Appendix 4
Chemical Selection

This report describes the process through which test chemicals for the EpiSensA validation study were selected.

The object of this validation study was to evaluate within- and between-laboratory reproducibility, as well as predictive capacity for skin sensitization potential (i.e., concordance with classification of sensitizers and non-sensitizers) of EpiSensA. As a complementary study, the validation management team (VMT) evaluated the predictability of the United Nations Globally Harmonized System of Classification and Labeling of Chemicals (UN GHS) with three classifications: Category 1A, Category 1B, and No category.

A pre-validation study (training) and validation studies (Phase I and Phase II testing) were conducted at three participating laboratories using the test chemicals shown in Table 1.

In addition, the chemical categories or physical state and chemical properties (e.g., solid, liquid, etc.) were included in the tables of these test chemicals in order to investigate the applicability domain.

Table 1. Breakdown of the EpiSensA validation study

Phase	Number of test substances	Number of repetitions	Examination	Date of experiment start
Training	4	2	Transferability (non-coded)	June, 2018
I-A	5	3	Within- and between-laboratory reproducibility (coded)	November, 2018
I-B	5	3	Within- and between-laboratory reproducibility (coded)	September, 2019
I-C	5	3	Within- and between-laboratory reproducibility (coded)	April, 2020
II	12	1	Between-laboratory reproducibility (coded)	November, 2020

1. Basis for chemical selection

The selection of test chemicals by the VMT was based on published papers on in vivo skin sensitization tests and validation studies for in vitro alternative assays on skin sensitization test methods.

1-1 Applied selection criteria

The applied selection criteria were as follows:

- Information on mode/site of action
- Quality and quantity of reference data (in vivo and in vitro testing)
- Availability of high-quality data derived from animal and (if available) human studies

- Coverage of a range of relevant chemical and product classes
- Information on interspecies variations (e.g., variabilities regarding assimilation of chemicals, metabolism, etc.)
- Coverage of a range of toxic effects and/or sensitizing potencies
- Information about indicating pre/pro-haptens
- Physical and chemical properties (and their suitability for experimental use as implied by their CAS No.)
- Single chemical entities or formulations known to be of high purity
- Commercial availability
- Cost

In the first phase of the selection procedure, the VMT identified and collected several existing lists of potential sensitizing chemicals in order to establish a primary database. These listings had originally been compiled by international experts as reference compounds for validation studies and other purposes. An extensive literature research was performed by the VMT to ensure that the selected chemicals fulfilled the selection criteria described above. Emphasis was placed on selecting chemicals of varying potencies (strong, weak, and no activity). In addition, it was decided that at least a third of the total substances should be non-sensitizers.

1-2 Chemical Acquisition, Coding, and Distribution

The assessment of between-laboratory transferability as well as of within- and between-laboratory reproducibility and predictive capacity was performed at all participating laboratories with coded chemicals. The VMT made provision for the need for additional testing at all participating laboratories. The coding was supervised by Japanese Center for the Validation of Alternative Methods (JaCVAM). JaCVAM was responsible for coding and distributing the test chemicals, references, and control reagents for the validation study.

1-3 Handling

Each participating laboratory was provided essential information on the test chemicals (e.g. physical state, weight or volume of sample, and storage instructions) by JaCVAM. Each laboratory was responsible for storing their respective chemicals in accordance with the storage instructions, and separately received sealed safety information, including Safety Data Sheets (SDS), describing hazard identification, exposure control, and personal protection for each chemical. The test chemicals were delivered directly to the study directors of each laboratory. The SDS were accessed only in the event of an accident, and the information was only disclosed on a need to know basis.

There were no accidents at the participating laboratories during the course of the validation study, and all of the laboratories returned the SDSs to JaCVAM in their sealed envelopes upon completion of the validation study. All unused test chemicals were disposed of in compliance with the rules and regulations of the participating laboratories upon completion of the validation study.

2. Pre-validation study

To evaluate transferability, the lead laboratory selected the four test substances shown in Table 2 for training.

Table 2. Test chemicals used for training

No.	Chemicals	CAS No.	GHS category
1	Bisphenol A diglycidyl ether	1675-54-3	1A
2	4-Nitrobenzyl bromide	100-11-8	1A
3	Clotrimazole	23593-75-1	1B
4	Cetrimide	57-09-0	No category

3. Validation study - Phase I testing

Fifteen test chemicals shown in Table 3 were selected by the VMT to evaluate within- and between-laboratory reproducibility. Three runs were performed, but the order of testing had no impact on the results. These chemicals were selected at the chemical selection meeting in Sumida, Japan, on July 5, 2018, in accordance with the applied chemical selection criteria. The chemicals were coded by JaCVAM and distributed to the test facilities.

Table 3. Test chemicals used for Phase I

Phase	No.	Chemicals	CAS No.	GHS category
I-A	1	Glyoxal 40% solution in water	107-22-2	1A
	2	Lauryl gallate	1166-52-5	1A
	3	Benzisothiazolinone	2634-33-5	1B
	4	Sodium lauryl sulfate	151-21-3	1B
	5	Diethyl phthalate	84-66-2	No category
I-B	1	DNCB	97-00-7	1A
	2	Ethyl acrylate	140-88-5	1B
	3	Hexane	110-54-3	No category
	4	Dextran	9004-54-0	No category
	5	Tween80	9005-65-6	No category
I-C	1	p-Phenylenediamine	106-50-3	1A
	2	Methyl heptine carbonate	111-12-6	1A
	3	Abietic acid	514-10-3	1B
	4	Farnesol	4602-84-0	1B
	5	Lactic acid	50-21-5	No category

4. Validation study - Phase II testing

The 12 chemicals shown in Table 4 were selected by the VMT to evaluate between-laboratory reproducibility and predictive capacity. These chemicals were selected at the chemical selection meeting in in Sumida, Japan, on July 5, 2018, in accordance with the applied chemical selection criteria. The chemicals were coded by JaCVAM and distributed to the participating laboratories.

Table 4. Test chemicals used for Phase II

No.	Chemicals	CAS No.	GHS category
1	Tetrachlorosalicylanilide	1154-59-2	1A
2	Isoeugenol	97-54-1	1A
3	2-Aminophenol	95-55-6	1A
4	Glutaraldehyde	111-30-8	1A
5	Lilial	80-54-6	1B
6	Methyl methacrylate	80-62-6	1B
7	Amyl cinnamic aldehyde	122-40-7	1B
8	Imidazolidinyl urea	39236-46-9	1B
9	Acetanisole	100-06-1	No category
10	1-Iodohexane	638-45-9	No category
11	propylene glycol	57-55-6	No category
12	Benzyl butyl phthlate	85-68-7	No category

References

OECD (2005). Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment. OECD Testing Series and Assessment Number 34. 281. ENV/JM/MONO(2005)14, pp 96, Paris, France: OECD. 282.

Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, Gribaldo L, Halder M, Hoffmann S, 283 Janusch Roi A, Prieto P, Sabbioni E, Scott L, Worth A, Zuang V (2004). A modular approach to 284 the ECVAM principles on test validity. ATLA 32, 467–472. 285.

