Toxicology in Vitro 29 (2015) 1816–1830

Contents lists available at ScienceDirect

Toxicology in Vitro

journal homepage: www.elsevier.com/locate/toxinvit

Optimization of the IL-8 Luc assay as an *in vitro* test for skin sensitization

Yutaka Kimura, Chizu Fujimura, Yumiko Ito, Toshiya Takahashi, Yoshihiro Nakajima, Yoshihiro Ohmiya, Setsuya Aiba *

Department of Dermatology, Tohoku University Graduate School of Medicine, Sendai 980-8574, Japan
Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Takamatsu, Kagawa 761-0395, Japan
Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba 305-8566, Japan

ARTICLE INFO

Article history:
Received 18 February 2015
Revised 25 June 2015
Accepted 6 July 2015
Available online 15 July 2015

Keywords:
Contact dermatitis
Sensitization test
Luciferase assay
Alternative method
Skin sensitization

ABSTRACT

We previously reported a dataset of the IL-8 Luc assay covering reference chemicals published by ECVAM, in which the effects of chemicals on IL-8 promoter activity were evaluated by an IL-8 reporter cell line, THP-G8 cells. To clarify its performance, we created another dataset of 88 sensitizers and 34 non-sensitizers. Simultaneously, to improve its performance, we changed the incubation time from 5 h to 16 h, deleted the criterion regarding the effects of N-acetylcysteine, and set an exclusion criterion for detergents. These modifications significantly improved its performance. In addition, we examined the following three criteria to judge chemicals as sensitizers: Criterion 1: Fold induction of SLO luciferase activity (FInSLO-LA) $\geqslant$ 1.4, Criterion 2: the lower limit of the 95% confidence interval of FInSLO-LA $\geqslant$ 1.0, Criterion 3: the intersection of criteria 1 and 2. Among them, Criterion 1 produced the best performance, demonstrating that the accuracy, sensitivity and specificity were 81%, 79%, and 90%, respectively. In addition, we found that the IL-8 Luc assay solubilizing chemicals with X-VIVO substantially improved its performance. Finally, the IL-8 Luc assay combined with DPRA and DEREK could improve substantially its performance. These data suggest that the IL-8 Luc assay is a promising test method to screen skin sensitizers.

1. Introduction

In the regulatory context, currently only data from animal experiments are acceptable to assess the skin sensitizing potential of chemicals. The European Union (EU) imposed an animal testing ban effective 2009 on both cosmetics products and their ingredients. This was accompanied by a concomitant marketing ban effective March 1, 2013 if animal tests were conducted after this date for the purpose of cosmetics legislation (Regulation (EC) No. 1223/2009 of the European Parliament and of the Council, 2009). On the other hand, under the European chemicals legislation REACH, skin sensitization data for any chemical registered under the European Chemicals Legislation (REACH, EC 1907/2006) is mandatory, and animal testing should only be performed as a last resort (http://echa.europa.eu/documents/10162/13639/alternatives_test_animals_2014_en.pdf).

Pushed by these ethical and legislative demands, various promising methods have been developed as alternative methods to detect the skin sensitizing potential of chemicals and four methods are currently undergoing formal validation at the European Centre for Validation of Alternative Methods (ECVAM). In 2012, the Organization for Economic Co-operation and Development (OECD) published the adverse outcome pathway (AOP) for skin sensitization (OECD, 2012) in which the key steps in the sensitization process are defined. According to the AOP, 4 methods are considered to target three different steps in the skin sensitization process: protein-binding/haptenization (e.g., the Direct Peptide Reactivity Assay, DPRA) (Gerberick et al., 2004), induction of the Kelch-like ECH-associated protein 1 (Keap-1)/nuclear factor (erythroid-derived 2)-like factor 2 (Nrf2) pathways in keratinocytes (e.g., the KeratinoSens™ assay) (Emter et al., 2010), and the activation of antigen presenting cells such as dendritic cell-like cell lines (The Myeloid U937 Skin Sensitization Test (U-SENS) (Piroird et al., 2015) or the human Cell Line Activation Test, h-CLAT) (Ashikaga et al., 2006; Sakaguchi et al., 2006). However, it is unlikely that a single assay will be sufficient to adequately assess the sensitization potential because of the complexity of the sensitization process (Bauch et al., 2012).

Dendritic cell activation is one of the key steps in sensitization indicated in the AOP for skin sensitization published by the OECD

* Corresponding author at: Department of Dermatology, Tohoku University Graduate School of Medicine, 1-1 Seiryo-machi, Sendai 980-8574, Japan.
E-mail address: saiba@med.tohoku.ac.jp (S. Aiba).

http://dx.doi.org/10.1016/j.tiv.2015.07.006
0887-2333/

(2012). In this step, CD54 and CD86 expression is augmented; an increase in interleukin-8 (IL-8) mRNA or IL-8 protein has been suggested as another biomarker for discriminating sensitizers from non-sensitizers in monocyte-derived dendritic cells (MoDCs) (Toebak et al., 2006), U937 cells (Python et al., 2007), or THP-1 cells (Mitjans et al., 2008, 2010; Nukada et al., 2008). IL-8 is well established as a potent chemotactic peptide for neutrophils, T lymphocytes, basophils (Leonard et al., 1990), and NK cells (Sebok et al., 1993). It was recently reported that human immature MoDCs express the IL-8 receptors CXCR1 and CXCR2, which are down-regulated in mature MoDCs (Gouwy et al., 2014) and that accordingly, human immature MoDCs are chemoattracted by IL-8 (Feijoo et al., 2005). It is impossible to demonstrate the exact role of IL-8 in contact hypersensitivity using IL-8 knockout mice because of the lack of a mouse counterpart of IL-8. Regardless, several studies suggest the importance of IL-8 in the DC activation step in the AOP for skin sensitization. Specifically, CCL2 that is coordinately regulated with IL-8 (Singha et al., 2014) plays a crucial role in dendritic cell maturation (Jimenez et al., 2010). On the other hand, Natsuaki et al. have reported DC clusters around macrophages in both the elicitation phase and the sensitization phase in murine contact sensitivity, suggesting the crucial role of CXCR2 expression on DCs in murine contact sensitivity (Natsuaki et al., 2014). Since CXCR2 is a receptor for IL-8 in humans (Marchese et al., 1995; Murphy and Tiffany, 1991), a murine counterpart of IL-8 produced by dermal macrophages may play a crucial role in murine contact sensitization.

In addition, it is now well-recognized that skin sensitization and chemical protein reactivity are linked. Although chemical sensitizers are extremely diverse in molecular weight and structure, most share electrophilic properties and possess intrinsic reactivity toward various amino acids containing nucleophilic heteroatoms (i.e., cysteine, lysine, histidine, arginine, and methionine). Indeed, a correlation between the reactivity of chemicals with cysteine or lysine residues in peptides and their sensitization potential has been demonstrated (Gerberick et al., 2007). Electrophiles can be detected by the Keap1-Nrf2 cellular sensor pathway implicated in the antioxidant response of the cell and recently reviewed by Itoh et al. (2010). Under normal conditions, Keap1 sequesters the transcriptional regulator nuclear Nrf2 in the cytoplasm, provoking its proteasomal degradation. In the presence of electrophiles, the highly reactive cysteine residues of Keap1 are modified, leading to the dissociation of Keap1 from Nrf2. Nrf2 translocates to the nucleus, forms heterodimers with small Maf proteins, and then induces the transcription of genes with an antioxidant response element (ARE) in their promoters (Holland and Fishbein, 2010). These genes code for proteins mostly involved in detoxification, such as heme oxygenase-1 (HO-1) and NADPH-quinone oxidoreductase 1 (Nqo1). Interestingly, Zhang et al. reported that the 5′ flanking region of the IL-8 gene has several areas homologous to the consensus ARE (ATGAC/TnnnGCA/); in addition, Nrf2 caused only a weak induction of IL-8 transcription but significantly increased the half-life of IL-8 mRNA (Zhang et al., 2005). These data suggested that the induction of IL-8 mRNA by haptens is regulated transcriptionally by p38 MAPK and post-transcriptionally by Nrf2. Therefore, the IL-8 Luc assay is a unique screening method for haptens since it detects their effects on p38 MAPK and Keap1-Nrf2.

We established the IL-8 reporter cell assay (IL-8 Luc assay) using a stable THP-1-derived IL-8 reporter cell line, THP-G8, which harbors SLO and SLR luciferase genes under the control of IL-8 and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) promoters, respectively (Takahashi et al., 2011). This previous study demonstrated that use of the IL-8 Luc assay to examine 35 chemicals, including reference chemicals published by ECVAM (referred to as the "ECVAM list"; (Casati et al., 2009), resulted in an overall test accuracy of 82%.

In the current study, we increased the number of chemicals for evaluation by the IL-8 Luc assay, explored the reason for false results, modified the procedure and protocol (including the criteria), examined the correlation of parameters between the IL-8 Luc assay and other test methods, and proposed the test battery system.

2. Materials and methods

2.1. Cells and culture

We previously established a reporter cell line, THP-G8 cells derived from the human acute monocytic leukemia cell line THP-1 cells containing stable luciferase orange (SLO) regulated by IL-8 promoter and stable luciferase red (SLR) by GAPDH promoter (Takahashi et al., 2011). THP-G8 cells were cultured in RPMI-1640 (Gibco, Carlsbad, CA) with antibiotic–antimycotic (Invitrogen, Carlsbad, CA) and 10% fetal bovine serum (FBS) (Biological Industries, Kibbutz Beit Haemek, Israel) at 37 °C with 5% CO_2.

2.2. Test chemicals and chemical treatment

We examined 122 chemicals that are listed with their Chemical Abstract Service (CAS) numbers in Table S1. All chemicals had been previously evaluated and classified with the LLNA (Gerberick et al., 2005). Seventy-two sensitizers were evaluated, including 8 extreme, 16 strong, 25 moderate, and 23 weak sensitizers, as classified by the local lymph node assay (LLNA); 28 non-sensitizers were also evaluated, one of which (sodium lauryl sulfate; SLS) was false positive in the LLNA. All the chemicals were purchased from Sigma–Aldrich, St. Louis, MO, at the highest available purity.

Water soluble chemicals were dissolved in distilled water at a concentration of 25 mg/ml, 50 mg/ml, or 100 mg/ml to determine the highest soluble concentration. Chemicals not soluble in water were dissolved in DMSO at 500 mg/ml. If they were not soluble at 500 mg/ml, the highest soluble concentration was determined by diluting the suspension from 500 mg/ml by a factor of two with DMSO. Sonication and vortex mixing were used if needed and the attempt to dissolve the chemical continued for at least 5 min. All dissolved chemicals were used within 4 h of being dissolved in distilled water or DMSO.

To examine the effects of FBS on the IL-8 Luc assay, three methodologies were examined to solubilize haptens: (1) solubilize in DMSO, and then dilute with RPMI-1640 containing 10% FBS, as used in the IL-8 Luc assay (DMSO/FBS); (2) solubilize in DMSO and then dilute with X-VIVO 15 (Lonza, Walkersville, MD) (DMSO/X-VIVO); and (3) solubilize in X-VIVO and then dilute with X-VIVO (X-VIVO/X-VIVO).

For water soluble chemicals, 11 serial dilutions were conducted using RPMI-1640 with 10% FBS diluting by a factor of 2, in the 1st experiment. In the 2nd, 3rd, or 4th experiment, 11 serial dilutions were conducted, diluting by a factor of 1.5. For water insoluble chemicals, 11 serial dilutions were conducted using DMSO as the solvent, diluting by a factor of 2 in the 1st experiment and by a factor of 1.5 in the 2nd, 3rd, and 4th experiments.

Based on the previous report (Saito et al., 2011; Takahashi et al., 2011), THP-G8 cells (5×10^4 cells/50 μl/well) in 96-well black plates (Greiner bio-one GmbH, Frickenhausen, Germany) were cultured for varying time periods. The optimum cell numbers at seeding were based on the previous reports. In some experiments, the cells were pretreated with 25 mM N-acetyl-L-cysteine (NAC) for 30 min.

2.3. IL-8 promoter-luciferase gene reporter assay

The luciferase reporter assay system was constructed using 2 luciferase genes, SLO and SLR, that emit orange and red light, respectively, with a single substrate. The activities of these luciferases can be measured simultaneously and quantitatively with optical filters. This system can rapidly and easily monitor multiple gene expression (Nakajima et al., 2005; Noguchi et al., 2008). In this study, luciferase activity was determined using a microplate-type luminometer with a multi-color detection system, Phelios (Atto Co., Tokyo, Japan), using the Tripluc® luciferase assay reagent (TOYOBO Co., Ltd., Osaka, Japan) according to the manufacturer's instructions.

Since some chemicals affected cell viability, we defined the parameter nSLO-LA to represent IL-8 promoter activity. This parameter is calculated by normalizing SLO luciferase activity (SLO-LA) to SLR luciferase activity (SLR-LA). We also calculated the inhibition index of SLR-LA (I.I.-SLR-LA) by dividing the SLR-LA of THP-G8 cells that were treated with chemicals by the SLR-LA of non-treated THP-G8 cells. The fold induction of IL-8 promoter activity (FInSLO-LA) was calculated by dividing the nSLO-LA of THP-G8 cells that were treated with chemicals by that of non-stimulated THP-G8 cells.

We further evaluated the suppressive effect of NAC co-treatment with each chemical by calculation of the inhibition index (I.I.). The I.I. was obtained by dividing the FInSLO-LA of THP-G8 cells stimulated with the chemical in the presence of NAC by the FInSLO-LA stimulated with the chemical alone, using the concentration of chemical at which the chemical induced the largest FInSLO-LA.

The parameters used in the IL-8 Luc assay are shown in Table 1.

Using these parameters, we defined the criteria to identify possible sensitizers. Each criterion is composed of two conditions. Condition 1 defines positive induction or negative induction of SLO-LA in each experiment and condition 2 defines sensitizers or non-sensitizers based on the repeated experiments. In our previous paper (Takahashi et al., 2011), chemicals that demonstrate FInSLO-LA $\geqslant$ 1.4 and I.I. $\leqslant$ 0.8 at the concentration of the chemical at which I.I.-SLR-LA is $\geqslant$0.2 in two or three of three different experiments are categorized as sensitizers and those that do not fulfill these criteria are classified as non-sensitizers. In this study, we optimized the IL-8 Luc assay by examining its performance using six criteria. The precise definition of each criterion is summarized in Table 2. Briefly, these six criteria are different from the original criterion in condition 2: chemicals are categorized as sensitizers when they fulfill condition 1 in two or three of three different experiments in the original criterion, while they are categorized as sensitizers when they fulfill condition 1 in two of two to four different experiments in the new six criteria. Among them, Criterion A, B, and C include the condition I.I. $\leqslant$ 0.8, while Criterion 1, 2, and 3 do not.

