Rapid Screening of Active Components with an Osteoclastic Inhibitory Effect in *Herba epimedii* Using Quantitative Pattern–Activity Relationships Based on Joint-Action Models

Xiao-Yan Yuan [1] iD, Meng Wang [2], Sheng Lei [2], Qian-Xu Yang [2,*] and Yan-Qiu Liu [3,*]

[1] Department of Analytical Chem istry of School of Pharmacy, Zunyi Medical University, Zunyi 563000, China; yuanxiaoyan815@163.com
[2] R & D Center, China Tobacco Yunnan Industrial Co., Ltd., Kunming 650231, China; wangmeng19871004@163.com (M.W.); leisheng19850928@163.com (S.L.)
[3] Institute (College) of Integrative Medicine, Dalian Medical University, Dalian 116044, China
* Correspondence: yangqianxu@dicp.ac.cn (Q.-X.Y.); yqliu@dmu.edu.cn (Y.-Q.L.); Tel.: +86-871-68328389 (Q.-X.Y.); +86-411-86110192 (Y.-Q.L.)

Abstract: Screening of bioactive components is important for modernization and quality control of herbal medicines, while the traditional bioassay-guided phytochemical approach is time-consuming and laborious. The presented study proposes a strategy for rapid screening of active components from herbal medicines. As a case study, the quantitative pattern–activity relationship (QPAR) between compounds and the osteoclastic inhibitory effect of *Herba epimedii*, a widely used herbal medicine in China, were investigated based on joint models. For model construction, standard mixtures data showed that the joint-action models are better than the partial least-squares (PLS) model. Then, the Good2bad value, which could reflect components' importance based on Monte Carlo sampling, was coupled with the joint-action models for screening of active components. A compound (baohuoside I) and a component composed of compounds with retention times in the 6.9–7.9 min range were selected by our method. Their inhibition rates were higher than icariin, the key bioactive compound in *Herba epimedii*, which could inhibit osteoclast differentiation and bone resorption in a previous study. Meanwhile, the half-maximal effective concentration, namely, EC_{50} value of the selected component was 7.54 μg/mL, much smaller than that of baohuoside I—77 μg/mL—which indicated that there is synergistic action between compounds in the selected component. The results clearly show our proposed method is simple and effective in screening the most-bioactive components and compounds, as well as drug-lead components, from herbal medicines.

Keywords: bioactive-component screening; *Herba epimedii*; herbal medicine; quantitative pattern–activity relationship

1. Introduction

The long-term clinical practice of herbal medicine confirms its importance and essential role in the healthcare system in China, especially in the prevention and treatment of chronic diseases [1–3]. However, the major obstacle to the development of herbal medicine is to determine the bioactive and toxic constituents in herbal medicines [4–6]. The bioactive constituents of the Chinese herbal medicines for a certain disease remain largely unknown, except a few examples including taxol from the Pacific yew for anticancer treatment, and artesunate from *Artemisia annua* L. for malaria treatment [7]. Bioassay-guided fractionation is the conventional phytochemical approach to screen active components from herbal medicine, but it is time-consuming and laborious [8–11]. Moreover, more than one

compound or component contributes to the activity of herbal medicines, and synergistic actions always exist among different constituents [12,13]. Rapid, effective and economical approaches should be developed for screening active components from herbal medicines.

The quantitative pattern–activity relationship (QPAR) approach has been proposed in recent years, which quantitatively correlates the chromatographic fingerprint and related bioactivity capacities of the samples by partial least-squares (PLS) [12,14,15], artificial neural network (ANN) [16] and other regression methods, where the quantitative pattern is the quantitative or semi-quantitative result by HPLC–UV, HPLC–MS, NMR spectroscopy and so on. Compared with conventional bioactive-component screening approaches, QPAR has great advantages due to being fast, efficient and allowing component-group screening. Now, PLS coupled with feature selection is the favorite method in QPAR [17–19]. The main problem of PLS lies in that it is a linear method. However, the relationship between component concentration and activity is not a linear correlation, but a sigmoid or hyperbolic correlation [20–22]. Therefore, feature selection based on PLS is an under-fitting method. Although ANN is a nonlinear method, it is inclined to over-fitting and it is hard to explain the relationship between each component.

It is well known that the relationship between component dose and activity could be fitted by the Hill equation in most cases [23]. Concentration addition (CA) [24] and response addition (RA) [25] are two basic joint-action models for modeling mixture activity based on the Hill equation, which is widely used for evaluating mixture activity in the pharmaceutical industry and in environmental risk assessment [26,27]. CA is used for components with similar mechanisms/modes of action (MoA) and RA for components with dissimilar MoA. However, the mixture is more complicated in herbal medicines, which may contain components with similar and dissimilar MoA influencing the final activity together. Thus, based on CA and RA, Olmstead proposed an integrated model for modeling a mixture containing components with similar and dissimilar MoA [28]. The Olmstead model preserves the obvious advantages of the clear relationships of components and a reasonable equation expression.

In this study, three joint-action models coupled with Monte Carlo sampling (MCS) were used for screening bioactive components in *Herba epimedii* for osteoclast growth inhibition. *Herba epimedii* has been used for thousands of years to treat osteoporotic conditions. Icariin, a flavonol glycoside, is one of the primary ingredients of *Herba epimedii*. Icariin has been shown to enhance osteogenic differentiation of rat bone-marrow stromal stem cells by increasing the transcriptional co-activators TAZ expression. Meanwhile, icariin protects against glucocorticoid-induced osteoporosis, increases the expression of the bone enhancer DEC1 and modulates the PI3K/Akt/GSK3b/b-catenin integrated signaling pathway [29,30]. Active components and compounds in *Herba epimedii* selected by the Olmstead model were prepared and validated. Results indicated that the joint-action models can reveal the latent components and compounds rapidly and accurately.

2. Results and Discussion

2.1. Model Comparison

Table 1 lists the 5%, 10%, 20%, 30%, 40% and 50% maximal effective concentration, namely, EC_5, EC_{10}, EC_{20}, EC_{30}, EC_{40} and EC_{50} values of six standards, which were determined in the RAW264.7 growth-inhibition test. The final concentration ratios of six standards through uniform design and the equivalent concentration method in mixtures 1–8 are summarized in Table S1. Figure 1 shows the model adequacy plot of six-standard mixtures by the joint-action and PLS models. The *x*-axis represents the measured osteoclastic inhibitory abilities of mixtures 1–8 under seven concentrations and the *y*-axis represents the predicted activities of PLS, CA, RA and the Olmstead model. The closer the R-square values were to 1, the closer the measured activities were to the predicted activities. The R-square values modeled by PLS, CA, RA and the Olmstead model were 0.548 (Figure 1A), 0.676 (Figure 1B), 0.642 (Figure 1C) and 0.857 (Figure 1D), respectively. It showed clearly that the

joint-action models, especially the Olmstead model, preserve an obvious advantage over the PLS model. Therefore, QPAR analysis based on the joint-action model may reveal a more accurate result.