Appendix 5
EpiSensA Dataset Including 136 Chemicals

Hazard prediction of EpiSensA compared to LLNA and human data

The predictive performance of the EpiSensA was evaluated for the ability to discriminate between sensitizers and non-sensitizers when compared to Local Lymph Node Assay (LLNA) and human data. Table 1 shows the dataset of 136 chemicals, including the LLNA, human, and EpiSensA data. The data shown in Table 2 are summarized for lipophilic chemicals (log Kow > 3.5, n = 69), hydrophilic chemicals (log Kow ≤ 3.5, n = 67), pre/pro-haptens (n = 37), and overall chemicals (n = 136). The EpiSensA was applicable to all 136 tested chemicals, demonstrating that the EpiSensA can be used to test a diverse range of chemicals. In addition, the EpiSensA had a sensitivity of 82.7%, an accuracy of 78.3% and a balanced accuracy of 73.7% for lipophilic chemicals. Regarding pre/pro-haptens, the EpiSensA demonstrated 97.3% sensitivity, and all pre/pro-haptens except for benzo[a]pyrene could be detected as positive. For all 136 chemicals, the EpiSensA had a high sensitivity (88.1%), accuracy (82.4%) and balanced accuracy (76.9%).

The predictive performance of the LLNA and EpiSensA for human data is shown in Table 3. For lipophilic chemicals, the EpiSensA predicted the human hazard with a sensitivity of 92.3%, and the performance was comparable to LLNA. In contrast, the EpiSensA and LLNA showed low specificities of 16.7% and 0%, respectively. In other words, all of the false negative chemicals against the human data were rated as positive in the LLNA. Regarding hydrophilic chemicals, the LLNA and EpiSensA showed 94.5% and 87.3% accuracy, respectively. For pre/pro-haptens, the LLNA and EpiSensA showed sensitivity of 100%. For all 80 chemicals, the EpiSensA showed similar predictive performance to the LLNA (accuracy of 77.5% and 81.3%, balanced accuracy of 73.2% and 77.3%, respectively).

Potency prediction of EpiSensA compared to LLNA and human data

For positive chemicals in the EpiSensA, the minimum estimated concentration (Min EC value) at which any of four marker genes exceeds the respective cut-off is used for potency classification. A test chemical is classified as strong or weak potency if the Min EC value is at ≤0.098% w/v or >0.098% w/v, respectively. If negative at the concentrations with over 80% cell viability, a test chemical is considered as a non-sensitizer.

Table 4 summarize the predictive performance of EpiSensA for GHS sub-categorization (Cat.1A, Cat.1B, NC) based on LLNA results of 136 chemicals. The EpiSensA had a potency accuracy of 71.3%, and the potency prediction of EpiSensA correlated positively with the GHS sub-categorization of LLNA (ρ=0.656 at Spearman rank-order correlation coefficient, κ=0.588 at Weighted Kappa). In addition, Table 5 and Table 6 summarize the predictive performance of EpiSensA and LLNA for human results based on GHS sub-categorization described in Annex 2 of OECD Guideline No.497 (OECD, 2021) or Basketter's classification (Basketter et al., 2014). The potency accuracy of EpiSensA for human results was 64.4% and similar to that of LLNA (64.4%). Furthermore, those potency prediction results correlated positively with the human results (ρ=0.659 at Spearman rank-order correlation coefficient, κ=0.532 at Weighted Kappa),

and the coefficients of correlation were comparable to those of LLNA for human results (ρ=0.677 at Spearman rank-order correlation coefficient, κ=0.541 at Weighted Kappa).

Tables 7 and 8 summarize the predictive performance of EpiSensA and kDPRA for GHS binary sub-categorization (GHS Cat.1A or not GHS Cat.1A) based on LLNA results of 72 chemicals which are available for all of EpiSensA, kDPRA, and Annex 2 of the OECD Guideline No.497 (OECD, 2021). kDPRA allows to distinguish GHS Cat.1A skin sensitizers from those not categorized as GHS Cat.1A (Cat.1B or NC). The sensitivity (the proportion of GHS Cat.1A chemicals that are correctly classified) of EpiSensA for LLNA results was 73.7%, and the sensitivity of kDPRA (84.2%) was slightly higher than that of EpiSensA. However, the difference is only two chemicals. In addition, Tables 9, 10, and 11 summarize the predictive performance for binary sub-categorization vs. human results based on Annex 2 of OECD Guideline No.497 or Basketter's classification (Basketter et al., 2014). The sensitivity of EpiSensA for human results was 66.7%, which is slightly higher than kDPRA (58.3%) and slightly lower than LLNA (75.0%). However, the difference is only one chemical. Therefore, the predictive performance of EpiSensA was comparable to that of kDPRA on the same set of chemicals.

