Molecular & Cellular Toxicology (2023) 19:635–656
https://doi.org/10.1007/s13273-023-00387-6

REVIEW

Evaluations of carcinogens from comparison of cancer slope factors: meta-analysis and systemic literature reviews

Kyung-Taek Rim[1]

Accepted: 2 August 2023 / Published online: 14 August 2023

Abstract

Background This study seeks to estimate whether a chemical has carcinogenic potential; and if it has carcinogenic activity, its carcinogenic efficacy in humans and experimental animals in terms of oral and inhalation slope factors.

Objective Target chemicals were selected by literature search using Google Scholar, PubMed, ScienceDirect, etc., among the chemicals set by the Ministry of Employment and Labor in Korea as existing chemicals, and the CSF of each chemical was determined using various sites and programs, including EPA Comptox Dashboard and VEGA Hub QSAR (ver. 1.2.3). The CSF value of each chemical obtained using the Comparative Toxicogenomics Database (CTD) was subjected to gene expression analysis for inhalation carcinogenicity according to CSF value priority estimation, and a database (chemical list) was made possible.

Results Based on KOSHA-MSDS, GHS classification, and reference values for the CSF of each chemical, they were classified and organized using the OncoLogic 9.0 program. The priority of inhalation carcinogenicity was estimated by comparison with gene expression and CSF values, especially those with large inhalation-related values, and carcinogenesis of priority chemicals for inhalation. All the contents were organized and presented in an Excel file, and the priority of inhalation carcinogenicity was estimated through comparison with gene expression, focusing on CSFs, especially those with large inhalation-related values.

Conclusion Based on the obtained CSF value, the gene expression analysis of each chemical and toxic gene expression analysis of the CTD, inhalation carcinogenicity priority was estimated and a DB (chemical list) was prepared according to the CSF value.

Keywords Carcinogens · Cancer slope factor · Meta-analysis · Systemic literature reviews

Introduction

Carcinogenicity is a very important hazardous property for the safety evaluation of chemicals, and in the last few decades, the development of quantitative structure–activity relationships (QSARs), together with in vitro or in silico tests, has become important for regulatory use. Currently, several classification models are available to predict carcinogenicity in murine, but few models quantitatively assess carcinogenicity in humans. The cancer slope factor (CSF), a parameter describing potential carcinogenicity used for human risk assessment, has never been modeled for both oral and inhalation exposure. Therefore, the need to characterize the effects of chemicals is considered a priority research area by all environmental and health-related institutions in many countries, evaluating chemical carcinogenicity based on the CSF, a key parameter in health risk assessment (Toma et al. 2020). Although several QSAR models have been proposed for this purpose, few models can quantitatively evaluate carcinogenicity.

Exposure to a chemical or mixture occurs in the environment, residence, and workplace, but diet, drugs, and lifestyle can also be important co-triggers (Li and Suh 2019). Adverse effects include chronic disease and cancer, which today is a major public health problem with huge incidence. Although the procedure is complex, costly, and time-consuming, animal models are the most widely used investigation method, and are in great demand (Madia et al. 2016).

✉ Kyung-Taek Rim
rim3249@gmail.com

[1] Inhalation Toxicity Research Center, Occupational Safety and Health Research Institute, Korea Occupational Safety and Health Agency, Daejeon, South Korea

Recently, various non-animal models have been proposed as alternative or complementary methods to evaluate carcinogenicity to reduce animal experiments, evaluation time, and cost; and these methods include in silico methods, such as QSAR models and expert systems (Golbamaki et al. 2016; Yamane et al. 2016). Most in silico carcinogenicity models are tools used to predict whether a chemical is carcinogenic in an animal model (Zhang et al. 2017). Many of these models have already been implemented as license-based or freely available software tools, but models for oral and inhalation slope factors (SF) used for the human risk assessment of environmental contaminants have not yet been developed (Raitano et al., 2018; Bossa et al. 2018). The SF is an upper bound estimate of the slope of the dose–response curve in the low-dose regimen for carcinogens, and is used to assess the lifetime increase in incidence. CSF is used to estimate cancer risk associated with exposure to carcinogens or potential carcinogens, with a 95% confidence limit for increased cancer risk due to lifetime exposure to a chemical by ingestion or inhalation (Basic Information about the Integrated Risk Information System 2023). Therefore, the higher the slope value, the higher the carcinogenic potential.

If the chemical is a known or probable carcinogen to humans, a toxicity value (i.e., a slope factor) is calculated that quantitatively defines the relationship between dose and response. Since risk at low exposure levels is difficult to measure by animal experiments or epidemiological studies, the establishment of a gradient factor is usually necessary to adapt the model to available data sets, and to extrapolate from the relatively high doses administered in the experiment (Risk Assessment for Carcinogenic Effects 2022).

The difference from previous studies is that in most environmental risk management studies, the CSF is used to calculate the excess carcinogenic risk to determine the level of risk to the human body, and the efficiency of the process of selecting chemicals for carcinogenic inhalation toxicity tests is improved by comparing their CSFs. Further, research to contribute to establishing a chemical selection system for a new inhalation carcinogenicity test has not yet been attempted.

