PHYTOMICS: A GENOMIC-BASED APPROACH TO HERBAL COMPOSITIONS

FIELD OF THE INVENTION

This invention relates to herbal compositions. More specifically, this invention provides tools and methodologies for improving the selection, testing, quality control and manufacture of herbal compositions, and to help guide the development of new herbal compositions and identify novel uses of existing herbal compositions.

BACKGROUND OF THE INVENTION

All publications and patent applications herein are incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Herbal medicine has been in use for centuries by people of Asia and Europe. In the United States (US), herbs have become commercially valuable in the dietary supplement industry as well as in holistic medicine. Approximately one third of the US population has tried some form of alternative medicine at least once (Eisenberg *et al.*, 1993, N. Engl. J. Med. 328:246-252).

Botanicals, including herbs, have also become a focal point for the identification of new active agents to treat diseases. Active compounds, derived from plant extracts, are of continuing interest to the pharmaceutical industry. For example, taxol is an antineoplastic drug obtained from the bark of the western yew tree. It is estimated that approximately 50 percent of the thousands of drugs commonly used and prescribed today are either derived from a plant source or contain chemical imitations of a plant compound (Mindell, E.R., 1992, *Earl Mindell's Herb Bible*, A Fireside Book).

Currently, a number of medicinal formulations, food supplements, dietary supplements and the like contain herbal components or extracts from herbs. Herbal

medicines have been used for treating various diseases of humans and animals in many different countries for a very long period of time (see, *e.g.*, I.A. Ross, 1999, *Medicinal Plants of the World, Chemical Constituents, Traditional and Modern Medicinal Uses, Humana Press*; D. Molony, 1998, *The American Associationof Oriental Medicine's Complete Guide to Chinese Herbal Medicine*, Berkely Books; Kessler et al., 1996, *The Doctor's Complete Guide to Healing Medicines*, Berkley Health/Reference Books); Mindell, supra).

Herbal Medicines. There are many branches of herbal medicine around the world, such as Ayurveda, Unani, Sida and Traditional Chinese Medicine (TCM). While modern Western medicine typically consists of administering a single chemical entity capable of intervening a specific biochemical pathway, each formula of TCM typically contains hundreds of chemical entities from several herbs which are designed to interact with multiple targets in the body in a coordinated manner. Although empirical practice contributed in a significant way to the herbal composition and prescription of these ancient herbal medicines, they are also supported, to a varying degree, by a set of theories which all are distinct from that of modern Western medicine in terms of anatomy, pharmacology, pathology, diagnosis treatment, etc. Among the different herbal medicine fields, TCM has developed a more complete set of theories over several centuries which have been well documented and practiced by local physicians caring for a huge population (>1.3 billion people) in greater China and in East Asia including Korea and Japan.

Western medicine generally uses purified compounds, either natural or synthetic, mostly directed towards a single physiological target. However, the compositions used in TCM are usually composed of multiple herbs and compounds which are aimed at multiple targets in the body based on unique and holistic concepts. TCM mainly used processed crude natural products, with various combinations and formulations, to treat different conformations resulting in fewer side effects. The great potential of TCM has yet to be realized for the majority of the world's people.

The herbs in a typical TCM prescription are assigned roles as the principal herb and the secondary herbs, including assistant, adjuvant and guiding herbs. The principal herb produces the leading effects in treating the cause or the main symptom of a disease. An assistant herb helps to strengthen the effect of the principal herb and produces leading effects in the treatment of the accompanying symptoms. There are three types of adjuvant herbs: 1) those that enhance the therapeutic effects of the principal and assistant herbs or treat tertiary symptoms, 2) those that reduce or eliminate the toxicity and other side effects of the principal and the assistant herbs and 3) those which act on complementary target tissues not specifically affected by the principal herb. A guiding herb directs the effect of other herbs to the affected site and/or coordinates and mediates the effects of the other herbs in the prescription or formulation. In contrast to most of the herbal medicines or supplements that consist of one or more parts of a single plant, the intended effects of TCM are directed at multiple tissues.

