



Systems toxicology assessment of a representative e-liquid formulation using human primary bronchial epithelial cells

Diego Marescotti^{1,*}, Carole Mathis¹, Vincenzo Belcastro, Patrice Leroy, Stefano Acali, Florian Martin, Rémi Dulize, David Bornand, Dariusz Peric, Emmanuel Guedj, Laura Ortega Torres, Matteo Biasioli, Matthieu Fuhrmann, Estela Fernandes, Felix Frauendorfer, Ignacio Gonzalez Suarez, Davide Sciuscio, Nikolai V. Ivanov, Manuel C. Peitsch, Julia Hoeng

PMI R&D, Philip Morris Products S.A., Quai Jeanrenaud 5, CH-2000 Neuchâtel, Switzerland

ARTICLE INFO

Keywords:

Flavoring substances
High-content screening
Systems toxicology
Electronic cigarettes
Bronchial epithelial cells

ABSTRACT

The development of reduced-risk products aims to provide alternatives to cigarettes that present less risk of harm for adult smokers. Responsible use of flavoring substances in these products may fulfill an important role in product acceptance. While most flavoring substances used in such products are also used by the food industry and are considered safe when ingested, their impact when inhaled may require further assessment. To aid in such an assessment, a three-step approach combining real-time cellular analysis, phenotypic high-content screening assays, and gene expression analysis was developed and tested in normal human bronchial epithelial cells with 28 flavoring substances commonly used in e-liquid formulations, dissolved individually or as a mixture in a base solution composed of propylene glycol, vegetable glycerin, and 0.6% nicotine. By employing this approach, we identified individual flavoring substances that potentially contribute greatly to the overall mixture effect (citronellol and alpha-pinene). By assessing modified mixtures, we showed that, although cytotoxic effects were found when assessed individually, alpha-pinene did not contribute to the overall mixture cytotoxicity. Most of the cytotoxic effect appeared to be attributable to citronellol, with the remaining substances contributing due to synergistic effects. We developed and used different scoring methods (Tox-Score, Phenotypic Score, and Biological Impact Factor/Network Perturbation Amplitude), ultimately enabling a ranking based on cytotoxicity, phenotypic outcome, and molecular network perturbations. This case study highlights the benefits of testing both individual flavoring substances and mixtures for e-liquid flavor assessment and emphasized the importance of data sharing for the benefit of consumer safety.

1. Introduction

As an alternative to cigarettes, a large number of electronic nicotine delivery systems (ENDS), including electronic cigarettes (e-cigarettes), have been developed and launched on the market. These systems heat a solution (e-liquid) to generate an aerosol, which is then inhaled by the user. All e-liquid solutions are composed of a mixture of propylene glycol (PG), vegetable glycerol (VG), water, and nicotine in variable proportions (referred to here as the base solution). Finally, a combination of different flavorings substances characterizing each unique product is added to complete the formulation.

Many of these flavor additives, usually found in food, have been extensively assessed and are “generally recognized as safe” (GRAS) [1]. However, GRAS™ certification by the Flavor Extracts Manufacturers

Association Expert Panel relates to an assessment of a flavor’s safety when ingested and not when it is inhaled [2]. Thus, for the majority of flavoring substances, the effects of inhaling these substances on the health of consumers will benefit from further study.

In recent years, we have developed and used a layered framework to systematically evaluate the potential impact of e-liquids and their corresponding aerosols using human *in vitro* models (to follow the 3Rs principles [3]) and a systems toxicology approach [4,5]. This assessment paradigm is intended to be an important complement to the required regulatory safety toxicological data and is in line with the Tox21 framework [6], which advocates the use of high-throughput methods to generate mechanistic data. For example, a recent study used regulatory safety toxicological tests to assess whether addition of flavoring substances to an unflavored product contributes to the toxicity of the

* Corresponding author at: PMI R&D, Philip Morris Products S.A., Quai Jeanrenaud 5, CH-2000 Neuchâtel, Switzerland.

E-mail address: diego.marescotti@pmi.com (D. Marescotti).

¹ Equally contributed to the work.

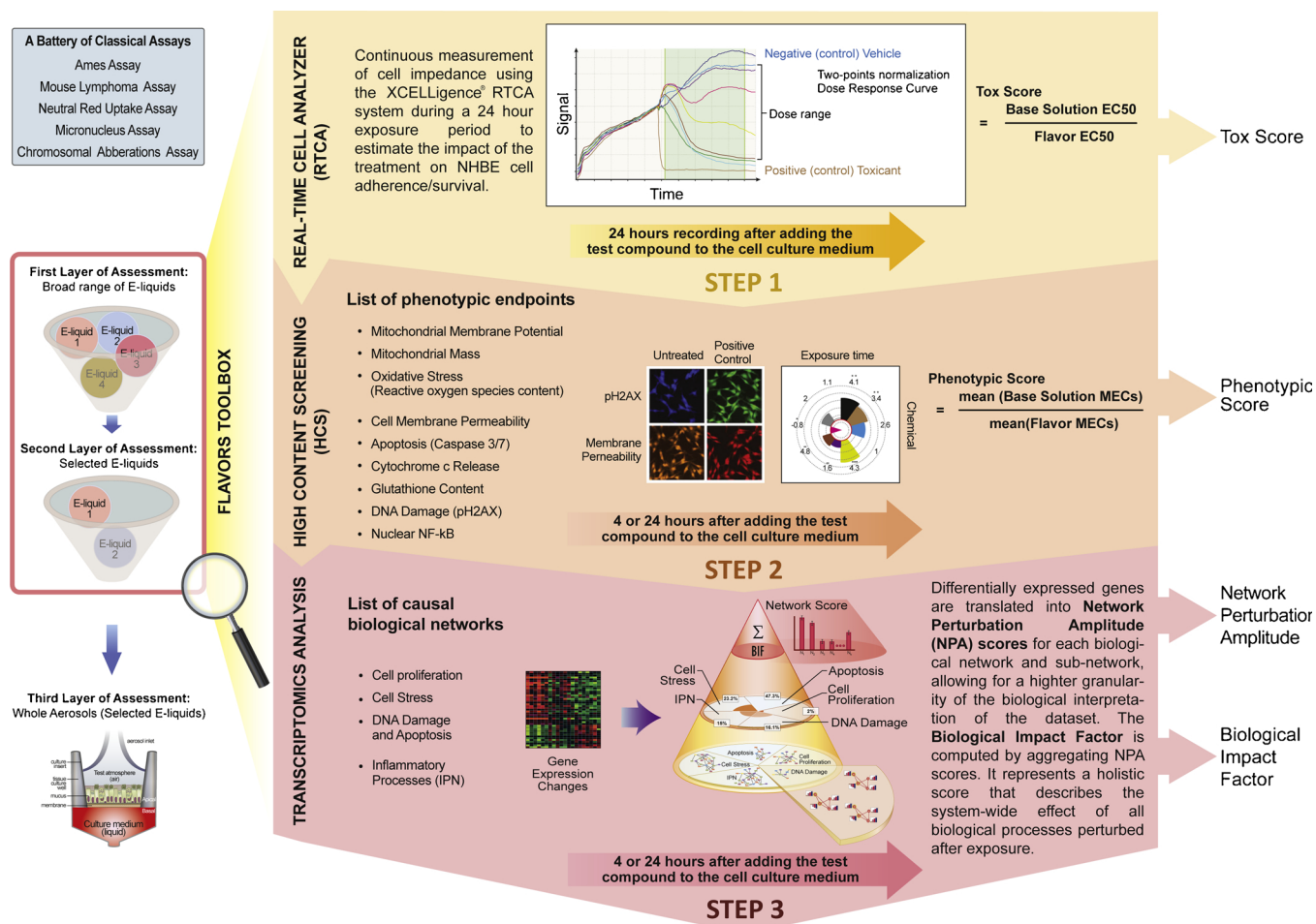


Fig. 1. The Flavor Toolbox: a three-step workflow to assess the toxicity of flavored solutions in NHBE cells. The Flavor Toolbox is a complementary approach to standard toxicity assays (e.g., Ames assay, mouse lymphoma assay, etc.). It is designed to screen a large number of e-liquids for potential toxicity prior to performing whole aerosol assessment on human organotypic tissue cultures. The Flavor Toolbox workflow comprises the following three steps: (i) STEP1, which quantifies the toxicity of the exposure using a real-time impedance-based measurement, expressed as Tox-Score; (ii) STEP 2, which measures and investigates the phenotypic impact of the exposure using HCS image analysis; and (iii) STEP 3, which combines transcriptomic data and computable biological networks in a systems toxicology approach. For each STEP, a computationally derived score (Tox-Score, Phenotypic Score, and NPA/BIF score) was derived to quantify the e-liquids' exposure effects.

gas-vapor phase [7]. Such tests can now also be conducted by using more advanced whole-aerosol exposure systems [8,9].

The toxicological assessment presented here is a case study performed in the context of the “Flavor Toolbox” extension introduced in our previous publication [5]. This assessment relies on a well-established three-pillar approach, which was also recently employed to investigate the effects of different base e-liquids. It proved useful in demonstrating increased toxicity upon nicotine addition to the PG/VG mixture and cytotoxic effects proportional to the PG content of the mixture [4].

In this study, (i) a real-time cellular analysis (RTCA) was performed to evaluate cellular cytotoxicity, (ii) a panel of imaging-based high-content screening (HCS) endpoints was used to analyze cellular phenotypic changes, and finally, (iii) a combination of gene expression analysis and computational models was used to gain a deeper understanding of changes occurring at the molecular level (Fig. 1).

For this case study, we generated a representative flavor mixture by grouping 180 flavor candidates into 25 distinct chemical groups based on structural similarities and potential metabolic and biological effects. At least one flavoring substance predicted to show the highest potential toxicological effect from each group was selected as the flavor group representative (FGR) based on *in silico* prediction and literature survey to generate the representative flavor mixture. By exposing normal human bronchial epithelial (NHBE) cells for at least 24 h to different dilutions of flavored e-liquids, including the complete mixture and individual flavoring

substances in base solution, we first estimated the effective concentration at which 50% cytotoxicity was reached (EC_{50}) for each flavored solution using real-time impedance-based cellular analysis (see STEP 1 in Fig. 1). We then computed a Tox-Score (ratio of EC_{50} base solution / EC_{50} flavored solution) to rank and evaluate each of the 28 flavoring substances individually. This enabled the identification of the flavoring substances contributing the most to the overall toxicity of the 28-flavor mixture. To confirm our hypothesis that only a small subset of flavoring substances (suspects) were primarily responsible for the toxicity of the 28-flavor mixture, we also tested additional mixtures where the main candidates were either omitted or added alone to the base solution.

In the second step, the phenotypic effects of the flavored e-liquids, including all mixtures and some of the individual flavoring substances, were further investigated in NHBE cells. In particular, a panel of HCS endpoints (see STEP 2 in Fig. 1) were evaluated after 4- and 24-h exposures. Finally, minimal effective concentrations (MEC) for each individual endpoint were calculated to quantify the overall impact observed with HCS through the computation of a Phenotypic Score (ratio of MEC base solution/MEC flavored solution).