It is necessary to select substances for carcinogenic inhalation toxicity test in a new aspect through comparison of their CSFs (as a carcinogenic coefficient, the carcinogenic potential) used in the hazard and risk assessment of chemicals, and efficient carcinogenesis. It was necessary to construct a database of the various aspects necessary to select the target chemicals for the inhalation toxicity test. In this study, I tried to estimate which chemicals are likely to be carcinogenic, and, if so, the carcinogenic efficacy for humans and laboratory animals by oral and inhalation slope factors can help evaluate this. Making a model version of these findings available free of charge would greatly aid health risk assessment by making it easier to screen for carcinogenic chemicals. By comparing CSFs centering on chemicals that have become social issues or published in various papers, I tried to build one of the most efficient working systems in the process of selecting chemicals for carcinogenicity inhalation toxicity tests.

Materials and methods

Comparison of the CSFs of chemicals contributed to the efficiency of the selection process of carcinogenic inhalation toxicity test target chemicals, and served to establish a new target selected system for inhalation carcinogenicity test. The list centers on chemicals that have become social issues, or that have been published in various papers; in doing so, I tried to contribute to the list of priority chemicals for carcinogenicity testing by the CSF value of each chemical.

Target chemicals were selected using literature search, such as Google Scholar, PubMed, ScienceDirect, etc., among the chemicals set by the Ministry of Employment and Labor in Korea as existing chemicals; and the CSF of each chemical was determined using various sites and programs, including EPA Comptox Dashboard and VEGA Hub QSAR (ver. 1.2.3). Values were searched and analyzed separately for oral and inhalation. VEGA Hub QSAR (v. 1.1.5) stands for "Virtual models for property Evaluation of chemicals within a Global Architecture", and is a download-based package developed and distributed by Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Laboratory of Environmental Chemistry and Toxicology, Italy.

Gene expression analysis of each chemical was used to obtain the CSF value using the Comparative Toxicogenomics Database (CTD), which was analyzed for inhalation carcinogenicity according to CSF value priority estimation, and a database (chemical list) was made possible.

Result

A population (of finally 960 chemicals) was selected based on chemicals that became social issues or were published in various papers, the chemicals with a circulation volume of 1,000 tons or more were prioritized as the primary DB, and the SMILES form of each chemical was entered for continued searching.

Gene expression related to each chemical substance searched on the CTD site (ctdbase.com) was summarized, and the CSF value for each chemical was found on the EPA Computational Toxicology site.

Using the VEGA Hub program (ver. 1.2.3), the CSF values for each chemical substance were predicted on an in silico basis, and each predicted value was classified into oral and inhalation. In addition, using the in silico

carcinogenicity classification model in the VEGA Hub and Protox-II programs (tox-new.charite.de/protox_II), the carcinogenicity of each chemical was predicted.

Based on KOSHA-MSDS, GHS classification of each chemical and carcinogen classification done by IARC, NTP, EPA, OSHA, ACGIH, NIOSH, etc., were referred. Additionally, reference values for the CSF of each chemical were classified and organized using the OncoLogic 9.0 program.

All of the above results were summarized and presented in an Excel file (as an appendix), and the priority of inhalation carcinogenicity was estimated by comparison with gene expression and CSF values, especially those with large inhalation-related values, and the carcinogenesis of priority chemicals for inhalation.

Table 1 shows the chemicals found to express cancer-related genes with a CSF value of 1 or more, expressed as VEGA in silico inhalation [1/(mg/kg-d), where d = day]. This table shows a total of 17 chemicals. Table 2 shows the chemicals with a CSF value of 1 or more, and gene expression with the carcinogenesis-related signaling pathway (Fig. 1), with a total of 44 chemicals being shown.

Table 3 shows the results expected to be carcinogens, excepting the false positives in Table 2, and there are a total of 11 chemicals.

Discussion

In this study, an integrated in silico approach was attempted for the evaluation of chemical carcinogenicity potential, including classification and models for inhalational and oral human carcinogenicity based on CSFs. The CSF, a parameter with potential carcinogenicity used for human risk assessment, has never previously been adapted for both inhalation and oral exposure. Cancer potency factor (CPF) or CSF is a parameter that is used during the quantitative risk assessment of a chemical or drug that is evaluated as a carcinogen. Cancer efficacy is measured as the slope of a straight line generated during linear extrapolation of the low-dose region in a chemical dose–response curve (Farris and Ray 2014).

In silico models are evolving toward integrating multiple perspectives, and this integration will allow better utilization of the available data and information to tackle more difficult tasks. Users may be interested in the application of these tools, the evaluation of specific chemicals, or the evaluation of a large group of chemicals, and VEGA's development approach best addresses these user needs, reducing the barriers between different approaches (Benfeati et al. 2019).

The oral slope factor (OSF) is used to quantitatively estimate the carcinogenic efficacy or risk associated with chemical exposure through the oral route (Kar et al. 2012). The overall risk associated with chemical exposure is determined by combining quantitative estimates of chemical exposure with the known effects. For chemicals that cause carcinogenicity, OSF and inhalation unit risk are used to estimate the risk associated with carcinogenicity or exposure by the oral or inhalation route, respectively (Rim 2020).

In this study, the population (of finally 960 chemicals) was selected based on substances that became social issues or were published in various papers, and the contents of gene expression related to each chemical substance searched on the CTD site were summarized. EPA Computational Toxicology was conducted focusing on searching for CSF values, such as finding CSF values for each chemical substance on the site, and organizing the contents to be searched. However, there were not many substances with those values presented, so we used the VEGA Hub program to conduct in silico analysis. The CSF values for each chemical substance based on this study were predicted, and each predicted value was divided into oral and inhalation, and the contents were summarized. In addition, using the in silico carcinogenicity classification model in the VEGA Hub and Protox-II programs, the carcinogenicity prediction of each chemical substance was summarized.

