


Evaluating Sufficient Similarity of Botanical Dietary Supplements: Combining Chemical and *In Vitro* Biological Data

Kristen R. Ryan,^{*,1} Madelyn C. Huang,^{*,1} Stephen S. Ferguson ,^{*} Suramya Waidyanatha,^{*} Sreenivasa Ramaiahgari,^{*} Julie R. Rice,^{*} Paul E. Dunlap,^{*} Scott S. Auerbach,^{*} Esra Mutlu,^{*} Tim Cristy,[†] Jessica Peirfelice,[†] Michael J. DeVito,^{*} Stephanie L. Smith-Roe,^{*} and Cynthia V. Rider^{*,2}

^{*}Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709; and [†]Battelle, Columbus, Ohio 43201

¹These authors contributed equally to this study.

²To whom correspondence should be addressed at Cynthia Rider at National Institute of Environmental Health Sciences, PO Box 12233, K2-12, Research Triangle Park, NC 27709. Fax: 919-541-1019; E-mail: cynthia.rider@nih.gov.

ABSTRACT

Botanical dietary supplements are complex mixtures with numerous potential sources of variation along the supply chain from raw plant material to the market. Approaches for determining sufficient similarity (ie, complex mixture read-across) may be required to extrapolate efficacy or safety data from a tested sample to other products containing the botanical ingredient(s) of interest. In this work, screening-level approaches for generating both chemical and biological-response profiles were used to evaluate the similarity of black cohosh (*Actaea racemosa*) and *Echinacea purpurea* samples to well-characterized National Toxicology Program (NTP) test articles. Data from nontargeted chemical analyses and gene expression of toxicologically important hepatic receptor pathways (aryl hydrocarbon receptor [AhR], constitutive androstane receptor [CAR], pregnane X receptor [PXR], farnesoid X receptor [FXR], and peroxisome proliferator-activated receptor alpha [PPAR α]) in primary human hepatocyte cultures were used to determine similarity through hierarchical clustering. Although there were differences in chemical profiles across black cohosh samples, these differences were not reflected in the biological-response profiles. These findings highlight the complexity of biological-response dynamics that may not be reflected in chemical composition profiles. Thus, biological-response data could be used as the primary basis for determining similarity among black cohosh samples. Samples of *E. purpurea* displayed better correlation in similarity across chemical and biological-response measures. The general approaches described herein can be applied to complex mixtures with unidentified active constituents to determine when data from a tested mixture (eg, NTP test article) can be used for hazard identification of sufficiently similar mixtures, with the knowledge of toxicological targets informing assay selection when possible.

Key words: sufficient similarity; complex mixtures; herbal supplements; black cohosh; *Echinacea purpurea*.

Botanical dietary supplements are widely used in the United States, with approximately 18% of people reportedly taking non-vitamin, nonmineral dietary supplements (Clarke *et al.*, 2015). The botanical industry continues to expand, with over 26 000

botanical ingredient products listed (NIH, 2019) and sales surpassing \$8 billion in the United States in 2017 (Smith *et al.*, 2018). Despite this growth, there is very little safety information for many botanical products. Furthermore, the complexity of

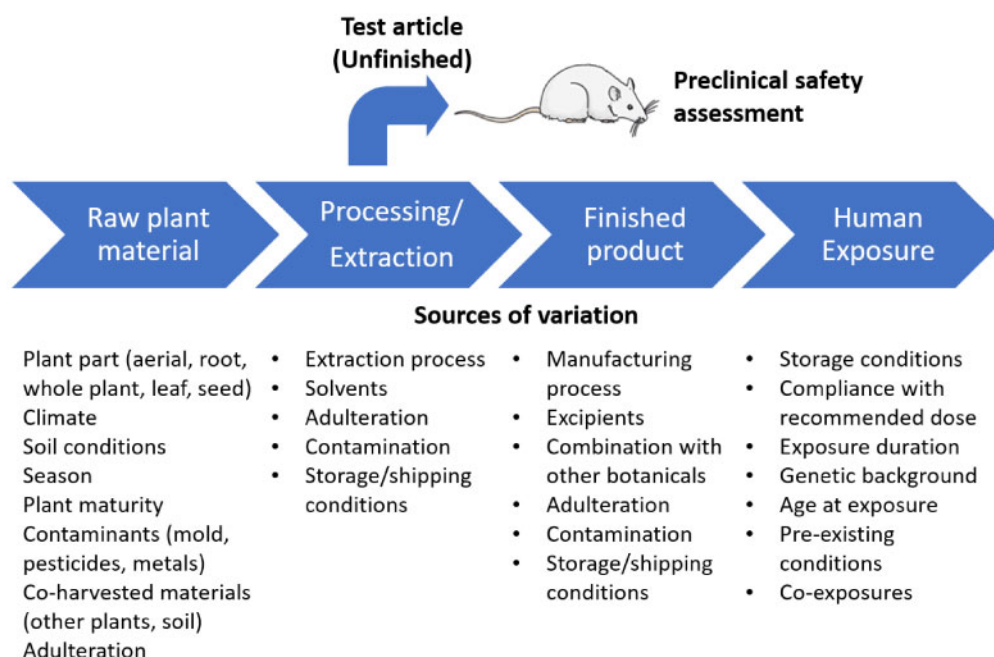


Figure 1. Potential sources of variation in botanical dietary supplements. In addition, this figure highlights that most preclinical safety evaluations (ie, in rodents) of botanicals are conducted with unfinished materials rather than formulated products.

botanical dietary supplement products presents a significant challenge in understanding their safety (Shipkowski et al., 2018).

Botanicals have a diverse array of chemical constituents, and they often contain a proportionally high unidentified fraction. In many cases, the biologically active constituents are unknown. Additionally, there are numerous sources of variation in marketed botanical products, which are compounded across the supply chain from raw material to the consumer (Figure 1). Because a comprehensive assessment of toxicity using repeat-dose animal testing is resource- and time-intensive, it is not feasible to conduct safety studies for each botanical product on the market. Furthermore, a large volume of material is required for animal studies, therefore, unfinished product (bulk material that serves as the source for finished products) is often used rather than the finished products (eg, commercially available tablet or capsule) that people typically consume. The unexplored assumption is that the differences across products and between finished and unfinished samples are not great enough to warrant toxicity testing of each individual product (Figure 1).

Similarity in both chemical composition and biological response is important in understanding how representative a selected test article is and if data generated on the test article are applicable to other products. This concept, termed “sufficient similarity,” is derived from risk assessment guidance for complex mixtures (USEPA, 2000). Sufficient similarity refers to a determination that 2 mixtures (eg, related botanical products) are similar enough in composition that toxicity data from one can be used to estimate the risk associated with the other, ie, a read-across approach for complex mixtures. Previously, we evaluated sufficient similarity across *Ginkgo biloba* extract products (Catlin et al., 2018). *Ginkgo biloba* extract was an ideal candidate to pilot this strategy because its chemical and toxicological effects are well characterized (NTP, 2013; van Beek and Montoro, 2009). In that study, 26 samples consisting of unfinished products, National Institute of Standards and Technology (NIST) standard reference materials (SRM), and finished products were compared with the National Toxicology Program

(NTP) test article using nontargeted chemical analysis, quantification of marker constituents (ie, chemicals used to identify the authenticity and quality of a sample), 2 liver-based *in vitro* assays, and a 5-day *in vivo* assay for investigating bioactivity responses. Based on an integrative assessment of the data streams, we found that 16 samples (62%) were sufficiently similar to the NTP test article, 7 samples (27%) were different and 3 samples (12%) were of intermediate classification. Here, we explore application of these methodologies to black cohosh (*Actaea racemosa*; syn. *Cimicifuga racemosa*) and *Echinacea purpurea* (commonly called purple coneflower). Black cohosh (marketed for gynecological health) and *E. purpurea* (used for immune system support) are popular botanicals, representing the sixth and second top selling botanical supplements in mainstream markets, respectively, in 2017 (Smith et al., 2018). Both products are representative of the botanical dietary supplement space because they have large unidentified fractions, insufficient associations between marker constituents and biological response, and inadequate toxicity data. In these 2 case studies, multiple samples, representing a breadth of unfinished material and finished products in the market, were compared with well-characterized NTP test articles that are currently being used in *in vivo* toxicity evaluations.

Our approach for establishing sufficient similarity in black cohosh and *E. purpurea* samples utilized data from nontargeted chemical analyses and an *in vitro* human hepatocyte model. These outcomes are rapid, cost-effective, and can be used to screen numerous samples. More specifically, we chose to use nontargeted chemical analysis as it does not require *a priori* knowledge of toxic constituents and provides high-content information ideal for use in global pattern identification. The *in vitro* bioassay selected for use in these case studies was also intended to be nonspecific and applicable to botanical ingredients with unknown toxicological effects, with the assumption that the 5 genes reflecting nuclear receptor activity (eg, aryl hydrocarbon receptor [AhR], constitutive androstane receptor [CAR], pregnane X receptor [PXR], farnesoid X receptor [FXR],

and peroxisome proliferator-activated receptor alpha [PPAR α] in the liver would provide adequate biological space to allow for recognition of different response patterns. The rationale for including the human hepatocyte nuclear receptor assay was that (1) the liver is often a target for botanical toxicity (Avigan et al., 2016; Roytman et al., 2018), (2) the liver can potentially serve as a sentinel by reflecting diseases in other systems (Edwards and Wanless, 2013; Shimizu, 2008), and (3) assessing nuclear receptor activation in hepatocytes is a well-established practice in preclinical drug safety evaluation for identifying potential for drug-drug interactions (Sinz et al., 2008). Future work will compare results from the “nonspecific” bioassay used here with results from assays that reflect known biological activity of the botanical ingredients (ie, genotoxicity with black cohosh [Mercado-Feliciano et al., 2012; Smith-Roe et al., 2018]), and immune modulation with *Echinacea* (Matthias et al., 2008; Sullivan et al., 2008).