Table 1
Parameters used in the IL-8 Luc assay.

Parameters	Description
SLO-LA	SLO luciferase activity regulated by IL-8 promoter
SLR-LA	SLR luciferase activity regulated by G3PDH promoter
nSLO-LA	SLO-LA/SLR-LA
I.I.-SLR-LA	SLR-LA of THP-G8 treated with chemicals/SLR-LA of non-treated THP-G8
FInSLO-LA	nSLO-LA of THP-G8 cells treated with chemicals/nSLO-LA of non-stimulated THP-G8 cells
I.I.	FInSLO-LA of THP-G8 cells stimulated with the chemical and NAC/FInSLO-LA stimulated with the chemical alone

2.4. Real-time monitoring of luciferase activity of THP-G8 cells after chemical stimulation

To clarify the time-dependent change of SLO-LA and SLR-LA of THP-G8 cells after chemical stimulation, THP-G8 was suspended with RPMI1640 supplemented with 10% FBS, 0.1 mM D-luciferin 25 mM HEPES/HCl (pH 7.0), and plated onto 35 mm dish at 2×10^6 cells/dish. After 30 min, 1.6 or 0.8 μg/ml DNCB, 1.6 or 0.8 μg/ml 4-NBB was added to the culture. Bioluminescence was in real-time under a 5% CO_2 atmosphere at 37 °C using the dish-type luminometer AB2500 Kronos (ATTO, Tokyo, Japan).

2.5. Statistics

To demonstrate statistical significance, representative data from at least three independent experiments for each analysis is shown. A one-way ANOVA test followed by Dunnett's post hoc test was used to evaluate statistical significance. p values <0.05 were considered statistically significant. We performed Pearson's product–moment correlation analysis to determine the strength of a correlation.

3. Results

3.1. Determination of the optimal incubation time

The performance of the IL-8 Luc assay, when we first reported (Takahashi et al., 2011), was an accuracy of 86%, a sensitivity of 83%, and a specificity of 90% for 35 chemicals including ECVAM list

Table 2
Criteria used in the IL-8 Luc assay.

Criteria		Condition 1	Condition 2
Original	Sensitizer	FInSLO-LA $\geqslant$ 1.4 and I.I. $\leqslant$ 0.8 at the concentration of the chemical at which I.I.-SLR-LA is $\geqslant$0.2	Fulfill the condition 1 in 2 or 3 of 3 different experiments
Original	Non-sensitizer		Do not fulfill the condition 1 in 2 of 3 different experiments
A	Sensitizer	FInSLO-LA $\geqslant$ 1.4 and I.I. $\leqslant$ 0.8 at the concentration of the chemical at which I.I.-SLR-LA is $\geqslant$0.2	Fulfill the condition 1 in 2 of 2–4 different experiments
B	Sensitizer	FInSLO-LA $\geqslant$ 1.4 and I.I. $\leqslant$ 0.8 at any concentrations	
C	Sensitizer	FInSLO-LA $\geqslant$ 1.4 and I.I. $\leqslant$ 0.8 at the concentration of the chemical at which I.I.-SLR-LA is $\geqslant$0.05	
1	Sensitizer	FInSLO-LA $\geqslant$ 1.4 at the concentration of the chemical at which I.I.-SLR-LA is $\geqslant$0.05	
2	Sensitizer	The lower limit of the 95% confidence interval of FInSLO-LA $\geqslant$ 1.0	
3	Sensitizer	1 and 2	
Either criteria	Non-sensitizer		Do not fulfill the condition 1 in 3 of 3–4 different experiments

(Casati et al., 2009). When we conducted ring trials with 3 different laboratories using 10 coded chemicals, they revealed the accuracy and the intralaboratory reproducibility were not necessarily high enough to conduct further validation studies. Indeed, the accuracy of each laboratory was 80% by Laboratory A, 60% by Laboratory B, and 63% by Laboratory C, respectively, and the intralaboratory reproducibility was 70%. Therefore, we first explored the reason for the poor response of the IL-8 Luc assay to some potent sensitizers, such as DNCB. In our previous study, we examined the time course of FInSLO-LA of THP-G8 cells after LPS stimulation and found the maximum induction of FInSLO-LA between 4 and 7 h after stimulation. In this study, we re-examined the time course of FInSLO-LA after the treatment with strong sensitizers, DNCB and 4-NBB.

We stimulated THP-G8 cells with different concentrations of DNCB or 4-NBB for different time periods and FInSLO-LA was measured (Fig. 1a and b). DNCB and 4-NBB significantly augmented FInSLO-LA dose-dependently from 5 h to 24 h and from 4 h to 24 h after stimulation, respectively. The maximum induction by DNCB was observed between 9 h and 12 h at the concentration of 1.19 μg/ml and 2.67 μg/ml. On the other hand, the maximum induction by 4-NBB was observed between 8 h and 10 h at the concentration of 1.19 μg/ml and at 10 h and 16 h at the concentration of 2.67 μg/ml.

Next, we monitored SLO-LA of THP-G8 cells for 1 min at intervals of 19 min during DNCB or 4-NBB treatment under a dish type luminometer (Fig. 2a and b). Consistent with the results obtained by measuring luciferase activity intermittently, real-time monitoring also demonstrated the increase in FInSLO-LA from 5 h to 24 h by both 0.8 μg/ml and 1.6 μg/ml of DNCB and 4-NBB with maximum induction at 12 h and 10–12 h, respectively.

These results demonstrated that the optimal incubation period with chemicals is around 10 h. This is not practical because one IL-8 Luc assay requires more than 12 h, including general preparation, plating the cells, applying the chemicals and measuring luciferase activity using a luminometer. Therefore, from the practical standpoint, we compared FInSLO-LA between 6 h incubation and 16 h incubation in Figs. 1 and 2. For both DNCB and 4-NBB, intermittent measurement and real-time monitoring indicated that FInSLO-LA at the optimal concentration was much higher in 16 h incubation than in 6 h incubation. This was confirmed by stimulating THP-G8 cells with different concentrations of DNCB and 4-NBB for 6 h and 16 h (data not shown).

Finally, we evaluated the ECVAM list of chemicals by the IL-8 Luc assay with a 16 h incubation period and with the IL-8 Luc assay with a 6 h incubation period and compared the results. As shown in Table 3, most sensitizers increased FInSLO-LA in the 16 h incubation IL-8 Luc assay more than in the 6 h incubation IL-8 Luc assay. In addition, the data from the 16 h incubation period moved isoeugenol from the non-sensitizer category to the sensitizer category, increasing accuracy to 94%.

3.2. Comparison between the criterion with I.I.-SLR-LA ⩾ 0.2 and that with I.I.-SLR-LA ⩾ 0.05

After increasing the incubation time from 6 h to 16 h, we examined 89 chemicals that were used as a data set of h-CLAT (Ashikaga et al., 2010) and evaluated their skin sensitization potential (Table S1). The dataset for the IL-8 Luc assay was created following examination of the IL-8 Luc assay's performance using the six criteria described in Table 2. When we examined its performance using Criterion A, which used the same condition 1 as the original criterion, Cooper statistics of the IL-8 Luc assay for these 89 chemicals yielded an accuracy of 69%, a sensitivity of 59%, and a specificity of 92%, suggesting that the IL-8 Luc assay using the current criterion produce false negative results for a considerable number of sensitizers tested. Closer examination of the data showed that most of the treatments that produced false negative results increased FInSLO-LA more than 1.4 at the concentrations providing I.I.-SLR-LA < 0.2 (data not shown). SLR-LA corresponds with promoter activity of GAPDH gene. GAPDH mRNA is a ubiquitously expressed at moderately abundant levels. It is frequently used as an endogenous control for quantitative real-time polymerase chain reaction because, in some experimental systems, its expression is constant at different times and after various experimental manipulations (Edwards and Denhardt, 1985; Mori et al., 2008; Winer et al., 1999). Indeed, in this study, all but two (benzocaine and methylisothiazolinone) of the 122 chemicals examined did not increase SLR-LA, and most dose-dependently decreased SLR-LA at toxic concentrations, suggesting that SLR-LA can act as an internal control to indicate cell number and viability. Our previous study demonstrated that I.I.-SLR-LA is more sensitive in detecting dying cells than the percentage of PI-excluding cells, and cells showing less than 0.2 of I.I.-SLR-LA retained more than 80% of PI-excluding cells. In this study, we further examined the correlation between the percentage of PI-excluding cells and I.I.-SLR-LA, and confirmed that THP-G8 cells treated with chemicals whose I.I.-SLR-LA showed ⩾0.05 maintained more than 75% of the PI-excluding cells (Fig. 3). We therefore examined the performance of the IL-8 Luc assay using Criterion C that included I.I.-SLR-LA ⩾ 0.05 instead of I.I.-SLR-LA ⩾ 0.2 in the condition 1 and obtained Cooper statistics for these 89 chemicals of 74% accuracy, 69% sensitivity, and 89% specificity.

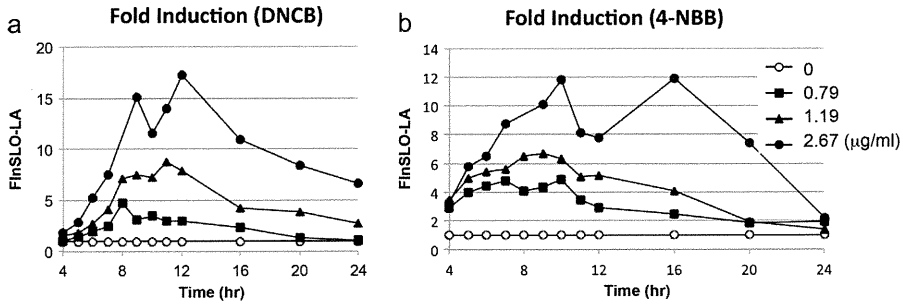

Fig. 1. Determination of the optimal incubation time – time course study for the IL-8 Luc assay. THP-G8 cells were stimulated with the indicated dose of DNCB (a) or 4-NBB (b) for various time periods, and luciferase activity was measured using a microplate-type luminometer.

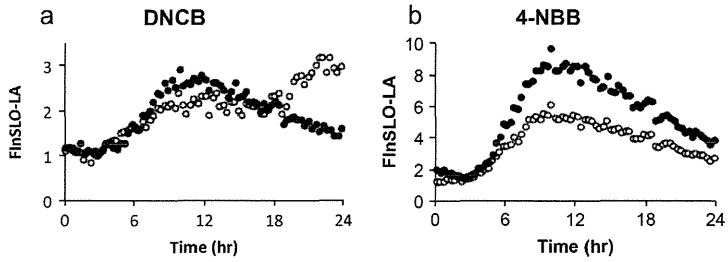

Fig. 2. Determination of the optimal incubation time – real time monitoring of IL-8 luciferase activity. THP-G8 cells were suspended in RPMI-1640 supplemented with 10% FBS, 0.1 mM D-luciferin, 25 mM HEPES/HCl (pH 7.0) and plated onto 35 mm dishes at 2×10^6 cells/dish. After 30 min, 0.8 or 1.6 μg/ml DNCB (a), or 0.8 or 1.6 μg/ml 4-NBB (b), was added to the culture. Bioluminescence was continuously recorded at intervals of 19 min under a 5% CO_2 atmosphere at 37 °C using a dish-type luminometer. Open circles: 0.8 μg/ml, closed circles: 1.6 μg/ml.

3.3. Comparison between the criterion with and without the response to NAC treatment

Despite having accepted the I.I.-SLR-LA ⩾ 0.05 criterion, the IL-8 Luc assay still produced significant numbers of false negative results. Therefore, we further modified the criterion. Although most haptens react with cysteine residues, there may be some exceptions. For example, Gerberick et al. reported that phthalic anhydride and trimellitic anhydride significantly reacted with glutathione and lysine peptides, but not with cysteine peptides in their direct peptide reactivity assay (DPRA) (Gerberick et al., 2007). We hypothesized that removing the condition that I.I. is ⩽0.8 (Criterion 1) would increase the accuracy and specificity of the IL-8 Luc assay. Indeed, this modification significantly improved the performance of the IL-8 Luc assay and significantly improved accuracy and sensitivity (i.e., accuracy of 78% and sensitivity of 77%) but decreased the specificity to 74% (Table S1 and Table 4), when we examined the 122 chemicals (9 extreme, 17 strong, 34 moderate, and 28 weak sensitizers, and 34 non-sensitizers as classified by the LLNA (Gerberick et al., 2005)) as a dataset for h-CLAT (Ashikaga et al., 2010; Nukada et al., 2012; Takenouchi et al., 2013).

3.4. Comparison between the criterion with FInSLO-LA ⩾ 1.4, the criterion with the lower limit of the 95% confidence interval of FInSLO-LA ⩾ 1.0, and their combination

We empirically determined the condition of FInSLO-LA ⩾ 1.4 (Takahashi et al., 2011), similar to the condition used in h-CLAT (Sakaguchi et al., 2006). In contrast, each independent repetition was statistically evaluated in the KeratinoSens assay (Emter et al., 2010). We likewise tried to evaluate each IL-8 Luc assay experiment statistically. When we evaluated 122 chemicals using the criterion of the lower limit of the 95% confidence interval of FInSLO-LA ⩾ 1.0 (Criterion 2) did not necessarily improve the performance of the IL-8 Luc assay (accuracy of 74%, sensitivity of 80%, and 59% of specificity). We also examined the performance of the combination of Criterion 1 and Criterion 2 (Criterion 3). Both the accuracy and sensitivity of Criterion 3 was inferior to those of Criterion 1, although the specificity of Criterion 3 and Criterion 1 was equal (Table S1 and Table 4).