This study simultaneously considers the CSF value used in the method of multiplying the lifetime exposure by the carcinogenic potential to find the excess carcinogenic risk in both the expression of genes, and the hazard and risk assessment of chemicals. As a new attempt to select a target substance for a toxicity test, it was intended to be used effectively. On the other hand, in VEGA Hub QSAR, when the result is negative but the result is statistically positive, it is termed a "false positive"; and when the result is negative, even though it is statistically positive, it is termed a "false negative". In this study, carcinogen was predicted by the CSF values, but it was judged that it would be possible to distinguish false positives depending on whether the experimental value was a carcinogen. Sensitivity and specificity are concepts to describe the accuracy of a test for reporting with or without a condition. The terms 'sensitivity' and 'specificity' were introduced in 1947 by Jacob Yerushalmy, a biostatistician (Yerushalmy 1947). Sensitivity (true positive rate) represents the probability of a conditionally positive when it is positive, while specificity (true negative rate) represents the probability of a conditionally negative when it is indeed negative.

Table 4 shows the changes in sensitivity and specificity in predicting carcinogenicity through VEGA Hub QSAR. When only carcinogenicity was predicted through the QSAR, the sensitivity was 53.85%, but when CSF was additionally considered, it increased to 58.82%; and when carcinogenic gene expression was additionally considered, it increased to 72.73%. In addition, when only carcinogenicity was predicted through QSAR, the specificity was 44.32%, but when CSF was additionally considered, it increased

Table 1 Priority substances for inhalation carcinogenicity

Chemicals (CAS No.)	GENE Expression	CSF-EPA CompTox	VEGA in silico inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf. Model	False positive (FP)/ False negative (FN)	ECHA Tox. Test rest
(Z)-9-Octadecenamide (CAS No. 301–02-0)	GJB1 GJB2 COL1A1 AHR CASP3 CASP9 CHAT GABRB3 PPARG	No cancer slope factor	3.57	Carcinogen	FP	–
(Z)-9-Octadecenoic acid (CAS No. 112–80-1)	TNF INS IL1B PPARA CPT1A ALB IL6 PLIN2 SREBF1 PPARG	No cancer slope factor	3.59	Carcinogen	–	–
(Z,Z)-9,12-Octadecandienoic acid; Linoleic acid (CAS No. 60–33-3)	PPARA PPARG TNF ABCA1 PTGS2 CASP3 IL6 AGER NFE2L2 ANXA5	No cancer slope factor	3.03	Carcinogen	FP	–
2,2-Bis(4'-glycidyloxyphenyl)propane (CAS No. 1675–54-3)	PPARG TNF VEGFA AGT ADIPOQ NTN1 RELA IL6 ACHE KLF5	No cancer slope factor	10.72	Carcinogen	FP	Carcinogenicity: via oral route (target organ): digestive: cecum Carcinogenicity: via dermal route (target organ): digestive: liver

Table 1 (continued)

Chemicals (CAS No.)	GENE Expression	CSF-EPA CompTox	VEGA in silico inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf. Model	False positive (FP)/ False negative (FN)	ECHA Tox. Test rest
4-(α,α-Dimethylbenzyl)phenol (CAS No. 599–64-4)	ESR1 AR NR1I2 BAX BCL2 ESRRG TFF1 THRB ABCB1 AR.L	No cancer slope factor	4.95	Carcinogen	FP	Data not provided by the registrant
Aluminum oxide; Alumina (CAS No. 1344–28-1)	IL1B TNF SPP1 IL6 VCAM1 CXCL8 EGFR SELE BCL2L1 CASP3	No cancer slope factor	3.96	Carcinogen	FN	Chronic toxicity, inhalation, rat: carcinogenicity: NOAEC ≥ 75 mg/m^3 as aluminum oxide The weight of evidence does not support a carcinogenic effect from exposure to aluminum oxide
Ammonium sulfate (CAS No. 7783–20-2)	HMOX1 IFNG PGR TNF	No cancer slope factor	1.81	Carcinogen	FP	No evidence of a carcinogenic potential was observed in a combined chronic toxicity/carcinogenicity study with rats closely following the requirements of OECD TG 453. Data on the purity of the test substance are lacking; however, since no adverse effects were observed, this is not considered to affect the evaluation of the carcinogenic potential of ammonium sulfate in an adverse manner
Biphenyl-4,4'-diol (CAS No. 92–88-6)	ESR1 AR PTGS2 RHEB	No cancer slope factor	30.19	Carcinogen	FP	–
Cellulose, methyl ester; Methylcellulose (CAS No. 9004–67-5)	ALB PPARA PPARB PPARG NR1I3	No cancer slope factor	4.26	Carcinogen	–	–

Table 1 (continued)

Chemicals (CAS No.)	GENE Expression	CSF-EPA CompTox	VEGA in silico inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf. Model	False positive (FP)/ False negative (FN)	ECHA Tox. Test rest
Dichromium trioxide (CAS No. 1308–38-9)	APBA1 BAG1 BAX BCL2 BMP2 BMP4 C3 CASP10 CASP3 CAT	No cancer slope factor	7.54	Carcinogen	FN	A number of published carcinogenicity studies have been performed with chromium (III) oxide. Studies are largely non-standard, using different routes of administration, but are consistently negative. Similarly, a number of studies performed using other chromium (III) compounds are consistently negative
Diiron trioxide (CAS No. 1309–37-1)	BAX TNF CAT DDIT3 IL6 PARP1 SOD2 ANLN BCL2 BCL2L11	No cancer slope factor	3.96	Carcinogen	FN	Long-term inhalation, oral and dermal carcinogenicity studies are not available. Seven different types of iron oxides were examined for carcinogenic properties after intratracheal instillation and intraperitoneal injection tests in rats, which represent particularly sensitive methods for local carcinogenic effects of particles/fibers. The total doses lay in the range of maximum tolerance
Dodecanoic acid (CAS No. 143–07-7)	HSTRPA CXCL8 CYP2C9 RELA CYP4Z1 NOD2 PCNA TJP1 ADH5 AKT1	No cancer slope factor	3.67	Carcinogen	–	–
Hexadecanoic acid (CAS No. 57–10-3)	TNF INS IL1B CASP3 IL6 CPT1A PPARA CASP7 INS1 NOS2	No cancer slope factor	4.63	Carcinogen	–	–