MATERIALS AND METHODS

Procurement and Selection of Black Cohosh and *E. purpurea* Samples for Chemical and Biological Analysis

Multiple black cohosh and *E. purpurea* (Table 1 and Supplementary Tables 1 and 2) samples and reference materials were procured. These included various unfinished products (ie, bulk material that serves as the source for finished products), a limited number of finished products (commercially available tablets or capsules), and reference materials. Currently, NIST SRM are not available for black cohosh or *E. purpurea*. Therefore, extract reference materials (XRM) and vouchered botanical reference materials (VBRM) were purchased from ChromaDex (Irvine, California; Supplementary Tables 1 and 2). Although all samples were assessed in nontargeted chemical analysis, only a subset of samples were evaluated in the *in vitro* human hepatocyte nuclear receptor assay. The intent of the *in vitro* assay was to provide a measure of biological activity for test article selection purposes, therefore only unfinished samples were included.

A total of 17 unfinished black cohosh samples, including the NTP test article, were purchased from 8 suppliers (BC1, BC A-P; Table 1 and Supplementary Table 1). Ten finished products (BC Q-Z; Table 1 and Supplementary Table 1) were obtained from 10 manufacturers and contained varied amounts (20–600 mg) of black cohosh according to their labels. Reference material for black cohosh and other cohoshes commonly found as adulterants in black cohosh were purchased from ChromaDex for comparison in this study: black cohosh root XRM (BC AA), Chinese cohosh root VBRM (*Cimicifuga dahurica*; BC AB), red cohosh root VBRM (*Actaea rubra*; BC AC), and yellow cohosh root VBRM (*Actaea podocarpa*; BC AD). All the samples were stored at -20°C .

A similar approach was taken for the procurement of *E. purpurea* unfinished samples, finished products, and reference materials (details provided in Supplementary Table 2). A total of 13 unfinished samples of *E. purpurea*, including the NTP test article, were purchased from 8 suppliers (EP1, EP A-L; Table 1 and Supplementary Table 2). Finished products (EP M-Q; Table 1 and Supplementary Table 2) were obtained from 5 suppliers and contained varied amounts (400–1000 mg) of *E. purpurea* according to their labels. A variety of reference materials were purchased from ChromaDex to better understand whether any observed differences could be related to the plant part or adulteration with other *Echinacea* species: *E. purpurea* root XRM (EP R), *E. purpurea* root VBRM (EP S), *E. purpurea* leaf and stem VBRM

Table 1. Summary of Botanical Products Used for Sufficient Similarity Assessment

Botanical	Materials	Sample IDs
Black cohosh ^a	1 National Toxicology Program test article	BC 1
	16 unfinished products	BC A-P
	10 finished products	BC Q-Z
	4 botanical reference materials ^b	BC AA-AD
<i>Echinacea purpurea</i> ^c	1 National Toxicology Program test article	EP 1
	12 unfinished products	EP A-L
	5 finished products	EP M-Q
	6 botanical reference materials ^d	EP R-W

^aSamples procured from 19 different suppliers.

^bBlack Cohosh Root XRM, Chinese Cohosh Root VBRM, Red Cohosh Root VBRM, and Yellow Cohosh Root VBRM.

^cSamples procured from 14 different suppliers.

^d*Echinacea purpurea* root extract XRM, *E. purpurea* root, *E. purpurea* flowers, *E. purpurea* leaf and stem, *Echinacea pallida* root, and *Echinacea angustifolia* root.

Abbreviations: XRM, extract reference material; VBRM, vouchered botanical reference material.

(EP T), and *E. purpurea* flower VBRM (EP U). Reference material from 2 other *Echinacea* species were also included: *Echinacea angustifolia* root XRM (EP V) and *Echinacea pallida* root VBRM (EP W). All the samples were stored at -20°C .

Nontargeted Chemical Analysis of Black Cohosh and *E. purpurea* Samples by High-Performance Liquid Chromatography Coupled to a Charged Aerosol Detector

Amounts of samples required to achieve 0.4 g of black cohosh in the final sample were extracted with methanol:water (80:20) to prepare 40 mg/ml black cohosh in the final extract, with the exception of BC U where the final concentration was 20 mg/ml (due to an unintended measurement error during sample preparation). Despite this error, we included sample BC U in the analysis to provide a reference point for a sample with notably lower black cohosh mass, which is a possible outcome of economic adulteration (when less expensive plant material or filler is added to botanical dietary supplement products). For finished products, label claims of the amount of black cohosh present were used to arrive at the starting amount. The Chinese, red, and yellow cohosh VBRM samples were prepared similarly. All samples were extracted by vortexing for 30 s and sonicating for 30 min. Samples were then centrifuged at approximately 3000 rpm for 5 min and the supernatants were analyzed using an Agilent (Santa Clara, California) high-performance liquid chromatograph (HPLC) coupled to a charged aerosol detector (CAD) using a Phenomenex (Torrance, California) Aqua C18 column (250 \times 4.6 mm, 5 μm). Mobile phases A (10% aqueous formic acid) and B (acetonitrile) were used at 1 ml/min and with a linear gradient (% B): 5–15, in 15 min with a 5-min hold, 15–30 in 15 min, 30–40 in 15 min, 40–50 in 45 min, and 50–95 in 5 min.

The preparation process for *E. purpurea* samples was slightly different than that for black cohosh based on an in-house comparison of solvent systems and conditions. A 0.5 g aliquot of bulk extract, bulk root powder, finished products (eg, capsules) or 5 g of VBRM of *E. purpurea* samples as listed in Table 1 were extracted as follows. To all samples, ethanol:water:trifluoroacetic acid (TFA) (60:40:0.1) was added such that the final extraction volume was approximately 20 ml and samples were extracted by vortexing for 5 min, sonicating for 20 min, followed by

rotating end over end at 70 rpm overnight (16–20 h). Samples were then centrifuged at approximately 1600 rpm for 5 min and the supernatants were collected. Supernatants were diluted to 25 ml with the same extraction solvent and filtered through a 0.45- μ m filter. Samples were analyzed using the same HPLC system as used for the analysis of black cohosh but with a Phenomenex Gemini C18 column (250 \times 4.6 mm, 5 μ m). Mobile phases A (0.1% aqueous TFA) and B (0.1% TFA in acetonitrile) were used at 1 ml/min and with a linear gradient (% B): 10–18 in 9 min, 18–30 in 5 min, 30–80 in 31 min, and 80–100 in 2 min.

Chromatograms were aligned using SpecAlign v2.4.1 (University of Oxford, England). For black cohosh, the chromatograms were rescaled to shift the negative baselines to zero and cropped to 5–97 min to remove the large void peak at the front and column cleaning peak at the end. The data were binned into 0.2-s bins to reduce the effective data rate from 25 points per second to 5 points per second and to provide some smoothing. The chromatograms were aligned, and the baselines adjusted to the same level. For *E. purpurea*, the chromatograms were cropped to 3.7–46 min to remove the large void peak at the front and column cleaning peak at the end. The data were not binned due to sufficient data acquisition rate. The peaks in the chromatograms were then aligned. Processed chromatograms were then exported as CSV files for similarity analysis.

Gene Expression for Major Hepatic Receptor Signaling Pathways in Sandwich Cultures of Primary Human Hepatocytes Exposed to Black Cohosh and *E. purpurea* Samples

A subset of black cohosh (BC 1, BC A-J, and BC AA-AD) and *E. purpurea* (EP 1, EP A-H, EP J-L, EP R, and EP V-W) samples were evaluated in sandwich cultures of primary human hepatocytes (SC-PHHs) (Supplementary Tables 1 and 2). These cultures, versus monolayer culture methods, have been shown to extend the longevity of primary hepatocytes *in vitro*, and to improve the predictive utility for modeling drug metabolism, liver enzyme induction, hepatic transport, and biliary excretion clearance in humans (Hewitt et al., 2007; Swift et al., 2010).

Cell culture. William's E medium (WEM), collagen I-coated 96-well plates, GlutaMAX Supplement, HEPES buffer, dexamethasone, and penicillin and streptomycin antibiotics were obtained from Life Technologies/Thermo (Carlsbad, California). Serum-free hepatocyte culture supplement ITS+, Matrigel, and BioCoat plates were obtained from Corning (Tewksbury, Massachusetts). SC-PHHs (from a 47-year-old Caucasian female; lot: HUM4080; Lonza [Walkersville, Maryland]) were prepared in 96-well collagen type I-coated BioCoat plates by plating SC-PHHs (according to the manufacturer's protocol) at a seeding density of approximately 50 000 cells/well. Briefly, 4-h post-plating, MP100 plating medium was removed and replaced with maintenance medium (WEM supplemented with ITS+, GlutaMax, 15 mM HEPES, 100 nM dexamethasone, and penicillin-streptomycin). Plated SC-PHHs were allowed to attach to the collagen type I-coated 96-well plates for approximately 5 h in a humidified incubator at 5% CO₂ and 37°C. SC-PHHs were overlaid with Matrigel (0.35 mg/ml) to form sandwich cultures in ice-cold culture maintenance medium. Cultures were maintained for approximately 4 days with daily renewal of cell culture maintenance medium and exposure compounds in humidified cell culture incubators at 37°C and 5% CO₂. Supplemental Figure 1 shows a representative image of vehicle-treated SC-PHH for 72 h.

Botanical exposures. Samples were prepared by adding 240 mg black cohosh or *E. purpurea* per ml 80:20 ethanol:water or 20:80

ethanol:water, respectively, based on preliminary efforts to optimize solubility. Samples were centrifuged at approximately 20 000 rcf for 10 min. Supernatants were transferred to glass amber vials. Black cohosh samples were further diluted 1:4 to equilibrate ethanol concentrations, then serially diluted in 20:80 ethanol:water, prior to a final 1:10 dilution in cell culture medium. SC-PHHs were exposed for 72 h to botanical samples (8 noncytotoxic concentrations, third-log spacing, triplicate) to evaluate gene expression of sentinel markers of hepatic receptor activation. Positive control chemicals were exposed at third-log spacing (analogous to botanical samples) and included omeprazole (0–100 μ M; AhR activation), phenobarbital (0–1000 μ M; CAR activation), rifampicin (0–10 μ M; PXR activation), chenodeoxycholic acid (0–100 μ M; FXR activation), and fenofibric acid (0–200 μ M; PPAR α activation). These chemicals were purchased from Sigma-Aldrich (St Louis, Missouri) and stock solutions (500 \times) were prepared, with a final concentration of 0.2% dimethylsulfoxide (DMSO).