Table 1. EC$_{5-50}$ [a] values of six standards in the RAW264.7 growth inhibition test.

Standards	Name	EC$_5$	EC$_{10}$	EC$_{20}$	EC$_{30}$	EC$_{40}$	EC$_{50}$
1	epimedin A	8.11	41.00	238.09	766.50	1998.72	4816.47
2	epimedin B	7.14	19.86	60.32	126.22	231.21	402.92
3	epimedin C	0.02	0.52	18.53	199.07	1394.17	8318.58
4	icariin	0.91	3.24	12.80	31.91	67.48	134.17
5	baohuoside I	13.77	21.32	34.26	46.95	60.79	77.06
6	icartin	1.90	4.73	12.73	24.56	42.09	69.02

[a] EC5, EC$_{10}$, EC$_{20}$, EC$_{30}$, EC$_{40}$ and EC$_{50}$ mean the 5%, 10%, 20%, 30%, 40% and 50% maximal effective concentration of standards.
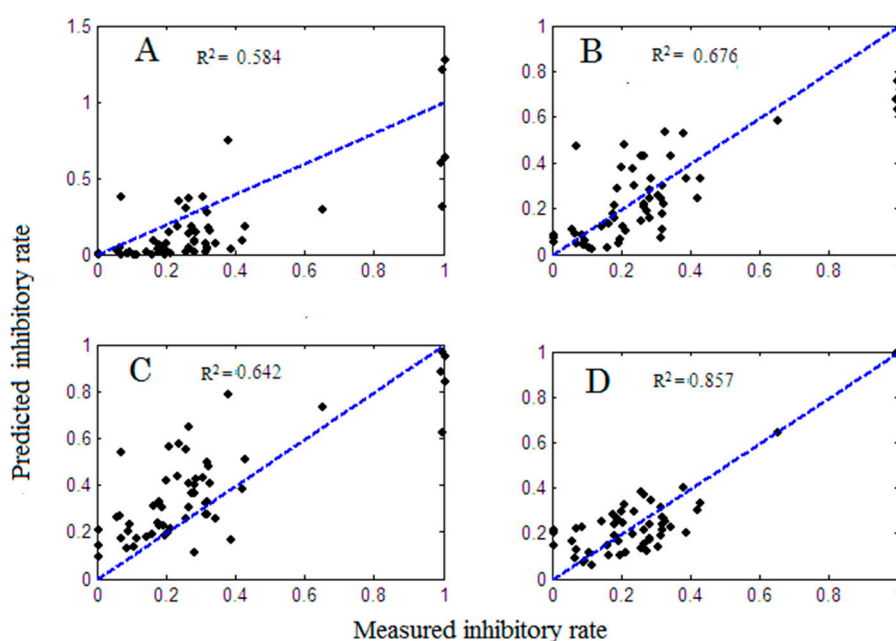

Figure 1. Model adequacy plot of six-standard mixtures by partial least-squares (PLS, **A**), concentration addition (CA, **B**) model, response addition (RA, **C**) model and the Olmstead model (**D**).

2.2. Chromatographic Profile of Herba epimedii and Feature Extraction

Figure 2 shows the overlapped total ion chromatograms (TICs, Figure 2A) and liquid chromatogram diode array detector (DAD) data (Figure 2B) of 66 *Herba epimedii* paired samples after baseline correction and alignment by ChromP. The preparation method of 66 pairwise samples is described in Section 3.2.1. The theoretical basis of QPAR lies in the fact that the varied quantity of bioactive compounds will influence the activity results of the herbal samples. As a result, component content variance is necessary for the QPAR study. For paired samples, the coefficient of variation (CV), namely, the relative standard deviation (RSD) of the TIC and DAD of compounds in all retention times, are shown in Supplementary Materials Figure S1. For the TIC, the CVs of compounds' area in all samples varied from 0.119 to 1.323, as shown in Figure S1A, indicating a useful variation for model construction. For DAD data, the CVs of most of the compounds' areas in all samples were less than 10, although some compounds' areas reached 20, as shown in Figure S1B. By using ChromP, there were totals of 73 and 61 features extracted from the DAD and TIC, respectively. Compared with XCMS and MZmine, ChromP tends to extract the visual and unduplicated features just from the chromatogram. Thus, it is a more-suitable global search method.
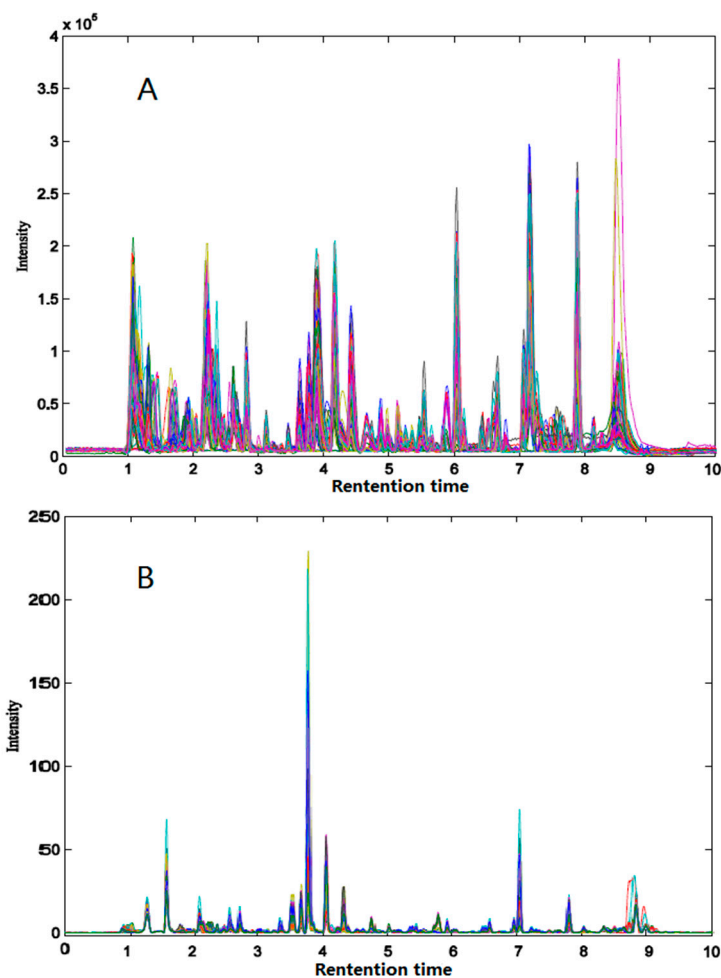
Figure 2. Overlapped total ion chromatograms (**A**) and liquid chromatograms (**B**) of paired samples after baseline correction and alignment by ChromP. The chromatograms of different paired samples show different colors in the diagram.