Table 1. EpiSensA dataset including 136 chemicals

No.	Chemical name	CAS No.	LogKow	Pre/pro[ij]	LLNA		Human			EpiSensA		
					GHS category	Ref.	Hazard	Potency category	Ref.	Vehicle	Hazard prediction	Potency prediction
1	Benzo[a]pyrene	50-32-8	**5.99**	X	1A	a				AOO	N	NS
2	Oxazolone	15646-46-5	1.51		1A	b	P		d	AOO	P	Strong
3	Chlorothalonil	1897-45-6	**3.66**		1A	b				AOO	P	Strong
4	3-Methylcatechol	488-17-5	1.58	X	1A	a				AOO	P	Strong
5	Bandrowski's Base	20048-27-5	0.74	X	1A	b				50% EtOH	P	Strong
6	Tetrachlorosalicylanilide	1154-59-2	**5.87**		1A	b	P	1A	b	AOO	P	Strong
7	4-Nitrobenzylbromide	100-11-8	2.7		1A	b				AOO	P	Strong
8	Dicyclohexyl carbodiimide	538-75-0	**6.83**		1A	a				AOO	P	Weak
9	Benzoyl peroxide	94-36-0	3.43		1A	b	P	1B	b	AOO	P	Weak
10	2,4-Dinitrochlorobenzene	97-00-7	2.27		1A	b	P	1A	b	AOO	P	Strong
11	1,4-Dihydroquinone	123-31-9	1.03	X	1A	b	P	Cat.3	e	50% EtOH	P	Strong
12	Chlorpromazine hydrochloride	69-09-0	**3.69**	X	1A	a				DW	P	Strong
13	Fluorescein isothiocyanate	3326-32-7	**4.69**		1A	a				AOO	N	NS
14	p-Phenylenediamine	106-50-3	-0.39	X	1A	b	P	1A	b	AOO	P	Strong
15	Hexyl salicylate	6259-76-3	**5.06**		1A	b	N	NC	b	AOO	P	Weak
16	Lauryl gallate	1166-52-5	**6.21**	X	1A	b	P	Cat.2	e	AOO	P	Weak
17	Propyl gallate	121-79-9	1.79	X	1A	b	P	Cat.2	e	50% EtOH	P	Strong
18	p-tert-Butylphenyl 1-(2,3-epoxy)propylether	3101-60-8	**3.52**		1A	a				AOO	P	Strong
19	2,5-Diaminotoluene sulfate	615-50-9	0.16	X	1A	b	P	Cat.2	e	DW	P	Strong
20	Chloroatranol	57074-21-2	**3.50**		1A	a	P	Cat.1	e	AOO	P	Strong
21	2-Aminophenol	95-55-6	0.60	X	1A	b	P	Cat.2	e	AOO	P	Strong
22	Cobalt chloride	7646-79-9	0.85		1A	a	P		d	DW	P	Weak
23	CD3	25646-71-3	-3.09	X	1A	b				DW	P	Weak
24	Glyoxal	107-22-2	-1.66		1A	b	P	1A	b	DW	P	Strong
25	Metol	55-55-0	2.34	X	1A	b	P	Cat.3	e	50% EtOH	P	Strong
26	Methyldibromoglutaronitrile	35691-65-7	1.63	X	1A	b	P	Cat.2	e	AOO	P	Strong
27	Dinocap	39300-45-3	**6.49**		1A	a				AOO	P	Weak
28	Cinnamic aldehyde	104-55-2	1.82		1A	b	P	1A	b	AOO	P	Strong
29	1-Naphthol	90-15-3	2.69	X	1A	b				AOO	P	Weak
30	Isoeugenol	97-54-1	2.65	X	1A	b	P	1B	b	AOO	P	Strong
31	4-Amino-m-cresol	2835-99-6	0.79	X	1A	b				AOO	P	Strong
32	Bisphenol A-diglycidyl ether	1675-54-3	**3.84**		1A	b	P	Cat.3	e	AOO	P	Weak
33	3-Dimethylamino propylamine	109-55-7	-0.45	X	1B	b	P	Cat.2	e	AOO	P	Weak
34	trans-2-decenal	3913-71-1	**3.55**		1B	a				AOO	P	Weak
35	2-Mercaptobenzothiazole	149-30-4	2.86		1A	b	P	1B	b	AOO	P	Weak
36	Benzyl salicylate	118-58-1	**4.31**		1B	b	N	Cat.5	e	AOO	P	Weak
37	Tetramethylthiuram disulfide	137-26-8	1.7		1B	b	P	1B	b	AOO	P	Strong
38	3-Aminophenol	591-27-5	0.24	X	1B	b				AOO	P	Weak
39	Diethylenetriamine	111-40-0	-2.13	X	1B	b	P	1A	b	DW	P	Weak
40	Ethylene diamine	107-15-3	-1.62	X	1B	b	P	Cat.3	e	DW	P	Weak
41	3-Propylidenephthalide	17369-59-4	2.03	X	1B	b	P	1B	b	AOO	P	Weak
42	Sodium lauryl sulfate	151-21-3	1.69		1B	b	N	NC	b	DW	P	Weak
43	2-Nitro-1,4-phenylenediamine	5307-14-2	0.55	X	1A	b	P	Cat.2	e	AOO	P	Strong
44	Bourgenal	18127-01-0	**3.94**		1B	b	P	Cat.4	f	AOO	P	Weak
45	5-Methyl-2-phenyl-2-hexenal	21834-92-4	**3.77**		1B	a				AOO	P	Weak
46	Farnesol	4602-84-0	**5.77**	X	1B	b	P	1B	b	AOO	P	Weak
47	Clotrimazole	23593-75-1	**6.26**		1B	a				AOO	P	Weak
48	5-Chlorosalicylanilide	4638-48-6	**3.94**		1B	a				AOO	P	Strong
49	alpha-Phellandren	99-83-2	**4.62**		1B	a				AOO	P	Weak
50	Embramine hydrochloride	13977-28-1	**4.45**		1B	a				AOO	P	Weak
51	2-Methoxy-4-methylphenol	93-51-6	1.88	X	1B	b	P	Cat.2	f	AOO	P	Weak
52	Squaric acid	2892-51-5	-0.44		1B	b				50% EtOH	N	NS
53	5-Amino-2-methylphenol	2835-95-2	0.79	X	1B	b				AOO	P	Weak
54	Resorcinol	108-46-3	1.03	X	1B	b	P	Cat.4	e	AOO	P	Weak
55	4-Chloroaniline	106-47-8	1.72	X	1B	a				AOO	P	Weak
56	Damascone	23726-91-2	**4.42**		1B	a				AOO	P	Strong
57	Undec-10-enal	112-45-8	**4.12**		1B	b				AOO	P	Weak
58	Dihydroeugenol	2785-87-7	2.87	X	1B	b				AOO	P	Weak
59	12-Bromo-1-dodecanone	3344-77-2	**5.11**		1B	a				AOO	P	Weak
60	Tocopherol	10191-41-0	**12.2**		1B	a	N	Cat.6	e	AOO	N	NS
61	Squalene	111-02-4	**14.12**		1B	a				AOO	N	NS
62	Ethylhexyl acrylate	103-11-7	**4.09**		1B	b				AOO	P	Weak
63	1-Bromohexane	111-25-1	**3.63**		1B	b				AOO	P	Weak
64	2-Methylundecanal	110-41-8	**4.67**		1B	b				AOO	P	Weak
65	Hexyl cinnamic aldehyde	101-86-0	**4.82**		1B	b	N	Cat.5	e	AOO	P	Weak
66	Abietic acid	514-10-3	**6.46**	X	1B	b	P	Cat.3	e	AOO	P	Weak
67	Trifluralin	1582-09-8	**5.31**		1B	a				AOO	P	Weak
68	Lillial	80-54-6	**4.36**		1B	b	P	1B	b	AOO	P	Weak
69	Salicylic acid	69-72-7	2.24		1B	b	N	Cat.6	e	AOO	N	NS
70	Citral	5392-40-5	3.45		1B	b	P	Cat.3	e	AOO	P	Strong