For example, a well-known TCM recipe, "Ephedra Decoction" used for treating asthma is composed of ephedra, cinnamon twig, bitter apricot kernel and licorice. Ephedra, is the principal herb, which expels cold, induces diaphoresis and facilitates the flow of the Lung Qi to relieve asthma, the main symptom. Cinnamon twig, as the assistant herb, enhances ephedra's induction of diaphoresis and warms the Channels to ensure the flow of Yang Qi for reducing headache and pantalgia. Bitter apricot kernel, as the adjuvant herb, facilitates the adverse flow of the Lung Qi and strengthens the asthma relief by ephedra. Licorice as the guiding herb moderates the effects of both ephedra and cinnamon to ensure a homeostasis of the vital Qi. While each of the four herbs clearly exhibits its respective activity, they complement as well as supplement each other when they are combined. In practice, the principal herb can be prescribed with one or more secondary herbs, depending on the symptoms at a patient's presentation (Prescriptions of Traditional Chinese Medicine, Chapter One, pp10-16, E. Zhang, editor in Chief, Publishing House, Shanghai University of Traditional Chinese

Medicine, 1998).

The main theories of TCM that guide the treatment of sickness with herbal medicine and other means, such as acupuncture, are 1) the theory of Yin and Yang, 2) the theory of Five Elements, 3) the theory of Viscera and Bowels, 4. the theory of Qi, Blood and Body Fluid, and 5) the theory of Channels and Collaterals.

In TCM, the first important aspect of making the proper diagnosis is to ascertain whether the disease is Yin or Yang. For example, those patients who have a fever, are thirsty, constipated or have a rapid pulse condition are of Yang character. Those individuals who have an aversion to cold, are not thirsty, and diarrhea and a slow pulse condition are of Yin character. The property, flavor and function of herbs can also be classified according to Ying and Yang theory. For example, herbs of cold and cool nature belong to Ying, while herbs which are warm and hot in nature belong to Yang. Herbs with sour, bitter and salty flavor belong to Ying, while herbs with pungent, sweet and bland flavor belong to Yang. Herbs with astringent and subsiding function belong to Yin, while herbs with dispersing, ascending and floating function belong to Yang. In TCM, the principles of treatment are based on the predominance or weakness of Yin and Yang. Herbs are prescribed according to their property of Ying and Yang and their function for restoring the imbalance of the Ying and Yang. In so doing, the benefit of treatment is achieved.

According to the theory of Five Elements there are five basic substances that constitute the material world (*i.e.*, wood, fire, earth, metal and water). In TCM, this theory has been used to explain the physiology and pathology of the human body and to guide clinical diagnosis and treatment. Herbal physicians have applied the laws of generation, restriction, subjugation and reverse restriction of the five elements to work out many effective and specific treatment regimens, such as reinforcing earth to generate metal (strengthening the function of the spleen to benefit the lung), replenishing water to nourish wood (nourishing the essence of the kidney to benefit the liver), supporting earth to restrict the wood (supplementing the function of the spleen

to treat the hyperactivity of the liver), and strengthening water to control fire (replenishing the essence of the kidney to treat hyperactivity of the heart). Specifically, the property of some herbs is assigned to each of the five Elements for the purposes of guiding the prescription of a TCM recipe.

In TCM, the internal organs of the human body are divided into three groups: five Viscera (the Heart, the Liver, the Spleen, the Lung and the Kidney), Six Bowels (the Gall Bladder, the Stomach, the large Intestine, the Small Intestine, the Urinary Bladder, and the Triple Warmer), the Extraordinary Organs (the Brain, the Medulla, the Bone, the Blood Vessel, the Gall Bladder, and the Uterus). In TCM, the Viscera or the Bowel are not only anatomic units, but are also concepts of physiology and pathology about interactions between different organs. For example, the heart also refers to some of the mental functions and influence functions of blood, hair, tongue and skin. Ying-Yang and the Five Elements influence the interactions among these Viscera, Bowels and Organs. The complexity of interplay of the theories is used to explain the pathology of diseases to which herbs are prescribed, as discussed below.