Table 1 (continued)

Chemicals (CAS No.)	GENE Expression	CSF-EPA CompTox	VEGA in silico inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf. Model	False positive (FP)/ False negative (FN)	ECHA Tox. Test rest
Lithium carbonate; Lithane (CAS No. 554–13-2)	ALAD CASP3 CAT GSR CXCL8 INS1 PFKFB2 PLA2G4A PTH ABCE-1	No cancer slope factor	8.22	Carcinogen	–	–
Melamine (CAS No. 108–78-1)	TGFB1 FN1 CCL2 IL6 VCAM1 CLU HAVCR1 BAD BAX CASP3	No cancer slope factor	2.76	Carcinogen	–	Two-year carcinogenicity studies were performed in rat and mouse by the US National Toxicology Program (NTP). Statistically significant increases in the incidence of transitional-cell carcinoma and combined incidences of transitional-cell carcinoma and papilloma in the urinary bladder were observed in male rats exposed to 4,500 ppm melamine (ca. 263 mg/kg bw/d), but not when exposed to 2,250 ppm melamine. With one exception, urinary bladder stones were observed in male rats that had transitional-cell carcinomas. Female rats did not develop tumors, even when exposed up to 9,000 ppm
Octadecanoic acid (CAS No. 57–11-4)	ALB IL1B PTGS2 TNF AKT1 INS SCD1 CSF2 CYP3A4 ADIPOQ	No cancer slope factor	5.36	Carcinogen	–	–

Table 1 (continued)

Chemicals (CAS No.)	GENE Expression	CSF-EPA CompTox	VEGA in silico inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf. Model	False positive (FP)/ False negative (FN)	ECHA Tox. Test rest
Tetrabutyl tin (CAS No. 1461–25-2)	CGB3 PPARG CYP17A1 HSD17B1 INSL3 LHCGR SCARB1 CYP19A1 STAR	No cancer slope factor	3.72	Carcinogen	–	–

Chemicals with a CSF value of 1 or higher and their oncogenic gene expression (17 chemicals) are shown

to 86.15%; and when carcinogenic gene expression was additionally considered, it slightly decreased to 80.56% (Table 4).

This is an indicator that when selecting a substance to be tested for carcinogenicity by considering its carcinogenic potential together with QSAR, it is possible to distinguish true negative, as well as true positive, show a significant improvement. Whereas it is not possible to find the expression of genes related to carcinogenesis in all chemicals, it is judged that additional consideration and research on methods for improving sensitivity and specificity using QSAR, etc., are necessary.

As for the expected effect and utilization plan of this study, it contributes to the selection of priority chemicals for efficient inhalation carcinogenicity, and a new attempt was made by estimating the CSF value using computational toxicology and toxicogenomics in chronic/carcinogenic inhalation toxicity. This CSF value can be used as a new frame for selecting test chemicals for these inhalation tests.

It was considered necessary to establish a DB in various aspects, such as the selection of chemicals to be tested for carcinogenicity in a new aspect through the comparison of CSF (as a carcinogenic potential) used in the hazard and risk assessment of chemicals. By comparing the CSF values that have become social issues or published in various papers, I sought to contribute to the list of chemicals subject to carcinogenicity testing. Based on the obtained CSF value, gene expression analysis of each chemical, and toxic gene expression analysis of the CTD, inhalation carcinogenicity priority estimation, and a DB (a chemical list) were made according to the CSF value. All the contents were organized and presented in an Excel file, and the priority of inhalation carcinogenicity was estimated through comparison with gene expression, focusing on CSFs, especially those with large inhalation-related values.

In this study, the change in sensitivity and specificity in predicting carcinogenicity through VEGA Hub QSAR when only carcinogenicity was predicted through the same QSAR was 53.85%, but when CSF was additionally considered, it increased to 58.82%; when the expression of oncogenes was additionally considered, it further increased to 72.73%. In addition, when only carcinogenicity was predicted through QSAR, the specificity was 44.32%; but when CSF was additionally considered, it increased to 86.15%; and when carcinogenic gene expression was additionally considered, it slightly decreased to 80.56%. This is an indicator that when selecting a substance to be tested for carcinogenicity by considering its carcinogenic potential together with QSAR, it is possible to distinguish true negative, as well as true positive, in predicting carcinogenicity. When the expression of carcinogenesis-related genes was also considered, the identification of true positives increased further, but the identification of

