Early Diagnosis of Productivity Through a Clock Gene Promoter Activity Using a Luciferase Bioluminescence Assay in *Arabidopsis thaliana*

Hirokazu Fukuda¹, Takatoshi Ichino², Takao Kondo³ and Haruhiko Murase⁴  

¹ Department of Applied Life Sciences, Graduate School of Life and Environmental Sciences, Osaka Prefecture University, Sakai 599–8531, Japan  
² Department of Intelligent Systems, School of Biology-Oriented Science and Technology, Kinki University, Kinokawa, Wakayama 649-6493, Japan  
³ Division of Biological Science, Graduate School of Science, Nagoya University, Chikusa-ku, Nagoya, Aichi 464–8601, Japan  

(Received November 25, 2008; Accepted April 22, 2011)

A method for early diagnosis of plant productivity using a luciferase bioluminescence assay at the early stage of cultivation was developed for transgenic *Arabidopsis thaliana* carrying the CCA1::LUC construct, in which the promoter of the CCA1 clock gene had been fused to a modified firefly luciferase (LUC) gene. Because the intensity of luciferase bioluminescence is proportional to the CCA1 promoter activity and the expression level of LUC, the dynamics of both can be investigated by measuring the bioluminescence. In this paper, plant biomass was considered a measure of plant production and LUC protein a measure of foreign protein production. The leaf area and bioluminescence of seedlings at the early stage of cultivation were investigated as indexes for early diagnosis of productivity; leaf area was a reflection of morphological information, and bioluminescence, molecular information. Bioluminescence showed higher correlation to productivity based on biomass and LUC protein than the leaf area under various growth conditions, indicating that bioluminescence is a better index for early diagnosis. In addition, we developed a general statistical method for selecting superior seedlings based on this early diagnosis and a general theory for determining an optimal parameter set for maximizing profit. This early diagnostic methodology appears suited to enhance the quality of the products in plant factories.

Keywords: circadian clock, circadian resonance, medical molecular farming, plant factory, selection of seedling

INTRODUCTION

Early diagnosis predicts susceptibility to injury, diseases and productivity potential in plants (Zhang et al., 2008; Lihua et al., 2008; Lopez-Morales et al., 2008). Using predictive information in the cultivation strategy, the cost of production can be reduced and productivity increased. In computer-controlled plant production systems such as plant factories, the optimization of the production process is realized by diagnosis of the physiological status of the plant and the use of a plant response model (Morimoto and Hashimoto, 2000; Hashimoto et al., 2004; Ushada et al., 2007; Okayama et al., 2008a; Ondimu and Murase, 2008). In recent years, the molecular informa-
tion including mRNA and protein concentrations have been also used. Because such molecular information show the state of target molecular production process, they seem very effective, in particular, for pharmaceuticals production using the genetically modified plants (Hiatt et al., 1989; Daniell et al., 2001; Joh et al., 2005; Kim et al., 2006; Sun et al., 2006; Li et al., 2007; Lim et al., 2011). Okayama et al. (2008b, 2009) developed a simulation model for heterologous protein production in transgenic lettuce, for which the molecular parameters of the model were determined experimentally. This model provided the optimum cultivation cycles of soluble protein production for a head of lettuce grown under various temperature conditions. However, the general theory for early diagnosis of productivity using the molecular information has not been addressed yet.