3.5. The factors that produce false negative or positive results in the IL-8 Luc assay (1) – physical properties

The IL-8 Luc assay conducted according to Criterion 1 produced 9 false negative results among 88 haptens determined by LLNA. To clarify the underlying reason for these false negative results, we first compared two physical properties of haptens judged by the IL-8 Luc assay to be sensitizers (true positive) and those judged to be non-sensitizers (false negative): molecular weight, and Log Ko/*w* or water solubility. The results could not demonstrate statistically significant differences (Fig. 4). Indeed, 8 sensitizers had Log Ko/*w* values above 3.5. The accuracy and sensitivity of these chemicals by the IL-8 Luc assay were both 87.5%.

3.6. The factors that produce false negative or positive results in IL-8 Luc assay (2) – the effects of FBS

The effects of FBS in the culture medium were considered next. Recently, several researchers have demonstrated that most sensitizers can bind to both FBS and cellular proteins, although the distribution of covalent binding to cellular or FBS protein varies depending on the hapten (Divkovic et al., 2005; Hopkins et al., 2005; Saito et al., 2013). These reports suggested that the amount of reactive electrophiles in haptens to bind to nucleophiles of cellular protein might be reduced in the presence of FBS.

Therefore, we examined whether a reduction in FBS concentration in the IL-8 Luc assay can reduce false negative results. First, THP-G8 cells were stimulated with oxazolone, which was judged as a non-sensitizer by the IL-8 Luc assay. Three methodologies were examined to solubilize oxazolone: (1) solubilize in DMSO, and then dilute with RPMI-1640 containing 10% FBS, as used in the IL-8 Luc assay (DMSO/FBS); (2) solubilize in DMSO and then dilute with X-VIVO (DMSO/X-VIVO); and (3) solubilize in X-VIVO and then dilute with X-VIVO (X-VIVO/X-VIVO) (Fig. 5a–c). As we have repeatedly demonstrated, oxazolone diluted with DMSO/FBS did not induce significant induction of FInSLO-LA at the concentration at which I.I.-SLR-LA showed ⩾0.05. In contrast, oxazolone diluted with X-VIVO/X-VIVO significantly and dose-dependently induced FInSLO-LA at the concentration at which I.I.-SLR-LA showed ⩾0.10. Oxazolone diluted with DMSO/X-VIVO significantly induced FInSLO-LA at the concentration at which I.I.-SLR-LA showed ⩾0.6, but significant induction was observed only at this single concentration. These results clearly demonstrated that the dilution of oxazolone with X-VIVO significantly improved the response of THP-G8 cells. Furthermore, contrary to our expectation, even solubilization of oxazolone with DMSO did not necessarily improve THP-G8 response.

Clearly, FBS perturbed the response of THP-G8 cells for oxazolone; we therefore next examined whether changing the culture medium during treatment with the chemicals from RPMI-1640 with 10% FBS to X-VIVO improved the response of THP-G8 cells to haptens. The results clearly showed that the response of THP-G8 cells to oxazolone became far weaker in X-VIVO than that RPMI-1640 with 10% FBS (data not shown).

Table 3
Comparison of the performance between IL-8 Luc assay (16 h) and IL-8 Luc assay (6 h).

Chemical	LLNA	The IL-8 Luc assay (6 h) Takahashi et al. in Toxicol. Sci.								The IL-8 Luc assay (16 h)							
		1st		2nd		3rd		Positive exp	Decision	1st		2nd		3rd		Positive exp	Decision
		FInSLO-LA	I.I.	FInSLO-LA	I.I.	FInSLO-LA	I.I.			FInSLO-LA	I.I.	FInSLO-LA	I.I.	FInSLO-LA	I.I.		
Oxazolone	Sensitizer	1	N.D.	1	N.D.			0	Non-sens	1.43	1.15	1.29	1.03			0	Non-sens
4-NBB	Sensitizer	3.9	0.27	4.4	0.21			2	Sens	7.4	0.14	6.63	0.14			2	Sens
DNCB	Sensitizer	2.3	0.44	1.8	0.53			2	Sens	5.83	0.15	10.38	0.08			2	Sens
MDGN	Sensitizer	1.6	0.65	2.7	0.5			2	Sens	3.68	0.28	1.87	0.47			2	Sens
Glyoxal	Sensitizer	1.7	0.77	1.4	0.76			2	Sens	2.97	0.71	3.76	0.3			2	Sens
2-MBT	Sensitizer	1.9	0.7	1.9	0.76			2	Sens	2.29	0.78	6.36	0.56			2	Sens
Cinnamal	Sensitizer	2	0.54	2.7	0.42			2	Sens	5.09	0.19	8.05	0.15			2	Sens
TMTD	Sensitizer	1.7	0.63	3.6	0.44			2	Sens	7.28	0.14	3.95	0.28			2	Sens
PPD	Pre/pro hapten	1.5	0.66	2.3	0.57			2	Sens	1.61	0.61	1.76	0.6			2	Sens
Isoeugenol	Pre/pro hapten	1.8	0.79	1.8	0.86	1.55	1.04	1	Non-sens	2.94	0.74	4.07	0.44			2	Sens
Eugenol	Pre/pro hapten	2.1	0.64	1.7	0.77			2	Sens	4.85	0.69	2.72	1.17	1.69	0.72	2	Sens
Cinnamic alcohol	Pre/pro hapten	2.1	0.47	2.3	0.76			2	Sens	7.68	0.54	8.52	0.54			2	Sens
Glycerol	Irritant	1.5	0.82	1.4	0.83			0	Non-sens	1.12	0.95	1.03	1.25			0	Non-sens
Salicylic acid	Irritant	1.3	1.27	1.1	0.98			0	Non-sens	1	1	1	1			0	Non-sens
Lactic acid	Irritant	1.2	0.82	1.1	1.32			0	Non-sens	1.35	1.95	1.54	2.18			0	Non-sens
SLS	Irritant	3.1	0.88	4.6	0.86			0	Non-sens	2.84	1.18	4.49	1.04			0	Non-sens

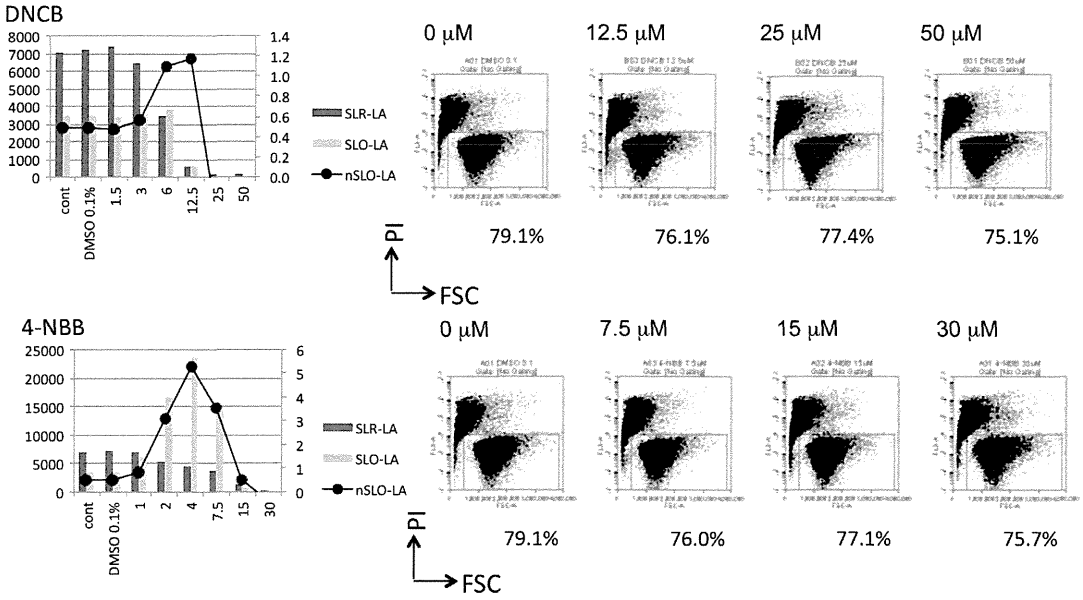

Fig. 3. Comparison between criterion with I.I.-SLR-LA ⩾ 0.2 and that with I.I.-SLR-LA ⩾ 0.05. Cell viability was determined by a PI exclusion assay using flow cytometry. THP-G8 cells were stimulated with the indicated dose of DNCB or 4-NBB. In this PI exclusion assay, THP-G8 cells after chemical treatment were mixed with 30 μg/ml of PI, and the live cells (which are not permeable to PI) were counted using flow cytometry. The results of the luciferase assay are shown on the left.

Table 4
Comparison of the performance among different criterion.

Criteria	With I.I. criterion (I.I. ⩽ 0.8)			Without I.I. criterion		
	Without I.I.-SLR-LA criterion	I.I.-SLR-LA ⩾ 0.2	I.I.-SLR-LA ⩾ 0.05	Criterion 1	Criterion 2	Criterion 3
Chemical numbers	98	89	97	122	122	122
Accuracy	0.75	0.69	0.74	0.77	0.73	0.73
Sensitivity	0.69	0.59	0.69	0.77	0.80	0.72
Specificity	0.89	0.92	0.89	0.74	0.59	0.74

Next, we examined whether the other chemicals providing false negative results in the IL-8 Luc assay could be judged as sensitizers if diluted with X-VIVO/X-VIVO (Table 5). Interestingly, 5 of the 14 chemicals that showed false negative results were judged as sensitizers. To confirm the efficacy of solubilization of chemicals with X-VIVO, we re-evaluated the chemicals in the ECVAM list (Table 6). The IL-8 Luc assay using chemicals diluted with X-VIVO/X-VIVO increased the FInSLO-LA of all sensitizers, i.e., oxazolone, 4-NBB, DNCB, MDGN, eugenol, and PPD, and changed the judgment of oxazolone. As a result, the Cooper statistics of the IL-8 Luc assay using chemicals diluted with X-VIVO yielded an accuracy of 94%, a sensitivity of 100%, and a specificity of 75%. These data suggested that the IL-8 Luc assay diluted with X-VIVO can improve the accuracy and sensitivity, while it does not lower the specificity.

3.7. The factors that produce false negative or positive results in IL-8 Luc assay (3) – detergents

The IL-8 Luc assay conducted according to Criterion 1 produced 8 false positive results among 28 non-sensitizers, of which hexadecyltrimethylammonium bromide, benzalkonium chloride, Tween-80, and SLS are well-known detergents. It has been reported that the treatment of reconstructed human epidermis with detergents, such as sodium lauryl sulfate, triton, and benzalkonium chloride, increased IL-8 mRNA levels in and IL-8 release from the cells (Coquette et al., 1999). Moreover, White et al. have demonstrated that SLS induced early growth response-1 (EGR-1) depending on the activation of MEK1/p44/42 ERK and EGFR (White et al., 2011). In contrast to these observations on epidermal cells, several researchers examined IL-8 release by THP-1 cells and demonstrated the lack of IL-8 production when stimulated with SLS (Mitjans et al., 2008; Trompezinski et al., 2008). However, the concentration of SLS used in these studies was 30 μg/ml or less. Since the optimal concentration of SLS to induce FInSLO-LA is 50 μg/ml, and the concentration range that significantly augmented FInSLO-LA was narrow, the previous authors either did not stimulate THP-1 with the optimal concentration of SLS or there was a discrepancy between IL-8 release and IL-8 mRNA induction that corresponds with FInSLO-LA. Regardless, we propose not to evaluate detergents by the IL-8 Luc assay.

3.8. The factors that produce false negative or positive results in the IL-8 Luc assay (4) – relative human skin sensitizing potency

Recently, Basketter et al. (2014) collected data regarding the sensitizing potential of chemicals to human skin and classified 131 chemicals into 6 categories based on their relative human skin

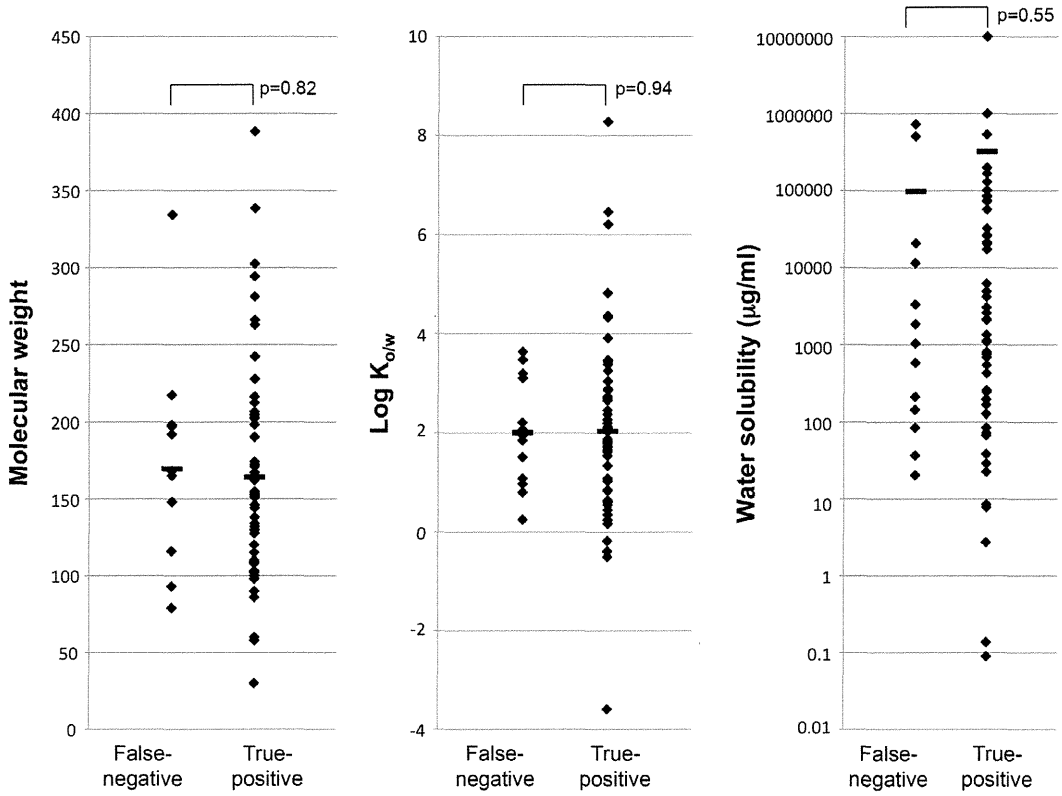

Fig. 4. Factors responsible for false negative or positive results in the IL-8 Luc assay (chemical properties). Sensitizers judged by the LLNA are divided into false negative and true positive groups according to the results of the IL-8 Luc assay. Molecular weight, LogKo/w, and water solubility of each chemical are plotted on the *y*-axis for each group. The mean and *p* value compared with Student's *t*-test are shown.

sensitizing potency, with category 1 being the most potent, category 5 being the least potent, and category 6 being true non-sensitizers. Comparison of the relative human skin sensitizing potency of these chemicals with their LLNA evaluation provided excellent correlation in that all sensitizers judged by relative human skin sensitizing potency are also classified as sensitizers by LLNA. However, there were some differences between the two categorization schemes. For example, although vanillin and benzalkonium chloride are classified as non-sensitizers by LLNA, relative human skin sensitizing potency included them in category 5, the least potent sensitizer. The IL-8 Luc assay also judged vanillin as sensitizers.