Table 1. *(continued)*

No.	Chemical name	CAS No.	LogKow	Pre/pro[ij]	LLNA		Human			EpiSensA		
					GHS category	Ref.	Hazard	Potency category	Ref.	Vehicle	Hazard prediction	Potency prediction
71	Eugenol	97-53-0	2.73	X	1B	b	P	1B	b	AOO	P	Weak
72	Benzyl benzoate	120-51-4	**3.54**		1B	b	N	NC	b	AOO	P	Weak
73	Phenyl benzoate	93-99-2	3.04		1B	b	P	1B	b	AOO	P	Weak
74	Benzyl cinnamate	103-41-3	**4.06**		1B	b	P	Cat.4	f	AOO	N	NS
75	Oxyfluorfen	42874-03-3	**5.21**		1B	a				AOO	N	NS
76	4-Allylanisole	140-67-0	3.47	X	1B	b				AOO	P	Weak
77	3-Chloro-4-(3-fluorobenzyloxy) nitrobenzene	443882-99-3	**4.44**		1B	a				AOO	N	NS
78	Dibutyl aniline	613-29-6	**5.12**	X	1B	b				AOO	P	Weak
79	Pentachlorophenol	87-86-5	**4.74**		1B	b	N	Cat.5	e	AOO	P	Strong
80	Cinnamic alcohol	104-54-1	1.84	X	1B	b	P	1B	b	AOO	P	Weak
81	alpha-Cetone	127-51-5	**4.84**		1B	b	N	NC	b	AOO	P	Weak
82	Undecylenic acid	112-38-9	**4.37**		1B	a	P		d	AOO	P	Weak
83	Cyclamen aldehyde	103-95-7	**3.91**		1B	b	P		g	AOO	P	Weak
84	Geraniol	106-24-1	3.47	X	1B	b	P	1B	b	AOO	P	Weak
85	Imidazolidinyl urea	39236-46-9	-8.28		1B	b	P	1B	b	DW	P	Weak
86	Iso E super	54464-57-2	**5.18**		1B	b	N	NC	b	AOO	P	Weak
87	Penicillin G	61-33-6	1.85		1B	b	P	1B	b	DW	P	Weak
88	1-Octen-3-yl-acetate	2442-10-6	**3.6**		1B	a	P	Cat.4	f	AOO	P	Weak
89	Linalool	78-70-6	3.38	X	1B	b	P	Cat.4	e	AOO	P	Weak
90	Butyl glycidyl ether	2426-08-6	1.08		1B	b	P	1A	b	AOO	P	Weak
91	Ethyleneglycol dimethacrylate	97-90-5	2.21		1B	b	P	Cat.4	e	AOO	P	Weak
92	2-Phenylethyl isovalerate	140-26-1	**3.97**		1B	a				AOO	P	Weak
93	1,1,3-Trimethyl-3-phenylindane	3910-35-8	**5.91**		1B	a	N	Cat.5	f	AOO	P	Weak
94	Citronellol	106-22-9	**3.56**		1B	b	N	NC	b	AOO	P	Weak
95	Isopropyl myristate	110-27-0	**7.17**		1B	b	N	Cat.5	e	AOO	N	NS
96	Bis-GMA	1565-94-2	**4.94**		1B	b				AOO	P	Weak
97	Tridecane	629-50-5	**6.73**		1B	a				AOO	N	NS
98	Limonene	5989-27-5	**4.83**	X	1B	b	N	Cat.5	e	AOO	P	Weak
99	Aniline	62-53-3	1.08	X	1B	b	P	1B	b	AOO	P	Weak
100	1-Iodohexane	638-45-9	**4.05**		NC	b				AOO	P	Weak
101	Xylene	1330-20-7	3.09		1B	c	N	Cat.6	e	AOO	N	NS
102	Acetanisole	100-06-1	1.75		NC	b	N		d	AOO	N	NS
103	α-Amylcinnamyl alcohol	101-85-9	**4.35**		1B	b	P	Cat.4	e	AOO	P	Weak
104	Benzyl butyl phthalate	85-68-7	**4.84**		NC	b				AOO	N	NS
105	1-Bromobutane	109-65-9	2.65		NC	a	N		h	AOO	N	NS
106	1-Butanol	71-36-3	0.84		NC	b	N	Cat.6	e	AOO	N	NS
107	Carbonic acid dioctyl ester	1680-31-5	**7.11**		NC	a				AOO	N	NS
108	Chlorpyrifos	2921-88-2	**4.66**		NC	a				AOO	N	NS
109	Clofibrate	637-07-0	**3.62**		NC	b				AOO	P	Weak
110	Decamethylcyclopentasiloxane	541-02-6	**7.93**		NC	a				AOO	N	NS
111	Dibutyl phthalate	84-74-2	**4.61**		NC	b				AOO	N	NS
112	Diethyl phthalate	84-66-2	2.65		NC	b	N	Cat.6	e	AOO	P	Weak
113	Diethyl toluamide	134-62-3	2.26		NC	b	N	Cat.6	e	AOO	P	Weak
114	Dioctyl ether	629-82-3	**6.94**		NC	a				AOO	N	NS
115	Equol	531-95-3	**3.67**		NC	a				AOO	P	Weak
116	Erucamide	112-84-5	**8.44**		NC	a				AOO	N	NS
117	Glucose	50-99-7	-2.43		NC	a	N	Cat.6	e	DW	N	NS
118	Glycerol	56-81-5	-1.65		NC	b	N	Cat.6	e	DW	N	NS
119	Cetrimide	57-09-0	3.18		NC	a	N	Cat.5	e	50% EtOH	N	NS
120	Hexane	110-54-3	3.29		NC	b	N	NC	b	AOO	N	NS
121	Hydrocortisone	50-23-7	1.62		NC	c	N	NC	b	AOO	N	NS
122	4-Hydroxybenzoic acid	99-96-7	1.39		NC	b	N		d	AOO	P	Weak
123	1-Iodooctadecane	629-93-6	**9.94**		NC	a				AOO	N	NS
124	Isopropanol	67-63-0	0.28		NC	b	N	Cat.5	e	AOO	N	NS
125	Lactic acid	50-21-5	-0.65		NC	b	N	Cat.6	e	DW	N	NS
126	(+)-trans-p-Menth-2-ene	5113-93-9	**4.7**		NC	a				AOO	P	Weak
127	Methyl salicylate	119-36-8	2.6		NC	c	N	Cat.5	e	AOO	P	Weak
128	Octanoic acid	124-07-2	3.03		NC	b	N	Cat.6	e	AOO	P	Weak
129	Propylene glycol	57-55-6	-0.78		NC	b	N	NC	b	DW	N	NS
130	Quinoxyfen	124495-18-7	**5.69**		NC	a				AOO	N	NS
131	Retapamulin	224452-66-8	**4.95**		NC	a				AOO	P	Weak
132	Rifamycin SV sodium salt	14897-39-3	**5.04**		NC	a				AOO	N	NS
133	Sulfanilamide	63-74-1	-0.55		NC	b	N	1B	b	50% EtOH	P	Weak
134	4'-Trifluoromethylbiphenyl-4-carbaldehyde	90035-34-0	**4.44**		NC	a				AOO	P	Weak
135	Vanillin	121-33-5	1.05		NC	b	N	Cat.5	e	AOO	N	NS
136	Zinc mercaptobenzothiazole	155-04-4	**5.02**		NC	a				AOO	N	NS

a: Median-like location parameter (Hoffmann et al., 2018) calculated based on Urbisch et al., 2015, Jaworska et al., 2015, and NICEATM LLNA database, 2013.
b: OECD Guideline 497
c: Hoffmann et al., 2018.
d: Urbisch et al., 2015.
e: Basketter et al., 2014.
f: Api et al., 2017.
g: Basketter et al., 2005.
h: Bauch et al. 2012.
i: Urbisch et al., 2016
j: Casati et al., JRC Technical reports, 2016.

Table 2. Hazard predictive performance of EpiSensA for the 136 chemicals tested by the LLNA

I. Lipophilic chemicals (N=69)		III. Pre/pro-haptens (N=37)	
No. of chemicals	69	No. of chemicals	37
Sensitivity (%)	82.7	Sensitivity (%)	97.3
Specificity (%)	64.7	Specificity (%)	-
Accuracy (%)	78.3	Accuracy (%)	-
Balanced accuracy (%)	73.7	Balanced accuracy (%)	-
II. Hydrophilic chemicals (N=67)		IV. Overall (N=136)	
No. of chemicals	67	No. of chemicals	136
Sensitivity (%)	93.9	Sensitivity (%)	88.1
Specificity (%)	66.7	Specificity (%)	65.7
Accuracy (%)	86.6	Accuracy (%)	82.4
Balanced accuracy (%)	80.3	Balanced accuracy (%)	76.9

Table 3. Hazard predictive performance of the LLNA and EpiSensA for the 80 chemicals which human test was performed.

	LLNA	EpiSensA		LLNA	EpiSensA
I. Lipophilic chemicals (N=25)			III. Pre/pro-haptens (N=23)		
No. of chemicals	25	25	No. of chemicals	23	23
Sensitivity (%)	100.0	92.3	Sensitivity (%)	100.0	100.0
Specificity (%)	0.0	16.7	Specificity (%)	-	-
Accuracy (%)	52.0	56.0	Accuracy (%)	-	-
Balanced accuracy (%)	50.0	54.5	Balanced accuracy (%)	-	-
II. Hydrophilic chemicals (N=55)			IV. Overall (N=80)		
No. of chemicals	55	55	No. of chemicals	80	80
Sensitivity (%)	100.0	100.0	Sensitivity (%)	100.0	97.9
Specificity (%)	85.7	66.7	Specificity (%)	54.5	48.5
Accuracy (%)	94.5	87.3	Accuracy (%)	81.3	77.5
Balanced accuracy (%)	92.9	83.3	Balanced accuracy (%)	77.3	73.2

Table 4. Potency predictive performance of EpiSensA for the 136 chemicals tested by the LLNA

		EpiSensA		
		Strong	Weak	NS
LLNA	1A	22	10	2
	1B	5	52	10
	No category	0	12	23

Overall accuracy 71.3%
Overall overprediction 12.5%
Overall underprediction 16.1%

Performance (N=136)	1A (N=34)	1B (N=67)	NC (N=35)
Correct classification (%)	68%	78%	63%
Underpredicted (%)	29% (1B); 6% (NC)	15% (NC)	NA
Overpredicted (%)	NA	7% (1A)	0% (1A); 34% (1B)

Table 5. Potency predictive performance of EpiSensA for the 73 chemicals which human test was performed.