Gene expression analysis. For gene expression assays, established (inventoried, validated) TaqMan assays for CYP1A2 (AhR), CYP2B6 (CAR), CYP3A4 (PXR), ABCB11 (FXR), and HMGCS2 (PPAR α) were performed. Cells were lysed with RLT solution from RNeasy-96 kits. Total RNA was isolated on a QiaVac vacuum manifold per the manufacturer's instructions. Isolated RNA was characterized by NanoDrop spectrophotometer (NanoDrop, Wilmington, Delaware), and RNA concentrations were adjusted to 200 ng/30 μ l for reverse transcription incubations. A High Capacity cDNA Reverse Transcription Kit (Life Technologies) was used to synthesize cDNA from 200 ng of total RNA in 60- μ l reactions using random hexamers, following the manufacturer's protocol. TaqMan universal PCR Master Mix and dNTPs were obtained from Applied Biosystems (Life Technologies) and used per the manufacturer's protocol. TaqMan gene expression assays were performed in 384-well plates with a QuantStudio7 instrument (Applied Biosystems, Waltham, Massachusetts) at 10- μ l reaction volumes containing 1 μ l of cDNA per manufacturer's protocol. Assay data were analyzed as per Riedel et al. (2014) using the mean of 3 endogenous control genes (PSMB6, β -actin, and GAPDH) for loading normalization. Because botanical samples were tested at different concentration ranges based on observed cytotoxicity in preliminary range-finding evaluations (data not shown), gene expression data from beyond the tested ranges were imputed from nearest-neighbor measured exposure levels to enable a common exposure range for sample comparisons. The relative fold mRNA content was log-transformed and fit with a nonlinear regression model (4-parameter logistic) using GraphPad Prism Version 7.05 (La Jolla, California). Next, the area under the curve (AUC) analysis was performed in GraphPad Prism to reflect the magnitude and potency of responses for each botanical sample. The net AUC was used for similarity analyses.

Sufficient Similarity Determination

In separate analyses, peak intensity over time from the nontargeted chemical analyses (ie, chemical similarity) and the net AUC for all 5 genes examined in SC-PHHs (ie, biological-response similarity) were used for clustering. Hierarchical clustering analyses were performed in JMP 13 (SAS, Cary, North Carolina) using the Ward method without data standardization and visualized using dendrograms and constellation plots generated in JMP. Within each data stream (nontargeted chemistry and *in vitro* gene expression assay in SC-PHHs), the following set of rules were used to determine similarity of each sample to the

NTP test article. If a sample was in the same cluster as the NTP test article, it was categorized as “similar.” If a sample was in the farthest cluster from the NTP test article-containing cluster, it was categorized as “different.”

Integration of Data Streams

Integration of chemistry and biology data was achieved using a visual interval evaluation method previously described in [Catlin et al. \(2018\)](#) using TIBCO Spotfire Analyst 7.8.0 (TIBCO Software, Palo Alto, California). In this approach, biological similarity calls are viewed in the context of chemical similarity, which is presented not as a similarity call, but as a continuous measure of divergence relative to the NTP test article. With this approach, nontargeted chemical analysis data were visualized by converting the peak intensity data for each sample (BC A-AD or EP A-W) into a distance value (Pearson’s r calculated in Partek Genomics Suite version 6.6; St Louis, Missouri) from the NTP test articles (BC 1 or EP 1). These distance values were plotted along a line which began with the NTP test articles (BC 1 or EP 1) and subsequent samples moving from left to right were increasingly chemically divergent from sample 1. The biological similarity determinations for each sample assessed in the SC-PHH assay were then superimposed on the chemistry line plot. The color corresponds to the similarity call from the hierarchical clustering analysis of gene expression in SC-PHHs, with black indicating the NTP test article, green indicating “similar” and red indicating “different” biological-response activity compared with BC 1 or EP 1. The gray color indicates botanical samples which were evaluated in the nontargeted chemical analysis, but not assessed in the SC-PHH assay.

RESULTS

Case Study 1: Black Cohosh

Chromatograms showing the nontargeted chemical analysis of black cohosh samples are arranged as NTP test article (BC 1), unfinished samples (A-P), finished products (Q-Z), and reference materials (AA-AD) ([Figure 2](#), left). A subset of 4 chromatograms (BC 1, AA, Q, and R) were enlarged to highlight peak similarities and differences among samples, with arrows drawing attention to peaks that appear to be unique to a subset of samples represented by BC Q and BC R ([Figure 2](#), right).

Peak intensities from the nontargeted chemical analysis were clustered using hierarchical clustering, which can be visualized as a dendrogram ([Figure 3A](#)) or a constellation plot ([Figure 3B](#)). Clustering produced 4 groups, where the group number was determined by default settings in the analysis program. Samples in the group containing BC 1 were classified as chemically “similar” to the NTP test article (21 samples). In fact, BC 1 was most closely clustered with the black cohosh root XRM (BC AA). As seen in the enlarged chromatographs in [Figure 2](#) (right), BC 1 (NTP test article) and BC AA (black cohosh reference material) contained many similar peaks and had similar peak intensities. The 3 other groups were considered chemically “different” from BC 1 and contained 9 black cohosh samples. For example, enlarged chromatograms of BC Q and BC R have distinct peaks (arrows) that are not observed in that of BC 1 ([Figure 2](#), right). Out of the 10 finished products, 5 (50%) were in the “similar” cluster and 5 (50%) in the “different” cluster.

Concentration-related changes in gene expression for the 15 evaluated black cohosh samples (BC 1, A-J, and AA-AD) are plotted in [Figure 4](#), and AUC values for each sample exposure profile are listed in [Supplementary Table 3](#). Induction of gene

expression relative to vehicle control is reflected by a positive slope in nonlinear regression analysis of gene expression values, and suppression is indicated by a negative slope in regression analysis of gene expression values. Most (14 out of 15) black cohosh samples resulted in elevated expression of CYP2B6 mRNA ([Figure 4B](#)), which suggests that the CAR/PXR pathways were activated following exposure to this botanical class. However, none of the black cohosh samples reached the maximum fold change achieved with the positive control, phenobarbital. Additionally, all black cohosh samples, except BC AD (yellow cohosh VRBM), increased CYP1A2 mRNA content ([Figure 4C](#)), which suggests activation of the AhR pathway in this liver culture model. mRNA levels of CYP3A4, HMGCS2, and ABCB11, which represent PXR, PPAR α , and FXR pathways, respectively, were divergent across black cohosh samples ([Figs. 4A, D, and E](#)). Many of the other black cohosh samples were associated with slight elevations of expression in these genes. Although a variety of responses were observed, black cohosh exposures did not generally elicit proportionally large increases in mRNA content compared with reference positive control agonists, suggesting these botanical samples are less likely to cause clinically relevant liver enzyme induction, and likely limits the discriminating utility of the hepatocyte nuclear receptor assay for black cohosh samples ([Figure 4](#)).

Hierarchical clustering of AUC data ([Supplementary Table 3](#)) from 5 genes revealed 2 distinct groups of biological-response patterns, which can be visualized in [Figure 5A](#) and the constellation plot in [Figure 5B](#). The cluster containing BC 1 was classified as biologically “similar” to the NTP black cohosh test article and consisted of 9 unfinished black cohosh samples (BC A, B, D, E, F, G, H, I, and J), the black cohosh root XRM (BC AA), and reference materials for other cohosh species (BC AB and AC). Samples in the other cluster were classified as biologically “different” from the NTP test article and consisted of one unfinished black cohosh sample (BC C) and the yellow cohosh VRBM (BC AD).

For data interpretation, 14 black cohosh samples tested in both the chemical and biological screening assays were compared with the NTP test article (BC 1). Nine samples (64%) were classified as “similar” to BC 1 in both the nontargeted chemical analysis and the gene expression analysis in SC-PHHs ([Figure 6](#)). These included the black cohosh reference material (BC AA) and 2 other cohoshes, the Chinese (BC AB) and red (BC AC) cohosh reference material. Five samples (36%) were classified as “different” from BC 1 by either chemical similarity analysis or biological-response similarity analysis. No samples (0%) were consistently found to be different from BC 1 in both chemical and biological similarity analyses.

Case Study 2: *E. purpurea*

As in the black cohosh case study, chromatograms showing the nontargeted chemical analysis of *E. purpurea* samples are arranged as NTP test article (EP 1), unfinished products (A-L), finished products (M-Q), and reference materials (R-W) ([Figure 7](#), left). A subset of 4 chromatograms (EP 1, S, G, and P) were enlarged to highlight peak similarities and differences among samples, with arrows drawing attention to peaks that appear to be unique to a subset of samples represented by EP G and EP P ([Figure 7](#), right).