2.3. Similarity Analysis of Extracted Features

The overlapped UV spectrum from 200 to 400 nm of each feature is shown in Figure S2. Features extracted from the TICs and liquid chromatograms were integrated together, so that all features belonging to the TICs were also included within the liquid chromatograms. In Figure S2, it is obvious that some features exhibit similar UV spectra, indicating that they could be grouped in a cassette in the Olmstead model. Figure 3 displays a UV spectrum correlation heat-map of 73 features extracted from DAD data. Most of the features have a correlation coefficient of more than 0.9, as shown in red in Figure 1, which should be clustered into one group. The features with correlation coefficients between 0.6–0.8 and 0.3–0.5, marked with purple and blue in Figure 3, respectively, should be clustered into another two groups. Apparently, 73 features should be clustered into three groups. Most of those compounds show strong correlations near to 1. This is reasonable for those compounds from species tending to share the same basic structure. At the same time, a small number of compounds display strong correlations with each other but poor correlations with others, which may indicate another category of component. Figure S3 shows an additional proof of compound grouping: principle component analysis (PCA) scatter plot of the UV spectrum of 73 features. By the k-means algorithm, 73 features were tentatively divided into categories one, two, three and four. The results showed that compounds should be clustered into three categories, according to the UV spectra.
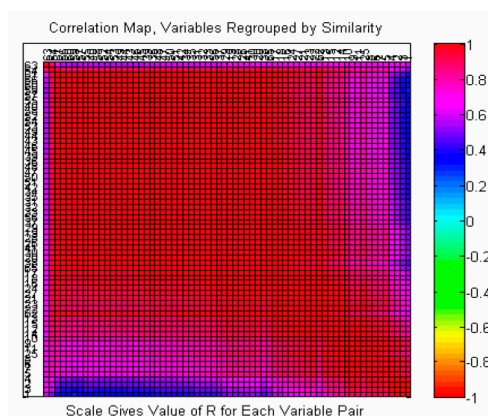
Figure 3. UV spectrum correlation heat-map of 73 features extracted from diode array detector (DAD) data.

2.4. Inhibition Activity to RAW264.7

The activity heat-map of paired *Herba epimedii* samples are shown in Figure 4. The activity of different paired samples varied greatly. The maximum, minimum, mean and standard deviation of activity was 0.74, 0.09, 0.33 and 0.14, respectively. The different activities lie in the variability of bioactive-component quantity in different paired samples, which is the basis of QPAR.
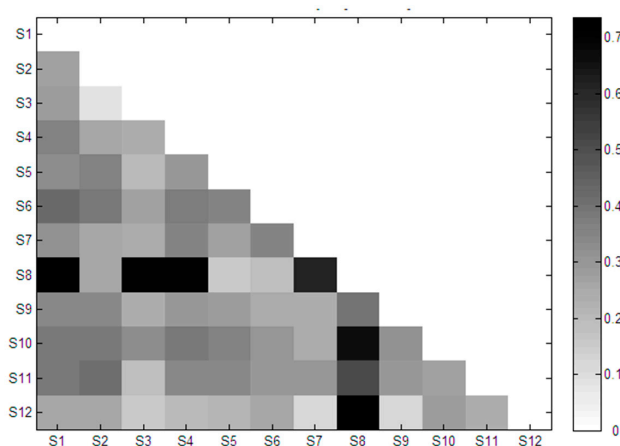

Figure 4. The heat-map of osteoclastic inhibitory activities of paired *Herba epimedii* samples. S1–S12, mean sample 1–sample 12 of *Herba epimedii*.

2.5. Results of Data Processing

The Good2bad analysis was performed on the liquid chromatogram data through CA, RA and the Olmstead models. Figure 5 shows the selected Good2bad values under different models. Generally, the top 1–3 features were selected as latent bioactive compounds. Table 2 lists the selected latent bioactive compounds by different models. In Table 2, the retention time (RT) of compound **56** and **67** selected by the Olmstead model was 6.57 min and 7.9 min, respectively. Compound **56** was identified as sagittatoside A and compound **67** was identified as baohuoside I [31]. The activity of sagittatoside A was validated by Zhang et al. [32]. Baohuoside I was simultaneously screened by the CA and the Olmstead models. Meanwhile, it was also screened by the controlled PLS model. Results indicated baohuoside I may be a latent bioactive compound. The retention times of most of the selected compounds were around 6.9–7.9 min, which indicated that these components may be latent bioactives.
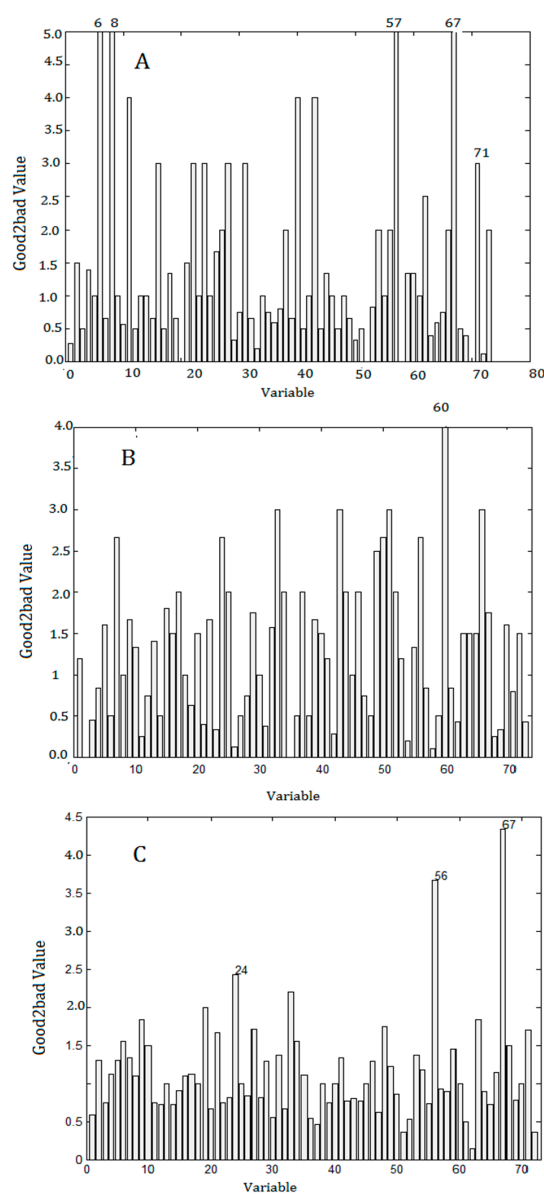
Figure 5. Good2bad values performed on the liquid chromatogram data through concentration additionmodel (CA, **A**), response addition model (RA, **B**) and the Olmstead model (**C**).

Table 2. DAD and TIC data of selected compounds by different models.

No. [a]	M−	RT (min)	Data Source	Model	Compounds
6	448.1861	1.68	DAD	CA	NA [b]
8	609.0456	1.80	DAD	CA	NA
11	564.4197, 627.4082	2.39	TIC	RA/CA	NA
33	661.2183, 724.2124	5.12	TIC	RA	Icarisoside B
50	659.2383, 722.2349	7.16	TIC	RA	2″-O-Rhamnosyl icariside II
56	675.2338, 738.2297	6.94	DAD	Olmstead	Sagittatoside A
57	659.2383, 722.2349	7.04	DAD	CA	2″-O-Rhamnosyl icariside II
60	717.2413, 780.2374	7.50	DAD	RA	NA
67	513.1794, 576.1794	7.90	DAD	Olmsted/CA/PLS	Baohuoside-I

[a] Number of selected features from 73 features extracted from DAD; [b] NA means unknown compound.