		EpiSensA		
		Strong	Weak	NS
Human	GHS 1A or Human Cat.1, 2	11	5	0
	GHS 1B or Human Cat.3, 4	5	22	1
	GHS No category or Human Cat.5, 6	1	14	14

Overall accuracy 64.4%
Overall overprediction 27.4%
Overall underprediction 8.2%

Performance (N=73)	1A (N=16)	1B (N=28)	NC (N=29)
Correct classification (%)	69%	79%	48%
Underpredicted (%)	31% (1B); 0% (NC)	4% (NC)	NA
Overpredicted (%)	NA	18% (1A)	3% (1A); 48% (1B)

Table 6. Potency predictive performance of the LLNA for the 73 chemicals which human test was performed.

		LLNA		
		Strong	Weak	NS
Human	GHS 1A or Human Cat.1, 2	12	4	0
	GHS 1B or Human Cat.3, 4	6	21	1
	GHS No category or Human Cat.5, 6	1	14	14

Overall accuracy 64.4%

Overall overprediction 28.8%

Overall underprediction 6.8%

Performance (N=73)	1A (N=16)	1B (N=28)	NC (N=29)
Correct classification (%)	75%	75%	48%
Underpredicted (%)	25% (1B); 0% (NC)	4% (NC)	NA
Overpredicted (%)	NA	21% (1A)	3% (1A); 48% (1B)

Table 7. Potency predictive performance of EpiSensA for the 72 LLNA data which are available for both EpiSensA and kDPRA

		EpiSensA	
		Strong	Not strong
LLNA	GHS 1A	14	5
	Not GHS 1A	3	50

Sensitivity 73.7%

Specificity 94.3%

Accuracy 88.9%

Balanced accuracy 84.0%

Table 8. Potency predictive performance of kDPRA for the 72 LLNA data which are available for both EpiSensA and kDPRA

		kDPRA[a]	
		GHS 1A	Not GHS 1A
LLNA	GHS 1A	16	3
	Not GHS 1A	3	50

Sensitivity 84.2%

Specificity 94.3%

Accuracy 91.7%

Balanced accuracy 89.3%

a: Natsch et al., 2020

Table 9. Potency predictive performance of EpiSensA for the 57 Human data which are available for both EpiSensA and kDPRA

		EpiSensA	
		Strong	Not strong
Human	GHS 1A or Human Cat.1, 2	8	4
	Not GHS 1A or Human Cat.1, 2	6	39

Sensitivity 66.7%
Specificity 86.7%
Accuracy 82.5%
Balanced accuracy 76.7%

Table 10. Potency predictive performance of kDPRA for the 57 Human data which are available for both EpiSensA and kDPRA

		kDPRA[a]	
		GHS 1A	Not GHS 1A
Human	GHS 1A or Human Cat.1, 2	7	5
	Not GHS 1A or Human Cat.1, 2	8	37

Sensitivity 58.3%
Specificity 82.2%
Accuracy 77.2%
Balanced accuracy 70.3%

a: Natsch et al., 2020

Table 11. Potency predictive performance of LLNA for the 57 Human data which are available for both EpiSensA and kDPRA

		LLNA	
		GHS 1A	Not GHS 1A
Human	GHS 1A or Human Cat.1, 2	9	3
	Not GHS 1A or Human Cat.1, 2	6	39

Sensitivity 75.0%
Specificity 86.7%
Accuracy 84.2%
Balanced accuracy 80.9%

References

Api, A. M., Parakhia, R., O'Brien, D., Basketter, D. A. (2017). Fragrances categorized according to relative human skin sensitization potency. *Dermatitis, 28*(5), 299-307. doi: 10.1097/DER.0000000000000304

Basketter, D. A., Clapp, C., Jefferies, D., Safford, B., Ryan, C. A., Gerberick, F., Dearman, R.J., Kimber, I. (2005). Predictive identification of human skin sensitization thresholds. *Contact Dermatitis, 53*(5), 260-267. doi: 10.1111/j.0105-1873.2005.00707.x

Basketter, D. A., Alépée, 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*(1), 11-21. doi: 10.1097/DER.0000000000000003

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. *Regulatory Toxicology and Pharmacology, 63*(3), 489-504. doi: 10.1016/j.yrtph.2012.05.013

Casati, S., Aschberger, K., Asturiol, D., Basketter, D., Dimitrov, S., Dumont, C., Karlberg, A.T., Lepoittevin, P., Patlewicz, G., Roberts, D.W., Worth, A. (2016). Ability of non-animal methods for skin sensitisation to detect pre- and pro-haptens. Report and Recommendations of an EURL ECVAM Expert Meeting. Joint Research Centre Technical Reports, 1-30, doi:10.2788/01803

Hoffmann S, Kleinstreuer N, Alépée N, Allen D, Api AM, Ashikaga T, Clouet E, Cluzel M, Desprez B, Gellatly N, Goebel C, Kern PS, Klaric M, Kühnl J, Lalko JF, Martinozzi-Teissier S, Mewes K, Miyazawa M, Parakhia R, van Vliet E, Zang Q, Petersohn D. (2018). Non-animal methods to predict skin sensitization (I): the Cosmetics Europe database. *Critical Reviews in Toxicology, 48*(5), 344-358, doi: 10.1080/10408444.2018.1429385

Jaworska, J. S., Natsch, A., Ryan, C., Strickland, J., Ashikaga, T., Miyazawa, M. (2015). Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: a decision support system for quantitative weight of evidence and adaptive testing strategy. *Archives of Toxicology, 89*(12), 2355-2383. doi: 10.1007/s00204-015-1634-2.

Natsch, A., Haupt, T., Wareing, B., Landsiedel, R., Kolle, S. N., (2020). Predictivity of the kinetic direct peptide reactivity assay (kDPRA) for sensitizer potency assessment and GHS subclassification. *ALTEX, 37(4)*, 652-664. doi: 10.14573/altex.2004292.

NICEATM LLNA Database (2013).
https://ntp.niehs.nih.gov/pubhealth/evalatm/resources-for-test-method-developers/refchem/index.html

OECD. (2021). Guideline No. 497, Guideline on Defined Approaches for Skin

Sensitisation, adopted 14th June.

Urbisch, D., Mehling, A., Guth, K., Ramirez, T., Honarvar, N., Kolle, S., Landsiedel, R., Jaworska, J., Kern, P.S., Gerberick, F., Natsch, A., Emter, R., Ashikaga, T., Miyazawa, M., Sakaguchi, H. (2015). Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regulatory Toxicology and Pharmacology, 71*(2), 337-351. doi: 10.1016/j.yrtph.2014.12.008

Urbisch, D., Becker, M., Honarvar, N., Kolle, S.N., Mehling, A., Teubner, W., Wareing, B., Landsiedel, R. (2016). Assessment of pre- and pro-haptens using nonanimal test methods for skin sensitization. *Chemical Research in Toxicology, 29*(5), 901-913. doi: 10.1021/acs.chemrestox.6b00055

Appendix 6
Comparability of Equipment

1. Background and Objectives

In the EpiSensA, the marker gene expression is measured by quantitative PCR using Real-Time PCR systems. It is possible that different results may be obtained by using different Real-Time PCR systems. As this will affect between-laboratory reproducibility and given that the participating laboratories used different Real-Time PCR systems during this validation study, it is necessary to confirm that comparable results can be obtained by using different Real-Time PCR systems. To investigate this, the comparability of Real-Time PCR systems was evaluated as follows.