The last step of the flavoring substances assessment workflow was to obtain a deeper understanding of the biological perturbations triggered by exposure to the 28-flavor mixture compared with those of the base solution alone using a combination of transcriptomic data and computable biological networks (see STEP 3 in Fig. 1). Perturbed biological

networks following exposure can be quantified using Network Perturbation Amplitude (NPA) and Biological Impact Factor (BIF) scoring methods [10–12]. This systems toxicology approach is a powerful way to investigate the exposure effects on pathways involved in cell proliferation, cell stress, DNA damage/apoptosis, and inflammatory responses and can help to identify potential biomarkers that could be used for future risk assessment.

2. Materials & methods

2.1. In vitro model

NHBE cells were purchased from Lonza (Basel, Switzerland) and were derived from a healthy donor, a 60-year-old Caucasian, alcohol-using (i.e., consuming alcohol, but not in a diseased state) male with no history of smoking (lot number 0000140733). The provider certified that the cells tested negative for mycoplasma, bacteria, yeast, fungi, HIV-1, hepatitis B, and hepatitis C. For all experiments described here, we used cells between passages 5 and 8. Cells were cultured as previously described [4].

2.2. Base solution and flavor mixture design

The 28 flavoring substances tested were derived from a list of more than 180 flavor ingredients (which do not include known carcinogens, mutagens, reproductive toxicants or respiratory sensitizers) considered for use in Philip Morris International's Reduced-Risk Products² and cigarettes (see Table 1). They were grouped into 25 families based on structural similarities, as defined by the European Food Safety Authority [13]. Natural extracts and essential oils were not included due to the variable composition inherent to their production.

The flavoring substances selected belonged to a number of different chemical classes, such as carboxylic acids, esters, ketones, or terpenes. Flavoring substances within a given chemical group were expected to share some metabolic and biological behaviors [13]. Therefore, based on physicochemical properties (volatility), available/predicted toxicological data (i.e., lethal dose 50, predicted lethal concentration 50, predicted ocular irritancy, predicted developmental toxicity, and predicted chronic lowest-observed-adverse-effect level), and the estimated usage level for the different flavoring substances, one or more FGR was selected for each chemical group.

Even though the total proportion of flavoring substances were reported to be in the range of 1–4% in many e-liquids [14], in this study we tested solutions with up to 5.7% flavor content. The 28 flavoring substances assessed as a mixture or as a single compounds in base solution were all of synthetic origin, of food-grade quality, and had at least 95% purity (see Table S1 for corresponding references and suppliers). Flavoring substances were dissolved in PG (BASF, Lampertheim, Germany) to prepare the flavor stocks, which were then added to a base solution composed of 41% PG, 38% VG (Emery Oleochemicals GmbH, Düsseldorf, Germany), and 0.6% nicotine (Siegfried, Zofingen, Switzerland). The 0.6% nicotine concentration in the base solution was in the range found in many e-liquids, as reported by Tierney et al. [14].

Phosphate-buffered saline (PBS) (Sigma-Aldrich International GmbH, St Gallen, Switzerland) was used to complete the e-liquid mixtures, and the PBS content was increased to compensate for decreased flavor volumes when individual flavoring substances or alternative mixtures were prepared. Each e-liquid solution was further diluted in cell culture medium, and the pH of the final solution at the condition used for exposure was assessed to ensure that it was within the physiological range (7.2–7.6). Considering the limited stability of butyric

acid, ethyl maltol, and isobutyl alcohol, the flavored stock solutions containing these ingredients were used within five days. All stock solutions were stored at 4 °C in the dark. Dilutions of the e-liquids in culture medium were freshly prepared before starting cell exposure, as shown in Fig. S1. It is important to note that although a wide range of e-liquid dilutions was assessed, only a single e-liquid flavor concentration was evaluated with the present approach. Despite this limitation, we still consider this approach to be relevant, because (i) the study was intended to be a case study for mixture assessment, and (ii) the individual flavoring substances assessment was not intended to provide an indication of non-toxic levels.

2.3. RTCA-based data generation and analysis

NHBE cells were seeded into E-Plate View 96-well tissue culture plates (ACEA Biosciences, San Diego, CA, USA) at a density of 7,200 cells per well in 100 µL of culture medium and incubated at 37 ± 1 °C in a humidified incubator with $5.0 \pm 0.5\%$ CO₂ for 24 ± 1 h. Cells were then exposed (in triplicate wells within the same plate) to seven different dilutions of each flavored solution. Positive (carbonyl cyanide m-chlorophenyl hydrazine, CCCP, C2759 or Triton X-100, T8787, from Sigma-Aldrich, St. Louis, MO, USA) and negative (cell culture medium) controls were included in each experiment. At least two independent experiments were performed for each solution. Impedance was measured continuously using an xCELLigence RTCA® system (ACEA Biosciences) from the time point of cell seeding and for at least 24 h of exposure to various mixtures, such as the 28-flavor mixture, individually flavored e-liquids, and the corresponding base solution. The resulting values were normalized to both the negative control and the most toxic condition within the plate (two-point normalization) to obtain the relative cell index (%) using the following formula:

$$\text{Cell index (\%)} = \left[1 - \left[\frac{F - V}{P - V} \right] \right] \times 100,$$

where F is the area under the curve (AUC) for the exposure condition at a defined concentration, V is the AUC for the negative control, and P is the AUC for the positive control, which is always measured beginning 1 h after adding the e-liquid solutions to the cells and for at least 24 h of exposure. All AUC values were normalized by the signal measured just prior to adding the test solution. The resulting value represents the percentage of cell index normalized to the vehicle control. A cell index equal to 100% was assigned to the negative control, because it did not induce any cytotoxicity. A cell index equal to 0% was assigned to the most toxic condition within the plate.

A two-parameter Hill equation was then used to fit the dose-response curve and extrapolate the EC₅₀:

$$Y = S_0 + \frac{S_{Inf} - S_0}{1 + \left(\frac{EC_{50}}{10^C} \right)^n}$$

where S_0 is the fitted activity level at zero concentration of the test compound, S_{Inf} is the fitted activity level at infinite concentration of the test compound, n is the Hill coefficient for the curve, EC_{50} is the concentration at which the activity reaches 50% of its maximum level, and C is the concentration in logarithmic units corresponding to the values on the x-axis of the dose-response plot.

2.4. Tox-Score calculation and statistical comparison

To minimize the variability due to biological cell responses, we scored the toxic effect of each flavoring substance by performing a paired analysis based on the flavor-specific shift in EC₅₀. The Tox-Score was defined as the ratio between the EC₅₀ of the base solution and the EC₅₀ of the flavored e-liquid (Fig. S2) to ensure that the decrease in EC₅₀ following addition of a toxic flavoring substance to the base solution was reflected by an increase in Tox-Score.

² "Reduced-Risk Products" or "RRPs" is the term we use to refer to products that present, are likely to present, or have the potential to present less risk of harm to smokers who switch to these products versus continued smoking.

Table 1

Family, chemical class, flavoring substance name, CAS registry number, and concentration of the 28 flavoring substances tested in the e-liquid solutions (in ppm and molarity).

Family	Chemical class	Flavoring substance name*	CAS registry number	Concentrations	
				ppm	Molarity (mM)
XXII	Ketone	2-acetylpyridine	1122-62-9	840	7.5
XXIII	Ketone	2-acetylthiazole	24295-03-2	16	0.2
III	Ester	Allyl hexanoate	123-68-2	280	1.6
XXV	Terpene	Alpha-pinene	80-56-8	480	3.0
I	Carboxylic acid	Butyric acid	107-92-6	5600	61.3
VII	Terpenoid, ketone	L-carvone	6485-40-1	1200	7.7
XVII	Alcohol	Cinnamyl alcohol	104-54-1	504	3.9
IV	Terpenoid, alcohol	Citronellol	106-22-9	4800	26.3
XXIV	Diketone	Cyclotene	80-71-7	1756	15.7
I	Ester	Ethyl acetate	141-78-6	5600	57.3
XIX	Aromatic heterocycle	2-ethyl-3,5-dimethylpyrazine	13925-07-0	160	1.1
I	Ester	Ethyl formate	109-94-4	5600	69.6
X	Ketone, alcohol	Ethyl maltol	08-11-4940	8160	58.2
XIII	Terpenoid, ether	Eucalyptol	470-82-6	720	4.3
XIV	Ester	Eugenyl acetate	93-28-7	1440	7.5
XI	Aromatic heterocycle	Furaneol	3658-77-3	1320	10.3
VIII	Lactone	Gamma-valerolactone	108-29-2	3000	31.5
XX	Phenol	Guaiacol	90-05-1	107	1.0
V	Ketone	Heptan-2-one	110-43-0	326	2.3
II	Alcohol	Isobutyl alcohol	78-83-1	544	5.9
VI	Terpene, alcohol	Linalool	78-70-6	2400	13.5
VII	Terpene, ketone	Menthone	89-80-5	1200	6.9
XXI	Ester, aniline	Methyl anthranilate	134-20-3	288	2.2
IX	Diketone	3-methyl-2,4-nonanedione	113486-29-6	22	0.1
XVIII	Ester, phenol	Methyl salicylate	119-36-8	2320	17.9
XV	Aldehyde, sulfide	3-(methylthio) propionaldehyde	3268-49-3	88	0.9
XVI	Ketone, phenol	4-(para-hydroxyphenyl)-butan-2-one	5471-51-2	4800	29.2
XII	Alcohol	Phenethyl alcohol	60-12-8	2840	32.1

Abbreviation: ppm, parts per million. * Ordered alphabetically.

Tox-Score comparisons were performed using *t*-tests on the log-transformed half maximal effective concentrations (EC₅₀), paired by experimental day. In the event a substance was evaluated twice during an experimental day, the averaged EC₅₀ was used.

2.5. HCS data generation

NHBE cells were seeded into black, clear-bottom 96-well tissue culture plates at a density of 37,500 cells per cm² (12,000 cells per well) in 100 µL of culture medium. The cells were incubated for 24 h in the culture medium and then exposed (in three replicate wells) to increasing concentrations of the e-liquid solutions containing the different flavoring substances. A corresponding base solution was included as a reference. The cells were exposed for 30 min (NF-κB endpoint only), 4 h, and 24 h before performing the HCS assays. In parallel, appropriate positive controls at three different concentrations were used for each endpoint (Table S2). Dimethyl sulfoxide (CASRN 67-68-5; purity > 99.9%; ref. D8418, Sigma-Aldrich) was used as the vehicle for all positive control treatments at a final concentration of 0.5%. Cell count was recorded in all assays using Hoechst 33342 staining (Life Technologies, Carlsbad, CA, USA), except for the oxidative stress assay, where DRAQ5™ dye (Biostatus, Shephed, UK) was used. Following staining of the NHBE cells, fluorescence was analyzed by image acquisition with a Thermo Fisher Cellomics® ArrayScan VTI High Content Screening Reader (Thermo Fisher Scientific, Waltham, MA, USA) and vHCS™:View software (Thermo Fisher Scientific). Twenty fields were imaged per well using a 10× wide field objective, as previously described [15].

2.6. HCS data processing

HCS assays generate multiple fluorescence readouts that are measured simultaneously. The fluorescence image analyses were stored in the database linked to the GladiaTOX package [16]. GladiaTOX is an R

package that extends the ToxCast analysis pipeline package [17]. As a quality check (QC), positive control results were first normalized against those of their corresponding vehicle and analyzed. Data (end-points) that did not pass the positive control-based quality criteria were masked. Raw data that passed the QC were normalized to vehicle as:

$$\log_2 \left(\frac{i}{Veh} \right),$$

where *i* is the measured raw signal value of a well, and *Veh* is the median of the measured signal values for the vehicle wells on a plate.