Table 2 Priority substances for inhalation carcinogenicity

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
(Z)-9-Octadecenamide (CAS No. 301-02-0)	GJB1 GJB2 COL1A1 AHR CASP3 CASP9 CHAT GABRB3 PPARG	3.57	Carcinogen	FP	hsa05200 Pathways in cancer - Homo sapiens (human) (3) [Cancer network viewer] hsa05016 hsa05417
(Z)-9-octadecenoic acid (CAS No. 112-80-1)	TNF INS IL1B PPARA CPT1A ALB IL6 PLIN2 SREBF1 PPARG	3.59	*Carcinogen*	–	hsa04932 hsa04936 hsa04931 hsa05200
(Z,Z)-9,12-Octadecandienoic acid; Linoleic acid (CAS No. 60-33-3)	PPARA PPARG TNF ABCA1 PTGS2 CASP3 IL6 AGER NFE2L2 ANXA5	3.03	Carcinogen	FP	hsa05417 hsa05010 has05022 hsa05200 Pathways in cancer-Homo sapiens (human) (5) [Cancer network viewer] hsa04932
2,2-Bis(4'-glycidyloxyphenyl)propane (CAS No. 1675-54-3)	PPARG TNF VEGFA AGT ADIPOQ NTN1 RELA IL6 ACHE KLF5	10.72	Carcinogen	FP	hsa04932 hsa04933 hsa05200 Pathways in cancer- Homo sapiens (human) (5) [Cancer network viewer]

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
4-(α,α-Dimethylbenzyl) phenol (CAS No. 599-64-4)	ESR1 AR NR1I2 BAX BCL2 ESRRG TFF1 THRB ABCB1 AR.L	4.95	Carcinogen	FP	hsa05200 Pathways in cancer - Homo sapiens (human) (4) [Cancer network viewer]
Aluminum oxide; alumina (CAS No. 1344-28-1)	IL1B TNF SPP1 IL6 VCAM1 CXCL8 EGFR SELE BCL2L1 CASP3	3.96	*Carcinogen*	FN	hsa05417 hsa04933 hsa05163 hsa04668 hsa05144 hsa05200
Biphenyl-4,4'-diol (CAS No. 92-88-6)	ESR1 AR PTGS2 RHEB	30.19	Carcinogen	FP	hsa05200
Cellulose, methyl ester; Methylcellulose (CAS No. 9004-67-5)	ALB PPARA PPARB PPARG NR1I3	4.26	*Carcinogen*	–	hsa05200
Dichromium trioxide (CAS No. 1308-38-9)	APBA1 BAG1 BAX BCL2 BMP2 BMP4 C3 CASP10 CASP3 CAT	7.54	*Carcinogen*	FN	hsa05200

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
Diiron trioxide (CAS No. 1309-37-1)	BAX TNF CAT DDIT3 IL6 PARP1 SOD2 ANLN BCL2 BCL2L11	3.96	*Carcinogen*	FN	hsa05200
Dodecanoic acid (CAS No. 143-07-7)	HSTRPA CXCL8 CYP2C9 RELA CYP4Z1 NOD2 PCNA TJP1 ADH5 AKT1	3.67	*Carcinogen*	–	hsa05200
Hexadecanoic acid (CAS No. 57-10-3)	TNF INS IL1B CASP3 IL6 CPT1A PPARA CASP7 INS1 NOS2	4.63	*Carcinogen*	–	hsa05200
Lithium carbonate; lithane (CAS No. 554-13-2)	ALAD CASP3 CAT GSR CXCL8 INS1 PFKFB2 PLA2G4A PTH ABCE-1	8.22	*Carcinogen*	–	hsa05200

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
Melamine (CAS No. 108-78-1)	TGFB1 FN1 CCL2 IL6 VCAM1 CLU HAVCR1 BAD BAX CASP3	2.76	*Carcinogen*	–	hsa05200
Octadecanoic acid (CAS No. 57-11-4)	ALB IL1B PTGS2 TNF AKT1 INS SCD1 CSF2 CYP3A4 ADIPOQ	5.36	*Carcinogen*	–	hsa05200
Tetrabutyl tin (CAS No. 1461-25-2)	CGB3 PPARG CYP17A1 HSD17B1 INSL3 LHCGR SCARB1 CYP19A1 STAR	3.72	*Carcinogen*	–	hsa05200
1,4:3,6-Dianhydro-D-glucitol (CAS No. 652-67-5)	BCHE NOS3 ACHE CASP3 CYP1A1 EDN1 G6PD NT5E PAFAH1B1 PRKG1	1.57	Non-carcinogen	–	hsa05200

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
2,2,4-Trimethyl-1,3-pentanediolester diisobutyrate (CAS No. 6846-50-0)	ESR1 ESR2	2.87	Non-carcinogen	FN	hsa05200
2,6-Dimethylhepta-2,5-dien-4-one (CAS No. 504-20-1)	HMOX1 FOS JUN ALDOA CYP2A5 CYP2B10 NFE2L2 ORM1 STAT3 ACVR2B	1.19	Non-carcinogen	–	hsa05200
2-Phenoxyethanol (CAS No. 122-99-6)	AR PGR	1.88	Non-carcinogen	FN	hsa05200
4-Vinyl-1-cyclohexene; 4-ethenylcyclohexene (CAS No. 100-40-3)	CYP2E1 AHR CYP2B1 ACHE ALAD CASP3 CAT GST KEAP1 P38A	1.13	Non-carcinogen	–	hsa05200
Benzaldehyde (CAS No. 100-52-7)	ALDH3A1 CYP2A6 ALDH1A1 ALDH2 SLC2A1 AKR1B8 ALDH1A7 ALDH3B1 AR MT2	1.13	Non-carcinogen	–	hsa05200