2. Method and Results

2-1. Study 1 - conducted by the lead laboratory

The technique used for real time PCR was evaluated at each laboratory using the same cDNA prior to the initiation of the validation study.

2-1-1. Test method

1. Two types of cDNA that had been synthesized by the lead laboratory (one was synthesized from RhEs exposed to 6.25% w/v bisphenol A diglycidyl ether (BADGE), and the other was from RhEs exposed to 3.13% w/v clotrimazole) was sent to the participating laboratories.
2. The marker gene expression was measured by each laboratory. The Real-Time PCR systems used in each laboratory are shown below in Table 1.
3. The results were compared.

Table 1. Real-Time PCR systems used in each laboratory.

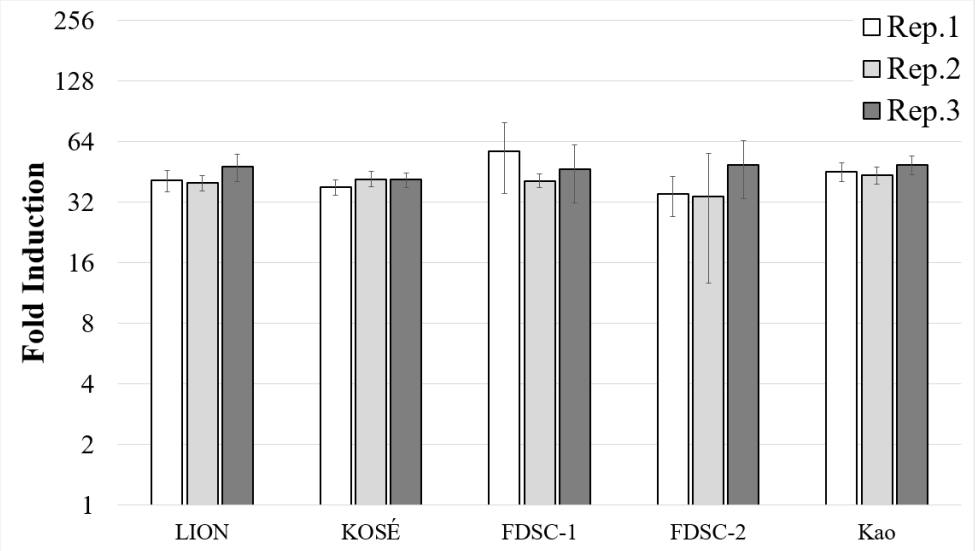
Laboratory	Real-Time PCR system	Maker
LION	CFX96 Real-Time PCR Detection System	Bio-Rad Laboratories
KOSÉ	CFX Connect Real-Time PCR Detection System	Bio-Rad Laboratories
FDSC	ABI PRISM 7900HT machine	Thermo Fisher Scientific
Kao	QuantStudio 3/5 Real-Time PCR System	Thermo Fisher Scientific

2-1-2. Result

Figure 1 shows the mean fold induction of 6.25% w/v BADGE obtained by each Real-Time PCR system, and Figure 2 shows the mean fold induction of 3.13% w/v clotrimazole obtained by each Real-Time PCR system. Two results are shown for the “ABI PRISM 7900HT” machine in Figures 1 and 2 because two operators participated from Hatano Research Institute, Food and Drug Safety Center (FDSC). All results were obtained from the same cDNA, and both cDNAs were measured three times.

Figure 1. Fold induction of 6.25% w/v BADGE obtained by each Real-Time PCR system

ATF3



GCLM

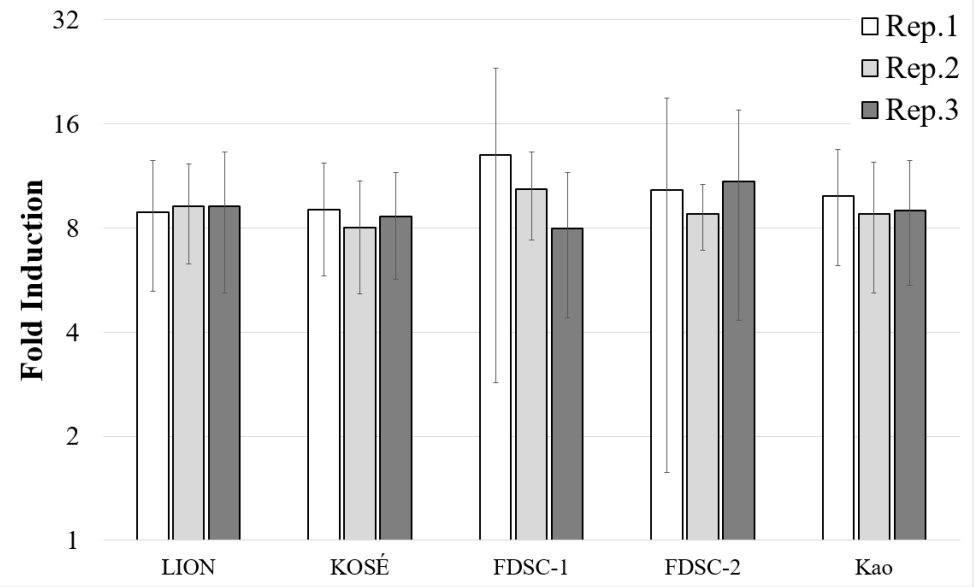
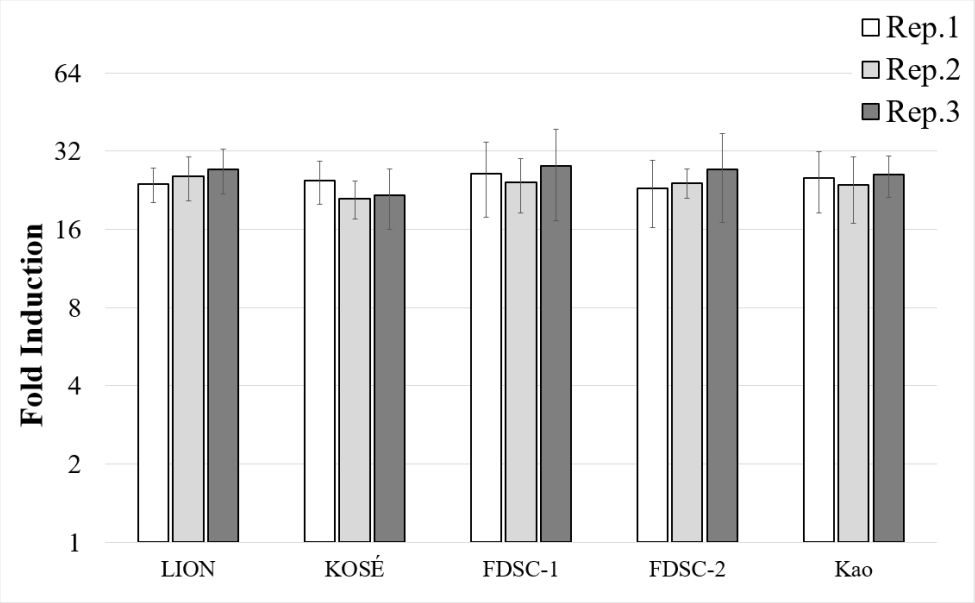