In this study, we have focused on molecular information about the expression of a pivotal circadian clock gene, *CCA1 (CIRCADIAN CLOCK-ASSOCIATED1)* (McClung, 2006). The circadian clock generates the endogenous oscillation with about a 24-h period, called circadian rhythm, and regulates various physiological processes such as photosynthesis, metabolism, growth and flowering (Harmer et al., 2000; Dodd et al., 2005; McClung, 2006; Ni et al., 2009). Therefore, by monitoring *CCA1* expression, we can obtain the information about various clock-regulated physiological statuses. The expression of *CCA1* can be monitored using a luciferase bioluminescence assay nondestructively and noninvasively (Nakamichi et al., 2004; Fukuda et al., 2007; Fukuda et al., 2008). We have chosen a transgenic *Arabidopsis thaliana CCA1::LUC*, in which a modified firefly luciferase gene has been fused to the *CCA1* promoter. The luciferase gene is expressed at the same time as *CCA1*, generating the luciferase protein LUC. LUC promptly reacts with luciferin provided in the nutrient medium, emits bioluminescence, and then is degraded. Because the intensity of this bioluminescence is proportional to the *CCA1* promoter activity and the production rate of LUC, the dynamics of both can be investigated by measuring the bioluminescence. The total bioluminescence is proportional to the total net of LUC production, namely, total net of a foreign protein production. In addition, *CCA1* promoter has the properties required for the foreign protein production, for example, widespread expression in the whole body, stable rhythmic-expression under various experimental conditions, and controllability of expression by lighting (Fukuda et al., 2008). Therefore, *CCA1::LUC* plant can be considered as a model plant for foreign protein production under the control of the clock-regulated promoter.

In this study, a method for early diagnosis of plant productivity through the clock gene promoter activity at the early stage of cultivation was developed in transgenic *Arabidopsis thaliana CCA1::LUC*. In order to verify the generality of our developed method, we have investigated the productivity under various light conditions. We also developed a general theory for determining an optimal parameter set for maximizing profit.

**MATERIALS AND METHODS**

*Plant material and growing conditions*

Experiments were carried out using transgenic *Arabidopsis thaliana* (Columbia accession, Col) *CCA1::LUC* with the clock gene *CCA1* fused to a modified firefly luciferase gene (Nakamichi et al., 2004). The seeds of *CCA1::LUC* were obtained from a redifferentiated shoot of an established cell line (Nakamichi et al., 2004) using a ordinary redifferentiation (Re’dei, 1962). *CCA1::LUC* plants were grown on gellan gum-solidified Murashige and Skoog plant salt mixture medium (Wako Co., Ltd., Osaka, Japan) with 2% (w/v) sucrose in each aseptic dish (40 mm in diameter) under 12 h light: 12 h dark cycles using 80 μmol m⁻² s⁻¹ of fluorescent white light at 22 ± 1°C for 14 d. About 1 d before starting to monitor bioluminescence, 1 mM luciferin solution (500 μl) dissolved in water containing 0.002% Triton X-100 was injected into the medium using a micropipette (final luciferin concentration: about 0.1 mM).
BIOLUMINESCENCE ASSAYS

Bioluminescence assays were carried out with a monitoring system developed by Kondo et al. (1993). Using this system, bioluminescence was detected from each plate by a photomultiplier tube (PMT) (Hamamatsu H7360-01MOD; Hamamatsu Photonics KK, Japan) enclosed in a light-tight box. Each plate was on a turntable that was rotated under the PMT sequentially every 20 min under control of a computer. Therefore, the plant in each plate was exposed to the dark period for 4 min every 20 min (at least 1.5 min of darkness allows chlorophyll fluorescence to decay). The bioluminescence monitoring system was in a temperature-controlled chamber (MIR-553, SANYO Electric Co., Ltd.) at 22.0±0.5°C.

Light conditions during the measurement of bioluminescence

During the 48 h from the start of measurement, seedlings were under 12 h light: 12 h dark cycles using 80 µmol m⁻² s⁻¹ of red LED light (λp=660 nm), in order to reset the circadian rhythm. After exposure to these light conditions, seedlings were cultivated under several different light conditions, i.e., continuous light (LL) and 10 h light: 10 h dark cycles (LD20), 12 h light: 12 h dark cycles (LD24), and 14 h light: 14 h dark cycles (LD28) using red LEDs (λp=660 nm) with a light intensity of 80 µmol m⁻² s⁻¹.

Analysis of visible leaf area, bioluminescence, and LUC production

The visible leaf area A of a seedling was estimated from the seedling image (Fig. 1a). The integrated values of bioluminescence from 24 h to 48 h after starting measurement was used as the bioluminescence B at seedling stage, indicating the production rate of LUC at the seedling stage. The integrated value of bioluminescence from 48 h to 288 h of measurement was used as the total bioluminescence during cultivation (Fig. 1b, c). The total bioluminescence during cultivation is considered an indicator for the net LUC production L during cultivation, because the intensity of bioluminescence is proportional to the production rate of LUC. After measurement, the aerial part of the plant was detached and dried under 80°C for 24 h, then the dry aerial biomass W was measured.