Peak intensity data from the nontargeted chemical analysis were clustered using hierarchical clustering and visualized using a dendrogram ([Figure 8A](#)) or a constellation plot ([Figure 8B](#)). For *E. purpurea*, clustering produced 3 groups. One group contained 12 samples that were in the same cluster containing EP 1

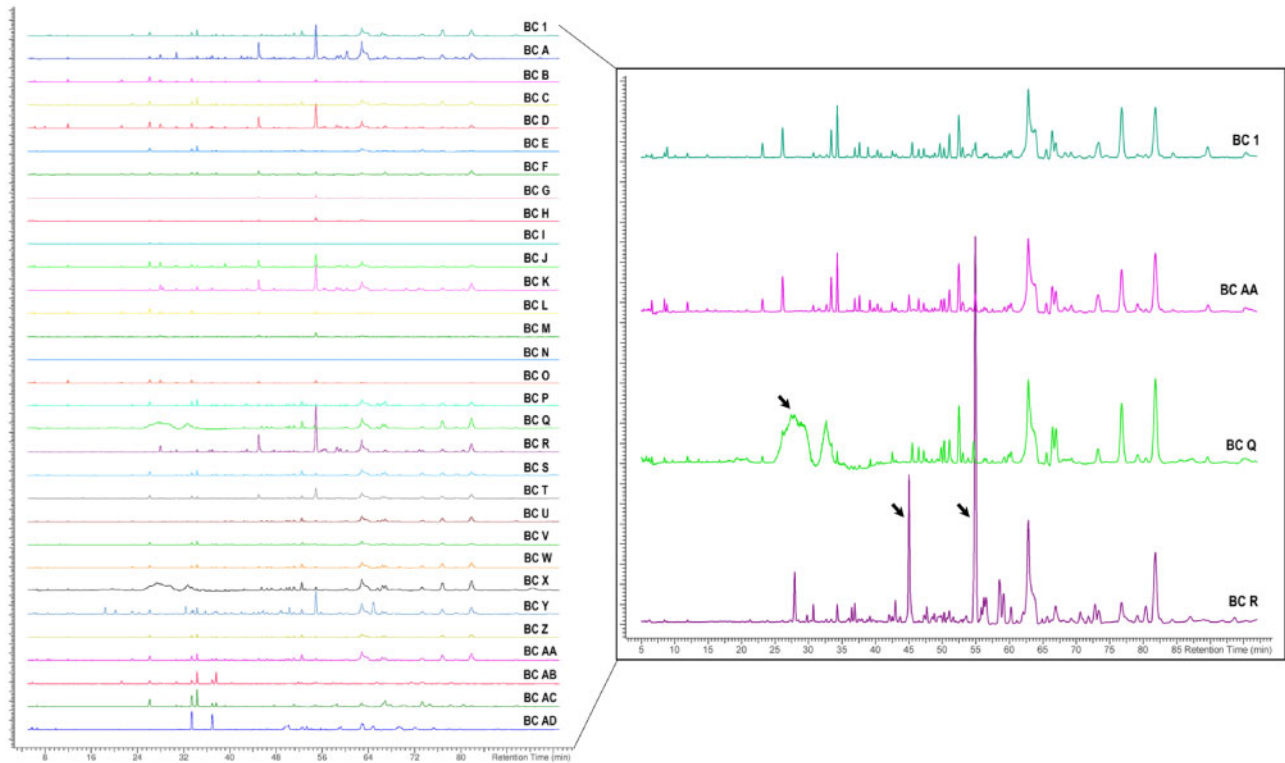


Figure 2. Nontargeted chemical analyses of black cohosh (BC) samples. Chromatograms (left) are showing aligned peak intensity versus time data for all BC samples. For illustrative purposes, 4 sample chromatograms are enlarged (right) to demonstrate the similarities or differences (black arrows) in chemical profiles of samples BC AA, BC Q, and BC R compared with the NTP designated test article, BC 1.

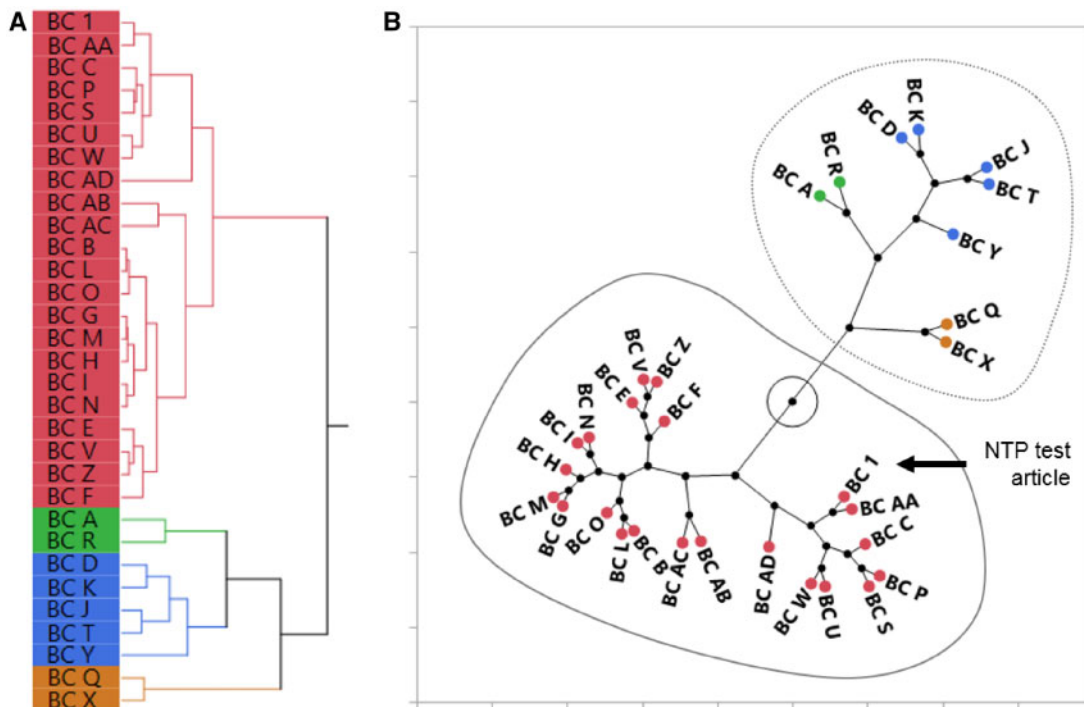


Figure 3. Sufficient similarity evaluation of BC samples using nontargeted chemical analyses. Aligned data from chromatograms were used for hierarchical clustering of BC samples using Ward's method. A dendrogram (A) and constellation plot (B) were generated from the clustering analysis. BC samples in the same cluster of the constellation plot as the NTP test article (BC 1) were determined to be chemically "similar" (circled in solid line). BC samples circled by the dotted line were classified as "different" from BC 1.

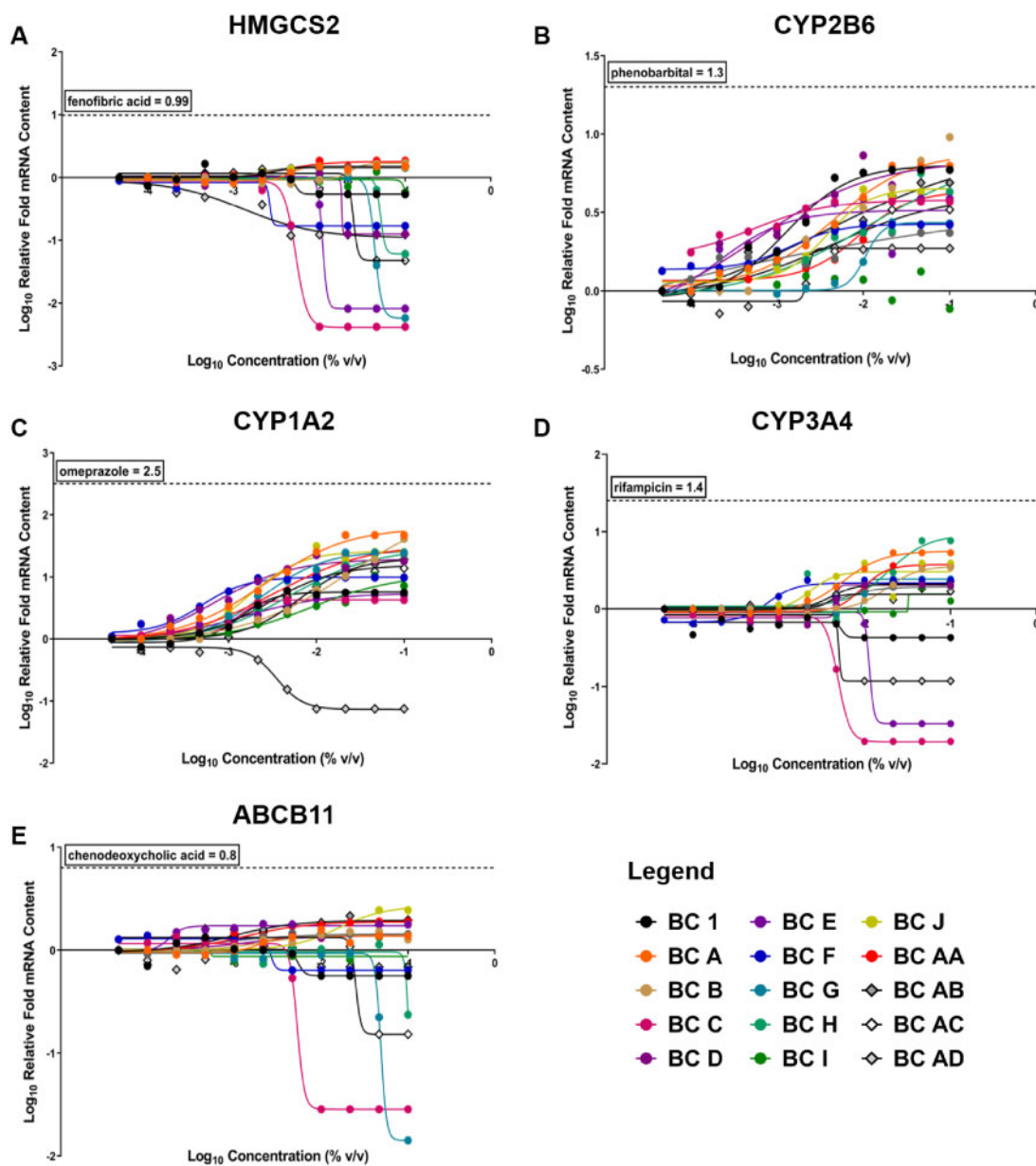


Figure 4. Concentration-response curves of gene expression for 15 BC samples in sandwich cultures of primary human hepatocytes (SC-PHs). Nuclear receptor activation of (A) peroxisome proliferator-activated receptor alpha (PPAR α) (HMGCS2), (B) constitutive androstane receptor (CAR) (CYP2B6), (C) aryl hydrocarbon receptor (AhR) (CYP1A2), (D) pregnane X receptor (PXR) (CYP3A4), and (E) farnesoid X receptor (FXR) (ABCB11) were evaluated at increasing concentrations of each sample. Human clinical activator/ positive control maximum responses for each receptor are shown by the horizontal dotted line on each graph. A legend is provided for all 15 samples tested, BC 1 (black dot) is the NTP test article.

and were classified as chemically “similar.” These included the reference material for *E. purpurea* root (EP S), flower (EP U), and leaf and stem (EP T). The similar peak profiles between EP 1 and EP S can be seen in the enlarged chromatograms (Figure 7, right). The other 2 groups were classified as chemically “different,” totaling 11 samples that included the 2 different *Echinacea* species (EP V and W). For example, enlarged chromatograms of EP G and EP P have distinct additional peaks (arrows) that are not observed in the EP 1 chromatogram (Figure 7, right). Four out of the 5 finished products were classified as chemically different.