2.6. Validation of Selected Components and Compounds

The latent bioactive components and compounds were validated by the cell assay depicted in Section 3.4. Components **1**, **2**, **3** and **4** were obtained from the extract of *Herba epimedii* on a HPD resin column as depicted in Section 3.6.5. Component **3** included most of the selected compounds with retention times around 6.9–7.9 min, which indicated that component **3** may be a latent bioactive component. Figure 6 is the RAW264.7 inhibition rate of the four components. It shows clearly that component **3** has an activity far superior to that of the other components. The inhibition rate of component **3** was almost 81.2%, while the inhibition rates of other components were below 7%. This indicated that compounds in component **3** may be the main bioactive chemicals in *Herba epimedii*. Meanwhile, the activity of component **3** was also higher than that of icariin, the key bioactive chemical in *Herba epimedii* that could inhibit osteoclast differentiation and bone resorption in a previous study [33] after 12 and 24 h incubation, respectively (Figure 7). Figure 8 shows the concentration-response curve (CRC) of component **3**'s inhibition of RAW264.7 growth, where its EC_{50} value is 7.54 µg/mL.
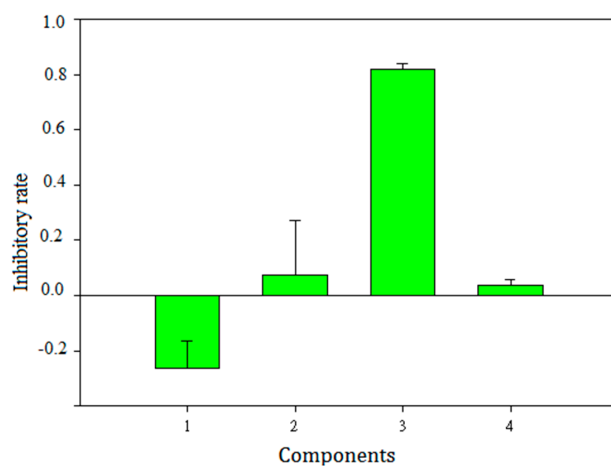

Figure 6. Inhibitory rates of different components **1**, **2**, **3** and **4** to RAW264.7 growth. Components **1**, **2**, **3** and **4** were obtained by separating the extract of *Herba epimedii* on a HPD resin column eluting with 10%, 50%, 80% and 100% ethanol, respectively.
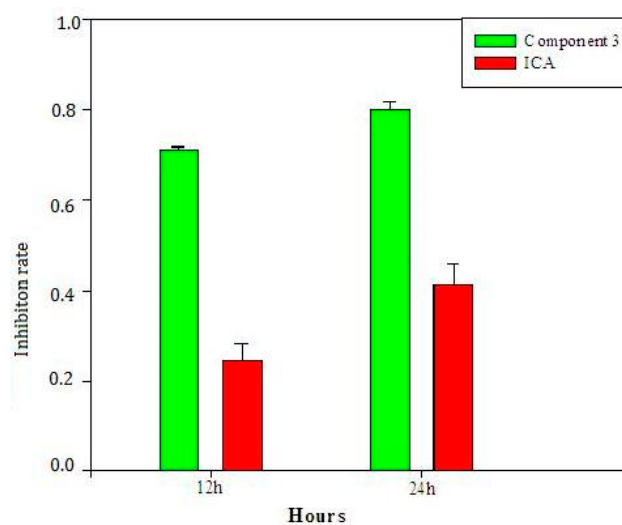

Figure 7. Comparison of activities of component **3** and icariin after 12 and 24 h incubation.
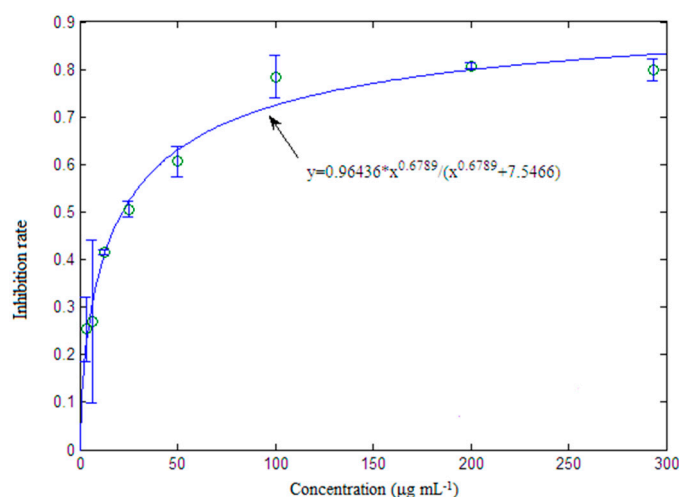
Figure 8. Concentration-response curve (CRC) of component **3** and its inhibition of RAW264.7 growth.

Activities of baohuoside I and icariin have been compared in our previous work [34], where the maximum activity of baohuoside I could reach 91% and icariin reached 60%. Meanwhile, in the osteogenic differentiation test, induced nitric oxide synthase (iNOS) and nitric oxide (NO) production were greatly promoted after adding baohuoside I [35]. All these data indicate that baohuoside I preserves higher activity for osteoclast growth inhibition than icariin.

By comparison of component **3** with baohuoside I, although the maximum activity of baohuoside I was larger than component **3**, the EC_{50} value of component **3** was 7.54 µg/mL, much lower than that of baohuoside I, 77 µg/mL. This indicated that there may be a synergistic effect of compound combinations, such as baohuoside I, sagittatoside A, 2''-*O*-rhamnosyl icariside II, or other unidentified compounds in component **3** [36].

3. Materials and Methods

3.1. Reagents and Materials

HPLC-grade acetonitrile and 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide were purchased from Sigma-Aldrich (St. Louis, MO, USA). HPLC-grade water was prepared using a Milli-Q water purification system (Millipore, Boston, MA, USA). The mouse macrophage pre-osteoclastic RAW264.7 cells were obtained from the Chinese Academy of Sciences (Shanghai, China). Standards of the compounds epimedin A (**1**), epimedin B (**2**), epimedin C (**3**), icariin (**4**), baohuoside I (**5**) and icartin (**6**) were isolated in our laboratory (purity > 98% based on the peak areas of HPLC in this assay). Their structures were unequivocally identified based on UV, IR and NMR spectroscopic analysis. Twelve *Herba epimedii* leaf samples were purchased from Zhi Lin Large Pharmacy Chain Co. Ltd in different locations of China and were identified by Dr. Yujing Zhang from the Department of Pharmacognosy of the School of Pharmacy, Zunyi Medical University (Zunyi, China). Table S2 lists the detailed locations of each pharmacy. The samples were dried in an oven for 2 h before use.