Dose-response curves were fitted on the normalized data using three function families: constant, Hill, and gain-loss (see Fig. S3A) functions. Constraints on the monotonicity of the curve (i.e., increasing or decreasing) for each endpoint are defined in Table S2 (up = positive, down = negative). The best-fitting model of the three function families (minimizing the Akaike information criteria) was retained, and an MEC (see Fig. S3B), defined as the intersection of the fitted curve with the noise band (grey band in Fig. S3B), was calculated. The noise band was computed as three times the baseline median absolute deviation of vehicle responses. MEC values are not available for the constant model (the modeled activity never intersects the noise band).

MEC values were extracted for flavored and flavorless matrices. Finally, MEC ratios for flavored e-liquids against their corresponding base solution (the base solution tested in the same set of experiments) were computed as follows:

$$MEC \text{ ratio} = \frac{\text{mean}(\text{Base Solution MECs})}{\text{mean}(\text{Flavor MECs})}$$

Missing MEC values were replaced by the value for the maximum tested concentration (see Fig. S4A). When both the numerator and denominator were missing, no imputation was possible, and consequently, the ratio was not computed.

2.7. Phenotypic score calculation

Phenotypic Scores are toxicological indicator values computed for each flavoring substance or flavor mixture as the median MEC value across endpoints and time points. The Phenotypic Score expresses the overall impact of e-liquid exposures across the list of HCS endpoints tested. Phenotypic Score ratios were computed as the ratio between the Phenotypic Score of the base solution and that of the flavored e-liquid or individual flavoring substance.

2.8. Microarray data generation and processing

Total RNA was isolated from untreated NHBE cells or from NHBE cells exposed to a range of dilutions of the 28-flavor mixture or the base solution for 4 h or 24 h. Cells were lysed in RLT buffer (QIAGEN, Hilden, Germany), and RNA was extracted using an RNeasy Micro kit (QIAGEN, Hilden, Germany). For each exposure condition (no treatment, base solution, and 28-flavor mixture), four wells were pooled for RNA extraction. RNA concentration was measured using a NanoDrop ND1000 spectrophotometer (Thermo Fisher Scientific), and its quality was verified using an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). RNA integrity number values ranged from 7.2 to 10 (average 9.87). For mRNA analysis, 50 ng of total RNA were processed as described in the NuGEN Ovation RNA Amplification System V2 protocol (NuGEN, San Carlos, CA, USA). For hybridization, the Genechip® Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA, USA) was used. Three biological replicates (three independent experiments using the same e-liquid and base solution on different days) were collected for each concentration and exposure time tested. All RNA samples within each experiment were processed in the same batch. According to the experimental design, a block randomization for the RNA extraction step was performed using the exposure plates, exposure duration, and experimental replicate as blocking factors. Raw CEL files were background corrected, normalized, and summarized using frozen robust multiarray analysis [18]. Background correction and quantile normalization were used to generate gene expression values from all arrays passing QC checks using the custom Chip Description File environment HGU133-Plus2_Hs_ENTREZG v16.0 [19]. A log-intensities plot, normalized-unscaled standard error plot, relative log expression plot, and pseudo images were generated for QC checks using Bioconductor packages (AffyPLM; Bioconductor, Seattle, WA, USA) [20,21].

2.9. Gene expression analysis

For each treatment group and its associated control, a model for estimating the treatment effect was fitted using LIMMA [22]. The p -values for each estimated effect were adjusted across genes using the Benjamini-Hochberg false discovery rate (FDR) method. Differentially expressed genes (DEG) were defined as the set of genes for which the FDR was below 0.05 (see Table S3). The Affymetrix datasets are available from ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>), accession number E-MTAB-7748.

2.10. Biological network analysis

The NPA method was previously reported [10,23,24]. Briefly, the methodology aims to contextualize transcriptome profiles (flavored vs. base solution) by combining the alteration of gene expression into differential node values (i.e., one value for each node of a causal network model). The network models represent molecular mechanisms across a wide range of biological processes relevant to human respiratory physiology, including cell fate, cell stress, cell proliferation, and inflammation [25]. The networks used to evaluate the response of the NHBE cells to the flavored mixture and to the base solution exposure were all representative of biological mechanisms that can take place in NHBE cells. Relevant network models used for the analysis in this study are listed on

the left side of Fig. 6. The BIF integrates the various NPA scores in a measure of the overall biological impact across all network models. This BIF value allows for an assessment of the exposure effects in an objective, systematic, and quantifiable manner [11,12,24].³

3. Results

3.1. STEP 1: Tox-Score determination for a representative flavor mixture and individual flavoring substances

We used impedance-based RTCA to assess the toxic effects of the 28-flavor mixture compared with those of the corresponding base solution. A range of different concentrations (from 0.0625% to 4.0% v/v) was assessed for both conditions, and dose-response curves were generated to evaluate the effects of the added flavoring substances. As previously shown [4], both PG and VG mostly exert their effect by inducing a dose-dependent increase in osmolality, which appeared to be the main cause of cytotoxicity (Fig. S5). If the addition of flavoring agents did not increase the osmolality of the solution, the observed increase in cytotoxicity over that induced by the base solution would be then attributed to the flavoring agents.

As depicted in Fig. 2A, a statistically significant decrease in EC₅₀ compared with the base solution was found for the 28-flavor mixture ($p < 0.01$), leading to a calculated Tox-Score of 9.84 (see Table 2). This indicated that the addition of the flavor mixture enhanced the inherent toxicity of the base solution.

To understand the contribution of individual flavoring substances to the toxic effect of the mixture, each flavoring substance was assessed individually at the same concentration used in the mixture. All Tox-Scores are reported together with their corresponding p -value in Fig. 2B and Table 2. As a positive control, we also assessed diacetyl, a known toxic flavored substance, which gave a Tox-Score of 3.61. Citronellol and alpha-pinene had the highest Tox-Score values (2.95 and 2.07, respectively) and were thus identified as potential main contributors to the toxic effect of the flavor mixture. The remaining flavoring substances had Tox-Scores ≤ 1.6 .

To investigate whether citronellol and alpha-pinene did indeed contribute to the observed toxicity of the 28-flavor mixture, three new flavor mixtures were generated and assessed in NHBE cells using RTCA: (i) a flavor mixture without both alpha-pinene and citronellol, (ii) a flavor mixture without alpha-pinene, and (iii) a flavor mixture without citronellol. As shown in Fig. 2C, when citronellol was removed from the mixture, either alone (when compared with the 28-flavor mixture: $p < 0.01$) or with alpha-pinene (when compared with the 28-flavor mixture: $p < 0.05$), the toxic effect of the mixture decreased by more than 60% (see Table 2). Surprisingly, when only alpha-pinene was removed from the flavor mixture, the Tox-Score remained high (12.79), even above that of the 28-flavor mixture (9.84), but this was not statistically significant ($p > 0.3$). Additionally, the Tox-Scores of the citronellol-free mixture with or without alpha-pinene did not differ markedly, with values of 3.50 and 3.82, respectively. We also tested the impact of the base solution with both citronellol and alpha-pinene at the same concentration used in the 28-flavor mixture or in the individual flavoring substance assessment on NHBE cells. The Tox-Score values of citronellol alone (2.95) or citronellol together with alpha-pinene (3.09) were similar. Together this evidence suggests that, although cytotoxic when tested individually, addition of alpha-pinene to the flavor mixture did not influence the cytotoxicity of the tested mixtures. In fact, the Tox-Score did not decrease when the flavoring substance was removed. Finally, the flavor mixture lacking both alpha-pinene and citronellol induced a higher Tox-Score (3.82) than the mixtures containing the individual flavoring substances (≤ 1.6). This finding illustrates the

³ **Supplementary figures and tables** All supplementary figures and tables are available on figshare.com: doi:10.6084/m9.figshare.8010335

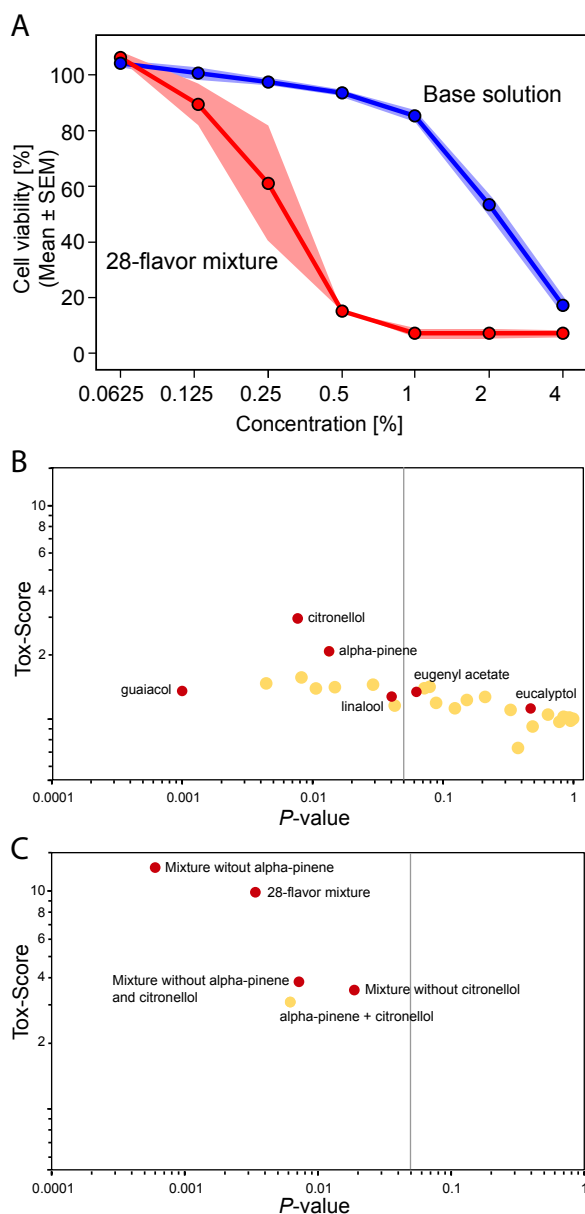


Fig. 2. (A) RTCA-based cell viability dose response curves for the 28-flavor mixture (red dots) and the corresponding base solution (blue dots). Green line corresponds to 50% of cell viability and dotted lines indicate extrapolated EC_{50} values. (B, C) Tox-Scores (y-axis) are shown as a function of their corresponding p -values computed for (B) each individual flavoring substance present in the 28-flavor mixture and (C) the flavor mixtures without citronellol, alpha-pinene, or both after a 24-hour exposure. Red dot in b and c indicate flavoring substances and mixtures selected for subsequent HCS investigation. The vertical line indicates a p -value of 0.05. Each dot corresponds to one flavor solution, with those selected for subsequent HCS-based investigation shown by white dots. Abbreviation: SEM, standard error of the mean.

potential interaction effects that can occur when mixing flavoring substances. Together, these results suggest that (i) citronellol is one of the main contributors to the toxicity of the 28-flavor mixture, (ii) alpha-pinene did not contribute to cytotoxicity when used in the context of the tested mixtures (other flavoring substances may mask or dampen its effect), and (iii) the interaction between flavoring substances within a mixture may also impact the final toxicity of the flavor mixture, making the assessment of mixtures even more challenging.