3.9. The performance of the IL-8 Luc assay after considering the exclusion criterion and human sensitization potential, and using X-VIVO as a solvent

If we delete the data of hexadecyltrimethylammonium bromide, Tween-80, SLS and benzalkonium chloride and consider vanillin as a sensitizer, the performance of the IL-8 Luc assay was accuracy of 81%, sensitivity of 79%, and specificity of 90% in Criterion 1, accuracy of 77%, sensitivity of 80%, and specificity of 69% in Criterion 2, and accuracy of 77%, sensitivity of 73%, and specificity of 90 in Criterion 3 (Table 7). Furthermore, if we consider the results using X-VIVO as a solvent, the performance of the IL-8 Luc assay was accuracy of 90%, sensitivity of 90%, and specificity of 90% in Criterion 1, accuracy of 82%, sensitivity of 87%, and specificity of 69% in Criterion 2, and accuracy of 86%, sensitivity of 84%, and specificity of 90% in Criterion 3 (Table 8).

3.10. Correlation between the IL-8 Luc assay and other screening methods

Next, we examined the correlation of the parameters between the IL-8 Luc assay and other assays. Statistical analysis of the correlation between FInSLO-LA, and the percent depletion of peptides containing either lysine or cysteine in DPRA, provided no significant correlation (Fig. 6). Similarly, we also examined the correlation between the minimum concentration required to induce more than 1.4 of FInSLO-LA in the IL-8 Luc assay (IL-8 Luc assay minimum induction threshold (MIT)) and that required to induce more than 150% of CD86 augmentation (h-CLAT MIT (CD86 EC150)) (Fig. 7a) or more than 200% of CD54 in h-CLAT (h-CLAT MIT (CD54 EC200)) (Fig. 7b). The results demonstrated weak correlation between them (Fig. 7a and b). In contrast, there was no significant correlation between the IL-8 Luc assay MIT and EC 1.5 of KeratinoSens (Fig. 7c).

3.11. Score-based battery system

To improve the performance of the IL-8 Luc assay for predicting sensitizing potential and the potency of chemicals, we developed a

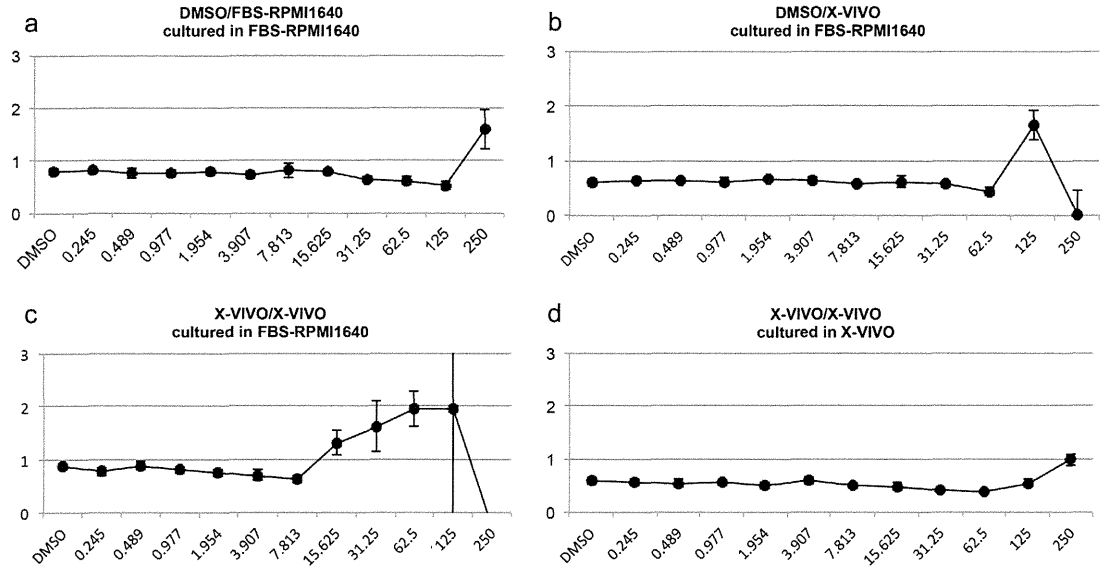

Fig. 5. The effects of FBS on the IL-8 Luc assay. To stimulate THP-G8 cells, we solubilized oxazolone in the following 3 ways: (a) solubilize in DMSO and then dilute with RPMI-1640 containing 10% FBS as used in the IL-8 Luc assay (DMSO/FBS); (b) solubilize in DMSO and then dilute with X-VIVO (DMSO/X-VIVO); and (c) solubilize in X-VIVO and then dilute with X-VIVO (X-VIVO/X-VIVO).

Table 5
Re-evaluation of chemicals demonstrating false negative results by the IL-8 Luc assay using X-VIVO as a solvent.

Chemicals	Experiments		Judgment
	1st	2nd	
Oxazolone	1.66	3.98	Sensitizer
Phthalic anhydride	1.09	1.60	
2-Hydroxyethyl acrylate	1.00	1.00	
Ethylenediamine	1.96	2.38	Sensitizer
Methyl-2-nonynoate	1.93	5.91	Sensitizer
3,4-Dihydrocoumarin	1.00	1.00	
Trimellitic anhydride	2.28	1.66	Sensitizer
1-Bromohexxane	1.07	1.05	
4-Allylanisole	4.67	19.71	Sensitizer
Benzocaine	1.11	1.00	
Ethyleneglycol dimethacrylate	4.98	4.36	Sensitizer
Penicillin G	1.35	1.00	
Pyridine	1.09	1.00	
Aniline	1.03	1.12	

test battery by assigning scores to the outcomes in each single test, based on the concept reported by Jowsey et al. (2006). We converted the results of the IL-8 Luc assay and DPRA into a score from 0 to 2 based on the sensitizing potency classification. In the IL-8 Luc assay, we first obtained the maximum value of FInSLO-LA (MAX FInSLO-LA) and the MIT for each chemical. MAX FInSLO-LA was the largest value of FInSLO-LA in all repeated experiments. The MIT was defined as the lowest value among the concentrations in all repeated experiments in which the chemical induced FInSLO-LA more than 1.4. MAX FInSLO-LA/MIT was calculated by dividing MAX FInSLO-LA by MIT and each chemical was given a score of 0, 1, or 2 based on the criterion shown in Table 9.

The mean ± SEM of MAX FInSLO-LA/MIT of chemicals in each group with different allergenicity was shown in Fig. 8a. Generally, the MAX FInSLO-LA/MIT values were higher in the group of chemicals containing more potent sensitizers. The data and the scoring system published by Nukada et al. (2013) and Jaworska et al. (2013) were used for DPRA and DEREK. The positive results in DPRA were classified as strong if the average depletion score of cysteine and lysine peptide was above 22.62% and weak if it was above 6.376% and less than or equal to 22.62%. Similarly, if the analysis outcome by DEREK was defined as probable or plausible, the test chemical was judged as a sensitizer. If defined as doubtful or no report, the test chemical was judged as a non-sensitizer. For the 103 test chemicals, the total battery score between 0 and 5 was then calculated by the sum of the individual scores. The box plot indicated the resulting scores split up for the five LLNA potency classifications (i.e., extreme, strong, moderate, weak, and not classified) (Fig. 8b). The median values in the box plot of the IL-8 Luc assay combined with DPRA and DEREK decreased with the corresponding LLNA potency classes with a better linear correlation (R^2 = 0.90) than that of the IL-8 Luc assay alone.

Table 6
Re-evaluation of the chemicals in the ECVAM List by the IL-8 Luc assay using X-VIVO as a solvent.

Chemicals	Experiments		Judgment
	1st	2nd	
Oxazolone	7.69	3.98	Sensitizer
4-NBB	10.03	8.77	Sensitizer
Glyoxal	3.50	2.23	Sensitizer
2-MBT	5.17	4.37	Sensitizer
DNCB	14.54	9.48	Sensitizer
MDGN	4.23	4.35	Sensitizer
Cinnamal	1.53	5.57	Sensitizer
TMTD	4.58	3.61	Sensitizer
PPD	4.30	2.48	Sensitizer
Isoeugenol	1.89	1.59	Sensitizer
Eugenol	2.57	2.52	Sensitizer
Cinnamic alcohol	7.24	7.13	Sensitizer
Glycerol	1.09	1.04	Non-sensitizer
Salicylic acid	1.18	1.00	Non-sensitizer
Lactic acid	1.00	1.08	Non-sensitizer
SLS	2.94	2.88	Sensitizer

Table 7
Performance of the IL-8 Luc assay after deleting the data for detergents and considering human sensitizing potential.

Criteria	IL-8 Luc assay			Modified IL-8 Luc assay		
	Criterion 1	Criterion 2	Criterion 3	Criterion 1	Criterion 2	Criterion 3
Chemical numbers	122	122	122	118	118	118
Accuracy	0.77	0.73	0.73	0.81	0.77	0.77
Sensitivity	0.77	0.80	0.72	0.79	0.80	0.73
Specificity	0.74	0.59	0.74	0.90	0.69	0.90

Modified IL-8 Luc assay: Deleting the data for detergents and considering human sensitizing potential.

Table 8
Performance of the IL-8 Luc assay after deleting the data for detergents, considering human sensitizing potential, and changing to X-VIVO as a solvent.

Criteria	IL-8 Luc assay			Modified IL-8 Luc assay		
	Criterion 1	Criterion 2	Criterion 3	Criterion 1	Criterion 2	Criterion 3
Chemical numbers	122	122	122	118	118	118
Accuracy	0.77	0.73	0.73	0.90	0.82	0.86
Sensitivity	0.77	0.80	0.72	0.90	0.87	0.84
Specificity	0.74	0.59	0.74	0.90	0.69	0.90

Modified IL-8 Luc assay: Deleting the data for detergents, considering human sensitizing potential, and changing to X-VIVO as a solvent.

Next, we set the positive criterion as a score of above 2 for hazard identification. Furthermore, to determine the sensitizing potency using a three-rank classification (strong, weak, and not classified) from the total battery scores, we set the strong score as 5, the weak score as 4, 3, and 2, and not classified as 1 and 0. Table 10 demonstrates the results of hazard identification and shows that the sensitivity, the specificity and the overall accuracy of the score-based battery system was 87.1%, 80.8%, and 85.4%, respectively, according to the Cooper statistics. All these parameters were improved compared to either DPRA or the IL-8 Luc assay alone for the examined chemical sets.

In the potency classification (Table 11), the strong class in this battery system included 14 of 23 extreme and strong sensitizers in the LLNA. Likewise, the weak class included 38 of 54 moderate and weak sensitizers in the LLNA. Thus, the accuracy in the potency prediction with the battery system was 70.9%, the over-prediction rate was 10.7%, and the under-prediction rate was 18.4%. These data suggested that the potential and potency of sensitizing chemicals were identified with good reliability by the score-based battery system.

3.12. Tiered system with h-CLAT and DPRA

We next developed a tiered approach, weighing the predictive performance of the IL-8 Luc assay and DPRA for the 103 evaluated chemicals. The IL-8 Luc assay provided high sensitivity and sufficiently detected extreme and strong sensitizers (Fig. 8a), indicating that the IL-8 Luc assay was a good first step for a tiered approach. The positive results in the IL-8 Luc assay were classified into two classes, strong or weak, based on Max-FInSLO-LA/MIT values. Then, to reliably predict weak and moderate sensitizers, which the IL-8 Luc assay failed to detect, we determined that DPRA was a good second step. The positive results in DPRA were classified into the weak class regardless of the average depletion score. If the chemical scored as negative in both tests, it was considered as not classified. As shown in Table 12, the tiered system provided a relatively high sensitivity of 96.1% (74 of 77 sensitizers) and an accuracy of 87.4% (90 of 103 test chemicals).

The potency classification by the tiered system is summarized in Table 13. All tested chemicals, except for 1 chemical classified as strong in the tiered system, were sensitizers in the LLNA, suggesting a positive predictivity of 98.0% in this class. Moreover, 16 of 19 chemicals categorized as not classified by the tiered system were non-sensitizers in the LLNA, suggesting a negative predictivity of 84.2%. The false negative rate in the weak class was 5.5% (3 of 54 sensitizers). Importantly, the tiered system with the IL-8 Luc assay and DPRA could detect 51 of 54 sensitizers classified as moderate and weak in the LLNA. The strong, weak, and non-classified classes have good correlation with the extreme/strong, moderate/weak, and non-classified class in the LLNA. Thus, the accuracy of potency prediction with the battery system was 62.1%, the over-prediction rate was 28.2%, and the under-prediction rate was 9.7%.

4. Discussion

In parallel with conducting the inter- and intra-laboratory reproducibility tests, we tried to improve the performance of the IL-8 Luc assay. Consequently, we found that the modification of the original protocol of the IL-8 Luc assay by increasing the incubation time, revising the lower limit of I.I.-SLR-LA, deleting the process to examine the effects of NAC, and setting the exclusion criterion could substantially improve the performance of the IL-8 Luc assay. In final, by taking these modifications into account, the performance of the IL-8 Luc assay was accuracy of 81%, sensitivity of 79%, and specificity of 90% in the evaluation of 122 chemicals.