3.2. Sample Preparation

3.2.1. Preparation of Standard Mixtures for Model Comparisons

For preparation of standard mixtures, a procedure was introduced as below: firstly, six standards were weighed accurately and dissolved in dimethyl sulfoxide (DMSO) to prepare stock solutions with concentrations of 400 µg/mL. The prepared standards solutions were diluted two times step by step to get seven concentrations. The EC_5, EC_{10}, EC_{20}, EC_{30}, EC_{40} and EC_{50} values of each

standard were calculated by the Hill equation according to the osteoclastic inhibitory abilities of standards 1–6 described in Section 3.4. Secondly, we composed mixtures 1–6 of six standards through uniform design [37]. A uniform design table of $U_7(7^6)$ (six factors and seven levels, Table S3) was used for composition of mixtures 1–6, where the standards indicate factors and EC_5–EC_{50} of each standard indicate levels. Thus, the $U_7(7^6)$ table is sufficient for a six-factors and seven-levels test. Then, we composed mixtures 7–8 of six standards through the equivalent-effect method [38]. Equivalent-effect means the concentration needed when an equivalent effect was elicited. In our test, EC_{50} and EC_5 were used for mixtures 7 and 8, respectively. Finally, as mentioned above, the composition design of six standards in mixtures 1–8 was established (Table 3). According to the EC_5 to EC_{50} values of each standard, the composition ratio of each standard in every mixture is available. Take ratio of standard 1 in mixture 1 as an example, which could be expressed as follows: ratio of standard 1 = Standard $1(EC_5)$/([Standard $1(EC_5)$ + Standard $2(EC_{10})$ + Standard $3(EC_{20})$ + Standard $4(EC_{30})$ + Standard $5(EC_{40})$ + Standard $6(EC_{50})$]). The total concentration of all standards in each mixture is 500 µg/mL. Then, the prepared solutions of mixtures 1–8 were diluted two times step by step to get seven concentrations, and their osteoclastic inhibitory abilities were also tested, as Section 3.4 depicted.

Table 3. Mixing ratio of six standards in each mixture.

Mixtures [b]	Standards (µg/mL) [a]					
	1	2	3	4	5	6
1	EC_5	EC_{10}	EC_{20}	EC_{30}	EC_{40}	EC_{50}
2	EC_{10}	EC_{30}	EC_{50}	EC_5	EC_{20}	EC_{40}
3	EC_{20}	EC_{50}	EC_{10}	EC_{40}	EC_5	EC_{30}
4	EC_{30}	EC_5	EC_{40}	EC_{10}	EC_{50}	EC_{20}
5	EC_{40}	EC_{20}	EC_5	EC_{50}	EC_{30}	EC_{10}
6	EC_{50}	EC_{40}	EC_{30}	EC_{20}	EC_{10}	EC_5
7	EC_{50}	EC_{50}	EC_{50}	EC_{50}	EC_{50}	EC_{50}
8	EC_5	EC_5	EC_5	EC_5	EC_5	EC_5

[a] Standards **1**, **2**, **3**, **4**, **5** and **6** were epimedin A, epimedin B, epimedin C, icariin, baohuoside I and icaritn, respectively; [b] mixtures 1–6 were composed of fixed concentration ratios of six standards through uniform design, and mixtures 7 and 8 were composed of fixed concentration ratios of six standards through the equivalent concentration method.

3.2.2. Preparation of *Herba epimedii* Pairwise Samples

Twelve dried *Herba epimedii* leaf samples mentioned in Section 3.1 were ground into a fine powder using a pulverizer prior to extraction. An amount of 2.0 g of each sample was extracted ultrasonically with 20 mL 80% ethanol twice for 20 min. The extract was combined and filtered, then condensed and purified by HPD macro-porous absorbent resin column chromatography to remove pigments and strong hydrophilic components. The eluate was condensed by evaporation under reduced pressure and freeze-dried to get dry powder for pairwise sample preparation.

2 mg of powder was weighed accurately and dissolved in 4 mL of DMSO in an ultrasonic bath to get a block solution. Pairwise samples were obtained by a fixed ratio of 1:1 of each block solution at 100 µL of aliquot to get component variations. Thus, 66 paired extract samples were gained for subsequent chromatographic analysis and cell activity tests.

3.3. Apparatus and Analytical Conditions

3.3.1. Chromatographic Analysis

All analyses were carried out on an Agilent Q-TOF 6520 mass spectrometer (Agilent, Santa Clara, CA, USA) equipped with a DAD detector, autosampler and column compartment. Chromatographic separations were performed on a Agilent Proshell 120 SB-C18 column (3.0 mm × 150 mm, 2.7 µm) by gradient elution using a mobile phase consisting of HPLC-grade acetonitrile (solvent A) and

ultrapure water (solvent B). The gradient elution was carried out as follows: 0–7.0 min, 25–55% A; 7.0–7.1 min, 55–100% A and 100% of solvent A holding for 2.9 min. The flow rate was kept at 0.5 mL/min. The column temperature was set to 45 °C and UV measurements were obtained at 267 nm. An aliquot of 2 μL of each sample block solution was injected onto the UPLC. Samples were filtered with 0.22 nm filter membrane prior to chromatographic analysis.

3.3.2. Q-TOF/MS Analysis

MS detection was performed using an Agilent Q-TOF 6520 mass spectrometer (Agilent, Santa Clara, CA, USA) equipped with an electrospray ionization (ESI) interface. The operating parameters were as follows: drying gas (N_2) flow rate, 8.0 L min^{-1}; drying gas temperature, 320 °C; nebulizer, 45 psig; vaporizer temperature, 350 °C; capillary voltage, 3500 V; corona current, 20 μA; skimmer voltage, 65 V; fragmenter voltage, 175 V. The system was operated under MassHunter Workstation software version B.0 2.00 (Agilent, Santa Clara, CA, USA). Each sample was analyzed in negative ion mode, and the spectra were recorded in the m/z range of 100–1500 for full-scan MS analysis.

3.4. Activity Test

The mouse macrophage pre-osteoclastic RAW264.7 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM). The RAW264.7 cells in exponential phase were seeded on 96-well plates (1×10^5 cells/well) and *Herba epimedii* paired extract samples were added to the plates at a final concentration of 100 μg/mL. The plates were incubated further for 24 h in a CO_2 incubator at 37 °C. Activity of inhibition to RAW264.7 cells was accessed by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, Sigma, St. Louis, MO, USA) assay. Inhibition rate was estimated by the following equation:

$$I = (A_t - A_b)/(A_c - A_b) \tag{1}$$

where I, A_t, A_b and A_c are inhibition rate, absorption of test sample, background (containing only DMEM) and control group (without drug) at 490 nm, respectively.