Concentration-related changes in gene expression for 15 *E. purpurea* samples (EP 1, A-H, J-L, R, and V-W) are plotted in Figure 9 and AUC values for each exposure profile are listed in Supplementary Table 4. Induction of gene expression relative to vehicle control is reflected by a positive slope in nonlinear regression analysis of gene expression values, and suppression is indicated by a negative slope in regression analysis of gene expression values. Overall, *E. purpurea* samples caused a concentration-related induction of CYP1A2 (except for EP W) and CYP2B6 mRNA (Figs. 9B and 9C), which indicates they may influence AhR and CAR pathways, respectively. Conversely, the

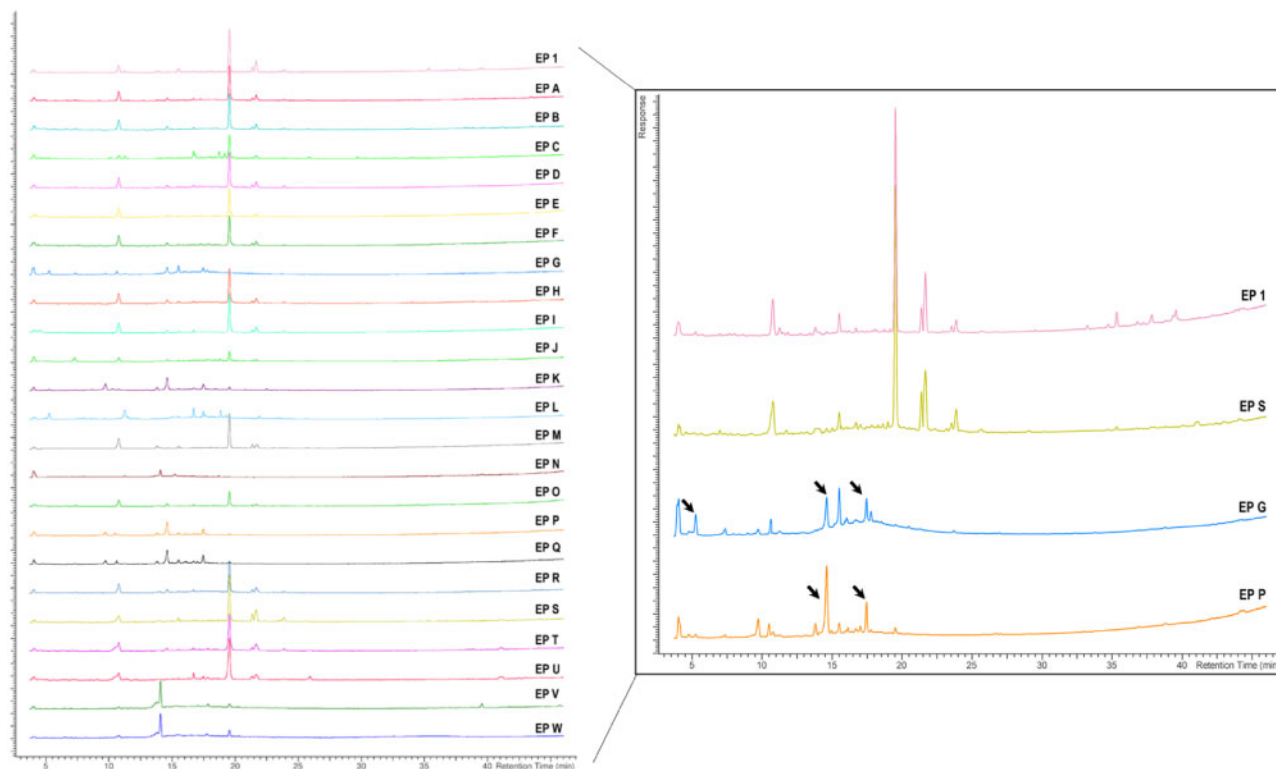


Figure 7. Nontargeted chemical analyses of *Echinacea purpurea* (EP) samples. Chromatograms (left) are showing aligned peak intensity versus time data for all EP samples. For illustrative purposes, 4 sample chromatograms are enlarged (right) to demonstrate the similarities or differences (black arrows) in chemical profiles of samples EP S, EP G, and EP P compared with the NTP designated test article, EP 1.

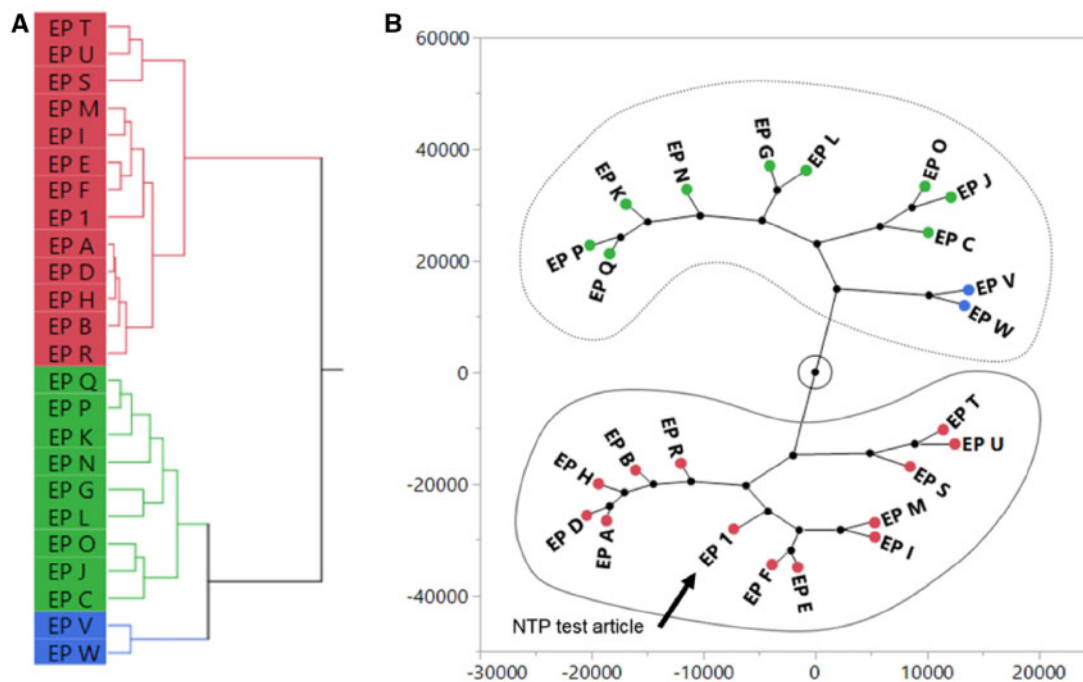


Figure 8. Sufficient similarity evaluation of *Echinacea purpurea* (EP) samples using nontargeted chemical analyses. Aligned data from chromatograms were used for hierarchical clustering of EP samples using Ward's method. A dendrogram (A) and constellation plot (B) were generated from the clustering analysis and EP samples in the same cluster of the constellation plot as the NTP test article (EP 1) were determined to be "similar" (circled in solid line). EP samples circled by the dotted line were classified as "different" from EP 1.