RESULTS AND DISCUSSION

Dry aerial biomass and net LUC production under various light conditions

Figure 2a and 2b show the dry aerial biomass W and the total net LUC production L, respectively. In the light: dark conditions, the maximization of both W and L has been observed at LD24. This means that LD24 is a better condition but LD20 and LD28 are worse ones for plant growth. This resonance-like phenomenon, called circadian resonance, has been reported by Dodd et al., 2005. In addition, there was high correlation between W and L under various light conditions (Fig. 2c). W and L were distributed in a wide range; these values differed by more than five-fold from the largest to the smallest plant (Fig. 2c). The distributions of W under various light conditions are shown in Fig. 3. The distribution curves were determined by the average and unbiased variance of W obtained from the experiment under the supposition that the distributions followed a Gaussian distribution. The validity of that supposition was confirmed by the Kolmogorov-Smirnov test. Because the differences in the average light intensity and average temperature for each plant were very small in our measurement system, it seems that the wide distribution was caused by the inherent individual differences of the seedlings. Similar distribution curves were obtained for L, although all plants were obtained from an established line and had same number of copies of CCA1::LUC. Therefore, the influence of individual differences cannot be ignored in plant and protein productions.

Increase in productivity by selective harvest

The average biomass of plants cultivated under LL <WLL> was larger than that under LD24 <WLDC> (Fig. 3). However, in plants cultivated under LD24, large plants whose biomass was larger
than $<W_{li}>$ were included (Fig. 3). Therefore, when only these large plants are selectively harvested, the productivity increases. The average biomass of only plants cultivated under LD24 with more than a certain biomass $W^*\$ is described as follows:

$$
\langle W^* \rangle = \frac{\int_{W^*}^{\infty} W \exp \left( -\frac{(W - <W_{LD24}>)^2}{2\sigma_{LD24}^2} \right) dW}{\int_{W^*}^{\infty} \exp \left( -\frac{(W - <W_{LD24}>)^2}{2\sigma_{LD24}^2} \right) dW}
$$

where $\sigma_{LD24}$ is the standard deviation of $W$ under LD24. $<W^*>$ exceeds $<W_{li}>$ above a critical value $W^*\$, as shown in Fig. 4. However, this increase of productivity is superficial effect and the actual profit is not improved.

Increase in productivity by selection of seedlings based on early diagnosis

Improvement of productivity by the selective harvest described above is an apparent effect provided by data handling after cultivation. Therefore, actual profit does not increase due to selective harvest. However, if we select only superior seedlings before cultivation and then cultivate
only them, true productivity improves and the actual profit increases. In the following, we address early diagnosis for selection of superior seedlings before cultivation and its effective increase on productivity.

First, we consider early diagnosis of productivity. Indexes that estimate productivity of seedlings are necessary for early diagnosis. Figure 5 shows the relationships between \( W \) and \( A \), and between \( W \) and \( B \). Because positive correlations were observed for both indexes under various light conditions, \( A \) and \( B \) can be used as indexes for early diagnosis. Under each of the conditions, the correlation coefficient \( R_{AB} \) between \( W \) and \( B \) was larger than the correlation coefficient \( R_{AW} \) between \( W \) and \( A \). Therefore, bioluminescence is a better index for early diagnosis of biomass than leaf area.

Figure 6 shows the relationships between the LUC production \( L \) and \( A \), and between \( L \) and \( B \). Because positive correlations were observed for both indexes under various light conditions, \( A \) and \( B \) can be used as indexes for early diagnosis. Under each of the conditions, the correlation coefficient \( R_{LB} \) between \( L \) and \( B \) was larger than the correlation coefficient \( R_{LA} \) between \( L \) and \( A \). Therefore, bioluminescence is also a better index for early diagnosis of LUC production than leaf area.