3.5. Theoretical Basis of Data Analysis

3.5.1. Joint-Action Models

For the QPAR study, the model- and feature-selection methods are two important factors when deciding the selection results, where the model links quantitative patterns and biological activity, and the feature-selection method is concerned with selecting subsets of relevant features.

$$\text{CA model}: EC_{x,mix} = \left(1 + \frac{1}{(\sum_{i=1}^{x} \frac{c_i}{EC_{50i}})^p}\right)^{-1}, \tag{2}$$

where $EC_{x,mix}$ is the effect of a mixture containing x chemicals, c_i is the concentration of chemical i, $EC50_i$ is the concentration of chemical i eliciting a half-maximal response, and p is the exponent of chemical 1 to x.

$$\text{RA model}: EC_{x,mix} = 1 - \prod_{i=1}^{x}(1 - EC_i), \tag{3}$$

where $EC_{x,mix}$ also means the effect of a mixture containing x chemicals and EC_i is the effect that chemical i elicits alone.

The Olmstead model is displayed below:

$$EC_{x,mix} = 1 - \prod_{i=1}^{m}\left(1 - \frac{1}{1 + (\sum_{i=1}^{x} \frac{c_i}{EC_{50i}})^{P_{\text{arg},i}}}\right), \tag{4}$$

where $EC_{x,mix}$ also means the effect of a mixture containing x chemicals, $P_{arg,i}$ is the average power of chemicals 1 to m (cassette m), which reflects the important index of cassette m to some extent, c_i is the concentration of the ith chemical and EC_{50i} is the half-maximal effective concentration of ith chemicals when singly used.

The Olmstead model is the combination of CA and RA models mentioned above. In this model, compounds with similar MoA are firstly grouped in a cassette as the CA model, and then compounds in different cassettes are combined as the RA model. Cassette number is a key factor for the joint-action model. When cassette number is 1, the Olmstead model is indeed a CA model. When cassette number is equal to compound number in the mixture, the Olmstead model is indeed an RA model. Cassette number is concerned with MoA similarity of each compound in the mixture. However, for herbal medicines, it is hard to determine the MoA of each constituent chemical. Therefore, for the Olmsted model, cassette number and compound cassette attribution should be decided first.

3.5.2. Cassette-Number Evaluation of the Olmstead Model

Absorption spectra could reflect a chemical structure to some extent [39]. The UV spectrum of a compound is related to the electromagnetic spectrum and electronic-transition information, and compounds with certain UV-wavelength absorption represent certain structural groups. If compounds have a similar structure, they will display similar MoA and UV spectra.

By UPLC coupled with a DAD or PDA detector, the UV spectrum of each compound in the mixture is available. The compounds are then grouped according to their UV spectrum by hierarchical clustering [40] and similarity analysis. Thus, the cassette information of the Olmstead model is established.

3.5.3. Good2bad Value Analysis Based on Monte Carlo Sampling

The main concept of Monte Carlo sampling (MCS) lies in it randomly and iteratively drawing a great number of subsets of the sample and feature direction together to construct N sub-models [41]. For each sub-model, interesting parameters are calculated for further analysis. The Good2bad value is proposed for selecting relevant variables based on MCS.

Figure S4 reflects the principle of the Good2bad value. Its main theory and calculation steps are listed below: N models were firstly constructed based on N subset after MCS, and the model error of each Olmstead model was calculated. Generally, prediction error and recognition error have a normal distribution. For a normal distribution error, the left tail part indicates models with small error (SEM), while the right tail part indicates models with large error (BEM), as shown in Figure S4A. A certain fraction, such as 1% or 5%, could be used to count SEM and BEM quantitatively, according to the balance of computational tolerance and variable selective frequency. Then, variable frequency in SEM and BEM were calculated, where the variable presents a high frequency in SEM indicating importance to the Olmstead model construction, and the opposite situation in BEM (Figure S4B,C). The variable importance could be expressed as the follow equation:

$$\text{Good2bad } (i) = \text{F}(i)\text{SEM}/\text{F}(i)\text{BEM}, \tag{5}$$

where F(i)SEM represents selected frequency of variable i in SEM, the same as F(i)BEM. The bigger the F(i)SEM and the smaller the F(i)BEM is, the more important variable i is.

Thus, as shown in Figure S4D, compounds with larger Good2bad values, such as variables 1, 2, 4 and 5, would be classified as relevant variables; on the other hand, variable 3, with a much smaller Good2bad value, indicates an irrelevant variable [42].

3.6. Bioactive-Component Screening Procedure and Parameter-Setting

3.6.1. Feature Extraction from Chromatographic Profile

XCMS [43], MZmine [44] or ChromP is used for signal deconvolution, noise filter, peak alignment and feature extraction. After feature extraction, a data matrix containing samples and features (chemical information) is available.

3.6.2. Obtain UV Spectrum of Each Compound and Cluster Analysis

The UV spectrum of each compound was extracted according to the retention time of extracted features. K-means was adopted for cluster analysis. The cassette of each compound, that is, the group number, was decided according to feature distribution in the principle component analysis (PCA) score plot.

3.6.3. Construct Quantitative Pattern–Activity Model Based on Subsets

Subsets were generated by MCS. The detailed parameter settings of MCS in our test were: iteration 1000; number of variable 0.05; number of sample 0.7. Each subset was modeled by the joint-action models, and parameters were estimated by the global search method [45]. Samples were selected in training and tested randomly. The cassette information of each compound was decided by the cluster analysis result. All the data were standardized by min–max normalization before modeling. For comparison, different groups and PLS methods were also used for bioactive-component selection.

3.6.4. Good2bad Analysis

After MCS, model errors of all subsets were analyzed, and Good2bad values for each feature were computed to reflect the importance of variables.

3.6.5. Verification

The selected latent bioactive components and compounds were prepared and verified by biological assay as depicted in Section 3.4. For latent active component, the extract of *Herba epimedii* was sampled on an HPD resin column, and eluted by 10%, 50%, 80% and 100% ethanol, and components **1**, **2**, **3** and **4** were obtained after condensation. The component **3** contained the latent active component by UPLC analysis. For preparation of the latent compound, an enzyme-hydrolysis method was established in our previous work [46]. Generally, the latent compound was prepared by dextranase hydrolyzing icariin at pH of 5.4, temperature of 40 °C and hydrolysis time of 3 h.

4. Conclusions

The currently used traditional Chinese medicine (TCM) (herbal medicine) is mostly prepared by a simple extraction, which contains hundreds of useless, unknown or even toxic components. Now, TCM simplification, also known as TCM redevelopment, is a key project for TCM modernization. For TCM simplification, the crucial problem is to decide which regions are bioactive, inactive or toxins in the chromatographic profile and the possible synergistic effect of compound combinations. Then, a directed preparation is needed.

By the QPAR method, the effective bioactive components could be revealed, giving an opportunity for TCM simplification. A QPAR method based on joint-action models was developed for screening bioactive components with RAW264.7 growth-inhibition activity in *Herba epimedii*. By combining the screening results of TIC and DAD quantitative data, a bioactive component and a compound (baohuoside I) were selected from *Herba epimedii*. As expected, their activity was superior to icariin, a key bioactive compound in *Herba epimedii* that exhibits an osteoclastic inhibitory effect. The EC_{50} value of the selected component is smaller than that of the selected compound (baohuoside I). Baohuoside I with retention time of 7.9 min was included in the selected component. Therefore, there is a synergistic

effect between compounds in the selected component. The results indicated that the joint-action model is more powerful than the PLS model in bioactive-component screening. The presented study proposes a strategy for rapid screening of active candidates from herbal medicines.