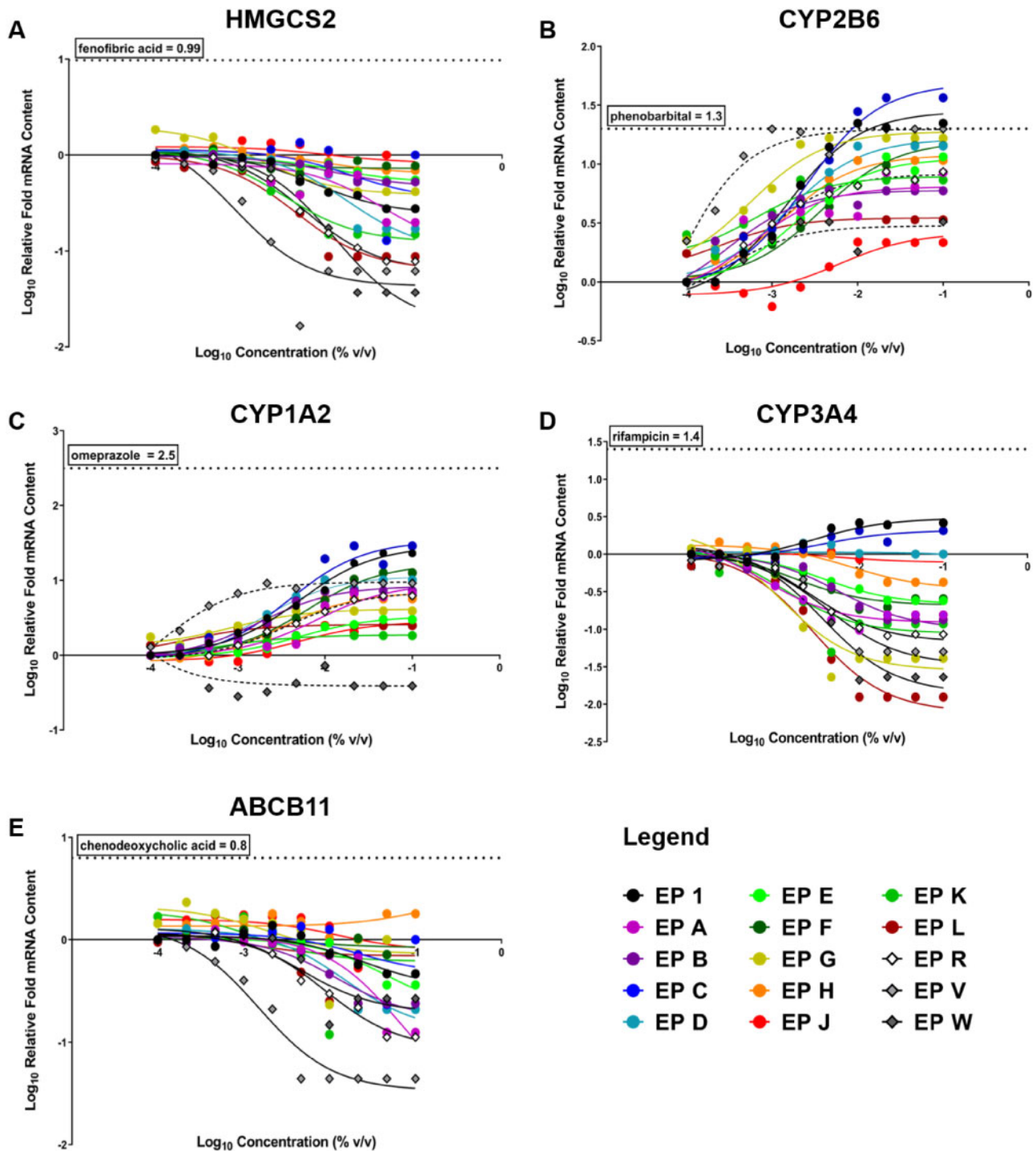


Figure 9. Concentration-response curves of gene expression for 15 *Echinacea purpurea* samples in SC-PHs. Nuclear receptor activation of (A) peroxisome proliferator-activated receptor alpha (PPAR α) (HMGCS2), (B) constitutive androstane receptor (CAR) (CYP2B6), (C) aryl hydrocarbon receptor (AhR) (CYP1A2), (D) pregnane X receptor (PXR) (CYP3A4), and (E) farnesoid X receptor (FXR) (ABCB11) were evaluated at increasing concentrations of each sample. Human clinical activator/positive control maximum responses for each receptor are shown by the horizontal dotted line on each graph. A legend is provided for all 15 samples tested, EP 1 (black dot) is the NTP test article.

DISCUSSION

Creating actionable comparisons of complex and variable mixtures is a key challenge in evaluating the safety of botanical dietary supplements. Approaches for determining sufficient similarity among botanical samples are useful in multiple

safety evaluation contexts. Prior to animal testing, sufficient similarity approaches can be used to identify a test article that resembles either a high-quality finished product or a certified reference material. Post-testing, a sufficient similarity assessment can be performed to compare a test article to related products in the marketplace in order to better understand the

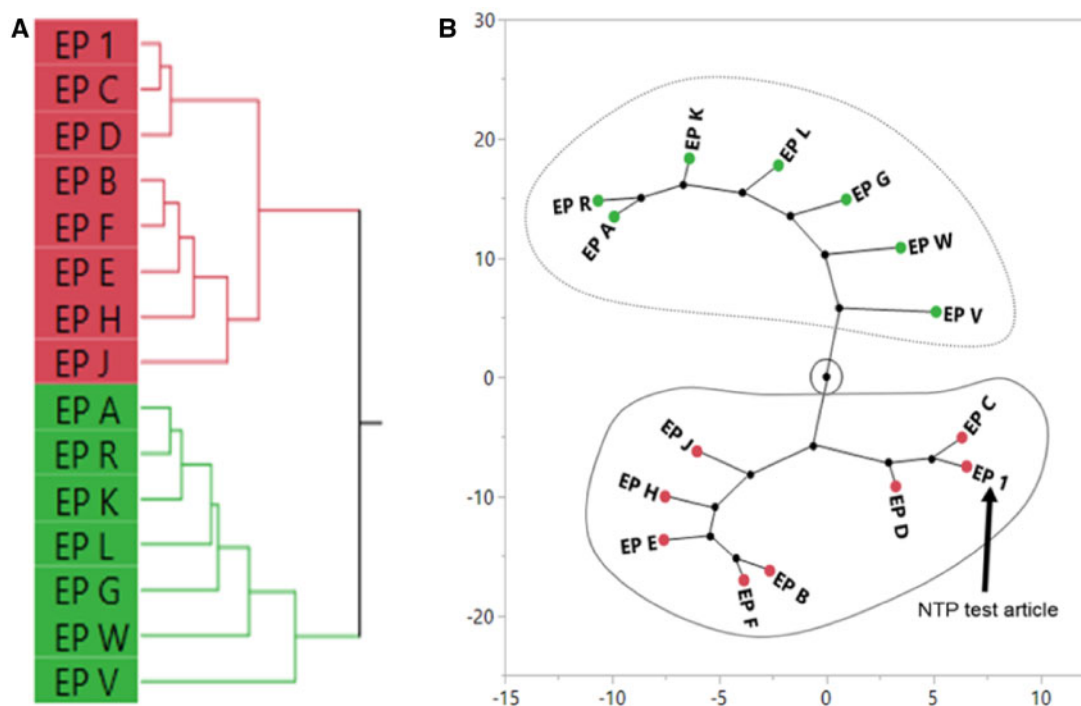


Figure 10. Sufficient similarity evaluation of *Echinacea purpurea* (EP) samples using the area under the curve of gene expression concentration-response curves in exposed sandwich-culture primary human hepatocytes. Hierarchical clustering of the area under the curve of gene expression concentration-response curves was done using the Ward's method (A). A constellation plot (B) was generated from the clustering analysis and EP samples in the same cluster as the NTP test article (EP 1) were determined to be "similar" (circled in solid line). EP samples circled by the dotted line were classified as "different" from EP 1.

Analysis	EP A	EP B	EP C	EP D	EP E	EP F	EP G	EP H	EP J	EP K	EP L	EP R	EP V	EP W
Untargeted chemistry	Black	Black	White	Black	Black	Black	Black	Black	Black	Black	Black	Black	Black	Black
SC-PHH Gene expression	Black	Black	Black	Black	Black	Black	Black	Black	Black	Black	Black	Black	Black	Black

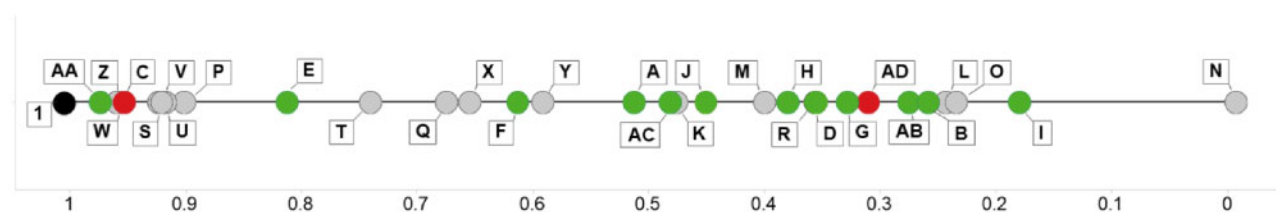
Figure 11. Summary of total sufficient similarity findings for *Echinacea purpurea* (EP) samples. Conclusions of sufficient similarity for the different data streams are shown. A black box indicates "similar" and a white box indicates "different" from the NTP test article. Only samples used in all analyses are presented. SC-PHH = sandwich culture of primary human hepatocytes.

relevance of toxicological findings to other commercially available products.

Black cohosh extract is currently being assessed for toxicity and carcinogenicity at the NTP. In 90-day toxicity studies, black cohosh extract induced hematological changes indicating mild anemia and chromosomal damage evidenced by an increase in micronuclei in peripheral red blood cells (Cora et al., 2017; Mercado-Feliciano et al., 2012). Also, an increase in liver weight (10–15%) was observed in mice and rats and some histopathological changes in the rat liver were noted in the top 2 black cohosh extract dose groups (Mercado-Feliciano et al., 2012). Interestingly, there have been many reports of idiosyncratic liver injury associated with black cohosh use over the years (Enbom et al., 2014; Mahady et al., 2008), along with persistent questions as to whether or not causality could be established (Teschke et al., 2009). Furthermore, lack of proper product identification and possible adulteration with other cohoshes (eg, yellow cohosh, red cohosh, and Chinese cohosh) has been suggested as a possible confounding factor in establishing causality of hepatotoxicity (Teschke et al., 2009). These circumstances highlight the need for methods to compare the chemical composition and biological-response activity of black cohoshes as part of a comprehensive safety assessment.