In addition, we here examined the difference in determining positive induction of SLO-LA between Criterion 1 and Criterion 2. Condition 1 used in Criterion 1 was determined empirically, while that used in Criterion 2 was based on the statistical significance. After examining the 122 chemicals, the concordance rate was 90%, which suggests that condition 1 used in Criterion 1, FInSLO-LA $\geqslant$ 1.4 at the concentration of the chemical at which I.I.-SLR-LA is $\geqslant$0.05, is statistically supported. We also accepted the new condition 2 to judge sensitizers based on the repeated experiments: chemicals are categorized as sensitizers when they fulfill condition 1 in two of two to four different experiments, while they were categorized as sensitizers when they fulfilled condition 1 in two or three of three different experiments in our previous report (Takahashi et al., 2011). In this new condition 2, some experiments showing positive induction are repeated only twice to judge sensitizers, while those showing negative induction are repeated three times to judge non-sensitizers. Therefore, to verify this condition, we repeated three experiments for 21 chemicals

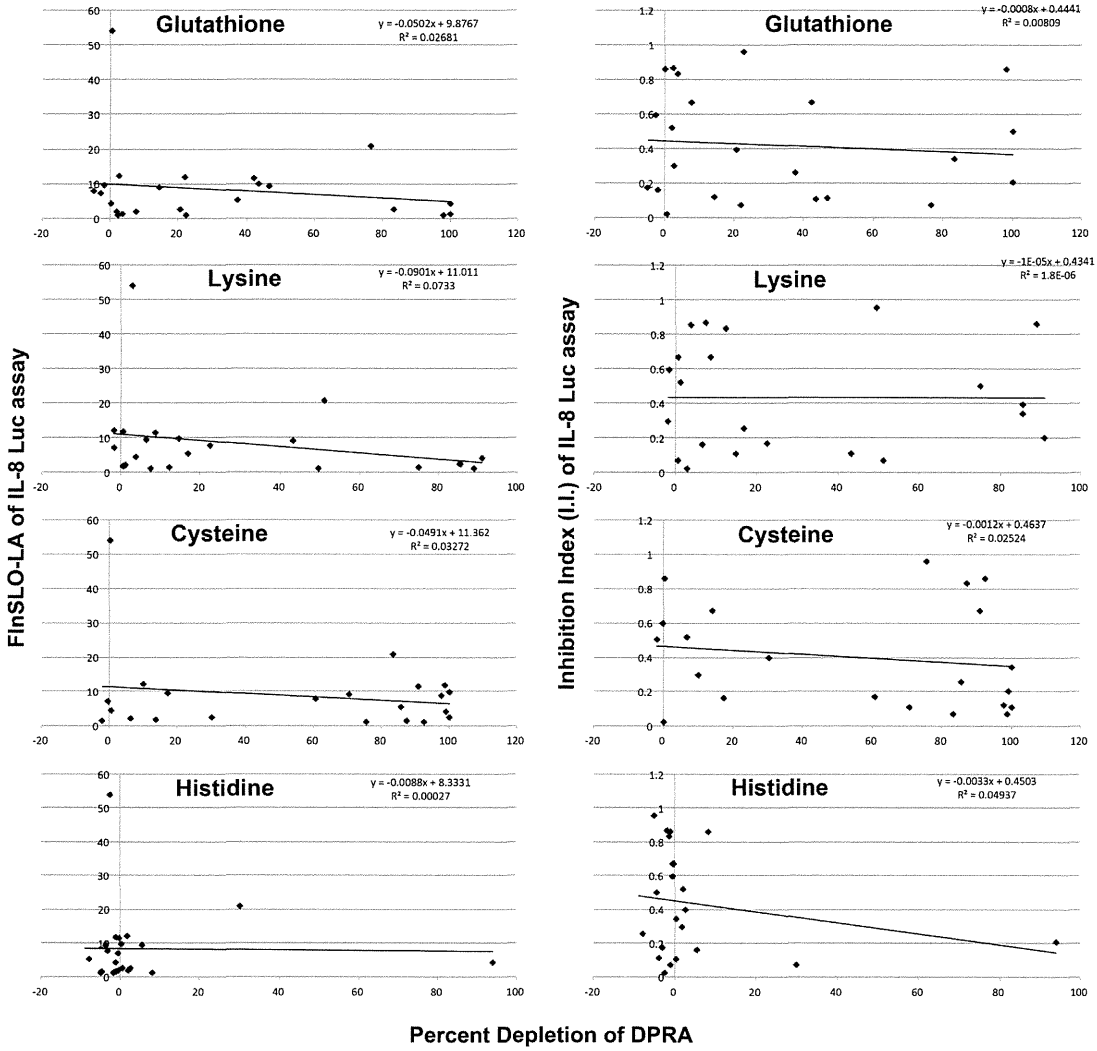

Fig. 6. Correlation between the IL-8 Luc assay and other screening methods (DPRA). The correlation between FlnSLO-LA and the percent depletion of peptides containing either lysine or cysteine in DPRA was examined.

that showed positive induction in the first two experiments. The results demonstrated that the third experiment also showed positive induction in all 21 chemicals, verifying the new condition 2.

In this study, we examined 4 detergents, i.e., hexadecyltrimethylammonium bromide, benzalkonium chloride, Tween-80 and SLS and found that these detergents significantly increased FlnSLO-LA, which suggested that they could stimulate IL-8 promoter activity. These data are consistent with the previous reports that have demonstrated IL-8 production by keratinocytes after the stimulation of various detergents (Coquette et al., 1999; White et al., 2011). Detergents can be categorized according to the charge present in the hydrophilic head (after dissociation in aqueous solution) into four primary groups: anionic, cationic, amphoteric (dual charge) and nonionic (Corazza et al., 2010). Quaternary ammonium compounds, such as hexadecyltrimethylammonium bromide and benzalkonium chloride, are cationic, while Tween-80 and SLS are nonionic and anionic, respectively. Therefore, our data suggested that detergents could stimulate IL-8 promoter activity irrespective of kinds of the charges present in the hydrophilic head. It is not necessarily clear how detergents induce IL-8 mRNA. In contrast to haptens that induce IL-8 mRNA expression by DCs or THP-1 cells depending on p38 MAPK, however, at least SLS has been demonstrated to induce IL-8 mRNA depending on the activation of MEK1/p44/42 ERK (White et al., 2011). Indeed, our previous study demonstrated NAC could not suppress FlnSLO-LA induced by Tween-80, SLS, and benzalkonium chloride, while it significantly attenuated it induced by most haptens (Takahashi et al., 2011). Therefore, it is plausible to include the criterion for examining the effects of NAC on FlnSLO-LA. However, since the IL-8 Luc assay including the criterion to examine the effects of NAC makes the test

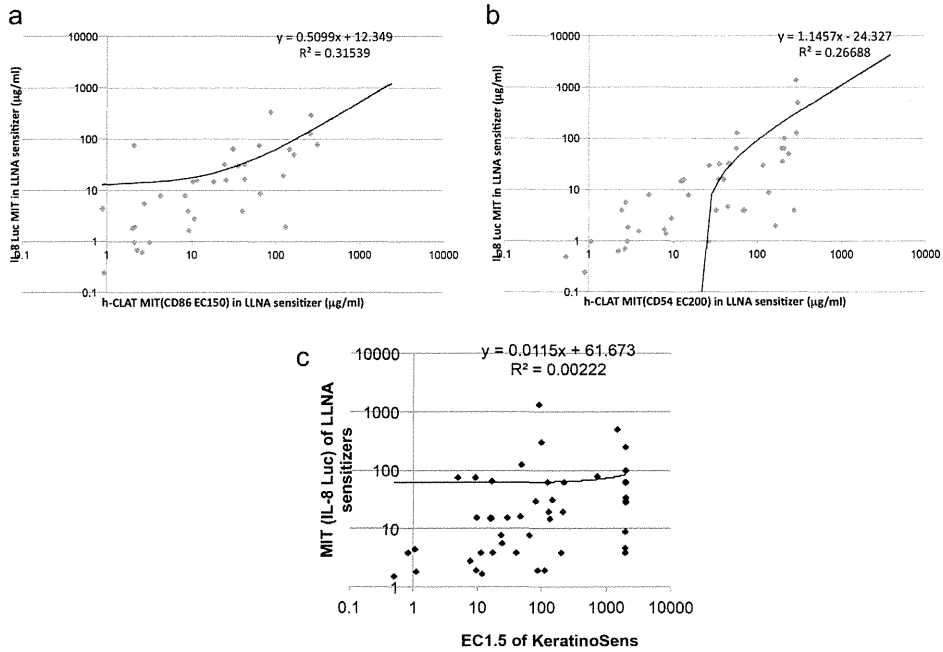

Fig. 7. Correlation between the IL-8 Luc assay and other screening methods (h-CLAT and KeratinoSens). The correlation between the minimum concentration required to induce more than 1.4 of FInSLO-LA in the IL-8 Luc assay (IL-8 Luc assay MIT) and that required to induce more than 150% of CD86 augmentation (h-CLAT MIT (CD86 EC150)) (a), that to induce more than 200% of CD54 in h-CLAT (h-CLAT MIT (CD54 EC200)) (b), and EC 1.5 of KeratinoSens (c).

Table 9
Conversion of the outcome in each single test into scores.

Score	MAX FInSLO-LA/minimum induction threshold (MIT) in IL-8 Luc assay	Avg. score in DPRA	DEREK or times
2	⩾0.1 (Strong positive)	>22.62% (Strong positive)	
1	0.01–0.1 (Weak positive)	6.376–22.62% (Weak positive)	Alert (Positive)
0	<0.01 (Negative)	<6.376% (Negative)	No alert (Negative)

The outcome of each single test was converted to a score, based on the previously reported concept by Jowsey et al. (2006).

method complicated and reduces intra- and inter-laboratory reproducibilities, we would like to include detergents in an exclusion criterion.

It is widely recognized that a single *in vitro* test is insufficient to replace animal testing and that integration of results from different *in vitro* tests, as well as *in silico* methods, is needed for prediction of the skin sensitization potential of chemicals, Since Jowsey et al. (2006) first proposed the integration framework based on the scoring system, weighing the evidence from structure–activity relationships in skin sensitization, penetration, peptide reactivity, and dendritic cell and T-cell activation to evaluate the sensitizing potential as well as the relative potency, a variety of test batteries integrated with different *in vitro* tests and/or *in silico* methods have been reported (Bauch et al., 2012; Jaworska et al., 2011; Natsch et al., 2009; Nukada et al., 2013; Tsujita-Inoue et al., 2014). Most of these approaches substantially improved the accuracy and sensitivity for the potential and potency prediction, compared with LLNA. Therefore, in this study, we also tried to combine the IL-8 Luc assay with other *in vitro* methods to test for skin sensitizing potentials.

Before determining the best combination of test methods, we first examined whether any of the parameters in the IL-8 Luc assay correlate with LLNA potency. Although we could not recognize significant correlation between MIT or FInSLO-LA of the IL-8 Luc assay and LLNA EC3 (data not shown), we found that the MAX FInSLO-LA/MIT values were higher in the group of chemicals with more potent sensitizers. Therefore, we decided to use the MAX FInSLO-LA/MIT with the battery approach.

When we tried score-based battery system and tiered battery system according to the procedure conducted by Nukada et al. (2013), both systems increased accuracy, 87.9% and 90.1%, respectively and the score-based system increased specificity while the tiered-system improved sensitivity. In particular, the chemicals judged as false negative by the IL-8 Luc assay due to the reactivity of chemicals with FBS can be judged as positive by DPRA except for ethylenediamine (Jaworska et al., 2013).

In addition to the yes/no prediction, *in vitro* approach to detect allergenicity of chemicals is required to predict the potency. However, the accuracy in the potency prediction with the scored battery system and the tiered battery system was 67% and 58.2%,

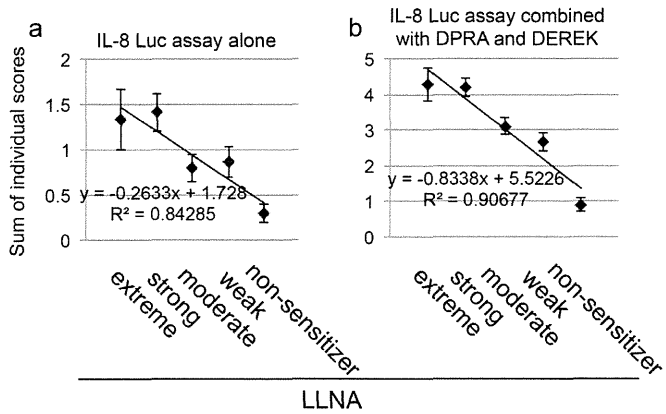

Fig. 8. Score-based battery system based on the IL-8 Luc assay. The mean ± SEM of MAX FInSLO-LA/MIT of chemicals in each group was plotted for different allergenicity groups in the LLNA after conversion of the raw data into a score based on Table 9 (a). The data and the scoring system published by Nukada et al. (2013) and Jaworska et al. (2013) were used for DPRA and DEREK. Then, for the 103 test chemicals, the total battery score between 0 and 5 was calculated by the sum of individual scores (b).

Table 10
Hazard identification by the score-based battery system based on the IL-8 Luc assay.

LLNA	Score-based battery system	
	Positive	Negative
Sensitizers (77)	67	10
Non-sensitizers (26)	5	21
	Sensitivity (%)	87.1
	Specificity (%)	80.8
	Accuracy (%)	85.4

respectively. It is still far from perfect. There is a significant spread for the quantitative data within individual potency classes. LLNA and human data correlate partly with each other, with an R^2 in log–log linear regression between LLNA and human data between 0.45 and 0.75 reported in different studies (ICCVAM, 2011). It means that EC3 in LLNA cannot correctly predict the potency classification of chemicals in humans. In addition, so far, it is not clear how the potency of haptens is determined. Most screening methods do not take T cell response into account. To improve the accuracy, it may need to construct a battery system with a screening method based on T cell response.