3.2. STEP 2: Phenotypic HCS results for the 28-flavor mixture and individual flavoring substances

To illustrate the second step of the flavoring substance assessment workflow depicted in Fig. 1, we selected the 28-flavor mixture, the two flavor ingredients with the highest Tox-Scores (alpha-pinene and citronellol), four other flavor ingredients with lower Tox-Scores (guaiacol, linalool, eugenyl acetate, and eucalyptol), and the three flavor mixtures in which alpha-pinene and citronellol were omitted individually or simultaneously (referred to as flavor mixture w/o alpha-pinene, flavor mixture w/o citronellol, and flavor mixture w/o alpha-pinene and citronellol). In addition, similar to the previous step, diacetyl was also included in the HCS assessment, and a phenotypic score of 4.36 was computed (Fig. S6).

NHBE cells were assessed after 30 min (only for the NF- κ B endpoint), 4 h, or 24 h of exposure for phenotypic changes using a battery of toxicity-related HCS endpoints to further characterize the flavor-specific effects compared with those of the base solution. The results of this comparison are shown in Fig. 3, with circular bar plots representing the MEC ratios (MEC base solution/MEC flavor mix) for each HCS endpoint (Fig. S4B). To further quantify and compare each flavor mixture, we calculated a Phenotypic Score (median MEC value across all HCS endpoints and time points for the base solution/median MEC value across all HCS endpoints and time points for the flavor mix), which, similarly to the Tox-Score, provides a quantification of the flavor's/flavors' additional effects when added to the base solution (Table 3).

When cells were exposed to the 28-flavor mixture, the first effect measured after 30 min of exposure was increased nuclear translocation of NF- κ B compared with the base solution, the potential downstream effects of which include the activation of inflammatory, cell survival, or apoptotic pathways [26,27]. After 4 h of exposure, the mitochondrial function of NHBE cells exposed to the flavor mixture was more impacted than that of cells exposed to the base solution, as evidenced by alterations in both mitochondrial mass and membrane potential. This effect was transient and not noted after the 24-h exposure. Signs of DNA damage, indicated by increases in H2AX phosphorylation (pH2AX), a marker of DNA double-strand breaks, were also observed following exposure to the flavor mixture at both time points. Oxidative stress (represented by the levels of reactive oxygen species [ROS] and intracellular glutathione [GSH]) increased in cells exposed to the 28-flavor mixture compared with cells exposed to the base solution for 24 h. Moreover, the detection of phosphorylated c-jun (p-c-jun), a marker of stress kinase pathway activation, in exposed NHBE cells as a marker of stress kinase pathway activation was in line with the observed increase in oxidative stress. Together, these effects may have led to both apoptotic and necrotic events, as suggested by an increase in cytochrome C release (after 24 h of exposure), caspase 3/7 activation (after 4 h of exposure), and cell membrane permeability (after 4 and 24 h of exposure), which correlated well with the high Tox-Score of the 28-flavor mixture obtained in the previous assessment step.

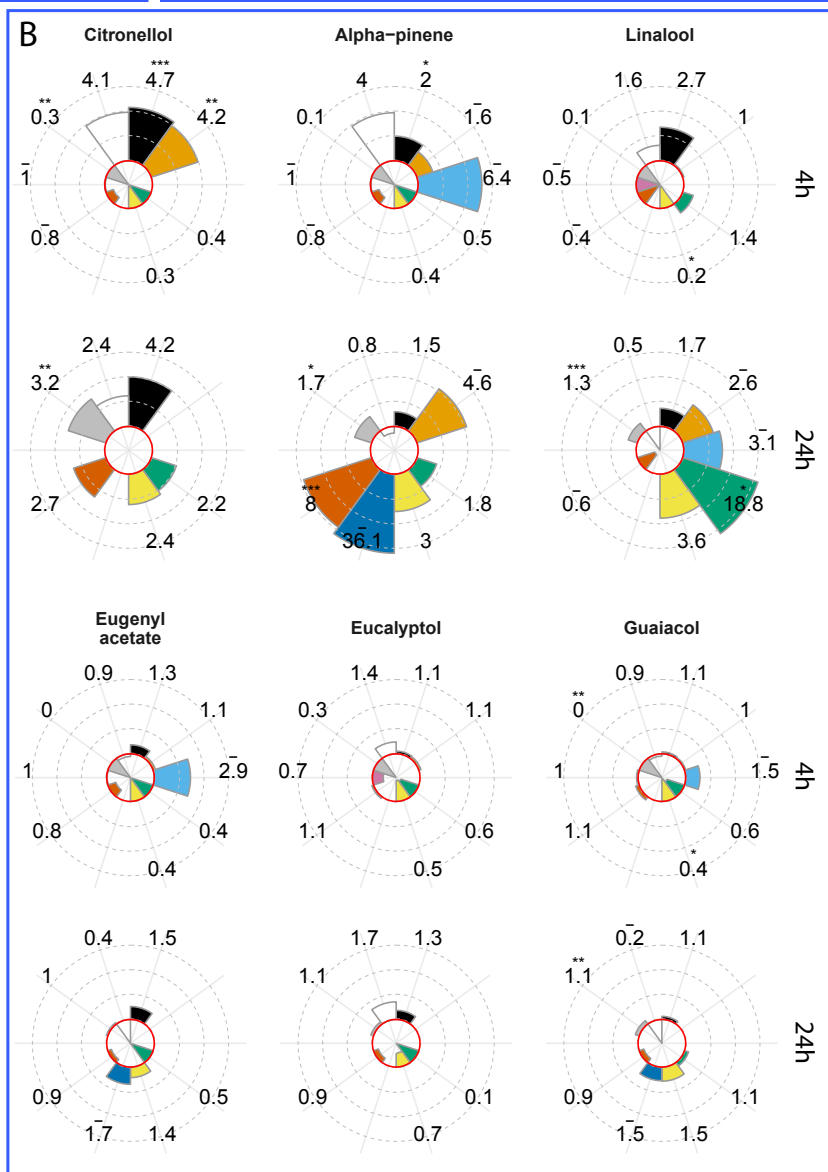
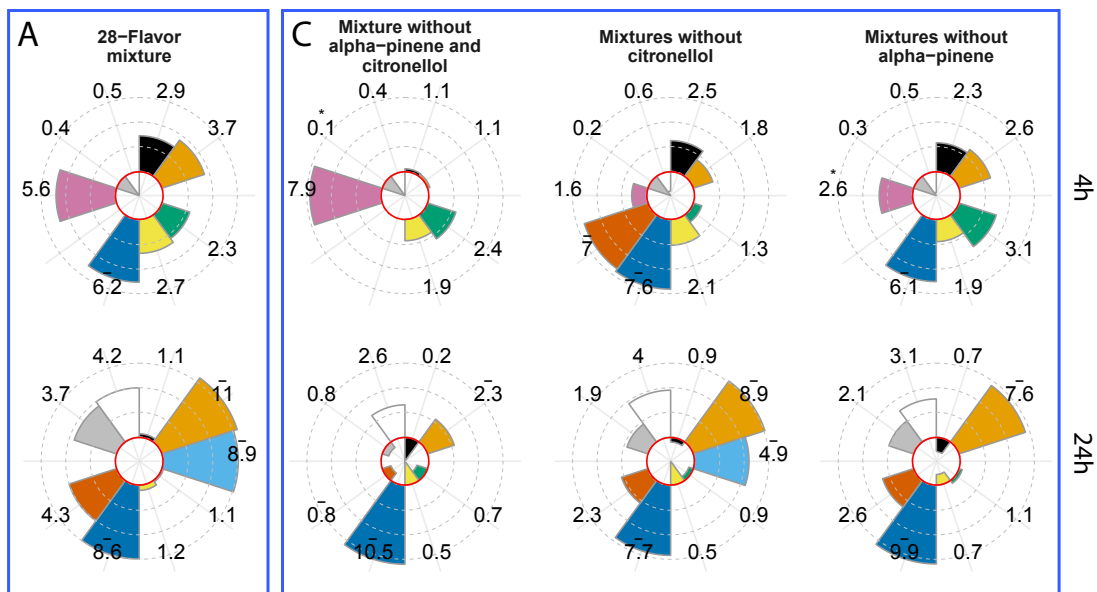
Exposure to the citronellol-only-flavored solution triggered some common effects compared with the exposure to the 28-flavor mixture (Fig. 3A). These included activation of caspase 3/7 and an increase in cell membrane permeability after 4 h of exposure or signs of oxidative stress (i.e., increased ROS, decreased GSH, and increased p-c-jun) after 24 h of exposure (Fig. 3B). Interestingly, these common phenotypic changes were not found in cells exposed to the flavor mixture without citronellol or without both alpha-pinene and citronellol (with the exception of ROS levels after 24 h of exposure) (Fig. 3C). These results suggest that citronellol is the main inducer of these effects in cells exposed to the 28-flavor mixture.

NHBE cells exposed to the alpha-pinene-only-flavored solution had increases in oxidative stress (indicated by increased ROS levels) and signs of apoptosis (supported by caspase 3/7 activation and cytochrome c release) after 4 h of exposure, which persisted after 24 h of exposure with greater DNA damage, increased cell membrane permeability, and

Table 2
Tox-Scores determined using RTCA for each exposure condition tested.

	Flavored solution	Tox-Score	LCL	UCL	N	P-value
Flavoring substances tested individually	2-acetylpyridine	1.191	0.937	1.513	3	0.0884
	2-acetylthiazole	1.469	1.315	1.641	3	0.0044
	Allyl hexanoate	1.568	1.314	1.869	3	0.0082
	Alpha-pinene	2.076	1.438	2.999	3	0.0134
	Butyric acid	0.731	0.222	2.405	3	0.3750
	L-carvone	1.002	0.485	2.092	3	0.9936
	Cinnamyl alcohol	1.024	0.653	1.608	3	0.8392
	Citronellol	2.948	1.726	5.037	4	0.0076
	Cyclotene	1.122	0.926	1.360	3	0.1232
	Ethyl acetate	1.049	0.720	1.529	3	0.6378
	2-ethyl-3,6-dimethylpyrazine	1.417	0.906	2.216	3	0.0787
	Ethyl formate	0.971	0.651	1.446	3	0.7779
	Ethyl maltol	1.228	0.831	1.815	3	0.1522
	Eucalyptol	1.119	0.648	1.932	3	0.4694
	Eugenyl acetate	1.340	0.963	1.865	3	0.0626
	Furaneol	1.104	0.792	1.539	3	0.3288
	Gamma-valerolactone	1.014	0.724	1.421	3	0.8714
	Guaiacol	1.352	1.297	1.409	3	0.0010
	Heptan-2-one	1.394	0.929	2.092	3	0.0719
	Isobutyl alcohol	0.922	0.610	1.393	3	0.4856
	Linalool	1.274	1.020	1.591	4	0.0404
	Menthone	1.016	0.523	1.975	3	0.9272
	Methyl anthranilate	1.448	1.097	1.910	3	0.0290
	3-methyl-2,4-nonanedione	1.390	1.200	1.610	3	0.0106
	Methyl salicylate	0.982	0.344	2.801	3	0.9481
	3-(methylthio) propionaldehyde	1.409	1.175	1.689	3	0.0148
4-(para-hydroxyphenyl)-butan-2-one	1.269	0.723	2.228	3	0.2096	
Phenethyl alcohol	1.156	1.012	1.320	3	0.0425	
Mixtures	28-flavor mixture	9.845	5.533	17.519	3	0.0034
	Flavor mixture w/o alpha-pinene	12.794	9.813	16.681	3	0.0006
	Flavor mixture w/o citronellol	3.501	1.654	7.410	3	0.0188
	Flavor mixture w/o citronellol and alpha-pinene	3.822	2.337	6.251	3	0.0072
	Citronellol and alpha-pinene	3.089	2.102	4.538	3	0.0062

Abbreviations: w/o, without; LCL, lower 95%-confidence limit; UCL, upper 95%-confidence limit; n, number of experimental repetitions.