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
Cumene (CAS No. 98-82-8)	KRAS CLU ETS1 GSTP1 MAP2K1 AKAP12 AREG CADM1 CAV1 CCN1	1.07	Non-carcinogen	–	hsa05200
Decanedioic acid (CAS No. 111-20-6)	INS ADH5 NFE2L2 STAT5B	8.2	Non-carcinogen	–	hsa05200
Diethanolamine (CAS No. 111-42-2)	CASP3 BAD BAX CCND2 CCNE1 CDKN1A CDKN1B E2F1 CXCL8 GSS	1.53	Non-carcinogen	–	hsa05200
Diisononyl 1,2-cyclohexanedicarboxylate (CAS No. 166412-78-8)	ESR1 ESR2 AR DCP2 FCRL6 FGF18 HIBCH HPGDS HS3ST3A1 IQSEC3	1.35	Non-carcinogen	FN	hsa05200
Dinonylphthalate (CAS No. 84-76-4)	ABCB4 PPARA PPARB PPARG RXRB	1.33	Non-carcinogen	FN	hsa05200

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
Divanadium pentaoxide (CAS No. 1314-62-1)	HBEGF STAT1 CXCL8 CXCL10 IFNB1 PRKCD CASP3 CD44 MAPK1 MAPK3	7.39	Non-carcinogen	–	hsa05200
Dodecane (CAS No. 112-40-3)	IL6 CXCL8 IL1A POMC TAC1 TNF	5.37	Non-carcinogen	–	hsa05200
Ethyl tetradecanoate; ethyl myristate (CAS No. 124-06-1)	AR ESR1 ESR2 NR3C1 PGR PPARG THRB	1.04	Non-carcinogen	FN	hsa05200
Ethylbenzene (CAS No. 100-41-4)	BAX BCL2 CASP3 CYCS CASP9 CYP2E1 HMOX1 ACE BMP8B CARMIL3	1.13	Non-carcinogen	FN	hsa05200

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
Lanthanum oxide (CAS No. 1312-81-8)	AMH BAX BCL2 CASP3 CDH2 OCLN TESMIN VIM ALB HSD3B1	3.96	Non-carcinogen	–	hsa05200
Maleic acid (CAS No. 110-16-7)	ACO2 CLDN2 HAVCR1 LCN2 NFE2L2 ABCA7 AGAP1 AKT2 B4GALNT4 C20ORF173	1.42	Non-carcinogen	FP	hsa05200
n-Butyl alcohol (CAS No. 71-36-3)	MAPK3 MAPK1 C5 PLD1 AKT1 PDGFB PLD2 PTGS2 AGT EGF	1.22	Non-carcinogen	–	hsa05200
Octamethylcyclotetra-siloxane (CAS No. 556-67-2)	BRCA1 ESR1 ATM ATR BRCA2 CHEK1 CHEK2 CYP2B1 CYP2B2 CYP3A2	3.18	Non-carcinogen	–	hsa05200

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
Pentane (CAS No. 109-66-0)	BAX CEBPB FN1 GNG12	7.54	Non-carcinogen	–	hsa05200
Propanoic acid; Propionic acid (CAS No. 79-09-4)	IL10 ASCL1 BCL2 CASP3 CAT IL6 LFNG PPARGC1A SIRT3 SLC16A1	11.17	Non-carcinogen	–	hsa05200
Propylene glycol (CAS No. 57-55-6)	TGFB1 IL6 MAPK1 MAPK3 ABCC2 ABCC3 ABCC4 ADH5 ANGPT1 ARHGEF26	11.17	Non-carcinogen	–	hsa05200
Tetraethoxy silane; tetraethoxysilicon (CAS No. 78-10-4)	BCL2 CREB3L1 AKT1 CASP3 CASP9	1.78	Non-carcinogen	–	hsa05200
Triethanolamine (CAS No. 102-71-6)	ALB CXCL8 IFNG IL2 KCNK18	1.67	Non-carcinogen	–	hsa05200

Table 2 (continued)

Chemicals (CAS No.)	Gene Expression	VEGA in silico Inhalation [1/(mg/kg-day)]	VEGA in silico Carc inhala classf.Model	Specificity/Sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
Triethylenetramine (CAS No. 112-24-3)	SOD1 APP SLC31A1 AGER CASP3 CXCL8 BACE1 S100B BCL2 PRNP	2.57	Non-carcinogen	–	hsa05200
Triethylphosphate (CAS No. 78-40-0)	CYP3A7 TP53 GJB1 LBFABP MT4 TTR TXN UGT1A9	1.12	Non-carcinogen	–	hsa05200
Tris(1-chloro-2-propyl) phosphate (CAS No. 13674-84-5)	CYP3A7 FABP1 NR1I2 SYN2A THRSP TP53 ACHE ADM ADM2 APLNR	4.52	Non-carcinogen	–	hsa05200
Tris(2-ethylhexyl) benzene-1,2,4-tricarboxylate (CAS No. 3319-31-1)	ESR1 AR ESR2 NR1I2 ESR1.L ESR2A ESR2B ESR2.L THRB	3.57	Non-carcinogen	FN	hsa05200

[a]FP, false positive; FN, false negative

Table 3 Substances expected to be carcinogens (11 chemicals), excepting the false positives in Table 2

