol, *N*-acetyl-glucosamine, potassium gluconate, capric acid, trisodium citrate and phenylacetic acid are negative. Strain HMF4667 (=NIBRBAC000498446) was isolated from fresh water wetland, located on Gwangju, Republic of Korea. The 16S rRNA gene sequence of strain HMF4667 has been deposited in the GenBank database under accession no. KY047398.



#### **Description of *Massilia terrae* HME8620**

Cells are Gram-stain-negative, non-flagellated and rod-shaped. Colonies are circular, raised, entire and brown-colored after incubation at 30 °C on R2A agar for 2 days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis and  $\beta$ -galactosidase (PNPG test) activity and assimilation of D-glucose, L-arabinose, D-mannose, *N*-acetyl-glucosamine, D-maltose, potassium gluconate and malic acid are positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase urease and gelatinase activities and assimilation D-mannitol, capric acid, adipic acid, trisodium citrate and phenylacetic acid are negative. Strain HME8620 (=NIBRBAC000002227) was isolated from fresh water reservoir of Hankuk University of Foreign Studies campus, Yongin, Republic of Korea. The 16S rRNA gene sequence of strain HME8620 has been deposited in the GenBank database under accession no. KC134363.

#### **Description of *Lysobacter olei* HME8633**

Cells are Gram-stain-negative, non-flagellated and rod-shaped. Colonies are circular, convex, entire and yellow-colored after incubation at 37 °C on R2A for 2 days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis and  $\beta$ -galactosidase (PNPG test) activity and assimilation of D-glucose, L-arabinose, D-mannose, *N*-acetyl-glucosamine, D-maltose, potassium gluconate and malic acid are positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase urease and gelatinase activities and assimilation D-mannitol, capric acid, adipic acid, trisodium citrate and phenylacetic acid are negative. Strain HME8633 (=NIBRBAC000002228) was isolated from fresh water reservoir of Hankuk University of Foreign Studies campus, Yongin, Republic of Korea. The 16S rRNA gene sequence of strain HME8633 has been deposited in the GenBank database under accession no. KC134364.

#### **Description of *Acinetobacter kookii* HME8514**

Cells are Gram-stain-negative, non-flagellated and ovoid rod-shaped. Colonies are circular, convex, entire and white-colored after incubation at 30 °C on R2A for 2 days. Oxidase is not

produced. With API 20NE strips, assimilation of capric acid and phenylacetic acid is positive, but nitrate reduction, indole production, glucose fermentation, aesculin hydrolysis and gelatinase activities and assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetyl-glucosamine, D-maltose, potassium gluconate, adipic acid, malic acid and trisodium citrate are negative. Strain HME8514 (=NIBRBAC000002215) was isolated from a fresh water stream located on Yongin, Republic of Korea. The 16S rRNA gene sequence of strain HME8514 has been deposited in the GenBank database under accession no. KC201357.

#### **Description of *Maribacter lutimaris* HME8202**

Cells are Gram-staining-negative and rod-shaped. Colonies are circular, convex, smooth and yellow-colored after incubation at 30 °C on marine agar for 2days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis and  $\beta$ -galactosidase (PNPG test) activity are positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase, urease and gelatinase activities and assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetyl-glucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid are negative. Strain HME8202 (=NIBRBAC000001868) was isolated from seawater of the Yellow Sea located on Buan, Republic of Korea. The 16S rRNA gene sequence of strain HME8202 has been deposited in the GenBank database under accession no. JQ420021.

#### **Description of *Maribacter litorisediminis* HME8336**

Cells are Gram-staining-negative and short rod-shaped. Colonies are circular, convex, smooth and yellow-colored after incubation at 30 °C on marine agar for 3days. Oxidase is produced. With API 20NE strips, nitrate reduction, aesculin hydrolysis and  $\beta$ -galactosidase (PNPG test) activity and assimilation of D-glucose and D-mannitol are positive, but indole production, glucose fermentation, arginine dihydrolase urease and gelatinase activities and assimilation of L-arabinose, D-mannose, *N*-acetyl-glucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid are negative. Strain HME8336 (=NIBRBAC000001873) was isolated from seawater of the Yellow Sea located on Buan, Republic of Korea. The 16S rRNA gene sequence of strain HME8336 has been deposited in the GenBank database under accession no. JQ762253.

#### **Description of *Bacillus algicola* HMF4132**

Cells are Gram-staining-positive, non-flagellated and ovoid rod-shaped. Colonies are circular, convex, smooth and yellow-colored after incubation at 30 °C on marine agar for 3 days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis, gelatinase and  $\beta$ -galactosidase (PNPG test) activities and assimilation of D-mannitol and trisodium citrate are positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase and urease activities and assimilation of D-glucose, L-arabinose, D-mannose, *N*-acetylglucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid and phenylacetic acid are negative. Strain HMF4132 (=NIBRBAC000497912) was isolated from a sand of seashore located on Jeju, Republic of Korea. The 16S rRNA gene sequence of strain HMF4132 has been deposited in the GenBank database under accession no. KT984000.