Figure 1. *(continued)*

DNAJB4



IL-8

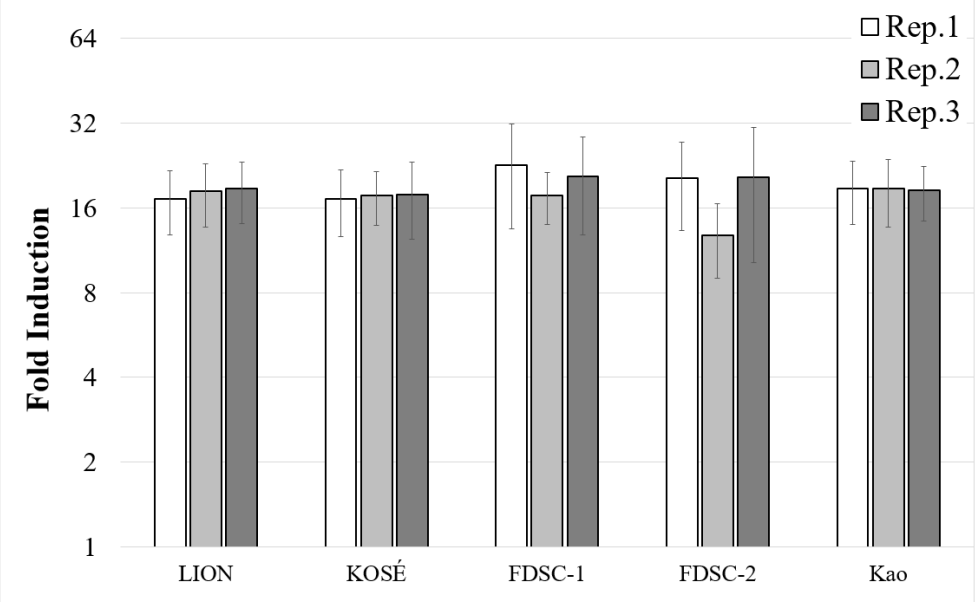
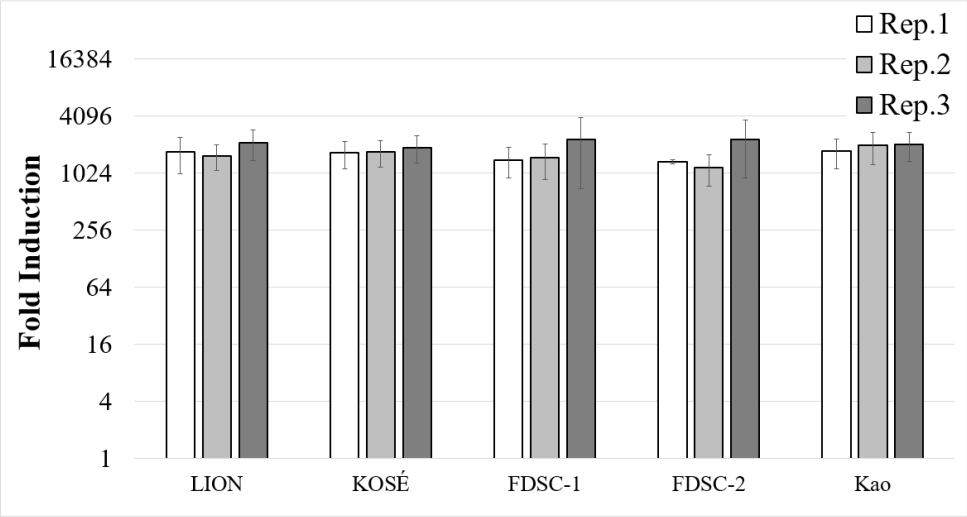


Figure 2. Fold induction of 3.13% Clotrimazole obtained by each Real-Time PCR system

ATF3



GCLM

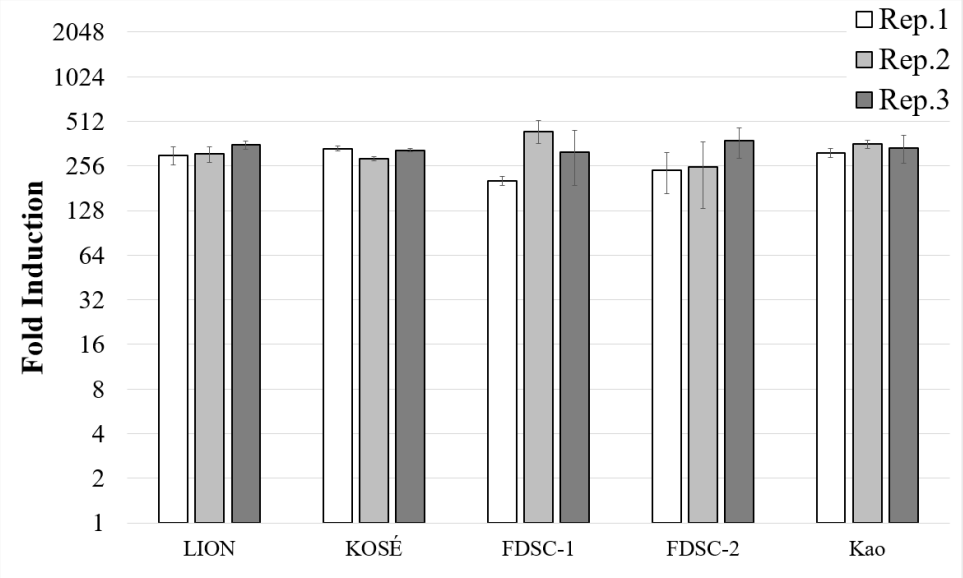
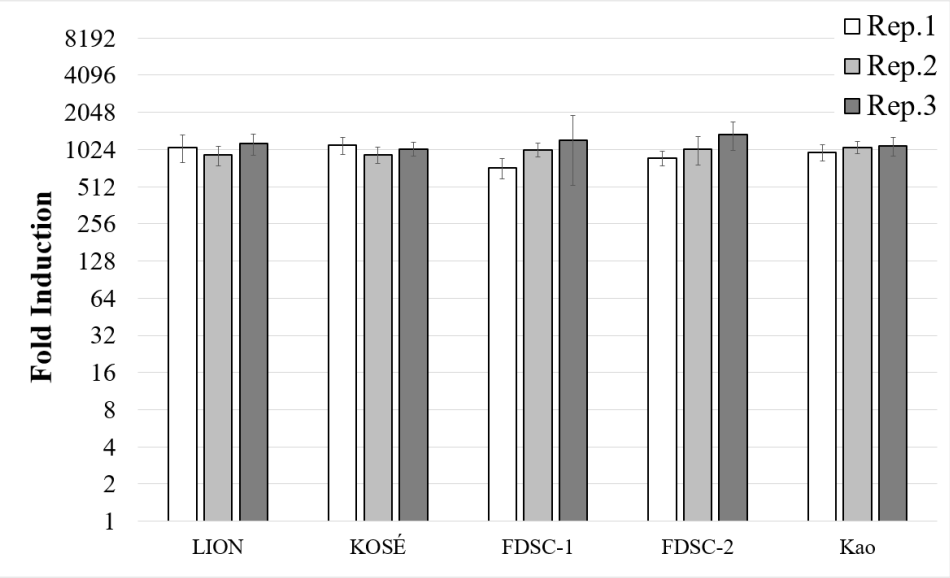
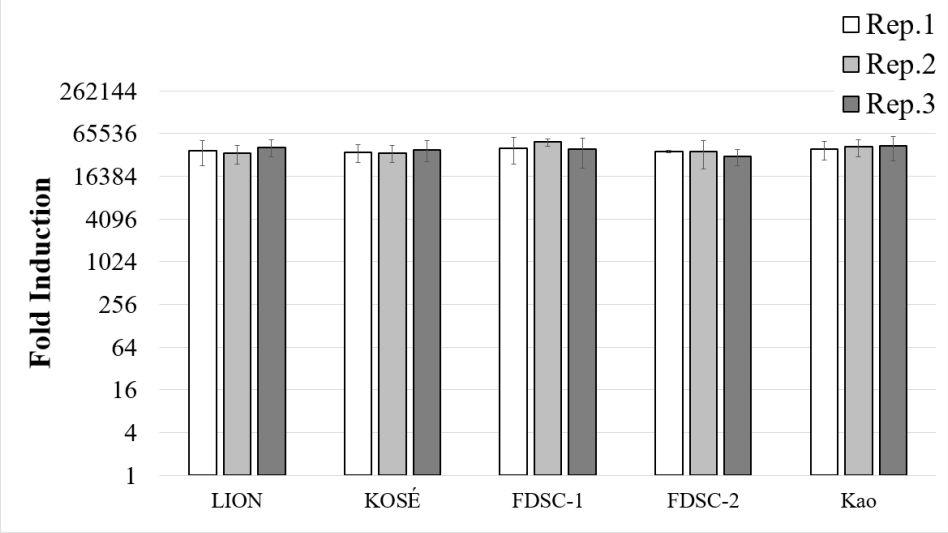