Next, we consider the improvement of productivity by the selection of seedlings based on early diagnosis. When the production \( P \) and the index \( I \) are distributed as Gaussian distributions, the average of production \( <P>_* \) that consists of only selected plants is described by:

\[
<P>_* = \frac{\int_{-\infty}^{\infty} y(I) \exp \left( -\frac{(I - \langle I \rangle)^2}{2\sigma_i^2} \right) dI}{\int_{-\infty}^{\infty} \exp \left( -\frac{(I - \langle I \rangle)^2}{2\sigma_i^2} \right) dI}
\]

where \( I_* \) is the threshold value of selection. \( y(I) \) is the regression line between \( P \) and \( I \) as follows:

\[
y(I) = \langle P \rangle + R \frac{\sigma_P}{\sigma_i} (I - \langle I \rangle)
\]

where \( \langle P \rangle \) is the average of production that consists of all plants, \( R \) is the correlation coefficient between \( P \) and \( I \), and \( \langle I \rangle \) is the average of the index. \( \sigma_P \) and \( \sigma_i \) are the estimators of the standard deviation of \( P \) and \( I \), respectively. In addition, we introduce \( Z \) according to

\[
Z = \frac{I - \langle I \rangle}{\sigma_i}
\]

Equation 1 can be normalized by \( Z \) as follows:

\[
\frac{<P>_*}{<P>} = \frac{\int_{-\infty}^{\infty} Z \exp \left( -\frac{Z^2}{2} \right) dZ}{\int_{-\infty}^{\infty} \exp \left( -\frac{Z^2}{2} \right) dZ}
\]

where \( Z_* \) is the normalized threshold value of selection according to:

\[
Z_* = \frac{I_* - \langle I \rangle}{\sigma_i}
\]

and \( <P>_*/<P> \) is the rate of improvement of productivity by the selection. Figure 7 shows \( <P>_*/<P> \) as a function of \( Z_* \) obtained from Eq. 2. \( <P>_*/<P> \) increased with \( R \). Therefore, bioluminescence, which has a higher correlation to production \( (R_{AW} \text{ and } R_{AB} > 0.7) \), can be used for early diagnosis. \( <P>_*/<P> \) depends on both \( R \) and \( \sigma_i/<P> \). Therefore, to further improve productivity, both \( R \) and \( \sigma_i/<P> \) should be large, as shown in Fig. 8. In addition to \( R \), an enhancement factor for early diagnosis is \( \sigma_i/<P> \).
Fig. 2  Dry aerial biomass $W$ (a) and net LUC production $L$ (b) of CCAI::LUC plant under various light conditions. (c) Correlation between $W$ and $L$. The lines are the regression lines between $W$ and $L$ in each condition. (a,b) In two-sample t tests comparing to LD24, *** and NS indicate significances at $p=0.01$, 0.05 and not significant, respectively. Error bars indicate SE ($n=38$ in LD24, $n=40$ in others).

Fig. 9  Average profit $Y$ as a function of the threshold of selection $Z^*$. (a) Dependence on $\beta$, the ratio of the average cost of raising seedling $C_s$ to the average cost of plant cultivation $C_p$. (b) Dependence on the correlation coefficient $R$ between the production $P$ and the index $Z$.

Fig. 8  Dependence of an increase in productivity $<P>_*/<P>$ on the correlation coefficient $R$ and the relative standard deviation $\sigma_P/<P>$. The numbers in the figure indicate $<P>_*/<P>$. The values are shown for $Z^*=2$.
Fig. 5  Relationship between dry aerial biomass and the investigated indexes under various light conditions. (a–d) Correlation between dry aerial biomass \( W \) and visible leaf area \( A \). (e–h) Correlation between \( W \) and bioluminescence \( B \). \( R_w \) and \( R_v \) indicate the correlation coefficients of \( W \) to \( A \) and \( B \), respectively.

Fig. 6  Relationship between net LUC production \( (10^5 \) bioluminescence counts plant \(^{-1}\) ) and the investigated indexes under various light conditions. (a–d) Correlation between net LUC production \( L \) and visible leaf area \( A \). (e–h) Correlation between \( L \) and bioluminescence \( B \). \( R_l \) and \( R_a \) indicate the correlation coefficients of \( L \) to \( A \) and \( B \), respectively.