Finally, in this study, we explored the factors that impair the performance of the IL-8 Luc assay. Among them, the impacts of FBS on the performance of the IL-8 Luc assay was not negligible. It is a well-known procedure to solubilize water insoluble chemicals in DMSO and dilute the solution with culture medium. This procedure was also employed by h-CLAT and KeratinoSens. Indeed, most of water insoluble haptens we examined in this study significantly increased FInSLO-LA after diluted by this procedure, although we must admit that the IL-8 Luc assay with this procedure produced considerable numbers of false negative results. Surprisingly, the judgment of 5 among 14 haptens in 122 chemical lists we examined was corrected by diluting chemicals with X-VIVO. Saito et al. have reported the similar observation, in which hapten-induced ROS production by THP-1 was significantly attenuated in the presence of FBS (Saito et al., 2013). It is conceivable that FBS suppresses the binding of some haptens with cysteine or lysine residues. Therefore, our novel procedure in which chemicals are diluted with X-VIVO may significantly improve the performance of the IL-8 Luc assay.

There are several advantages in the IL-8 Luc assay. At first, the culture of THP-G8 cells are relatively simple and does not use trypsin or EDTA because THP-G8 cells do not stick to the culture dishes. The second is its simple procedure. At first, chemicals in graded concentrations are added into 96-well culture plate as required in every *in vitro* test method. Then, the cells adjusted to the optimum concentration are seeded to each plate. After 16 h incubation, the plates are set in the luminometer. The process afterward is completely automated except calculating the obtained results in the predesigned Excel sheet. Therefore, the IL-8 Luc assay is considered as a test method that can significantly reduce human errors.

Moreover, the IL-8 Luc assay does not need the step to preculture or determine cell viability after chemical treatment. In the IL-8 Luc assay, since THP-G8 cells can present the promoter activities of IL-8 promoter and GAPDH, a well known house keeping gene, the information of the effects of chemicals on both IL-8 induction and cell viability is obtained simultaneously in each experiment. Therefore, even though 4 experiments are required, one set of experiments can be completed within 4 days. Therefore, the IL-8 Luc assay is a truly high-through method.

Finally, this study succeeded in optimizing the IL-8 Luc assay to predict allergenicity of chemicals and demonstrated that the IL-8

Table 11
Potency classification by the score-based battery system based on the IL-8 Luc assay.

LLNA	Score-based battery system		
	Strong	Weak	Not-classified
Extreme + strong (23)	14	9	0
Moderate + weak (54)	6	38	10
Non-sensitizer (26)	0	5	21
Over prediction rate (%)		10.7% (11/91)	
Under prediction rate (%)		18.4% (19/103)	
Accuracy (%)		70.9% (73/103)	

Table 12
Hazard identification by the tiered system based on the IL-8 Luc assay.

LLNA	Tiered system	
	Positive	Negative
Sensitizers (77)	74	3
Non-sensitizers (26)	10	16
	Sensitivity (%)	96.1
	Specificity (%)	61.5
	Accuracy (%)	87.4

Table 13
Potency classification by tiered battery system based on the IL-8 Luc assay.

LLNA	Tiered system		
	Strong	Weak	Not-classified
Extreme + strong (23)	16	7	0
Moderate + weak (54)	19	32	3
Non-sensitizer (26)	1	9	16
Over prediction rate (%)		28.2% (29/103)	
Under prediction rate (%)		9.7% (10/103)	
Accuracy (%)		62.1% (64/103)	

Luc assay is a promising *in vitro* alternative method. Furthermore, the battery approach of the IL-8 Luc assay combined with DPRA and DEREK could significantly improve the performance.

Conflict of Interest

The authors declare that there are no conflicts of interest.

Transparency Document

The Transparency document associated with this article can be found in the online version.

Acknowledgments

This work was supported in part by the New Energy and Industrial Development Organization (NEDO), by Grant-in-Aid for Challenging Exploratory Research of JSPS KAKENHI Grant Number 26670473, by Health Labor Sciences Research in Japan, and by Ministry Economy, Science and Industry in Japan, and Japanese Society for Alternatives to Animal Experiments.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.tiv.2015.07.006.

References

Ashikaga, T., Sakaguchi, H., Sono, S., Kosaka, N., Ishikawa, M., Nukada, Y., Miyazawa, M., Ito, Y., Nishiyama, N., Itagaki, H., 2010. A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA). ATLA 38, 275–284.

Ashikaga, T., Yoshida, Y., Hirota, M., Yoneyama, K., Itagaki, H., Sakaguchi, H., Miyazawa, M., Ito, Y., Suzuki, H., Toyoda, H., 2006. Development of an in vitro skin sensitization test using human cell lines: the human Cell Line Activation Test (h-CLAT). I. Optimization of the h-CLAT protocol. Toxicol. In Vitro 20, 767–773.

Basketter, D.A., Alepee, N., Ashikaga, T., Barroso, J., Gilmour, N., Goebel, C., Hibatallah, J., Hoffmann, S., Kern, P., Martinozzi-Teissier, S., Maxwell, G., Reisinger, K., Sakaguchi, H., Schepky, A., Tailhardat, M., Templier, M., 2014. Categorization of chemicals according to their relative human skin sensitizing potency. Dermatitis 25, 11–21.

Bauch, C., Kolle, S.N., Ramirez, T., Eltze, T., Fabian, E., Mehling, A., Teubner, W., van Ravenzwaay, B., Landsiedel, R., 2012. Putting the parts together: combining in vitro methods to test for skin sensitizing potentials. Regul. Toxicol. Pharmacol. 63, 489–504.

Casati, S., Aeby, P., Kimber, I., Maxwell, G., Ovigne, J.M., Roggen, E., Rovida, C., Tosti, L., Basketter, D., 2009. Selection of chemicals for the development and evaluation of in vitro methods for skin sensitisation testing. ATLA 37, 305–312.

Coquette, A., Berna, N., Vandenbosch, A., Rosdy, M., Poumay, Y., 1999. Differential expression and release of cytokines by an in vitro reconstructed human epidermis following exposure to skin irritant and sensitizing chemicals. Toxicol. In Vitro 13, 867–877.

Corazza, M., Lauriola, M.M., Zappaterra, M., Bianchi, A., Virgili, A., 2010. Surfactants, skin cleansing protagonists. J. Eur. Acad. Dermatol. Venereol. 24, 1–6.

Divkovic, M., Pease, C.K., Gerberick, G.F., Basketter, D.A., 2005. Hapten-protein binding: from theory to practical application in the in vitro prediction of skin sensitization. Contact Dermatitis 53, 189–200.

Edwards, D.R., Denhardt, D.T., 1985. A study of mitochondrial and nuclear transcription with cloned cDNA probes. Changes in the relative abundance of mitochondrial transcripts after stimulation of quiescent mouse fibroblasts. Exp. Cell Res. 157, 127–143.

Emter, R., Ellis, G., Natsch, A., 2010. Performance of a novel keratinocyte-based reporter cell line to screen skin sensitizers in vitro. Toxicol. Appl. Pharmacol. 245, 281–290.

Feijoo, E., Alfaro, C., Mazzolini, G., Serra, P., Penuelas, I., Arina, A., Huarte, E., Tirapu, I., Palencia, B., Murillo, O., Ruiz, J., Sangro, B., Richter, J.A., Prieto, J., Melero, I., 2005. Dendritic cells delivered inside human carcinomas are sequestered by interleukin-8. Int. J. Cancer 116, 275–281.

Gerberick, G.F., Ryan, C.A., Kern, P.S., Schlatter, H., Dearman, R.J., Kimber, I., Patlewicz, G.Y., Basketter, D.A., 2005. Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. Dermatitis 16, 157–202.

Gerberick, G.F., Vassallo, J.D., Bailey, R.E., Chaney, J.G., Morrall, S.W., Lepoittevin, J.P., 2004. Development of a peptide reactivity assay for screening contact allergens. Toxicol. Sci. 81, 332–343.

Gerberick, G.F., Vassallo, J.D., Foertsch, L.M., Price, B.B., Chaney, J.G., Lepoittevin, J.P., 2007. Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol. Sci. 97, 417–427.

Gouwy, M., Struyf, S., Leutenez, L., Portner, N., Sozzani, S., Van Damme, J., 2014. Chemokines and other GPCR ligands synergize in receptor-mediated migration of monocyte-derived immature and mature dendritic cells. Immunobiology 219, 218–229.

Holland, R., Fishbein, J.C., 2010. Chemistry of the cysteine sensors in Kelch-like ECH-associated protein 1. Antioxid. Redox Signal. 13, 1749–1761.

Hopkins, J.E., Naisbitt, D.J., Kitteringham, N.R., Dearman, R.J., Kimber, I., Park, B.K., 2005. Selective haptenation of cellular or extracellular protein by chemical allergens: association with cytokine polarization. Chem. Res. Toxicol. 18, 375–381.

ICCVAM, 2011. ICCVAM Test Method Evaluation Report: Usefulness and Limitations of the Murine Local Lymph Node Assay for Potency Categorization of Chemicals Causing Allergic Contact Dermatitis in Humans. <http://iccvam.niehs.nih.gov/docs/immunotox_docs/LLNA-pot/TMER.pdf>.

Itoh, K., Mimura, J., Yamamoto, M., 2010. Discovery of the negative regulator of Nrf2, Keap1: a historical overview. Antioxid. Redox Signal. 13, 1665–1678.

Jaworska, J., Dancik, Y., Kern, P., Gerberick, F., Natsch, A., 2013. Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice. J. Appl. Toxicol. 33, 1353–1364.

Jaworska, J., Harol, A., Kern, P.S., Gerberick, G.F., 2011. Integrating non-animal test information into an adaptive testing strategy – skin sensitization proof of concept case. ALTEX 28, 211–225.

Jimenez, F., Quinones, M.P., Martinez, H.G., Estrada, C.A., Clark, K., Garavito, E., Ibarra, J., Melby, P.C., Ahuja, S.S., 2010. CCR2 plays a critical role in dendritic cell maturation: possible role of CCL2 and NF-kappa B. J. Immunol. 184, 5571–5581.

Jowsey, I.R., Basketter, D.A., Westmoreland, C., Kimber, I., 2006. A future approach to measuring relative skin sensitising potency: a proposal. J. Appl. Toxicol. 26, 341–350.

Leonard, E.J., Skeel, A., Yoshimura, T., Noer, K., Kutvirt, S., Van Epps, D., 1990. Leukocyte specificity and binding of human neutrophil attractant/activation protein-1. J. Immunol. 144, 1323–1330.

Marchese, A., Heiber, M., Nguyen, T., Heng, H.H., Saldivia, V.R., Cheng, R., Murphy, P.M., Tsui, L.C., Shi, X., Gregor, P., et al., 1995. Cloning and chromosomal mapping of three novel genes, GPR9, GPR10, and GPR14, encoding receptors related to interleukin 8, neuropeptide Y, and somatostatin receptors. Genomics 29, 335–344.

Mitjans, M., Viviani, B., Lucchi, L., Galli, C.L., Marinovich, M., Corsini, E., 2008. Role of p38 MAPK in the selective release of IL-8 induced by chemical allergen in naive THP-1 cells. Toxicol. In Vitro 22, 386–395.

Mitjans, M., Galbiati, V., Lucchi, L., Viviani, B., Marinovich, M., Galli, C.L., Corsini, E., 2010. Use of IL-8 release and p38 MAPK activation in THP-1 cells to identify allergens and to assess their potency in vitro. Toxicol. In Vitro 24, 1803–1809.

Mori, R., Wang, Q., Danenberg, K.D., Pinski, J.K., Danenberg, P.V., 2008. Both beta-actin and GAPDH are useful reference genes for normalization of quantitative RT-PCR in human FFPE tissue samples of prostate cancer. Prostate 68, 1555–1560.

Murphy, P.M., Tiffany, H.L., 1991. Cloning of complementary DNA encoding a functional human interleukin-8 receptor. Science 253, 1280–1283.

Nakajima, Y., Kimura, T., Sugata, K., Enomoto, T., Asakawa, A., Kubota, H., Ikeda, M., Ohmiya, Y., 2005. Multicolor luciferase assay system: one-step monitoring of multiple gene expressions with a single substrate. Biotechniques 38, 891–894.

Natsch, A., Emter, R., Ellis, G., 2009. Filling the concept with data: integrating data from different in vitro and in silico assays on skin sensitizers to explore the battery approach for animal-free skin sensitization testing. Toxicol. Sci. 107, 106–121.

Natsuaki, Y., Egawa, G., Nakamizo, S., Ono, S., Hanakawa, S., Okada, T., Kusuba, N., Otsuka, A., Kitoh, A., Honda, T., Nakajima, S., Tsuchiya, S., Sugimoto, Y., Ishii, K.J., Tsutsui, H., Yagita, H., Iwakura, Y., Kubo, M., Ng, L.G., Hashimoto, T., Fuentes, J., Guttman-Yassky, E., Miyachi, Y., Kabashima, K., 2014. Perivascular leukocyte clusters are essential for efficient activation of effector T cells in the skin. Nat. Immunol. 15, 1064–1069.