#### **Description of *Maricaulis maris* HMF2457**

Cells are Gram-staining-negative, flagellated and rod-shaped. Colonies are circular, convex, smooth and pale pink-colored after incubation at 25 °C on marine agar for 3days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis and gelatinase activity are positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase, urease and  $\beta$ -galactosidase (PNPG test) activities and assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetylglucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid are negative. Strain HMF2457 (=NIBRBAC000498226) was isolated from lagoon located on Gangneung, Republic of Korea. The 16S rRNA gene sequence of strain HMF2457 has been deposited in the GenBank database under accession no. KP099953.

#### **Description of *Maricaulis maris* HMF6043**

Cells are Gram-staining-negative, flagellated and rod-shaped. Colonies are circular, convex, smooth and beige colored after incubation at 25 °C on marine agar for 3days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis and gelatinase activity are positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase,

urease and  $\beta$ -galactosidase (PNPG test) activities and assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetyl-glucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid are negative. Strain HMF6043 (=NIBRBAC000498449) was isolated from seawater of the South Sea located on Boseong, Republic of Korea. The 16S rRNA gene sequence of strain HMF6043 has been deposited in the GenBank database under accession no. KY047401.

#### **Description of *Litoreibacter ponti* HME8262**

Cells are Gram-staining-negative and rod-shaped. Colonies are circular, convex, entire and white-colored after incubation at 30 °C on marine agar for 3days. Oxidase is produced. With API 20NE strips, aesculin hydrolysis is positive, but nitrate reduction, indole production, glucose fermentation, arginine dihydrolase, urease, gelatinase and  $\beta$ -galactosidase (PNPG test) activities and assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, *N*-acetyl-glucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid are negative. Strain HME8262 (=NIBRBAC000001870) was isolated from seawater of the Yellow Sea located on Buan, Republic of Korea. The 16S rRNA gene sequence of strain HME8262 has been deposited in the GenBank database under accession no. JQ687345.

#### **Acknowledgements**

This work was supported by Hankuk University of Foreign Studies Research Fund in 2018 and the project on survey of indigenous species of Korea of the National Institute of Biological Resources (NIBR) under the Ministry of Environment (MOE).