Caspase 3/7 intensity
 Cytochrome C release
 Mitochondrial membrane potential
 Stress kinase (c-Jun)
 Glutathione content
 NF-kB nuclear content (30min)
 Oxidative stress (ROS)

Cell membrane permeability
 Mitochondrial mass
 DNA damage (pH2AX)

(caption on next page)

Fig. 3. Circular bar plots of HCS endpoint MEC ratios for (A) the 28-flavor mixture, (B) various individual flavoring substances with different Tox-Scores, and (C) flavor mixture without citronellol and/or alpha-pinene. Each MEC ratio (reported next to each segment of the circular chart for each HCS endpoint) was computed by dividing the mean base solution MEC (from n = 3 replicates) by the mean flavor mix MEC. A unilateral t-test was computed with null hypothesis: The base solution MEC mean is higher than the flavor mix MEC mean. The t-test p-values are reported as follows: *** < 0.001, ** < 0.01, and * < 0.05. The “-” sign on top of an MEC ratio denotes an imputed MEC value. Red circles correspond to an MEC of 1. Abbreviations: pH2AX, phosphorylated H2A histone family member X; NF-κB, nuclear factor kappa-light-chain-enhancer of activated B cells; ROS, reactive oxygen species.

Table 3
Phenotypic Scores and HCS endpoints counts for the six individual flavoring substances, 28-flavor mixture, and the three mixture variants that were evaluated using an HCS approach. Corresponding Tox-Scores are also shown.

Flavored solution	Phenotypic Score	HCS endpoints count	Tox-Score
Citronellol	3.16	12	2.95
28-flavor mixture	2.54	18	9.85
Alpha pinene	2.52	15	2.08
Mixtures w/o alpha-pinene	1.75	16	12.79
Mixtures w/o citronellol	1.44	19	3.50
Linalool	1.30	13	1.27
Mixtures w/o alpha-pinene & citronellol	1.14	14	3.82
Eucalyptol	1.09	14	1.12
Guaiacol	0.84	15	1.35
Eugenyl acetate	0.79	16	1.34

advanced activation of the stress kinase pathway (Fig. 3B). Many of these effects (e.g., activation of caspase 3/7 after 4 h of exposure and increased pH2AX, cell membrane permeability, and p-c-jun after 24 h of exposure) were also found when assessing the 28-flavor mixture (Fig. 3A). Interestingly, these common effects did not subside when cells were exposed to the flavor mixture without alpha-pinene (Fig. 3C). These results may explain why the flavor mixture without alpha-pinene did not show a lower Tox-Score than the 28-flavor mixture.

When comparing the computed Phenotypic Scores, we observed

that these solutions (mixture without alpha-pinene, 28-flavor mixture) were those with a high score, reflecting the Tox-Score observations (Table 3). Surprisingly, we found that citronellol alone had a higher impact than the 28-flavor mixture (3.16 and 2.54, respectively). However, it should be noted that a broader and therefore more complex cell response to the mixture was observed. In fact, when positive endpoint counts (endpoints for which a MEC was computed) are considered, the Phenotypic Score of the 28-flavor mixture is based on 18 endpoints, whereas only 12 endpoints were considered for the citronellol-only mixture (Table 3).

For the other individually tested flavoring substances, most of the obtained responses were smaller than those observed for citronellol and alpha-pinene (all Phenotypic Scores < 2.0) (Fig. 3B). Most of these results are in line with the low Tox-Scores found using the RTCA approach and could support the selection of a Tox-Score threshold as an indicator of responses detected by the HCS approach that are similar to those of the base solution. The only exceptions were the response to linalool and to the mixtures without both alpha-pinene and citronellol. Even though linalool had a low Tox-Score (1.27), its impact on NHBE cells in terms of HCS endpoints differed from that of the base solution. In fact, a Phenotypic Score of 1.30 was computed for linalool. Specifically, we noted activation of caspase 3/7 and an impact on mitochondrial health (indicated by changes in mitochondrial mass and membrane potential) after 4 h of exposure and increases in cell membrane permeability and cytochrome c release after 24 h of exposure. Conversely, the mixtures excluding both alpha-pinene and citronellol had a

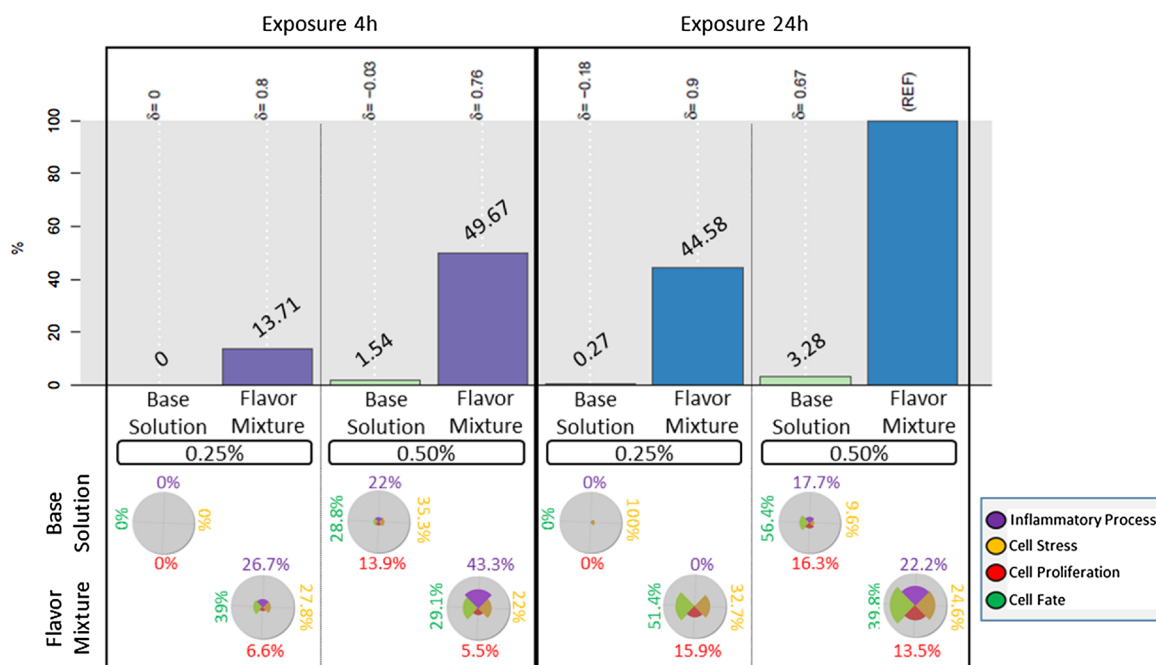


Fig. 4. Bar plots (upper panel) represent relative BIF (RBIF) values (indicated above each bar) for NHBE cells exposed to the 28 flavor mixture or base solution using Inflammatory Processes, Cell Stress, Cell Proliferation, and Cell Fate networks. The percentages indicate the relative biological impact derived from the cumulated network perturbations caused by the treatment relative to the reference (REF, defined as the treatment comparison showing the highest perturbation; REF = 100%). For each treatment comparison, the δ value (−1 to 1) indicates how similar the underlying network perturbations are with respect to the REF. A δ value of 0 indicates no similarity of the perturbed networks, a δ value of 1 indicates identical network perturbations, and a δ value of −1 indicates a completely opposite network responses. The star plots (lower panel) represent the contributions of each network family (Inflammatory Processes, Cell Stress, Cell Proliferation, and Cell Fate) to the RBIF for each treatment. The sum of the segment is equal to the BIF score for the treatment. Percentage values in black indicate the dilution tested.

very high Tox-Score and a low Phenotypic Score.

It is noteworthy that eugenyl acetate or guaiacol, which had a similar Tox-Scores (1.34 and 1.35, respectively) to linalool (1.27), did not show major changes in any HCS endpoint compared with the base solution at any of the time points investigated (Fig. 3B). These findings supported the notion that a functional assessment (e.g., using the panel of HCS endpoints) is useful in providing complementary data for assessing of the impact of flavoring substance exposure.

3.3. STEP 3: Biological impact of the 28-flavor mixture on the transcriptome of NHBE cells

The last step of the flavoring substances assessment workflow is based on a systems toxicology approach using computational network models to interpret transcriptomic data. This methodology provides a more detailed molecular understanding of biological network perturbations and also quantifies the molecular perturbations triggered by the exposure. As a case study, the effects of 4- and 24-h exposure to two dilutions of the 28-flavor mixture (0.25% and 0.50%) or to the base solution were investigated in NHBE cells in comparison with untreated cells (Fig. 4).

The highest RBIF value across all contrasts was used as REF and assigned an RBIF value of 100%. RBIF values in the other groups are expressed as a percentage of this maximum response. Overall, the RBIF values varied in dose- and time-dependent manners for both the 28-flavor mixture and the base solution. Indeed, the RBIF values were always higher at the highest dose and after 24 h of exposure.

The exposure to the base solution had a limited impact on NHBE cells, as indicated by the moderate fold changes and few DEGs compared with the untreated cell culture (Table S3). Adding the 28 flavors to the base solution increased the network perturbations, with the higher RBIF value (REF) observed after 24 h of exposure to the 0.50% dilution. Interestingly, all δ values for all flavor mixture conditions were positive and ≥ 0.76 . This suggests that the network perturbations were globally similar across all flavor mixture conditions assessed.

To gain a deeper insight into the exposure effect, RBIF values can be decomposed into their mechanistic components. Fig. 4 (lower panel) shows the biological processes that contributed most to the RBIF value at each time point for both the flavor mixture and base solution (versus negative control). The limited response to the exposure to the 0.50% base solution affected all four biological networks at both time points. Following the exposure to the flavor mixture, this effect was also observed at both time points and dilutions, with the exception of the exposure to the 0.25% dilution, which had no impact on the Inflammatory Processes Network after 24 h of exposure.

Each network can be further decomposed into several subnetworks describing more specific biological processes. The extent of perturbations in the subnetworks across all concentrations, time points, and exposure conditions tested is shown as a heatmap in Fig. 5.

The NPA scores for almost all subnetworks that were significantly perturbed by exposure to the base solution were very low (< 0.2) compared with the maximum effect observed for the flavor mixture. The most perturbed at 24 h was the osmotic stress subnetwork at the high base solution dilution. Measurement of the osmolality of this exposure condition indicated that it was still below the maximum physiological level of osmolality that the cells can tolerate (Fig. S5) before cell survival is affected. Even if the addition of the 28 flavors to the base solution induced a higher NPA score for the osmotic stress subnetwork than the base solution alone, the two statistical tests that assess the specificity of the response with respect to the biology described in the subnetwork were not significant.