Chemicals (CAS No.)	Gene expression	VEGA in silico inhalation [1/(mg/kg-day)]	VEGA in silico carc inhala classf.Model	Specificity/sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
(Z)-9-Octadecenoic acid (CAS No. 112-80-1)	TNF INS IL1B PPARA CPT1A ALB IL6 PLIN2 SREBF1 PPARG	3.59	*Carcinogen*	–	hsa04932 hsa04936 hsa04931 hsa05200
Aluminum oxide; Alumina (CAS No. 1344-28-1)	IL1B TNF SPP1 IL6 VCAM1 CXCL8 EGFR SELE BCL2L1 CASP3	3.96	*Carcinogen*	FN	hsa05417 hsa04933 hsa05163 hsa04668 hsa05144 hsa05200
Cellulose, methyl ester; Methylcellulose (CAS No. 9004-67-5)	ALB PPARA PPARB PPARG NR1I3	4.26	*Carcinogen*	–	hsa05200
Dichromium trioxide (CAS No. 1308-38-9)	APBA1 BAG1 BAX BCL2 BMP2 BMP4 C3 CASP10 CASP3 CAT	7.54	*Carcinogen*	FN	hsa05200
Diiron trioxide (CAS No. 1309-37-1)	BAX TNF CAT DDIT3 IL6 PARP1 SOD2 ANLN BCL2 BCL2L11	3.96	*Carcinogen*	FN	hsa05200
Dodecanoic acid (CAS No. 143-07-7)	HSTRPA CXCL8 CYP2C9 RELA CYP4Z1 NOD2 PCNA TJP1 ADH5 AKT1	3.67	*Carcinogen*	–	hsa05200

Table 3 (continued)

Chemicals (CAS No.)	Gene expression	VEGA in silico inhalation [1/(mg/kg-day)]	VEGA in silico carc inhala classf.Model	Specificity/sensitivity[a]	KEGG Mapper—Color (genome.jp/kegg/mapper/color.html) (genome.jp/pathway/hsa05200)
Hexadecanoic acid (CAS No. 57-10-3)	TNF INS IL1B CASP3 IL6 CPT1A PPARA CASP7 INS1 NOS2	4.63	*Carcinogen*	–	hsa05200
Lithium carbonate; lithane (CAS No. 554-13-2)	ALAD CASP3 CAT GSR CXCL8 INS1 PFKFB2 PLA2G4A PTH ABCE-1	8.22	*Carcinogen*	–	hsa05200
Melamine (CAS No. 108-78-1)	TGFB1 FN1 CCL2 IL6 VCAM1 CLU HAVCR1 BAD BAX CASP3	2.76	*Carcinogen*	–	hsa05200
Octadecanoic acid (CAS No. 57-11-4)	ALB IL1B PTGS2 TNF AKT1 INS SCD1 CSF2 CYP3A4 ADIPOQ	5.36	*Carcinogen*	–	hsa05200
Tetrabutyl tin (CAS No. 1461-25-2)	CGB3 PPARG CYP17A1 HSD17B1 INSL3 LHCGR SCARB1 CYP19A1 STAR	3.72	*Carcinogen*	–	hsa05200

[a]*FP* false positive, *FN* false negative

true negatives did not show much improvement. On the other hand, the expression of carcinogenesis-related genes cannot be found in all chemicals, so it is judged that additional consideration and research on this are necessary.