Noguchi, T., Ikeda, M., Ohmiya, Y., Nakajima, Y., 2008. Simultaneous monitoring of independent gene expression patterns in two types of cocultured fibroblasts with different color-emitting luciferases. BMC Biotechnol. 8, 40.
Nukada, Y., Miyazawa, M., Kosaka, N., Ito, Y., Sakaguchi, H., Nishiyama, N., 2008. Production of IL-8 in THP-1 cells following contact allergen stimulation via mitogen-activated protein kinase activation or tumor necrosis factor-alpha production. J. Toxicol. Sci. 33, 175–185.
Nukada, Y., Ashikaga, T., Miyazawa, M., Hirota, M., Sakaguchi, H., Sasa, H., Nishiyama, N., 2012. Prediction of skin sensitization potency of chemicals by human Cell Line Activation Test (h-CLAT) and an attempt at classifying skin sensitization potency. Toxicol. In Vitro 26, 1150–1160.
Nukada, Y., Miyazawa, M., Kazutoshi, S., Sakaguchi, H., Nishiyama, N., 2013. Data integration of non-animal tests for the development of a test battery to predict the skin sensitizing potential and potency of chemicals. Toxicol. In Vitro 27, 609–618.
OECD, 2012. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins, Part 1: Scientific Evidence. OECD Environment, Health and Safety Publications, Series on Testing and Assessment No. 168.
Piroird, C., Ovigne, J.M., Rousset, F., Martinozzi-Teissier, S., Gomes, C., Cotovio, J., Alepee, N., 2015. The Myeloid U937 Skin Sensitization Test (U-SENS) addresses the activation of dendritic cell event in the adverse outcome pathway for skin sensitization. Toxicol. In Vitro 29, 901–916.
Python, F., Goebel, C., Aeby, P., 2007. Assessment of the U937 cell line for the detection of contact allergens. Toxicol. Appl. Pharmacol. 220, 113–124.
Saito, K., Miyazawa, M., Nukada, Y., Sakaguchi, H., Nishiyama, N., 2013. Development of an in vitro skin sensitization test based on ROS production in THP-1 cells. Toxicol. In Vitro 27, 857–863.
Saito, R., Hirakawa, S., Ohara, H., Yasuda, M., Yamazaki, T., Nishii, S., Aiba, S., 2011. Nickel differentially regulates NFAT and NF-kappaB activation in T cell signaling. Toxicol. Appl. Pharmacol. 254, 245–255.
Sakaguchi, H., Ashikaga, T., Miyazawa, M., Yoshida, Y., Ito, Y., Yoneyama, K., Hirota, M., Itagaki, H., Toyoda, H., Suzuki, H., 2006. Development of an in vitro skin sensitization test using human cell lines: human Cell Line Activation Test (h-CLAT). II. An inter-laboratory study of the h-CLAT. Toxicol. In Vitro 20, 774–784.
Sebok, K., Woodside, D., Al-Aoukaty, A., Ho, A.D., Gluck, S., Maghazachi, A.A., 1993. IL-8 induces the locomotion of human IL-2-activated natural killer cells. Involvement of a guanine nucleotide binding (Go) protein. J. Immunol. 150, 1524–1534.
Singha, B., Gatla, H.R., Manna, S., Chang, T.P., Sanacora, S., Poltoratsky, V., Vancura, A., Vancurova, I., 2014. Proteasome inhibition increases recruitment of IkappaB kinase beta (IKKbeta), S536P-p65, and transcription factor EGR1 to interleukin-8 (IL-8) promoter, resulting in increased IL-8 production in ovarian cancer cells. J. Biol. Chem. 289, 2687–2700.
Takahashi, T., Kimura, Y., Saito, R., Nakajima, Y., Ohmiya, Y., Yamasaki, K., Aiba, S., 2011. An in vitro test to screen skin sensitizers using a stable THP-1-derived IL-8 reporter cell line, THP-G8. Toxicol. Sci. 124, 359–369.
Takenouchi, O., Miyazawa, M., Saito, K., Ashikaga, T., Sakaguchi, H., 2013. Predictive performance of the human Cell Line Activation Test (h-CLAT) for lipophilic chemicals with high octanol-water partition coefficients. J. Toxicol. Sci. 38, 599–609.
Toebak, M.J., Pohlmann, P.R., Sampat-Sardjoepersad, S.C., von Blomberg, B.M., Bruynzeel, D.P., Scheper, R.J., Rustemeyer, T., Gibbs, S., 2006. CXCL8 secretion by dendritic cells predicts contact allergens from irritants. Toxicol. In Vitro 20, 117–124.
Trompezinski, S., Migdal, C., Tailhardat, M., Le Varlet, B., Courtellemont, P., Haftek, M., Serres, M., 2008. Characterization of early events involved in human dendritic cell maturation induced by sensitizers: cross talk between MAPK signalling pathways. Toxicol. Appl. Pharmacol. 230, 397–406.
Tsujita-Inoue, K., Hirota, M., Ashikaga, T., Atobe, T., Kouzuki, H., Aiba, S., 2014. Skin sensitization risk assessment model using artificial neural network analysis of data from multiple in vitro assays. Toxicol. In Vitro 28, 626–639.
White, K.J., Maffei, V.J., Newton-West, M., Swerlick, R.A., 2011. Irritant activation of epithelial cells is mediated via protease-dependent EGFR activation. J. Invest. Dermatol. 131, 435–442.
Winer, J., Jung, C.K., Shackel, I., Williams, P.M., 1999. Development and validation of real-time quantitative reverse transcriptase-polymerase chain reaction for monitoring gene expression in cardiac myocytes in vitro. Anal. Biochem. 270, 41–49.
Zhang, X., Chen, X., Song, H., Chen, H.Z., Rovin, B.H., 2005. Activation of the Nrf2/antioxidant response pathway increases IL-8 expression. Eur. J. Immunol. 35, 3258–3267.

Send Orders for Reprints to reprints@benthamscience.ae

Combinatorial Chemistry & High Throughput Screening, 2015, *18*, 000-000 1

Simultaneous Multicolor Luciferase Reporter Assays for Monitoring of Multiple Genes Expressions

Yoshihiro Ohmiya*

DAILAB, Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan

Yoshihiro Ohmiya

Abstract: Multicolor luciferase reporter assays are powerful tools for simultaneously monitoring the expression of several genes. Dual-color luciferase reporter assays were established in 2004 and have been used in broad range of application from basic researches to applications in clinical research, e.g., drug screening and alternative toxicological tests. The advantage of this method is that it can be used to simultaneously monitor sequential and complex changes in the expression of multiple genes from the single-cell level to the tissue level. In this review, I focus on the simultaneous analysis of multiple genes using multicolor beetle luciferases.

Keywords: Beetle luciferase, bioluminescence, multicolor luciferase assay, imaging luciferase assay, reporter assay.

INTRODUCTION

Bioluminescence is a form of "cold light" in living organisms, based on the chemical reaction; that is, the measured light activity depends on the amount of enzyme, substrate, and cofactor(s). Bioluminescence reporters, such as luciferase (enzyme) and luciferin (substrate), are widely used for the analysis of various biological functions, such as protein-protein interactions, and gene expression [1-3]. The structural relationship between luciferin and luciferase is restricted in each bioluminescence reaction. Additionally, bioluminescence can be reproduced *in vitro* using a chemically synthesized luciferin reporter and a genetically produced luciferase enzyme. The bioluminescence system comprising these two components produces a unique light source used widely in the field of biotechnology for visualizing biological functions. Thus, luciferases are good candidate reporter enzymes in cell-based assays. Advanced biotechnologies using multicolor luciferases permit the quantitative analysis of multiple genes from the single-cell level to the tissue level using a luminometer/charged-coupled device (CCD) camera equipped with optical filters. Bioluminescent protein techniques are sensitive approaches for understanding cell physiology, whereas fluorescence imaging that use fluorescent proteins (e.g., green fluorescent protein [GFP]) as probes have contributed to the understanding of cell physiology researches. Especially, bioluminescent probes are highly sensitive, display good linearity, and are easy to perform compared with other fluorescent probes. Beetle luciferases are particularly suitable as reporter enzyme because they produce stable color lights ranging from green to red under any condition; in contrast, the light produced by firefly luciferases can change depending on pH or temperature. In this review, I discuss the simultaneous analysis of multiple genes using multicolor beetle luciferases.

*Address correspondence to this author at the Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki, Japan;
Tel: ??????????????????????????????; Fax: ??????????????????????????????;
E-mail: y-ohmiya@aist.go.jp

FIREFLY AND BEETLE BIOLUMINESCENCE

Bioluminescence is a well-known phenomenon. In the presence of firefly and beetle luciferases as a catalytic enzyme and ATP and Mg^{2+} as cofactors, the bioluminescence reactions produc photons (Fig. **1**) [4-6]. In light-emitting reactions, the emission maxima of firefly D-luciferin-type bioluminescence are found at around 535-630 nm (Fig. **2A**, **B**). Beetle bioluminescence is also known to produce much higher quantum yield (QY) values than other types of luminescence [7, 8]. The QY value of native firefly D-luciferin varies from 0.15 to 0.60 depending on the luciferase enzyme; for example, the green-yellow-emitting luciferase of the American firefly (*Photinus pyralis*) has a QY value of 0.41. Moreover, the QY value of the firefly luciferase is correlated with the luminescence peak wavelength. The red-emitting luciferase of the Brazilian railroad worm (Phrixothrix hirtus), which has the reddish peak wavelength, has the lowest QY (0.15), whereas the green-emitting luciferase of the Brazilian click beetle (Pyrearinus termitilluminans), whose luminescence spectrum has the greenish peak of all reported luciferases, possesses the highest quantum yield (0.61) [9].

Interestingly, beetle bioluminescence, including that of fireflies, can be classed into two large groups according to bioluminescent spectral sensitivity to pH: 1) pH-sensitive luciferases (fireflies) that exhibit a red color shift with increasing temperature or decreasing pH [4, 5]; and 2) pH insensitive (railroad worm and click beetle) luciferases that are not affected by these changes in pH and temperature [9, 10]. Additionally, the spectra of pH-insensitive beetle luciferases are narrower than those of pH-sensitive firefly luciferases, although many color mutants have still been constructed. For example, Fig. (**2**) shows the corrected luminescence spectra for pH-insensitive (Fig. **2A**) and -sensitive luciferases (Fig.

1386-2073/15 $58.00+.00

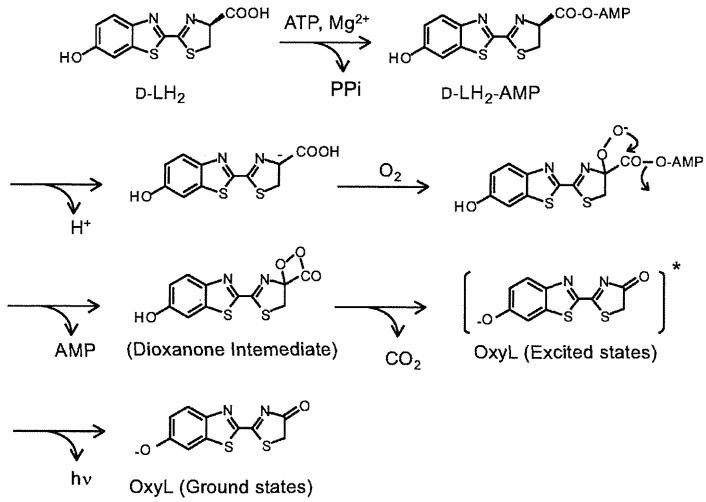

Fig. (1). Proposal mechanism of chemical reaction in firefly bioluminescence. Luciferin (D-LH_2) is converted into an adenylate (D-LH_2-AMP) in the presence of ATP, which is oxidized in the presence of oxygen, forming a peroxide intermediate by splitting of AMP. Decomposition of the intermediate *via* the dioxetanone intermediate is sufficiently energetic to produce the excited state of the oxyluciferin (OxyL) monomer or dianion, and then to produce the light from excited state to ground state.

2B). The spectra of the multicolor luciferase mutants are broader than those of pH-insensitive red and green light-emitting luciferases [9]. For the various beetle luciferases, the QY values and spectra of the bioluminescence reaction have been analyzed in order to explain their relationships.

The mechanisms mediating the different color emissions of beetle bioluminescence were proposed based on active site determined by the 3D-structure of the *Luciola cruciata* firefly luciferase [11]. For example, Hirano *et al.* investigated the spectroscopic properties of the phenolate anion of firefly luciferin and proposed that the excited luciferin as a light emitter was modulated by the polarity of the active-site environment of firefly luciferase and protonated basic moiety in the active site [12]. On the other hand, the emitter of firefly bioluminescence may relate to the lability of firefly oxyluciferin. Maltsev *et al.* explained that its lability is due to autodimerization of the coexisting enol and keto forms in a Mannich-type reaction based on NMR spectroscopy and X-ray crystallography data of a side product [13]. Nazivet *et al.* demonstrated that emitting light depends on the micro-environmental polarity at the phenolate/phenol of the benzothiazole fragment in oxyluciferin, and furthermore, denied that the color modulation of the emitting light depends on the size of the compact luciferase protein which is a cavity embedding the excited oxyluciferin molecule [14]. Nazivet *et al.* also demonstrated that based on a model of Cypridina oxyluciferin and coelenteramide, carbonyl group of firefly oxyluciferin or the different chemical environment of the dioxetanone is more rigid in the firefly bioluminescence

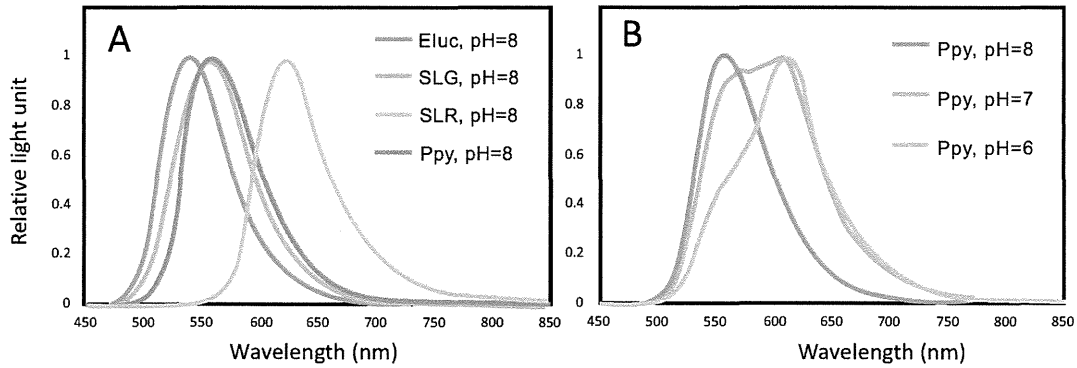

Fig. (2). Luminescence spectra of bioluminescence reactions for various beetle luciferases. Spectra were measured using PicaGene reagent at 24°C, corrected for spectral sensitivity of the sphere spectrometer, and normalized at each luminescence maximum [9]. Ppy; *Photinus pyralis* luciferase, ELuc, SLG, SLR (see Table **1**).

system [15]. However, the mechanisms mediating the pH sensitivity (or insensitivity) of the beetle luciferase enzyme remain unknown. In particular, the mechanism of pH insensitivity in the beetle luciferases is not clear because the 3D-structure of pH-insensitive luciferases has not been determined. Thus, improving our understanding of the mechanisms of beetle bioluminescence could facilitate the development of novel luciferase assays and biological applications of luciferases in a variety of research fields.