Various cell stress subnetworks, such as the xenobiotic metabolism response, oxidative stress, hypoxic stress, and NFE2L2 signaling, were consistently and significantly perturbed in dose- and time-dependent manners when cells were exposed to the flavor mixture. This was also observed for the Jak Stat, Hox, Hedgehog, growth factor, cell cycle, and

calcium subnetworks and for the senescence, autophagy, and apoptosis subnetworks. Conversely, some subnetworks were only significantly perturbed after 24 h of exposure, including mTOR, epigenetics, and endoplasmic reticulum stress, or only at the 4-h time point, such as necroptosis and nuclear receptors (only for the 0.50% flavor mixture) and PGE2 and Mapk (for both dilutions). The clock subnetwork was found to be significantly perturbed only when the NHBE cells were exposed to the 0.50% flavor mixture, with a time-dependent increase in the NPA score. For both inflammatory process subnetworks, the response following the exposure to the flavor mixture was similar, with a significantly high NPA score after exposure to the 0.50% dilution at both time points and a low NPA score after exposure to the 0.25% dilution (significant only at the 4-h time point). Finally, no significant effect was observed on the response to DNA damage, Wnt, Notch, and cell interaction subnetworks for any of the exposure conditions tested.

4. Discussion

Although many smokers are switching to ENDS and e-cigarettes thinking that they might be a safer alternative to cigarettes due to their lack of tobacco combustion [28–31], their effects on the health of consumers are still unclear [32–34]. Indeed, differently from cigarettes, there are no epidemiological data on their impact after long-term exposure, simply because they have only been available on the market for 10 years. Another challenge concerning their safety evaluation is the huge number of variables characterizing these products, including flavor composition and concentration, the actual device, the device specificity (e.g., type of battery and/or atomizer), and nicotine content. More than 7,700 flavoring substances were reported as available online in January 2014, with 242 new flavoring substances reportedly being added each month [35]. Some of these flavoring substances, such as diacetyl and/or acetyl propionyl (2,3-pentanedione), have already been reported to be harmful when inhaled because of an association with bronchiolitis obliterans [36,37]. Thus, the impact of most of these flavoring substances, alone or in a mixture, is still unclear, and additional experimental evidence should be generated to support the regulatory framework. Identification of potentially harmful flavoring substances (and their toxic doses) or combinations of flavoring substances will allow e-liquid manufacturers to minimize or even eliminate these hazards from their products, thereby making them less harmful for consumers.

Our study demonstrated that the Tox-Score of the artificial 28-flavor mixture decreased dramatically when citronellol, one of the most toxic flavoring substances evaluated individually, was removed from the mixture. This shows that it is valuable to screen single flavoring substances, rank them based on their toxicity, so that manufacturers can develop and/or produce e-liquids with non-toxic flavor composition and doses. Our approach illustrated another point with the example of the flavor mixture that did not contain alpha-pinene and citronellol, the two most toxic flavoring substances (as indicated by their high Tox-Scores and Phenotypic Scores). This modified mixture contained flavoring substances that all had a low Tox-Score (≤ 1.6). However, when tested as a mixture, the Tox-Score obtained was 3.82, suggesting that even if individual flavoring substances lead to low or no cytotoxicity, the same may not necessarily be true when they are used as a mixture. Therefore, it is not sufficient to assess individual flavoring substances, and this makes the evaluation of the final flavor mixture highly beneficial to the assessment.

We also noted that linalool, when tested individually using a real-time impedance-based assessment, showed a relatively low Tox-Score (1.27). On the other hand, when assessed using a panel of HCS toxicity endpoints, its Phenotypic Score (1.30) indicated additional toxicity compared with that of the base solution. This observation could be explained by the fact that the Tox-Score is computed considering cell viability over a period of 24 h, whereas for the Phenotypic Score, only limited and specific time points were evaluated. In the case of extremely

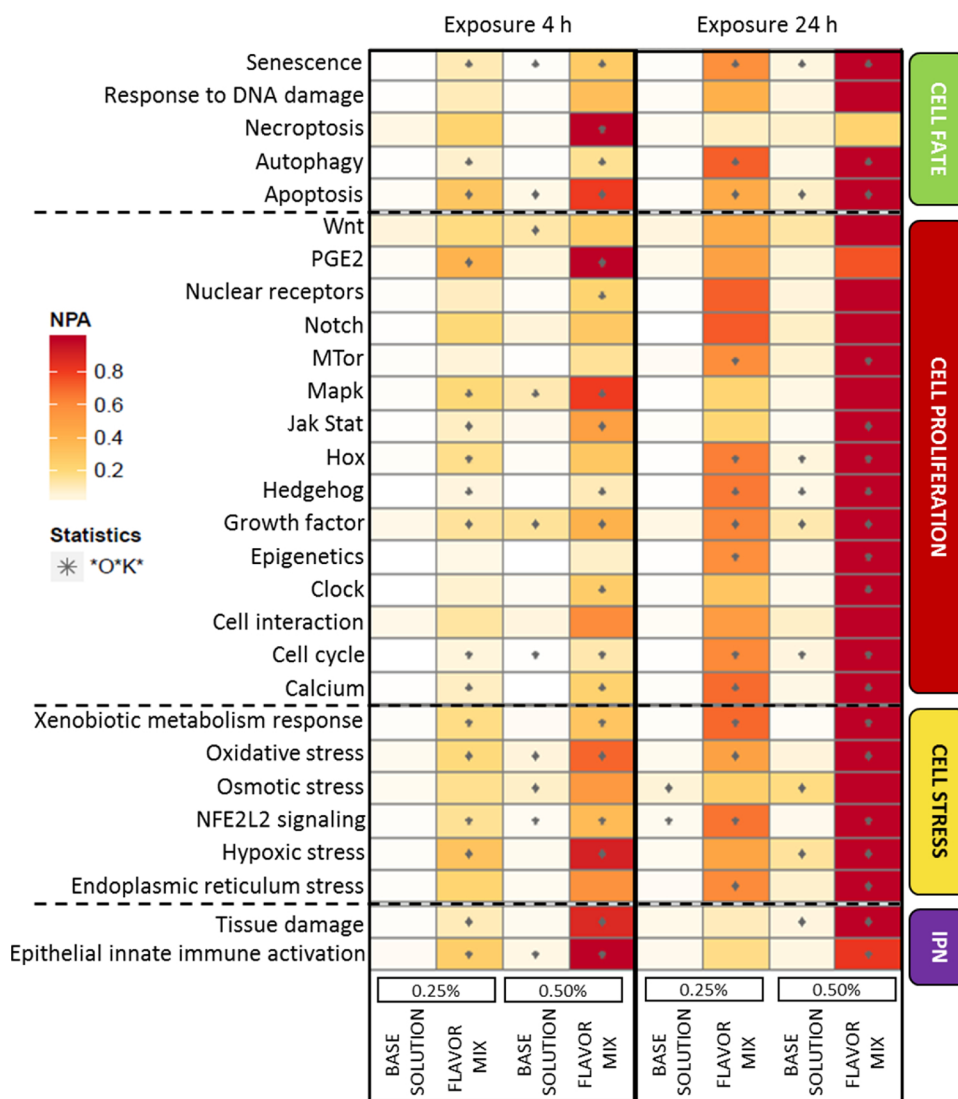


Fig. 5. NPA heatmap of the subnetworks impacted by the base solution alone or by the flavor mixture at two dilutions (0.25% and 0.50% v/v) at the 4- and 24-h time points. A network is considered perturbed if, in addition to the significance of the NPA score with respect to the experimental variation, the two companion statistics (O and K) that report the specificity of the NPA score with respect to the biology described in the network, are also significant (as indicated by an asterisk). The darker the color the stronger the perturbations. Abbreviation: IPN: Inflammatory Processes Network.

high cytotoxicity, this can become a limitation, as HCS endpoints may only be observable and quantifiable in a more limited time frame, which may require a time point adaptation. The potential discrepancy observed between the Tox-Score and the Phenotypic Score for particular mixture suggests that the impedance-based assessment is not sufficient to categorize some flavored solution as “non-toxic” compared with the base solution. It is therefore imperative to complement the RTCA approach with additional tests to understand what functional changes flavoring substances exposures can trigger in cells. The panel of HCS endpoints used here capture exposure effects associated with DNA damage, mitochondrial health, oxidative stress, or signs of apoptosis/necrosis/inflammation at different time points. Some of these endpoints can highlight early signs of cell stress responses linked to flavoring substances exposure that may not yet result in cytotoxicity and thus may not be measurable using the RTCA approach. As we are assessing the impact of acute exposure, it is of interest to investigate such early signs of cellular stress, because they could suggest that a particular flavoring substance can be more harmful when longer or chronic exposures occur.

In our previous article, we described the use of the first step of the assessment workflow (RTCA) to select flavoring substances for the second step (HCS), which is more time- and resource-consuming [5].

This approach requires the identification of an optimal selection criterion (e.g., a Tox-Score threshold), which presents an opportunity for researchers to determine a meaningful threshold following tests on a larger set of e-liquids. Then, only selected flavored solutions that are below this Tox-Score limit will be further assessed using the panel of HCS endpoints evaluated in the second step of the e-liquid assessment workflow to confirm their safety (at least in the context of acute exposure in NHBE cells).

Considering the diversity of different products on the market, it would be also interesting to assess whether the ranking of different flavoring substances based on Tox-Score and/or Phenotypic Score varies if the composition of the base solution (e.g., different PG/VG ratio or different concentration of nicotine) is changed.

The last step of this e-liquid assessment workflow is based on analysis of gene expression data. Using computable network models representing biological pathways relevant for lung cells, it is possible to quantify the impact of the exposure to flavored e-liquids and to obtain a mechanistic understanding of the effects occurring in the exposed cells. Indeed, we observed that adding the 28 flavoring substances to the base solution increased the network perturbations, as indicated by a higher BIF compared with the base solution alone. The detailed analysis of the network