## Reference

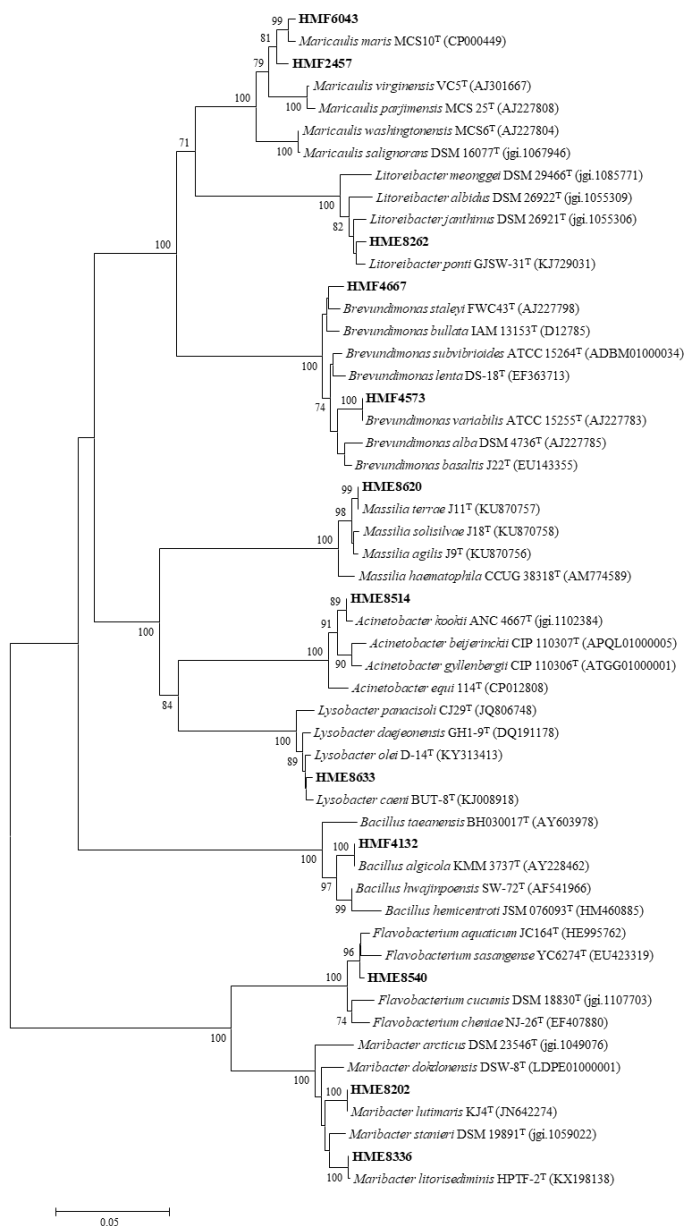
- Felsenstein, J. (1985).** Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 783-791.
- Gajigan, A. P., Yniguez, A. T., Villanoy, C. L., San Diego-McGlone, M. L., Jacinto, G. S. & Conaco, C. (2018).** Diversity and community structure of marine microbes around the Benham Rise underwater plateau, northeastern Philippines. *PeerJ* 6, e4781.
- Hall, T. A. (1999).** BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41, 95-98.
- Kerstens, K., De Vos, P., Gillis, M., Swings, J., Vandamme, P. & Stackebrandt, E. (2006).** Introduction to the Proteobacteria. In *The prokaryotes*, pp. 3-37. Edited by Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, & E. Stackebrandt. New York: Springer.
- Kirchman, D. L. (2002).** The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol Ecol* 39, 91-100.
- Kumar, S., Stecher, G. & Tamura, K. (2016).** MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33, 1870-1874.
- Lane, D. J. (1991).** 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics*, pp. 125-175. Edited by E. Stackebrandt & M. Goodfellow. Chichester: Wiley.
- Logue, J. B., Bürgmann, H. & Robinson, C. T. (2008).** Progress in the ecological genetics and biodiversity of freshwater bacteria. *BioScience* 58, 103-113.
- Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4, 406-425.
- Shklar-McAuley, I. A. & Ferroni, G. (1986).** Psychrophiles, psychrotrophs, and glucose mineralization in a lake with seasonal temperature fluctuations. *J Gen Appl Microbiol* 32, 383-391.
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994).** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673-4680.
- Woodman, M. E. (2008).** Direct PCR of intact bacteria (colony PCR). *Curr Protoc Microbiol* Appendix 3:Appendix 3D.

**Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H. & Chun, J. (2017).** Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* **67**, 1613-1617.

Characterization and identification of twelve bacterial strains  
isolated from fresh water and marine environments

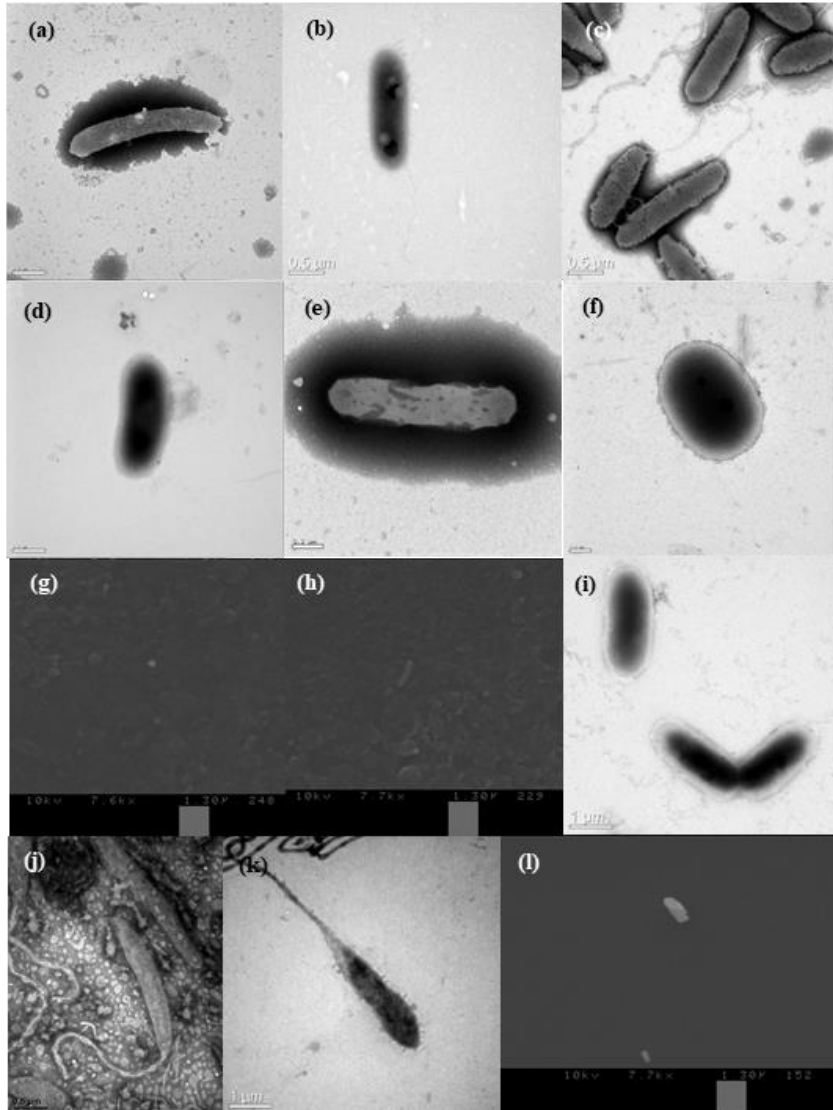
**Table 1.** Summary of isolated strains and their taxonomic affiliations.