CELL-BASED MULTICOLOR LUCIFERASE REPORTER ASSAYS

Fig. (**3**) demonstrates a simple luciferase reporter assay. In general, the reporter plasmid containing the luciferase gene plus the target promoter region of interest is transfected into target cells, and luciferase-expressing cells are lysed after an appropriate period, e.g., 1–2 days. We can measure the amount of expressed luciferase as a light signal *in vitro*, and we can estimate the target promoter activity as a light intensity. Table **1** summarizes the characteristic properties of bioluminescence systems based on commercially available luciferases from beetles, sea pansy, copepods, and ostracods. Commercially available luciferins comprise only three types: firefly D-luciferin, coelenterazine and *Cypridina* luciferin, although other luciferins, including dinoflagellate [16] and *Latia* luciferins [17], have been identified. In light-emitting reactions, the emission maxima of firefly D-luciferin-type, coelenterazine-type, and *Cypridina* luciferin-type bioluminescence are found at around 535–630, 460–480, and 460 nm, respectively. The molecular weights of luciferases vary widely (20–62 kDa), and their molecular structures, which originate from phylogenetically distant systems, belong to different super families. *Cypridina* [18], *Gaussia* [19], and *Metridia* [20] luciferases (Cluc, GLuc, and MetLuc, respectively) are secreted enzymes. The luciferin-luciferase reaction is triggered by adding luciferin, although the bioluminescence of firefly, click beetle and railroad worm luciferases requires ATP and magnesium ions as cofactors.

Advances in the luciferase assay system have an additional luciferase as an internal control reporter (dual-reporter assay), thereby minimizing inherent experimental variability that can undermine experimental accuracy, such as differences in the number and viability of cells used and the efficiency of cell transfection and lysis. Thus, the first generation of dual-reporter assays is a combination of firefly luciferase and sea pansy *Renilla* luciferase using firefly D-luciferin-type and coelenterazine-type bioluminescence. Of the luciferases identified to date, the firefly luciferase from *P. pyralis* is the commonly used bioluminescent reporter in commercial vectors. Only the expression of one gene or one target event can be monitored at a time, although this luciferase has been extensively used to monitor cellular events in cell-based assays and *in vivo* imaging [21-23].

Gene expression events are both complex and sequential because of the elaborated regulatory pathways found in living cells. Many researchers have sought new reporter assay systems, focusing on the characteristics of multicolor beetle luciferases. Advanced luciferase technology, involving progressives in both the luciferase and the detection equipment, as well as newly cloned luciferase genes, have allowed us to simultaneously monitor the expressions of multiple genes when luciferases are used that induce different color emission spectra in the catalysis of a single D-luciferin substrate. The advantages of beetle luciferases producing multiple colors are as follows: (1) the colors are separable with an optical filter; (2) the number of substrates is minimal; (3) the temperature dependences are similar; and (4) the half-lives are similar (Fig. **4**). Thus, these mixed emission spectra are measured simultaneously. Each intensities can be quantified by splitting them with optical filter(s) [24].

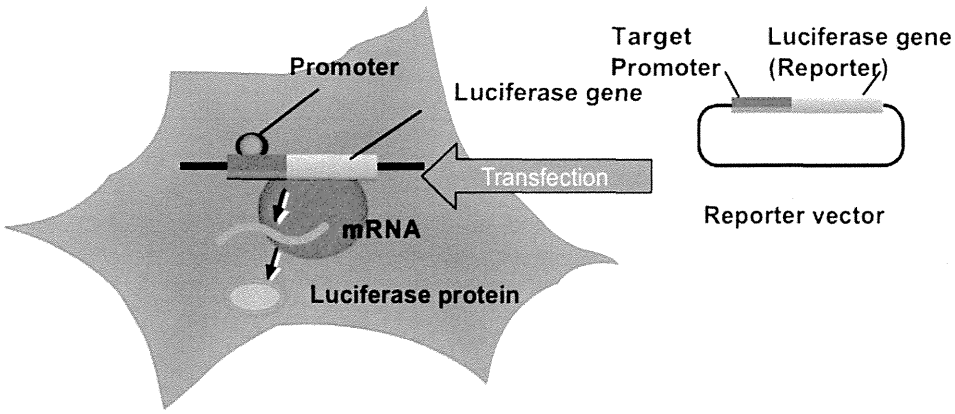

Fig. (3). Principle of a simple luciferase reporter assay. The reporter plasmid vector consists of the target promoter sequence and luciferase gene sequence. After transfection of the plasmid into target cells, the promoter region regulates the expression of luciferase gene in living cells. The expressed luciferase protein catalyzes a reaction with luciferin to produce light. In the transient transfection luciferase assay, luciferase-expressing cells are lysed for an appropriate period. The amount of expressed luciferase protein can be estimated from the light intensity which indicates the promoter activity in living cells. In this case, the promoter activity is normalized by cell numbers or cellular enzymatic activity.

Table 1. Summary of characteristic properties of commercialized bioluminescence system.

Organism	Gene Symbol	Luciferin	Mass (kDa)	λmax (nm)	Main Company
Non-Secreted					
Firefly	luc(+), luc2	firefly luciferin	61	562	Promega
Sea pansy	Rluc	coelenterazine	36	480	Promega
Click beetle (Jamaica)	CBGluc	firefly luciferin	60	537	Promega
Click beetle (Jamaica)	CBRluc	firefly luciferin	60	613	Promega
Click beetle (Brazil)	ELuc	firefly luciferin	61	638	TOYOBO
Railroad-worm	SLR	firefly luciferin	60	630	TOYOBO
Railroad-worm (Japan)	SLG	firefly luciferin	60	550	TOYOBO
Railroad-worm (Japan)	SLO	firefly luciferin	60	580	TOYOBO
Secreted					
Copepoda	GLuc	coelenterazine	20	480	NEB
Copepoda	MetLuc	coelenterazine	24	480	Clontech
Ostracod	Cluc	cypridinid luciferin	61	465	NEB

APPLICATION OF CELL-BASED MULTICOLOR LUCIFERASE REPORTER ASSAY

In the first example of dual-color luciferase assays in 2004, Kitayama *et al.* constructed a simple dual-reporter system monitored simultaneously two promoter activities in living cyanobacterial cells [24]. Two Phrixothrix railroad-worm luciferases [10] catalyzing the generation of different color emissions served as the dual reporters; each emissions was separated by interference filters to estimate the individual emission signals using photomultiplier tubes. Using this system, they clearly demonstrated the expression profiles between promoters in the same cells.

As a second example, Nakajima *et al.* developed a simultaneous monitoring system [25] in mammalian cell line using green light emitting luciferase (SLG, see Table **1**) and red light-emitting luciferase (SLR, see Table **1**) from the Phrixothrix railroad worm. The two spectral mixed emissions were divided using a > 600-nm long-pass filter, and the respective luciferase activities were calculated. Splitting the emissions with a long-pass filter is advantageous in that the emission loss is less than when interference filters are used. The linear response range of this system using cell extracts which are expressed the green and red light-emitting luciferases could be estimated to be more than two orders of magnitude.

Furthermore, Nakajima *et al.* established a tricolor reporter assay using SLG, orange-emitting luciferase (SLO, see Table **1**), and SLR [26]. Because all of these enzymes emit light with D-luciferin, their respective activities can be detected in a one-step reaction from a single sample. Using this method, they estimated that the linear response range of the system exceeds two orders of magnitude, although the low-threshold light intensities require one order of magnitude higher intensity than those estimated in the dual-color luciferase assay. Using this system, they introduced SLG as an internal control reporter and measured its activity separately from those of SLO for *Bmal1* and SLR for the part of *Bmal1* promoters for retinoic acid receptor-related orphan receptor α response element (RORE) (Fig. **5**). They demonstrated a simultaneous monitoring of the clock genes expressions *in vitro* and clarified the role of retinoic acid receptor-related orphan receptor α (RORα) in the transcriptional regulation of the clock genes *Bmal1* and RORE in mammalian cell.

Since 2005, many researchers have published papers describing the application of the dual-color luciferase assay. Branchini *et al.* developed a model system for the dual-color luciferase assay by combination of green light-(Val241Ile/Gly246Ala/Phe250Ser, λmax = 549 nm) and red (Ser284Thr, λmax = 615 nm) emitting firefly luciferase mutants. They divided the emissions using two band-pass filters and confirmed that the green light-emitting GST fusion protein could be measured over a 10,000-fold range from about 20 amol to 200 fmol when the amount (10 fmol) of red light-emitting GST fusion protein was kept constant in the crude cell extracts [27]. Moreover, based on the results, they generated thermostable firefly luciferase mutants and quantified both luciferase activities at 1.0 fmol in a mixture using the microplate luminometer format [28]. Using a similar approach, Michelini *et al.* simultaneously monitored the expression of two genes in cell extracts using two beetle luciferases (green light-emitting luciferase from *P. pyralis* [λmax = 560 nm] and red light-emitting luciferase from *Luciola italica* [λmax = 613 nm]), which emit light *via* reaction with firefly D-luciferin as the experimental reporter. The latter enzyme is a mutant of the wild-type form, displaying a red-shift emission and better thermostability [29].

Furthermore, in plant cells, Ogura *et al.* used reporter plasmids harboring the green light-emitting luciferase from the Jamaican click beetle *Pyrophorus plagiophthalamus* (CBG; λmax = 537 nm) connected with the chlorophyll *a/b* binding protein (*Cab*) promoter and the red light-emitting luciferase from *P. plagiophthalamus* (CBR; λmax = 613 nm) connected with the cauliflower mosaic virus promoter as a control. They co-injected these plasmids into plant cells

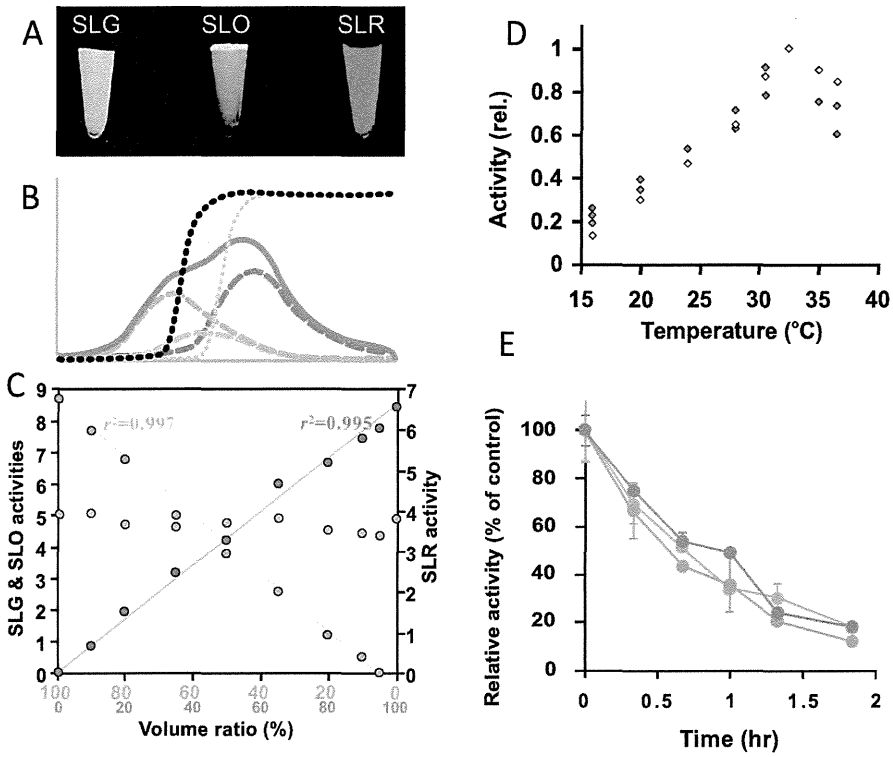

Fig. (4). Characteristic properties of SLG, SLO, and SLR. **A**; Bioluminescence color of SLG, SLO, and SLR. **B**; Bioluminescence spectra example (blue line) of mixture of SLG, SLO, and SLR and individual spectra of SLG (green dotted line), SLO (orange dotted line) and SLR (red dotted line) luciferases, and the transmission spectra of > 560 nm (O56, black dotted line) and > 600 nm (R60, gray dotted line) long-pass filters. **C**; Quantitative relationship among SLG (green circle), SLO (orange circles), and SLR (red circles) activities in a mixture of each proteins expressed in silkworm. Each samples were diluted with PicaGene Dual lysis buffer (Toyo B-net, JP) at the indicated volume ratio were mixed. The respective luciferase activities were measured with an AB2250 luminometer (ATTO, JP) after injecting PicaGene. RLU, relative light unit. **D**; Temperature dependence of SLG (green circle), SLO (orange circles), and SLR (red circles) activities of each proteins. Luminescence activities under different temperatures were measured as followings. 5 μL of each luciferases solution (0.1 μg/mL) at 4°C was mixed with 50 uL of Tripluc Luciferase Assay Reagent (TOYOBO, JP) that was pre-incubated at 16, 20, 24, and 28°C for 10 min. All the apparatus except for the luminometer at room temperature were also incubated at each temperatures in advance. **E**; Half-lives of PEST (rapid degradation sequence) sequence fused-SLG (green circle), -SLO (orange circles), and -SLR (red circles) activities. Functional half-life of PEST-fused SLG, SLO and SLR luciferases in NIH3T3 cells. Expression plasmids were independently transfected into NIH3T3 cells. Forty hours after transfection, the culture medium was replaced with DMEM supplemented with 10% FBS and 100 μM cycloheximide, and incubated for 30 min to block protein synthesis. After 30 min (time = 0), incubation was continued in the same medium. At the indicated periods, the cells were disrupted and measured their activities in Tripluc Luciferase Assay Reagent (TOYOBO, JP).

using the microprojectile bombardment system and demonstrated the simultaneous measurement of dual-color luciferase activities in extracts of spinach leaves [30]. Respective light intensities were measured simultaneously with two interference filters using a CCD camera and demonstrated the light-induced expression of the *Cab* promoter.

PRACTICAL APPLICATIONS OF CELL-BASED MULTICOLOR LUCIFERASE REPORTER ASSAYS

For the purposes of diagnostics, drug discovery and alternatives for animal models, the reliability and high-throughput features of reporter assays need to be improved. Assays must also save time and money, reduce the amount of sample needed, and facilitate the interpretation of data. The dual-color reporter assay is a simple method in which only one luciferin is used and the similar characteristic properties of luciferases are exploited; however, special equipment is required for measuring the different color emissions. The internal control reporter as a simultaneous dual-reporter assay can be optimized for inherent experimental reliability in order to improve experimental accuracy.

Indeed, to establish a practical drug screening, Davis *et al.* developed a high-throughput dual-color luciferase assay system using a 1,536-well plate format for screening small