Based on the nontargeted chemical analysis, we found that the differences in chemical composition observed between black cohosh samples could not be explained by the type of sample (eg, unfinished versus finished). In other words, unfinished samples did not resemble the NTP test article (unfinished) to a greater degree than finished samples (Figs. 2 and 3). Perhaps surprisingly, Chinese cohosh, red cohosh, and yellow cohosh VBRM were as similar to the NTP test article and the black cohosh XRM as some of the unfinished and finished samples labeled as black cohosh. Recent work by van Breemen and colleagues (Nikolic et al., 2015) has uncovered an alkaloid metabolome present in black cohosh that they posit could be responsible for biological activity. Furthermore, they suggest that alkaloids represent a minor yet potent class (Nikolic et al., 2015), which could help to explain the lack of correlation between the chemical and biological-response endpoints in the black cohosh case study (Figs. 6 and 12A). If minor constituents, such as select alkaloids, are responsible for the observed biological activity of black cohosh samples, the pattern of their occurrence in samples would be masked by the presence of more abundant chemical classes (ie, triterpene glycosides and phenolic acids), thereby precluding detection of the relationship between alkaloids and biological activity. Here, a nontargeted chemical analysis method (HPLC-CAD) was used to provide a rapid screen of a

A Black cohosh



B

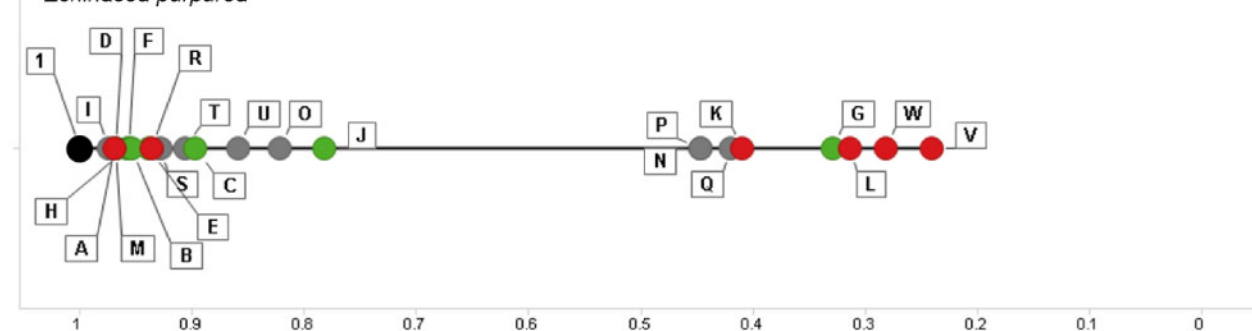
Echinacea purpurea

Figure 12. Distance line plots for black cohosh (A) and *Echinacea purpurea* (B) samples. Pearson's linear correlation coefficients were calculated in reference to the NTP test article sample (BC/EP 1) using data from the nontargeted chemistry analyses. The samples are plotted by distance from the reference with the test article set at 1 (far left of the plot). Further distance from the NTP test article (further right) indicates more chemical dissimilarity. Similarity determinations from the gene expression analysis in sandwich-culture primary human hepatocytes are indicated by color (green = similar; red = different). The black dot references the NTP test article.

broad range of constituents. It is important to note that selection of extraction solvents (eg, ethanol: water), analytical methods, and detection instruments could all influence the constituent profile. In the current case studies, literature reviews and preliminary chemical analyses were used to guide method selection to maximize constituent detection. However, more targeted and sensitive analytical methods for determining black cohosh constituents, including alkaloids, continue to be developed and should be considered for incorporation in future analyses (Bittner et al., 2016; Cicek et al., 2018; Jiang et al., 2011).

In a previous study, the black cohosh samples assessed in the hepatocyte assay were evaluated in an *in vitro* assay measuring micronuclei formation (Smith-Roe et al., 2018). In the micronuclei assay, all the tested samples, including the other cohosh species, induced chromosomal damage. In conjunction with the data presented in the current study, these findings indicate that the chemical composition differences observed in the nontargeted chemistry of black cohosh samples do not correlate with the biological-response activity observed in either the hepatocyte gene expression (Figure 12A) or the micronucleus assays (Smith-Roe et al., 2018). Because indications of chromosomal damage were also observed *in vivo* (Mercado-Feliciano et al., 2012), we conclude that the toxicity findings generated for the NTP test article are relevant for all samples that induce micronuclei formation *in vitro*, regardless of observed differences in chemical composition. Further work is required to identify the constituent(s) responsible for observed biological activity. Bioassay-guided fractionation (Roberts et al., 2019) coupled with the *in vitro* micronuclei assay could be a useful approach for identifying the toxic constituent(s) in cohosh samples. This type of approach has been successfully employed with black cohosh extract in a drug discovery context to identify the active constituent class responsible for modulating the processing of amyloid precursor protein (Findeis et al., 2012).

Echinacea purpurea is currently being evaluated by the NTP for adverse effects but toxicity targets have yet to be definitively identified. The effects of *Echinacea* on drug metabolizing enzymes have been noted in the literature, suggesting potential drug-botanical or botanical-botanical interactions (Meng and Liu, 2014). However, there are inconsistencies between different studies in terms of induction versus inhibition of cytochromes P450s with *Echinacea* treatment (Awortwe et al., 2015; Mooiman et al., 2014; Yale and Glurich, 2005). In this study, the chemical composition of *E. purpurea* samples seems to be divided into 2 distinct groups, in contrast to the black cohosh case study where there was no clear distinction. *Echinacea angustifolia* VRBM and *E. pallida* VRBM were classified as "different" from the NTP test article (Figs. 7 and 8). Although biological-response similarity to the NTP test article does not correlate perfectly with compositional similarity, there does appear to be some congruence between the 2 (Figs. 11 and 12B). The similarity analysis described here could be applied in selection of a test article for additional *in vivo* toxicity studies at the NTP and elsewhere. For example, samples B, D, E, F, and H would be good candidates for additional studies because they are similar to the NTP test article. Based on evaluation of the *G. biloba* extract case study (Catlin et al., 2018) and the black cohosh and *E. purpurea* case studies described here, we have decided to incorporate *in vitro* testing of candidate lots, as well as chemical and biological similarity evaluation into the NTP test article selection process for botanicals moving forward.

Decisions regarding data processing and analysis can dramatically impact the outcome of sufficient similarity determination processes and should be fit-for-purpose. Although there are many options available for comparing across large datasets (Kellogg et al., 2016), in the current case studies, the goal was to develop a simple, screening-level approach with broad application potential. Therefore, minimal data processing and easy-to-apply rules for determining similarity were favored. For

example, with the nontargeted chemistry data, we aligned peaks in the chromatogram to account for run-to-run drift, but we did not normalize the data, which is often done in authentication of botanical samples (Harnly et al., 2016). Although normalization can facilitate identification of qualitative differences between samples (ie, normalizing peak levels can improve detection of the presence/absence of peaks), it can obscure quantitative differences (ie, adjusting peak heights precludes the correlation between peak height and constituent quantity within the sample). Here, we were interested in both quantitative and qualitative differences between samples. The tradeoff for simplicity and ease-of-application is that the inherent lack of supervision could obfuscate the magnitude of difference among samples and oversimplify nuanced data. For example, weighting certain gene responses in the SC-PHH assay or filtering the responses based on a threshold (eg, 20% of positive control maximum), would likely change the similarity determination for some samples. Future efforts will be aimed at refining the methods described here by comparing outcomes based on different input data and threshold adjustments.

The case studies described here offer one approach for combining chemical composition data with biological-response data to evaluate the similarity of botanical samples to a well-characterized referent. Although the clear majority of botanical research comparing commercial products has been based on composition alone, inclusion of a biological activity measure can provide additional insight. In fact, in cases such as black cohosh where toxic effects/mechanisms are known (i.e., genotoxicity), we suggest that the biological-response data could drive decisions about test article selection and extrapolation of results to related products. Although knowledge of bioactive constituents and toxicological targets facilitates the design of sufficient similarity evaluations, as evidenced in the previous *G. biloba* extract case study (Catlin et al., 2018), the nontargeted chemical analysis and the SC-PHH gene expression assays used in these case studies are a reasonable starting place.

Several caveats and considerations should be noted in future application and further development of the methods described here. First, the SC-PHH gene expression assay covers a limited biological space, focusing on a single organ (i.e., liver) and a small set of gene targets within that organ. Biological-response activity measures that cover a greater range of signaling pathways (e.g., high-throughput transcriptomic approaches) and multiple tissue targets, should be explored in future sufficient similarity case studies to evaluate their utility in this context. Second, the determination of sufficient similarity described here is focused on the hazard characterization phase of risk assessment rather than understanding how *in vitro* concentrations relate to doses in preclinical animal studies or human exposure levels. There is important work being done in the area of *in vitro-in vivo* extrapolation that offers a path forward for extension of these approaches from a hazard characterization context to one of risk (Wambaugh et al., 2018). Although beyond the scope of the current effort, we assert that building bridges between human-based *in vitro* assays, *in vivo* animal toxicity studies, and human data from clinical trials, epidemiological studies, and adverse event reporting is critical to understanding the safety of dietary botanical ingredients.

SUPPLEMENTARY DATA

Supplementary data are available at Toxicological Sciences online.

DECLARATION OF CONFLICTING INTERESTS

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ACKNOWLEDGMENT

The authors are grateful to Troy Hubbard and Georgia Roberts (NIEHS) for their review of this manuscript.

FUNDING

This work was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences, Intramural Research project ZIA ES103316-04, and performed in part for the National Toxicology Program, National Institute of Environmental Health Sciences, National Institutes of Health, U.S. Department of Health and Human Services, under contract HHSN273201400027C (Battelle, Columbus, OH).