Supplementary Materials: The following are available online. Figure S1: CV of compounds' area in all *Herba epimedii* samples in total ions chromatogram (TIC, A) and liquid chromatogram (B). Figure S2: Overlapped UV spectrum of 73 compounds extracted from liquid chromatogram. Figure S3: Principle component analysis (PCA) score plot of UV spectrum of 73 features, Table S1: Final concentrations of 6 standards in mixtures 1–8, Table S2: Information of *Herba epimedii* leaf samples purchased in experiment.

Author Contributions: In this paper, X.-Y.Y., Q.-X.Y. and Y.-Q.L. conceived and designed the experiments; M.W. and S.L. M.W. contributed reagents/materials/analysis tools and performed the experiments; Q.-X.Y. analyzed the data; X.-Y.Y. and Q.-X.Y. wrote the paper. All authors approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jiang, M.; Zhang, C.; Cao, H.X.; Chan, K.; Lu, A.P. The role of Chinese medicine in the treatment of chronic diseases in China. *Planta Med.* **2011**, *77*, 873–881. [CrossRef] [PubMed]

2. Yang, X.X.; Gu, W.; Liang, L.; Yan, H.L.; Wang, Y.F.; Bi, Q.; Zhang, T.; Yu, J.; Rao, G.X. Screening for the bioactive constituents of traditional Chinese medicines—Progress and challenges. *RSC Adv.* **2017**, *7*, 3089–3100. [CrossRef]

3. Ghosh, N.; Ali, A.; Ghosh, R.; Das, S.; Mandal, S.C.; Pal, M. Chronic inflammatory diseases: Progress and prospect with Herbal Medicine. *Curr. Pharm. Des.* **2016**, *22*, 247–264. [CrossRef] [PubMed]

4. Li, J.W.H.; Vederas, J.C. Drug discovery and natural products: End of an era or an endless frontier? *Science* **2009**, *325*, 161–165. [CrossRef] [PubMed]

5. Brown, A.C. Kidney toxicity related to herbs and dietary supplements: Online table of case reports. Part 3 of 5 series. *Food Chem. Toxicol.* **2017**, *107 Pt A*, 502–519. [CrossRef]

6. Tarantino, G.; Pezzullo, M.G.; di Minno, M.N.D.; Milone, F.; Pezzullo, L.S.; Milone, M.; Capone, D. Drug-induced liver injury due to "natural products" used for weight loss: A case report. *World J. Gastroenterol.* **2009**, *15*, 2414–2417. [CrossRef] [PubMed]

7. Ng, M.C.H.; Lau, T.Y.; Fan, K.; Xu, Q.S.; Poon, J.; Poon, S.K.; Lam, M.K.; Chau, F.T.; Sze, D.M.Y. Prediction of Radix Astragali immunomodulatory effect of CD80 expression from chromatograms by quantitative pattern-activity relationship. *Biomed. Res. Int.* **2017**, *2017*. [CrossRef] [PubMed]

8. Koehn, F.E.; Carter, G.T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.* **2005**, *4*, 206–220. [CrossRef] [PubMed]

9. Rauf, A.; Uddin, G.; Khan, H.; Arfan, M.; Siddiqui, B.S. Bioassay-guided isolation of antibacterial constituents from Diospyros lotus roots. *Nat. Prod. Res.* **2016**, *30*, 426–428. [CrossRef] [PubMed]

10. Erenler, R.; Meral, B.; Sen, O.; Elmastas, M.; Aydin, A.; Eminagaoglu, O.; Topcu, G. Bioassay-guided isolation, identification of compounds from Origanum rotundifolium and investigation of their antiproliferative and antioxidant activities. *Pharm. Biol.* **2017**, *55*, 1646–1653. [CrossRef] [PubMed]

11. Zhang, X.; Qian, F.; Tan, J.J.; Guo, F.J.; Kulka, M.; Xu, J.W.; Li, Y.M. Bioassay-guided isolation of bisepoxylignans from the flower buds of Magnolia biondii Pamp and their antiallergic effects. *RSC Adv.* **2017**, *7*, 34236–34243. [CrossRef]

12. Zhao, C.X.; Yuan, D.L.; Chan, H.Y.; Chau, F.T.; Liu, L.Q.; Tan, X.Y.; Wang, X.M.; Liang, Y.Z. Rapid screening of antioxidant active constituents from Puerariae Lobatae Radix based on the investigation of quantitative patteren-activity relationship. *Curr. Anal. Chem.* **2015**, *11*, 265–272. [CrossRef]

13. Hu, X.Q.; Sun, Y.; Lau, E.; Zhao, M.; Su, S.B. Advances in synergistic combinations of Chinese herbal medicine for the treatment of cancer. *Curr. Cancer Drug Targets* **2016**, *16*, 346–356. [CrossRef] [PubMed]

14. Yuliana, N.D.; Khatib, A.; Verpoorte, R.; Choi, Y.H. Comprehensive extraction method integrated with NMR metabolomics: A new bioactivity screening method for plants, adenosine A1 receptor binding compounds in orthosiphon stamineus benth. *Anal. Chem.* **2011**, *83*, 6902–6906. [CrossRef] [PubMed]

15. Chau, F.T.; Chan, H.Y.; Cheung, C.Y.; Xu, C.J.; Liang, Y.; Kvalheim, O.M. Recipe for uncovering the bioactive components in herbal medicine. *Anal. Chem.* **2009**, *81*, 7217–7225. [CrossRef] [PubMed]

16. Cabrera, A.C.; Prieto, J.M. Application of artificial neural networks to the prediction of the antioxidant activity of essential oils in two experimental in vitro models. *Food Chem.* **2010**, *118*, 141–146. [CrossRef]

17. Cai, W.S.; Li, Y.K.; Shao, X.G. A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra. *Chemom. Intell. Lab. Syst.* **2008**, *90*, 188–194. [CrossRef]

18. Tistaert, C.; Dejaegher, B.; Heyden, Y.V. Chromatographic separation techniques and data handling methods for herbal fingerprints: A review. *Anal. Chim. Acta* **2011**, *690*, 148–161. [CrossRef] [PubMed]

19. Sun, X.D.; Dong, X.L. Improved partial least squares regression for rapid determination of reducing sugar of potato flours by near infrared spectroscopy and variable selection method. *J. Food Meas. Charact.* **2014**, *9*, 95–103. [CrossRef]

20. Chou, T.C.; Talalay, P. Quantitative analysis of dose-effect relationships: The combined effects of multiple drugs or enzyme inhibitors. *Adv. Enzym. Regul.* **1984**, *22*, 27–55. [CrossRef]

21. Chou, T.C. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol. Rev.* **2006**, *58*, 621–681. [CrossRef] [PubMed]