Figure 2. *(continued)*

DNAJB4



IL-8



2-2. Study 2 - conducted by FDSC

The FDSC changed Real-Time PCR systems after this validation study. When it was changed, the FDSC evaluated the comparability of Real-Time PCR systems.

2-2-1. Test method

1. cDNA was prepared from RhE models exposed to 0.78% w/v clotrimazole and 0.10% w/v 4-nitrobenzyl bromide (4NBB).
2. The marker gene expression was measured using the same cDNA on different Real-Time PCR systems. The Real-Time PCR systems and test condition used in this test are shown below in Table 2.
3. The results were compared.

Table 2. Real-Time PCR systems and test condition in study-2

Real-Time PCR system	Maker
LightCycler 480 System II	Roche Diagnostics
7900 HT Fast Real Time PCR System	Thermo Fisher Scientific

2-2-2. Result

Figure 3 shows the mean fold induction of 0.78% w/v clotrimazole obtained by each Real-Time PCR system, and Figure 4 shows the mean fold induction of 0.10% w/v 4NBB obtained by each Real-Time PCR system. All results were obtained by using the same cDNA, and both cDNAs were measured three times.

Figure 3. Fold induction of 0.78% w/v clotrimazole obtained by each Real-Time PCR system

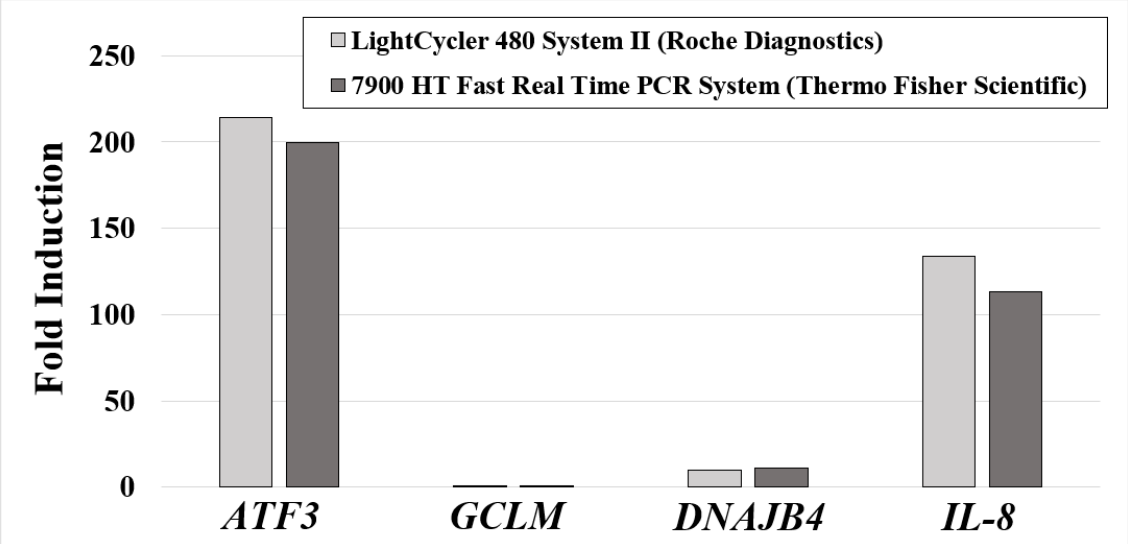
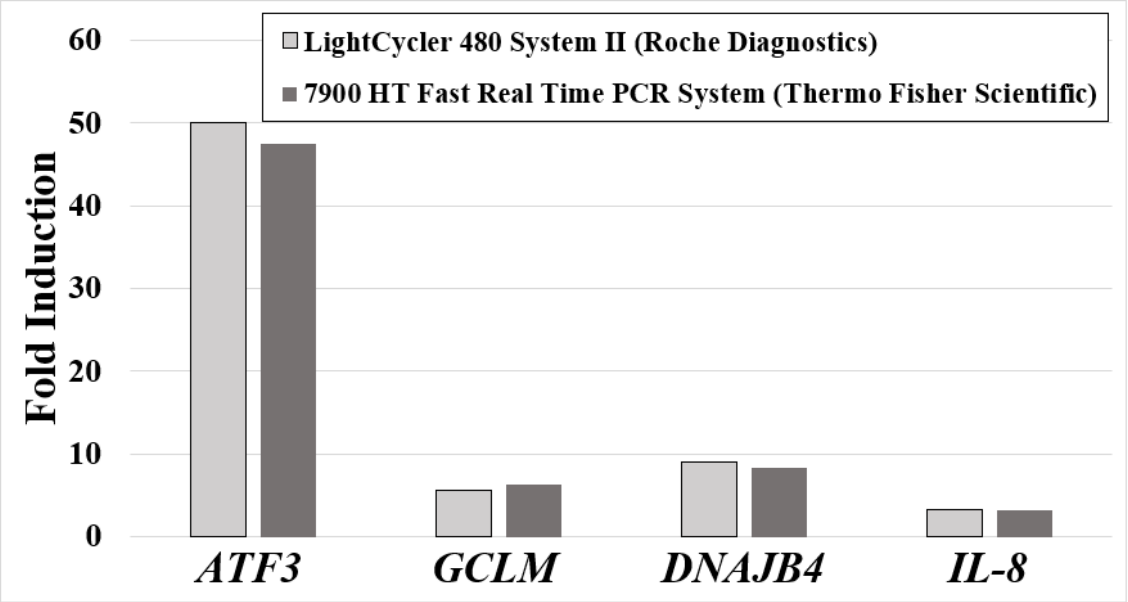


Figure 4. Fold induction of 0.10% w/v 4NBB obtained by each Real-Time PCR system



3. Discussion

Based on the above results, the results of Real-Time PCR systems which discussed in study 1 and 2 sections are comparable. Therefore, it is unlikely that different Real-Time PCR systems lead to different results.