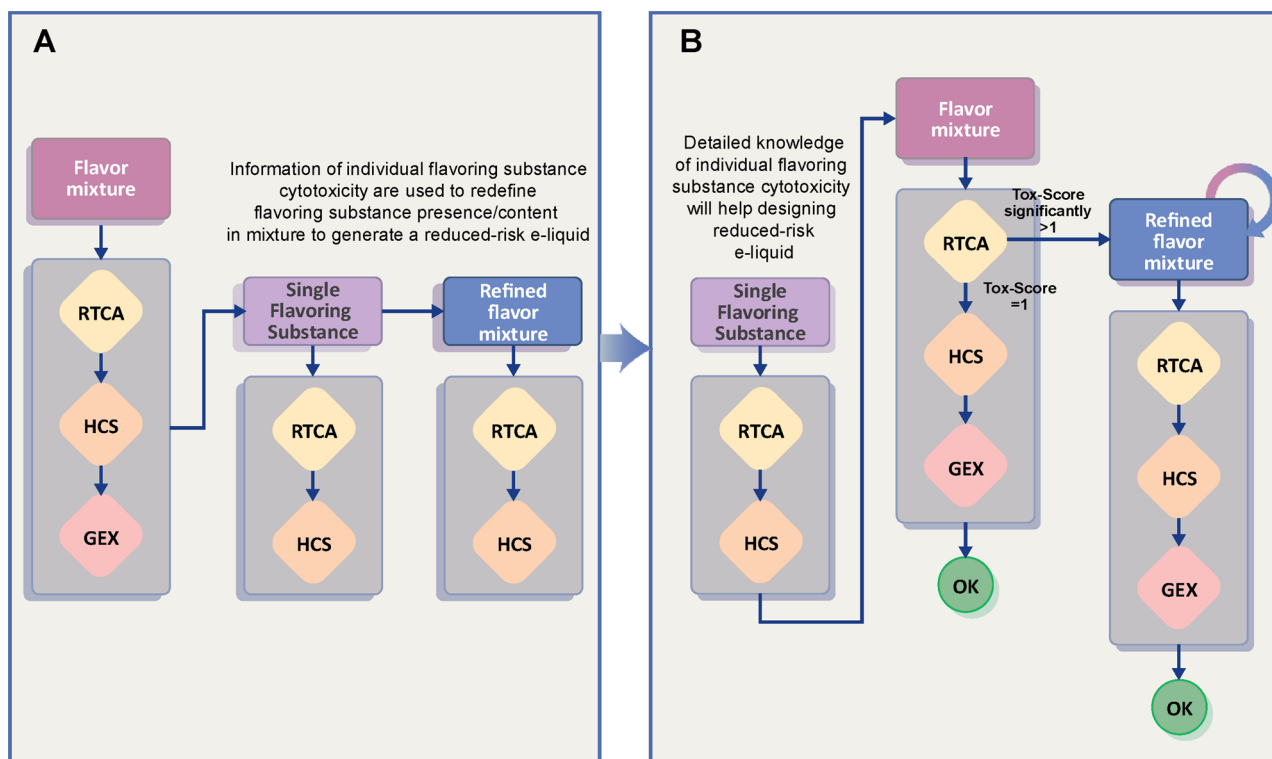


Fig. 6. Flavoring substances assessment workflow. This diagram shows how the flavor mixture assessment was performed in the present study (A) and how it will be readapted based on the key learnings (B). Abbreviations: RTCA, real-time cell analysis; HCS, high-content screening; GEX, gene expression.

perturbations at the subnetwork level was also performed and highlighted a dose- and time-dependent impact on various subnetworks (e.g., xenobiotic metabolism response, oxidative stress, or apoptosis subnetworks) specific to exposure to the 28-flavor mixture. Network analysis represents a sensitive tool that can highlight particular pathways perturbed by this exposure and could be used to investigate and develop lung toxicity biomarkers relevant for exposure to recently highlighted flavoring agents [38]. Although this systems toxicology approach may not be suitable for large-scale screening of flavoring substances, it offers a global view of the cellular changes that flavoring substances exposure can induce, thus constituting an important complement to the regulatory requirement for safety toxicological assessment.

Given the great numbers of flavoring substances used in e-liquids today, it is essential to develop new safety assessment strategies that can provide fast, cost-effective, accurate, and relevant toxicological data. As recently mentioned, the popularity of ENDS and their large diversity requires re-thinking of the more traditional ways of performing toxicological studies (e.g., traditional animal testing or post-market surveillance) [39]. Alternative methods based on high-throughput approaches and state-of-the-art technologies should be further developed, implemented, and validated. The approach that we are presenting here has great potential to improve our knowledge of the toxicity of flavoring substances and e-liquids used in ENDS or e-cigarettes. However, our results should be considered in the context of the applied experiments (i.e., in our case, an acute exposure of submerged cells derived from one particular donor). The choice of using NHBE cells as a test system has the following advantages for our described screening method: (1) These cells are appropriate for a high-throughput approach; (2) As primary cells, their response to the exposure might be more representative than those of immortalized cell lines; (3) Their human origin avoids species-translatibility issues; (4) They represent one of the major cell types that compose the human airway epithelium; and (5) Their use is in line with the 3Rs strategy to reduce, refine, and replace animals in research. Still, NHBE cells do not fully capture the complexity of the highly differentiated bronchial epithelium, which is composed of specialized cell types such as goblet, ciliated, and basal cells. Thus, as a second step of

assessment of flavoring substances, the impact of aerosolization on the chemical composition of flavored e-liquids has to be addressed by using organotypic tissue culture, which is a better model of the human airway epithelium [40], as explained in our previous publications [5,41,42].

Recently, Sassano et al. described a three-phase, 384-well, plate-based, high-throughput screening assay to rapidly screen multiple e-liquids [43]. Their approach was combined with the development of a publicly available, searchable website to disseminate their findings and allow better use of the large toxicological dataset they are producing. Like Sassano et al., we believe that data sharing is essential, especially for this timely and sensitive topic [44]. Being transparent and allowing stakeholders like the scientific community, industry, funding and regulatory agencies, non-profit entities, and the public to access raw data and information on the study design, data processing, and analysis is essential and will clearly benefit the scientific community and society at large. Indeed, sharing data will not only enable independent verification, re-analysis, and reuse but will also foster discussion to improve and potentially establish standardized methods for the toxicological assessment of inhaled flavoring substances. Using INTERVALS (<https://www.intervals.science>), an online platform developed by Philip Morris International Research and Development, all data collected during this study are made publicly available (doi: 10.26126/intervals.lwo6mb.1).

5. Conclusion

The approach described in this paper is based on three pillars: (1) RTCA, (2) evaluation of a panel of phenotypic HCS endpoints, and (3) gene expression analysis. For each pillar, computationally derived scores were developed and used to quantify (i) the toxicity (Tox-Score), (ii) the phenotypic impact (Phenotypic Score), and (iii) the transcriptomic/mechanistic effect (BIF/NPA) of an exposure. Using an artificial mixture of 28 flavoring substances dissolved in a base solution, we showed that this method is suitable for identifying candidate constituents that account for the toxicity of a mixture (like citronellol in our study) (Fig. 6A). Based on these findings, we aimed to decrease the

cytotoxicity of the 28-flavor mixture by omitting the two flavoring substances that appeared to have the highest cytotoxic effect when assessed individually. On one hand, the results for citronellol indicated that compound toxicity was maintained when it was added to the 28-flavor mixture, while on the other hand, alpha-pinene did not appear to substantially contribute to the cytotoxicity of the 28-flavor mixture. Additionally, and not unexpectedly, we found that the remaining flavoring substances, although scoring low when assessed individually, induced a significant increase in cytotoxicity when combined in a mixture. These findings clearly demonstrate that assessment of individual substances can only be partially informative of their potential contribution in a more complex mixture, as synergistic and antagonistic interactions may occur.

For this reason, we suggest a revised workflow (Fig. 6B) based on an initial single flavoring substance screen aimed at supporting any flavor mixture design. Assessing individual flavoring substances over a broad range of concentrations, as used in e-liquids, will help to define safe levels (i.e., the highest concentration at which no cytotoxic effect is observed). Once this data is known, a candidate minimal-risk mixture can be designed. To fully ensure that the generated mixture does not elicit additional cytotoxicity when compared with its corresponding flavor-free base solution, cytotoxicity assessments must be performed (pillar 1). If no cytotoxicity is found (threshold limit still to be defined), further evidence can be collected by completing the three-pillar assessment. However, if the mixture shows increased cytotoxicity, further refinement will be required to identify which flavoring substance, or combination of flavoring substances, is responsible for the observed effect. To do so, we propose to perform an iterative method based on a single flavoring substance exclusion strategy. With this design, each flavoring substance is individually removed, leading to the generation of a number of new mixtures equivalent to the number of flavoring substances. By assessing all of these mixtures, it should be possible to further identify which flavoring substance, or combination of flavoring substances, contributes to the increase in cytotoxicity. Once this is understood and a reduced risk mixture is obtained that still has an acceptable taste for the consumer, the assessment is then completed with the remaining pillars to further confirm the absence of undesired effects.

Authors contribution

D.M., C.M., I.G.S., J.H., F.M., P.L. and D.S. contributed to the conception and design of research.

D.M. and C.M. performed the data interpretation and drafted the manuscript.

V.B., P.L. and F.M. computed and analyzed the data.

S.A. and L.O.T. performed the *in vitro* experiments.

R.D., D.B., D.P., E.G. and N.V.I. generated the gene expression dataset.

M.B., M.F., E.F. and F.F. prepared the tested e-liquid solutions.

M.C.P., F.M. and J.H. edited and revised the manuscript.

M.C.P. and J.H. approved final version of manuscript.

Acknowledgements

Alain Sewer for the transcriptomic dataset submission to Array Express. Karsta Luettich, Anatoly Mazurov, and Sandra Schorderet Weber for scientific review of the manuscript. Karine Baumer for technical support. Elena Scotti, Gopala Reddy Sindhoora Bhargavi and Nicholas Karoglou for editorial review of the manuscript. Samantha Elmhurst for artwork support. Philip Morris International is the sole source of funding and sponsor of this research.

Conflict of Interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.toxrep.2019.11.016>.