*Increase in profit by selection of seedlings based on early diagnosis*

Next, we consider the improvement of profit by the selection of seedlings based on early diagnosis. The ratio \( \rho \) of selected seedlings to all seedlings is described as

\[
\rho = \frac{1}{n} \sqrt{2\pi} \int_{-\infty}^{\infty} \exp\left(-Z^2/2\right) dZ
\]
where $\rho$ is also the reciprocal of the average number $n$ of seedlings needed to obtain one superior seedling. If $Z^*$ is large, $<P^*/<P>$ is large but the cost of raising each seedling increases with an increase in $n$. Therefore, a balance between $<P^*/<P>$ and $n$ has to be considered for maximizing profit.

It is defined that the average cost of raising a seedling is $C_r$, the average cost of cultivating the plant (the cost of planting the seedling, cultivating, harvesting) is $C_s$, and the average selling price of the plant is $X$. By selection based on early diagnosis, the cost of raising the seedling becomes $nC_r$ and the average selling price of the plant becomes $<P^*/<P>X$. Therefore, the average profit per plant $Y$ is described as:

$$Y=\frac{<P^*/<P>}{X-C_r-nC_r}$$

Defining $\alpha$ and $\beta$ as

$$X=\alpha C_r$$
$$C_s=\beta C_r$$

yields

$$Y=C_r\left(\frac{<P^*/<P>}{X-C_r-nC_r} - 1 - \beta \right)$$  \hspace{1cm} (3)

Figure 9 shows $Y$ as a function of $Z^*$ obtained from Eq. 3. There is an optimal $Z^*$ for maximizing $Y$. The improvement of profit depends on the rate $\beta(\beta = C_s/C_r)$ (Fig. 9a). Moreover, the improvement of profit increases with the correlation coefficient $R$ (Fig. 9b). As selection based on early diagnosis becomes more effective, the cost of raising the seedling becomes relatively low and the index will have a high correlation to productivity. In case of the pharmaceutical production in genetically modified plant, the requirement of high correlation seems to be achieved by using bioluminescence which indicates the principal promoter activity.

Equations 2 and 3 are general equations that describe relationships between products and indexes. Therefore, these equations can be used for any product and any index; for example, the products are the functional ingredients in plant or the pharmaceutical proteins introduced by genetically modification, the index, morphological information or these promoter activities.

In this study, a method for early diagnosis of plant productivity at the seedling stage was developed for transgenic Arabidopsis thaliana CCA1::LUC. The dry aerial biomass and the luciferase protein were investigated as products. The visible leaf area and bioluminescence of seedlings were respectively used as a morphological index and a molecular index for early diagnosis of productivity. Bioluminescence showed higher correlation to both biomass and luciferase protein productivities than the leaf area, indicating that bioluminescence, a molecular index, is a better index for early diagnosis. To verify the generality of our developed method, we have also shown that our method of early diagnosis was useful under various light conditions, including the better light condition (LD24) and the worse light conditions (LD20 and LD28).

In addition, we developed a general statistical method for selecting seedlings based on early diagnosis of plant productivity and a general theory for determining an optimal parameter set for maximizing profit. In order to maximize profit, this theory requires the following:

1) use of an excellent index with a high correlation coefficient $R$ to the product,
2) low cost of raising a seedling, $C_r$,
3) large relative standard deviation of production, $\sigma_p/<P>$.

Requirement 1 seems to be achieved through an optimization approach using molecular information including bioluminescence, particularly when the products are pharmaceuticals. In relation to requirement 3, by using the molecular information to identify rare seedlings with high productivity,
we can base further efforts on these seedlings. The method for early diagnosis through an optimization approach using molecular information developed in this study will contribute to increasing the plant production in plant factories.

We are grateful to Dr. Nakamichi (RIKEN, Japan) for providing transgenic plants. This study was part of the NAIST Plant Factory Research Project supported by the Ministry of Economy, Trade and Industry of Japan (METI).

REFERENCES


Okayama, T., Okazumia, K., Murase, H. 2008b. A simulation model for heterologous protein production in