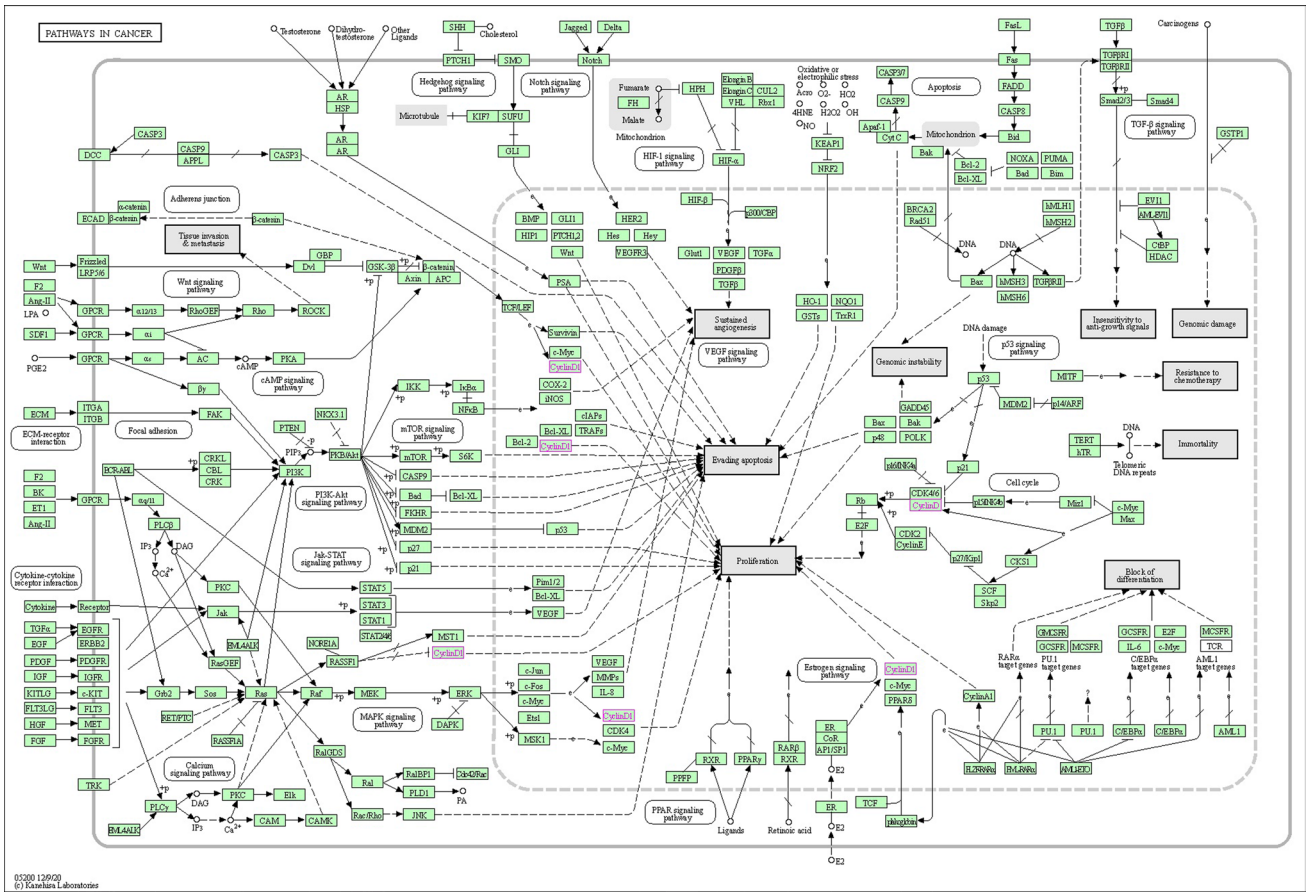

Fig. 1 Oncogenesis-related signaling pathway (hsa05200). Sourced from the Kyoto Encyclopedia of Genes and Genomes (KEGG), https://genome.jp/pathway/hsa05200. Adapted with permission

Table 4 Changes in VEGA Hub QSAR carcinogenicity predicted sensitivity and specificity

	Sensitivity (%)	Specificity (%)
Cancer prediction	14/(14 + 12) = 53.85	121/(121 + 152) = 44.32
+ CSF	10/(10 + 7) = 58.82	56/(56 + 9) = 86.15
+ Gene expression	8/(8 + 3) = 72.73	58/(58 + 14) = 80.56

This table shows the chemicals with a CSF value of 1 or higher and their gene expression in the signaling pathway related to carcinogenesis (44 chemicals)

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13273-023-00387-6.

Acknowledgements This study was supported by the Korea Occupational Safety and Health Agency (Ulsan, Republic of Korea), the Ministry of Employment and Labor (Sejong, Republic of Korea), and a Grant-in Aid for chemical research (2022).

Author contributions K-T Rim designed the experiments, analyzed the results, and wrote the manuscript.

Declarations

Conflict of interest K-T Rim declares that he has no conflict of interest regarding the contents of this article.

Ethical approval This article does not contain any studies with human participants or animals performed by the author, and it has been carried out following the institutional and national guidelines.

References

Basic Information about the Integrated Risk Information System. 2023. http://www.epa.gov/iris/basic-information-about-integrated-risk-information-system. Accessed 19 Apr 2023

Benfenati E, Roncaglioni A, Lombardo A, Manganaro A (2019) Integrating QSAR, read-across, and screening tools: the vegahub platform as an example. In: Advances in computational toxicology. Springer, Berlin, pp 365–381

Bossa C, Benigni R, Tcheremenskaia O, Battistelli CL (2018) (Q) SAR methods for predicting genotoxicity and carcinogenicity: scientific rationale and regulatory frameworks. In: Computational Toxicology. Springer, Berlin, pp 447–473

Farris FF, Ray SD (2014) Cancer potency factor. In: Encyclopedia of toxicology, vol 1, 3rd edn. Elsevier, Amsterdam, The Netherlands, pp 642–644

Golbamaki A, Benfenati E, Golbamaki N, Manganaro A, Merdivan E et al (2016) New clues on carcinogenicityrelated substructures derived from mining two large datasets of chemical compounds. J Environ Sci Health Part C 34:97–113

Kar S, Deeb O, Roy K (2012) Development of classification and regression based QSAR models to predict rodent carcinogenic potency using oral slope factor. Ecotoxicol Environ Saf 82:85–95

Li D, Suh S (2019) Health risks of chemicals in consumer products: a review. Environ Int 123:580–587

Madia F, Worth A, Corvi R (2016) Analysis of carcinogenicity testing for regulatory purposes in the European Union. European Commission, Luxembourg

Raitano G, Goi D, Pieri V, Passoni A, Mattiussi M et al (2018) Eco) toxicological maps: A new risk assessment method integrating traditional and in silico tools and its application in the Ledra River (Italy. Environ Int 119:275–286

Rim KT (2020) In silico prediction of toxicity and its applications for chemicals at work. Toxicol Environ Health Sci 12:191–202

Risk Assessment for Carcinogenic Effects. 2022. http://ww.epa.gov/fera/risk-assessment-carcinogenic-effects. Accessed 19 Apr 2023

Toma C, Manganaro A, Raitano G, Marzo M, Gadaleta D et al (2020) QSAR models for human carcinogenicity: an assessment based on oral and inhalation slope factors. Molecules 26:127

Yamane J, Aburatani S, Imanishi S, Akanuma H, Nagano R et al (2016) Prediction of developmental chemical toxicity based on gene networks of human embryonic stem cells. Nucleic Acids Res 44:5515–5528

Yerushalmy J (1947) Statistical problems in assessing methods of medical diagnosis with special reference to x-ray techniques. Public Health Rep 62:1432–1439

Zhang L, Ai H, Chen W, Yin Z, Hu H et al (2017) CarcinoPred-EL: novel models for predicting the carcinogenicity of chemicals using molecular fingerprints and ensemble learning methods. Sci Rep 7:2118