References

- [1] R. Smith, S.M. Cohen, J. Doull, V. Feron, J. Goodman, L. Marnett, T. Adams, GRAS flavoring substances 21, *Food Technol.* 57 (2003) 46–59.
- [2] J. Hallagan, The Safety Assessment and Regulatory Authority to Use Flavors - Focus on E-Cigarettes, Retrieved from (2015) www.femaflavor.org.
- [3] M. Balls, ATLA (Alternatives to Laboratory Animals): past, present and future, *Altern. Lab. Anim.* 38 (6) (2010) 437–441, <https://doi.org/10.1177/026119291003800606>.
- [4] I. Gonzalez-Suarez, D. Marescotti, F. Martin, E. Scotti, E. Guedj, S. Acali, M.C. Peitsch, In vitro systems toxicology assessment of nonflavored e-Cigarette liquids in primary lung epithelial cells, *Appl. In Vitro Toxicol.* 3 (1) (2017) 41–55, <https://doi.org/10.1089/aivt.2016.0040>.
- [5] A.R. Iskandar, I. Gonzalez-Suarez, S. Majeed, D. Marescotti, A. Sewer, Y. Xiang, J. Hoeng, A framework for in vitro systems toxicology assessment of e-liquids, *Toxicol. Mech. Methods* 26 (6) (2016) 389–413, <https://doi.org/10.3109/15376516.2016.1170251>.
- [6] D. Krewski, D. Acosta Jr., M. Andersen, H. Anderson, J.C. Bailar 3rd, K. Boelkelheide, L. Zeise, Toxicity testing in the 21st century: a vision and a strategy, *J. Toxicol. Environ. Health B Crit. Rev.* 13 (2–4) (2010) 51–138, <https://doi.org/10.1080/10937404.2010.483176>.
- [7] T. Le Godec, I. Crooks, K. Scott, C. Meredith, In vitro mutagenicity of gas-vapour phase extracts from flavoured and unflavoured heated tobacco products, *Toxicol. Rep.* 6 (2019) 1155–1163, <https://doi.org/10.1016/j.toxrep.2019.10.007>.
- [8] K. Fowler, W. Fields, V. Hargreaves, L. Reeve, B. Bombick, Development, qualification, validation and application of the Ames test using a VITROCELL(R) VC10(R) smoke exposure system, *Toxicol. Rep.* 5 (2018) 542–551, <https://doi.org/10.1016/j.toxrep.2018.04.003>.
- [9] B.M. Keyser, R. Leverette, M. Hollings, A. Seymour, L. Reeve, W. Fields, Investigation of multiple whole smoke dosimetry techniques using a vitrocell® VC10® smoke exposure system, *Toxicology Reports* (2019), <https://doi.org/10.1016/j.toxrep.2019.10.011>.
- [10] F. Martin, A. Sewer, M. Taliikka, Y. Xiang, J. Hoeng, M.C. Peitsch, Quantification of biological network perturbations for mechanistic insight and diagnostics using two-layer causal models, *BMC Bioinformatics* 15 (2014) 238, <https://doi.org/10.1186/1471-2105-15-238>.
- [11] F. Martin, T.M. Thomson, A. Sewer, D.A. Drubin, C. Mathis, D. Weisensee, M.C. Peitsch, Assessment of network perturbation amplitudes by applying high-throughput data to causal biological networks, *BMC Syst. Biol.* 6 (2012) 54, <https://doi.org/10.1186/1752-0509-6-54>.
- [12] T.M. Thomson, A. Sewer, F. Martin, V. Belcastro, B.P. Frushour, S. Gebel, M.C. Peitsch, Quantitative assessment of biological impact using transcriptomic data and mechanistic network models, *Toxicol. Appl. Pharmacol.* 272 (3) (2013) 863–878, <https://doi.org/10.1016/j.taap.2013.07.007>.
- [13] EFSA, Scientific Opinion - Guidance on the data required for the risk assessment of flavourings to be used in or on foods, *Efsa J.* 8 (6) (2010) 1623, <https://doi.org/10.2903/j.efsa.2010.1623>.
- [14] P.A. Tierney, C.D. Karpinski, J.E. Brown, W. Luo, J.F. Pankow, Flavour chemicals in electronic cigarette fluids, *Tob. Control* 25 (e1) (2016) e10–15, <https://doi.org/10.1136/tobaccocontrol-2014-052175>.
- [15] D. Marescotti, I. Gonzalez Suarez, S. Acali, S. Johnne, A. Laurent, S. Frentzel, M.C. Peitsch, High content screening analysis to evaluate the toxicological effects of harmful and potentially harmful constituents (HPHC), *J. Vis. Exp.* 111 (2016), <https://doi.org/10.3791/53987>.
- [16] V. Belcastro, S. Cano, D. Marescotti, S. Acali, I. Gonzalez Suarez, F. Martin, J. Hoeng, GladiaTOX: GLobal assessment of Dose-IndicAtor in TOXicology, *Bioinformatics* (2019).
- [17] D.L. Filer, P. Kothiyi, R.W. Setzer, R.S. Judson, M.T. Martin, Tcpl: the ToxCast pipeline for high-throughput screening data, *Bioinformatics* 33 (4) (2017) 618–620, <https://doi.org/10.1093/bioinformatics/btw680>.
- [18] M.N. McCall, B.M. Bolstad, R.A. Irizarry, Frozen robust multiarray analysis (fRMA), *Biostatistics* 11 (2) (2010) 242–253, <https://doi.org/10.1093/biostatistics/kxp059>.
- [19] M. Dai, P. Wang, A.D. Boyd, G. Kostov, B. Athey, E.G. Jones, F. Meng, Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data, *Nucleic Acids Res.* 33 (20) (2005) e175, <https://doi.org/10.1093/nar/gni179>.
- [20] B.M. Bolstad, R.A. Irizarry, M. Astrand, T.P. Speed, A comparison of normalization methods for high density oligonucleotide array data based on variance and bias, *Bioinformatics* 19 (2) (2003) 185–193, <https://doi.org/10.1093/bioinformatics/19.2.185>.
- [21] L. Gautier, L. Cope, B.M. Bolstad, R.A. Irizarry, Affy-analysis of Affymetrix GeneChip data at the probe level, *Bioinformatics* 20 (3) (2004) 307–315, <https://doi.org/10.1093/bioinformatics/btg405>.
- [22] G.K. Smyth, Linear models and empirical bayes methods for assessing differential expression in microarray experiments, *Stat. Appl. Genet. Mol. Biol.* 3 (2004), <https://doi.org/10.2202/1544-6115.1027> Article3.

- [23] J. Hoeng, R. Deehan, D. Pratt, F. Martin, A. Sewer, T.M. Thomson, M.C. Peitsch, A network-based approach to quantifying the impact of biologically active substances, *Drug Discov. Today* 17 (9-10) (2012) 413–418, <https://doi.org/10.1016/j.drudis.2011.11.008>.
- [24] J. Hoeng, M. Talikka, F. Martin, A. Sewer, X. Yang, A. Iskandar, M.C. Peitsch, Case study: the role of mechanistic network models in systems toxicology, *Drug Discov. Today* 19 (2) (2014) 183–192, <https://doi.org/10.1016/j.drudis.2013.07.023>.
- [25] S. Boue, M. Talikka, J.W. Westra, W. Hayes, A. Di Fabio, J. Park, J. Hoeng, Causal biological network database: a comprehensive platform of causal biological network models focused on the pulmonary and vascular systems, *Database (Oxford)*, 2015 (2015), <https://doi.org/10.1093/database/bav030>.
- [26] Y. Fan, J. Dutta, N. Gupta, G. Fan, C. Gelinas, Regulation of programmed cell death by NF-kappaB and its role in tumorigenesis and therapy, *Adv. Exp. Med. Biol.* 615 (2008) 223–250, https://doi.org/10.1007/978-1-4020-6554-5_11.
- [27] R. Piva, G. Belardo, M.G. Santoro, NF-kappaB: a stress-regulated switch for cell survival, *Antioxid. Redox Signal.* 8 (3-4) (2006) 478–486, <https://doi.org/10.1089/ars.2006.8.478>.
- [28] D. Azzopardi, K. Patel, T. Jaunky, S. Santopietro, O.M. Camacho, J. McAughey, M. Gaca, Electronic cigarette aerosol induces significantly less cytotoxicity than tobacco smoke, *Toxicol. Mech. Methods* 26 (6) (2016) 477–491, <https://doi.org/10.1080/15376516.2016.1217112>.
- [29] M.L. Goniewicz, J. Knysak, M. Gawron, L. Kosmider, A. Sobczak, J. Kurek, N. Benowitz, Levels of selected carcinogens and toxicants in vapour from electronic cigarettes, *Tob. Control* 23 (2) (2014) 133–139, <https://doi.org/10.1136/tobaccocontrol-2012-050859>.
- [30] G. Romagna, E. Alliffranchini, E. Bocchietto, S. Todeschi, M. Esposito, K.E. Farsalinos, Cytotoxicity evaluation of electronic cigarette vapor extract on cultured mammalian fibroblasts (ClearStream-LIFE): comparison with tobacco cigarette smoke extract, *Inhal. Toxicol.* 25 (6) (2013) 354–361, <https://doi.org/10.3109/08958378.2013.793439>.
- [31] M. Taylor, T. Carr, O. Oke, T. Jaunky, D. Breheny, F. Lowe, M. Gaca, E-cigarette aerosols induce lower oxidative stress in vitro when compared to tobacco smoke, *Toxicol. Mech. Methods* 26 (6) (2016) 465–476, <https://doi.org/10.1080/15376516.2016.1222473>.
- [32] F. Cervellati, X.M. Muresan, C. Sticozzi, R. Gambari, G. Montagner, H.J. Forman, G. Valacchi, Comparative effects between electronic and cigarette smoke in human keratinocytes and epithelial lung cells, *Toxicol. In Vitro* 28 (5) (2014) 999–1005, <https://doi.org/10.1016/j.tiv.2014.04.012>.
- [33] J.L. Fetterman, R.M. Weisbrod, B. Feng, R. Bastin, S.T. Tuttle, M. Holbrook, N.M. Hamburg, Flavorings in tobacco products induce endothelial cell dysfunction, *Arterioscler. Thromb. Vasc. Biol.* 38 (7) (2018) 1607–1615, <https://doi.org/10.1161/ATVBAHA.118.311156>.
- [34] C.F. McDonald, E-cigarettes for smoking cessation: current state of play, *Respirology* (2019), <https://doi.org/10.1111/resp.13630>.
- [35] S.H. Zhu, J.Y. Sun, E. Bonnevie, S.E. Cummins, A. Gamst, L. Yin, M. Lee, Four hundred and sixty brands of e-cigarettes and counting: implications for product regulation, *Tob. Control* 23 (Suppl 3) (2014), <https://doi.org/10.1136/tobaccocontrol-2014-051670> iii3-9.
- [36] K.E. Farsalinos, K.A. Kistler, G. Gillman, V. Voudris, Evaluation of electronic cigarette liquids and aerosol for the presence of selected inhalation toxins, *Nicotine Tob. Res.* 17 (2) (2015) 168–174, <https://doi.org/10.1093/ntr/ntu176>.
- [37] H.R. Park, M. O'Sullivan, J. Vallarino, M. Shumyatcher, B.E. Himes, J.A. Park, Q. Lu, Transcriptomic response of primary human airway epithelial cells to flavoring chemicals in electronic cigarettes, *Sci. Rep.* 9 (1) (2019) 1400, <https://doi.org/10.1038/s41598-018-37913-9>.
- [38] G. Kaur, T. Muthumalage, I. Rahman, Mechanisms of toxicity and biomarkers of flavoring and flavor enhancing chemicals in emerging tobacco and non-tobacco products, *Toxicol. Lett.* 288 (2018) 143–155, <https://doi.org/10.1016/j.toxlet.2018.02.025>.
- [39] T. Hartung, E-cigarettes and the need and opportunities for alternatives to animal testing, *ALTEX* 33 (3) (2016) 211–224, <https://doi.org/10.14573/altex.1606291>.
- [40] C. Mathis, C. Poussin, D. Weisensee, S. Gebel, A. Hengstermann, A. Sewer, M.C. Peitsch, Human bronchial epithelial cells exposed in vitro to cigarette smoke at the air-liquid interface resemble bronchial epithelium from human smokers, *Am. J. Physiol. Lung Cell Mol. Physiol.* 304 (7) (2013) L489–503, <https://doi.org/10.1152/ajplung.00181.2012>.
- [41] W.A. Hayes, R. Li, J. Hoeng, A. Iskandar, M.C. Peitsch, M.L. Dourson, New approaches to risk assessment of chemical mixtures, *Toxicol. Res. Appl.* 3 (2019) 1–10.
- [42] A.R. Iskandar, F. Zanetti, A. Kondylis, F. Martin, P. Leroy, S. Majeed, J. Hoeng, A lower impact of an acute exposure to electronic cigarette aerosols than to cigarette smoke in human organotypic buccal and small airway cultures was demonstrated using systems toxicology assessment, *Intern. Emerg. Med.* (2019), <https://doi.org/10.1007/s11739-019-02055-x>.
- [43] M.F. Sassano, E.S. Davis, J.E. Keating, B.T. Zorn, T.K. Kochar, M.C. Wolfgang, R. Tarran, Evaluation of e-liquid toxicity using an open-source high-throughput screening assay, *PLoS Biol.* 16 (3) (2018) e2003904, <https://doi.org/10.1371/journal.pbio.2003904>.
- [44] S. Boue, M. Byrne, A.W. Hayes, J. Hoeng, M.C. Peitsch, Embracing transparency through data sharing, *Int. J. Toxicol.* (2018), <https://doi.org/10.1177/1091581818803880> 1091581818803880.