22. Kovačević, S.Z.; Tepić, A.N.; Jevrić, L.R.; Podunavac-Kuzmanović, S.O.; Vidović, S.S.; Šumić, Z.M.; Ilin, Ž.M. Chemometric guidelines for selection of cultivation conditions influencing the antioxidant potential of beetroot extracts. *Comput. Electron. Agric.* **2015**, *118*, 332–339. [CrossRef]

23. Gadagkar, S.R.; Call, G.B. Computational tools for fitting the Hill equation to dose—Response curves. *J. Pharmacol. Toxicol. Methods* **2015**, *71*, 68–76. [CrossRef] [PubMed]

24. Drescher, K.; Boedeker, W. Assessment of the combined effects of substances: The relationship between concentration addition and independent action. *Biometrics* **1995**, *51*, 716–730. [CrossRef]

25. Feron, V.; Groten, J. Toxicological evaluation of chemical mixtures. *Food Chem. Toxicol.* **2002**, *40*, 825–839. [CrossRef]

26. Backhaus, T.; Arrhenius, Å.; Blanck, H. Toxicity of a mixture of dissimilarly acting substances to natural algal communities: Predictive power and limitations of independent action and concentration addition. *Environ. Sci. Technol.* **2004**, *38*, 6363–6370. [CrossRef] [PubMed]

27. Diaz, F.J.; Mcdonald, P.R.; Pinter, A.; Chaguturu, R. Measuring and statistically testing the size of the effect of a chemical compound on a continuous in vitro pharmacological response through a new statistical model of response detection limit. *J. Biopharm. Stat.* **2015**, *25*, 757–780. [CrossRef] [PubMed]

28. Olmstead, A.W.; Leblanc, G.A. Toxicity assessment of environmentally relevant pollutant mixtures using a heuristic model. *Integr. Environ. Assess. Manag.* **2005**, *1*, 114–122. [CrossRef] [PubMed]

29. Wei, Q.S.; He, M.C.; Chen, M.H.; Chen, Z.Q.; Yang, F.; Wang, H.B.; Zhang, J.; He, W. Icariin stimulates osteogenic differentiation of rat bone marrow stromal stem cells by increasing TAZ expression. *Biomed. Pharmacother.* **2017**, *91*, 581–589. [CrossRef] [PubMed]

30. Hu, J.H.; Mao, Z.; He, S.C.; Zhan, Y.R.; Ning, R.; Liu, W.; Yan, B.F.; Yang, J. Icariin protects against glucocorticoid induced osteoporosis, increases the expression of the bone enhancer DEC1 and modulates the PI3K/Akt/GSK3b/b-catenin integrated signaling pathway. *Biochem. Pharmacol.* **2017**, *136*, 109–121. [CrossRef] [PubMed]

31. Wang, Y.; Guo, Z.M.; Jin, Y.; Zhang, X.L.; Wang, L.; Xue, X.Y.; Liang, X.M. Identification of prenyl flavonoid glycosides and phenolic acids in Epimedium koreanum Nakai by Q-TOF-MS combined with selective enrichment on "click oligo (ethylene glycol)" column. *J. Pharm. Biomed. Anal.* **2010**, *51*, 606–616. [CrossRef] [PubMed]

32. Zhang, D.; Zhang, C.J.; Wang, X. Cytotoxic Effects of Flavonol Glycosides and Nonflavonoid Constituents of Epimedium koreanum on Primary Osteoblasts. *Pharm. Biol.* **2008**, *46*, 185–190. [CrossRef]

33. Hsieh, T.P.; Sheu, S.Y.; Sun, J.S.; Chen, M.H. Icariin inhibits osteoclast differentiation and bone resorption by suppression of MAPKs/NF-κB regulated HIF-1α and PGE(2) synthesis. *Phytomedicine* **2011**, *18*, 176–185. [CrossRef] [PubMed]

34. Yang, Q.X.; Liu, Y.Q.; Wang, L.; Xiao, H.B. Screening of bioactive components in epimedium for osteoporosis treatment by model population analysis. *Acta Pharm. Sin.* **2012**, *47*, 1205–1209.

35. Zhai, Y.K.; Chen, K.M.; Ge, B.F.; Ma, H.P.; Ming, L.G.; Cheng, G.Z. The changes of iNOS and NO in the osteogenic differentiation process of rat bone marrow stromal cells promoted by icariside II. *Acta Pharm. Sin.* **2011**, *46*, 383–389.

36. Liu, Y.Q.; Yang, Q.X.; Cheng, M.C.; Xiao, H.B. Synergistic inhibitory effect of Icariside II with Icaritin from *Herba epimedii* on pre-osteoclastic. *Phytomedicine* **2014**, *21*, 1633–1637. [CrossRef] [PubMed]

37. Wang, Y.; Fang, K. A note on uniform distribution and experimental design. *Kexue Tongbao* **1981**, *26*, 485–489.

38. Qin, L.T.; Liu, S.S.; Zhang, J.; Xiao, Q.F. A novel model integrated concentration addition with independent action for the prediction of toxicity of multi-component mixture. *Toxicology* **2011**, *280*, 164–172. [CrossRef] [PubMed]

39. Varmuza, K.; Karlovits, M.; Demuth, W. Spectral similarity versus structural similarity: Infrared spectroscopy. *Anal. Chim. Acta* **2003**, *490*, 313–324. [CrossRef]

40. Johnson, S.C. Hierarchical clustering schemes. *Psychometrika* **1967**, *32*, 241–254. [CrossRef] [PubMed]

41. Li, H.D.; Liang, Y.Z.; Xu, Q.S.; Cao, D.S. Model population analysis for variable selection. *J. Chemom.* **2010**, *24*, 418–423. [CrossRef]

42. Yang, L.; Yang, Q.X.; Yang, S.H.; Wang, J.J.; Hou, Y.; Wang, B.X.; Tang, Q.; Pan, X.J. Application of near infrared spectroscopy to detect mould contamination in tobacco. *J. Near Infrared Spectrosc.* **2015**, *23*, 391–400. [CrossRef]

43. Smith, C.A.; Elizabeth, J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **2006**, *78*, 779–787. [CrossRef] [PubMed]

44. Sales, C.; Cervera, M.I.; Gil, R.; Portolés, T.; Pitarch, E.; Beltran, J. Quality classification of Spanish olive oils by untargeted gas chromatography coupled to hybrid quadrupole-time of flight mass spectrometry with atmospheric pressure chemical ionization and metabolomics-based statistical approach. *Food Chem.* **2017**, *216*, 365–373. [CrossRef] [PubMed]

45. Ugray, Z.; Leon, L.; John, P.; Fred, G.; James, K.; Rafael, M. Scatter search and local NLP solvers: A multistart framework for global optimization. *Inf. J. Comput.* **2007**, *19*, 328–340. [CrossRef]

46. Yang, Q.X.; Wang, L.; Zhang, L.X.; Xiao, H.B. Baohuoside I production through enzyme hydrolysis and parameter optimization by using response surface and subset selection. *J. Mol. Catal. B Enzym.* **2013**, *90*, 132–138. [CrossRef]

Sample Availability: Not available.