Environment	Class	Order	Family	Genus	Strain ID	NIBR ID	Most closely related species	Similarity (%)	Isolation Source	Medium	Incubation condition
Fresh water	<i>Flavobacteria</i>	<i>Flavobacteriales</i>	<i>Flavobacteriaceae</i>	<i>Flavobacterium</i>	HME8540	NIBRBAC000002219	<i>Flavobacterium aquaticum</i>	98.8	Fresh water	R2A	30°C, 2d
					HMF4573	NIBRBAC000498442	<i>Brevundimonas variabilis</i>	100	Wetland	R2A	25°C, 3d
	<i>Alphaproteobacteria</i>	<i>Caulobacterales</i>	<i>Caulobacteraceae</i>	<i>Brevundimonas</i>	HMF4667	NIBRBAC000498446	<i>Brevundimonas staleyi</i>	98.8	Wetland	R2A	25°C, 3d
					<i>Betaproteobacteria</i>	<i>Burkholderiales</i>	<i>Oxalobacteraceae</i>	<i>Massilia</i>	HME8620	NIBRBAC000002227	<i>Massilia terrae</i>
	<i>Gammaproteobacteria</i>	<i>Lysobacteriales</i>	<i>Lysobacteraceae</i>	<i>Lysobacter</i>	HME8633	NIBRBAC000002228	<i>Lysobacter olei</i>	98.7	Fresh water	R2A	37°C, 2d
					<i>Pseudomonadales</i>	<i>Moraxellaceae</i>	<i>Acinetobacter</i>	HME8514	NIBRBAC000002215	<i>Acinetobacter kookii</i>	98.6
Marine	<i>Flavobacteria</i>	<i>Flavobacteriales</i>	<i>Flavobacteriaceae</i>	<i>Maribacter</i>	HME8202	NIBRBAC000001868	<i>Maribacter lutimaris</i>	100	Seawater	MA	30°C, 2d
					HME8336	NIBRBAC000001873	<i>Maribacter litorisediminis</i>	99.7	Seawater	MA	30°C, 3d
	<i>Bacilli</i>	<i>Bacillales</i>	<i>Bacillaceae</i>	<i>Bacillus</i>	HMF4132	NIBRBAC000497912	<i>Bacillus algicola</i>	100	Sand of sea shore	MA	30°C, 3d
					HMF2457	NIBRBAC000498226	<i>Maricaulis maris</i>	99.0	lagoon	MA	25°C, 3d
	<i>Alphaproteobacteria</i>	<i>Rhodobacteriales</i>	<i>Hyphomonadaceae</i>	<i>Maricaulis</i>	HMF6043	NIBRBAC000498449	<i>Maricaulis maris</i>	99.3	Seawater	MA	25°C, 3d
			<i>Rhodobacteraceae</i>		<i>Litoreibacter</i>	HME8262	NIBRBAC000001870	<i>Litoreibacter ponti</i>	98.9	Seawater	MA



**Figure 1.** Neighbor-joining phylogenetic tree based on 16S rRNA gene sequences, showing the relationship between isolated strains and their relatives. Bootstrap values (>70%) are shown. Bar, 0.05 substitutions per nucleotide position.





**Figure 2.** Transmission electron micrographs or scanning electron micrographs of cells. Strains: a, HME8540; b, HMF4573; c, HMF4667; d, HME8620; e, HME8633; f, HME8514; g, HME8202; h, HME8336; i, HMF4132; j, HMF2457; k, HMF6043; l, HME8262.



기초과학연구

제 47집

인쇄 2019년 2월 25일

발행 2019년 2월 25일

발행인  
운영위원  
인쇄처  
발행소  
주소

소장 권혁만  
정미연, 이태욱, 김태경, 박진균, 장태용  
외대문화사  
한국외국어대학교 기초과학연구소  
경기도 용인시 처인구 모현면 외대로 81  
449-791

<비매품>