REFERENCES

- Avigan, M. I., Mozersky, R. P., and Seeff, L. B. (2016). Scientific and regulatory perspectives in herbal and dietary supplement associated hepatotoxicity in the United States. *Int. J. Mol. Sci.* **17**, 331.
- Awortwe, C., Manda, V. K., Avonto, C., Khan, S. I., Khan, I. A., Walker, L. A., Bouic, P. J., and Rosenkranz, B. (2015). *Echinacea purpurea* up-regulates CYP1A2, CYP3A4 and MDR1 gene expression by activation of pregnane X receptor pathway. *Xenobiotica* **45**, 218–229.
- Bittner, M., Schenk, R., Springer, A., and Melzig, M. F. (2016). Economical, plain, and rapid authentication of *Actaea racemosa* L. (syn. *Cimicifuga racemosa*, Black Cohosh) herbal raw material by resilient RP-PDA-HPLC and chemometric analysis. *Phytochem. Anal.* **27**, 318–325.
- Catlin, N. R., Collins, B. J., Auerbach, S. S., Ferguson, S. S., Harnly, J. M., Gennings, C., Waidyanatha, S., Rice, G. E., Smith-Roe, S. L., Witt, K. L., et al. (2018). How similar is similar enough? A sufficient similarity case study with *Ginkgo biloba* extract. *Food Chem. Toxicol.* **118**, 328–339.
- Cicek, S. S., Girreser, U., and Zidorn, C. (2018). Quantification of the total amount of black cohosh cycloartanoids by integration of one specific (1)H NMR signal. *J. Pharm. Biomed. Anal.* **155**, 109–115.
- Clarke, T. C., Black, L. I., Stussman, B. J., Barnes, P. M., and Nahin, R. L. (2015). Trends in the use of complementary health approaches among adults: United States, 2002–2012. *Natl. Health Stat. Rep.* **10**, 1–16.
- Cora, M. C., Gwinn, W., Wilson, R., King, D., Waidyanatha, S., Kissling, G. E., Brar, S. S., Olivera, D., Blystone, C., and Travlos, G. (2017). A black cohosh extract causes hematologic and biochemical changes consistent with a functional cobalamin deficiency in female B6C3F1/N mice. *Toxicol. Pathol.* **45**, 614–623.
- Edwards, L., and Wanless, I. R. (2013). Mechanisms of liver involvement in systemic disease. *Best Pract. Res. Clin. Gastroenterol.* **27**, 471–483.
- Enbom, E. T., Le, M. D., Oesterich, L., Rutgers, J., and French, S. W. (2014). Mechanism of hepatotoxicity due to black cohosh (*Cimicifuga racemosa*): Histological, immunohistochemical and electron microscopy analysis of two liver biopsies with clinical correlation. *Exp. Mol. Pathol.* **96**, 279–283.

- Findeis, M. A., Schroeder, F., McKee, T. D., Yager, D., Fraering, P. C., Creaser, S. P., Austin, W. F., Clardy, J., Wang, R., Selkoe, D., et al. (2012). Discovery of a novel pharmacological and structural class of gamma secretase modulators derived from the extract of *Actaea racemosa*. *ACS Chem. Neurosci.* **3**, 941–951.
- Harnly, J., Chen, P., Sun, J. H., Huang, H. L., Colson, K. L., Yuk, J., McCoy, J. A. H., Reynaud, D. T. H., Harrington, P. B., and Fletcher, E. J. (2016). Comparison of flow injection MS, NMR, and DNA sequencing: Methods for identification and authentication of black cohosh (*Actaea racemosa*). *Planta Med.* **82**, 250–262.
- Hewitt, N. J., Lecluyse, E. L., and Ferguson, S. S. (2007). Induction of hepatic cytochrome P450 enzymes: Methods, mechanisms, recommendations, and in vitro-in vivo correlations. *Xenobiotica* **37**, 1196–1224.
- Jiang, B., Ma, C., Motley, T., Kronenberg, F., and Kennelly, E. J. (2011). Phytochemical fingerprinting to thwart black cohosh adulteration: A 15 *Actaea* species analysis. *Phytochem. Anal.* **22**, 339–351.
- Kellogg, J. J., Todd, D. A., Egan, J. M., Raja, H. A., Oberlies, N. H., Kvalheim, O. M., and Cech, N. B. (2016). Biochemometrics for natural products research: Comparison of data analysis approaches and application to identification of bioactive compounds. *J. Nat. Prod.* **79**, 376–386.
- Mahady, G. B., Dog, T. L., Barrett, M. L., Chavez, M. L., Gardiner, P., Ko, R., Marles, R. J., Pellicore, L. S., Giancaspro, G. I., and Sarma, D. N. (2008). United States Pharmacopeia review of the black cohosh case reports of hepatotoxicity. *Menopause* **15**, 628–638.
- Matthias, A., Banbury, L., Bone, K. M., Leach, D. N., and Lehmann, R. P. (2008). Echinacea alkylamides modulate induced immune responses in T-cells. *Fitoterapia* **79**, 53–58.
- Meng, Q., and Liu, K. X. (2014). Pharmacokinetic interactions between herbal medicines and prescribed drugs: Focus on drug metabolic enzymes and transporters. *Curr. Drug Metab.* **15**, 791–807.
- Mercado-Feliciano, M., Cora, M. C., Witt, K. L., Granville, C. A., Hejtmancik, M. R., Fomby, L., Knostman, K. A., Ryan, M. J., Newbold, R., Smith, C., et al. (2012). An ethanolic extract of black cohosh causes hematological changes but not estrogenic effects in female rodents. *Toxicol. Appl. Pharmacol.* **263**, 138–147.
- Mooiman, K. D., Maas-Bakker, R. F., Hendriks, J. J. M. A., Bank, P. C. D., Rosing, H., Beijnen, J. H., Schellens, J. H. M., and Meijerman, I. (2014). The effect of complementary and alternative medicines on CYP3A4-mediated metabolism of three different substrates: 7-Benzoyloxy-4-trifluoromethyl-coumarin, midazolam and docetaxel. *J. Pharm. Pharmacol.* **66**, 865–874.
- NIH (2019). National Institutes of Health, Office of Dietary Supplements, and National Library of Medicine Bethesda, MD. <https://www.dslid.nlm.nih.gov/dslid/index.jsp#contact>, Accessed March 26, 2019.
- Nikolic, D., Lankin, D. C., Cisowska, T., Chen, S. N., Pauli, G. F., and van Breemen, R. B. (2015). Nitrogen-containing constituents of black cohosh: Chemistry, structure elucidation, and biological activities. *Recent Adv. Phytochem.* **45**, 31–75.
- NTP (2013). *NTP Technical Report on the Toxicology and Carcinogenesis Studies of Ginkgo biloba Extract (CAS No. 90045-36-6) in F344/N Rats and B6C3F1/N Mice (Gavage Studies)*. Natl Toxicol Program Tech Rep Ser. **578**, 1–183. Accessed at https://ntp.niehs.nih.gov/ntp/htdocs/lt_rpts/tr578_508.pdf
- Riedel, G., Rudrich, U., Fekete-Drimusz, N., Manns, M. P., Vondran, F. W. R., and Bock, M. (2014). An extended delta CT-method facilitating normalisation with multiple reference genes suited for quantitative RT-PCR analyses of human hepatocyte-like cells. *PLoS One* **9**, e93031.
- Roberts, G. K., Gardner, D., Foster, P. M., Howard, P. C., Lui, E., Walker, L., van Breemen, R. B., Auerbach, S. S., and Rider, C. (2019). Finding the bad actor: Challenges in identifying toxic constituents in botanical dietary supplements. *Food Chem. Toxicol.* **124**, 431–438.
- Roytman, M. M., Poerzgen, P., and Navarro, V. (2018). Botanicals and hepatotoxicity. *Clin. Pharmacol. Ther.* **104**, 458–469.
- Shimizu, Y. (2008). Liver in systemic disease. *World J. Gastroenterol.* **14**, 4111–4119.
- Shipkowski, K. A., Betz, J. M., Birnbaum, L. S., Bucher, J. R., Coates, P. M., Hopp, D. C., MacKay, D., Oketch-Rabah, H., Walker, N. J., Welch, C., et al. (2018). Naturally complex: Perspectives and challenges associated with Botanical Dietary Supplement Safety assessment. *Food Chem. Toxicol.* **118**, 963–971.
- Sinz, M., Wallace, G., and Sahi, J. (2008). Current industrial practices in assessing CYP450 enzyme induction: Preclinical and clinical. *AAPS J.* **10**, 391–400.
- Smith, T., Kawa, K., Eckl, V., Morton, C., and Stredney, R. (2018). Herbal supplement sales in US Increased 8.5% in 2017, topping \$8 billion. *HerbalGram* **119**, 62–71.
- Smith-Roe, S. L., Swartz, C. D., Shepard, K. G., Bryce, S. M., Dertinger, S. D., Waidyanatha, S., Kissling, G. E., Auerbach, S. S., and Witt, K. L. (2018). Black cohosh extracts and powders induce micronuclei, a biomarker of genetic damage, in human cells. *Environ. Mol. Mutagen.* **59**, 416–426.
- Sullivan, A. M., Laba, J. G., Moore, J. A., and Lee, T. D. G. (2008). Echinacea-induced macrophage activation. *Immunopharm. Immunotoxicol.* **30**, 553–574.
- Swift, B., Pfeifer, N. D., and Brouwer, K. L. R. (2010). Sandwich-cultured hepatocytes: An in vitro model to evaluate hepatobiliary transporter-based drug interactions and hepatotoxicity. *Drug Metab. Rev.* **42**, 446–471.
- Teschke, R., Bahre, R., Genthner, A., Fuchs, J., Schmidt-Taenzer, W., and Wolff, A. (2009). Suspected black cohosh hepatotoxicity—Challenges and pitfalls of causality assessment. *Maturitas* **63**, 302–314.
- USEPA (2000). *Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures*, Vol. EPA 630/R-00/002. Accessed at https://archive.epa.gov/raf/web/pdf/chem_mix_08_2001-2.pdf.
- van Beek, T. A., and Montoro, P. (2009). Chemical analysis and quality control of *Ginkgo biloba* leaves, extracts, and phytopharmaceuticals. *J. Chromatogr. A* **1216**, 2002–2032.
- Wambaugh, J. F., Hughes, M. F., Ring, C. L., MacMillan, D. K., Ford, J., Fennell, T. R., Black, S. R., Snyder, R. W., Sipes, N. S., Wetmore, B. A., et al. (2018). Evaluating in vitro-in vivo extrapolation of toxicokinetics. *Toxicol. Sci.* **163**, 152–169.
- Yale, S. H., and Glurich, I. (2005). Analysis of the inhibitory potential of *Ginkgo biloba*, *Echinacea purpurea*, and *Serenoa repens* on the metabolic activity of cytochrome p450 3A4, 2D6, and 2C9. *J. Altern. Complem. Med.* **11**, 433–439.