The prescription of herbal medicine in TCM starts with the diagnosis, which consists of four main items: interrogation, inspection, auscultation and olfaction, pulse taking and palpation. During the interrogation phase, much information is gathered, including the characteristics of the main symptoms. For instance, if the main symptom is characterized by dull pain of epigastric region, which may be relieved by warming and pressing, this suggests the insufficiency of the Spleen-Yang. Soreness and weakness of the loins and knees, intolerance of coldness with cold extremities manifests a weakness of the Kidney-Yang. During inspection, observations are made for vitality, skin color and the general appearance and the condition of the tongue. For example, a pale complexion corresponds internally to the Lung of autumn, whose Qi is dry. This may occur when Yang Qi is lacking and the circulation of Qi and blood is impeded, or when the coldness in the channels and collaterals causes them to contract.

In TCM, it is from Qi, blood and body fluid that come energy needed by the

Viscera and Bowels, Channels and Collaterals, tissues and other organs for carrying-out their physiological functions; and on which the formation and metabolism of Qi, blood and body fluid depend. Prescriptions of TCM consider the herbal effects on Qi and blood for treatments.

TCM holds that Channels, Collaterals and their subsidiary parts are distributed over the entire body It is through them that herbs exert influence on pathological targets and achieve the improvement of sickness. For example, ephedra acts on the Channels of the Lung and Urinary Bladder so as to induce sweat for relieving asthma and promoting diuresis. As noted above, clinical applications of acupuncture are also guided by the theory of Channels and Collaterals.

In summary, while the nature or property of each herb in TCM may be assigned as Yin or Yang, and to one of the Five Elements, they act through Channels and Collaterals and are mediated via Qi, Blood and Fluid to yield therapeutic effects on targets, such as Viscera and Bowels. Pathogenic factors may be disguised as decoy through the very same systems of Channels and Collaterals to adversely affect the functions of Viscera and Bowels and thus cause sickness.

From the foregoing discussion, it is clear that the TCM terminology is as much of a philosophical concept as an anatomical one. For example, the Heart represents a host of tissues, organs or systems in the body that contribute to a function described in TCM. Thus, the concept of the Heart requires a multiple dimension data set to describe each concept of TCM. Once this is accomplished, a molecular holistic medicine can be developed.

U.S. Regulatory Process. In the US, dietary supplements (such as botanical products, vitamins and minerals, amino acids and tissue extracts) are regulated under the Dietary Supplement Health and Education Act of 1994 (the DSHE Act). This Act removed the ingredients of dietary supplements from regulation as food additives under the Federal Food, Drug, and Cosmetic Act. In addition, the DSHE Act requires that The Food and Drug Administration (FDA) bear the burden of proof that a marketed

dietary supplement presents a serious or unreasonable risk under the conditions of use on the label or as commonly consumed. Thus, there are currently no federal regulations that establish specific criteria for purity, identification and manufacturing procedures for dietary supplements. In addition, few published papers on herbal quality have resulted from the establishment of the Office of Alternative Medicine by Congress in 1992 (Angell *et al.*, 1998, N. Engl. J. Med. 339:839-841).

At the present time, the FDA must approve each one of the chemical entities in a drug composition or cocktail, and then clinical trials must be undertaken so as to obtain separate FDA approval for marketing the drug. This process is extremely tedious and costly. A molecular holistic medicine may require a less arduous evaluation since the previous use of a particular herbal composition as a botanical drug permits clinical trials with multiple chemicals at the outset (*i.e.*, clinical trials using the herbal composition or specific components of the herbal composition). Recently, the FDA has approved the testing of some herbal medicines in clinical trials as botanical drugs (FDA Guidance on Botanical Drugs, April, 1997). While these events represent a positive development for health care in general, it also raises important issues regarding the formulation, manufacturing and quality control of herbal medicines and dietary supplements, including the traditional Chinese medicines.

The multitude of relevant biological responses induced by the multiple chemicals in herbs are not currently available and will be increasingly important to support marketing approval by the FDA.

Herbal-based industries are coming under increasing pressure to upgrade their current practices (see, *e.g.*, Angell *et al.*, *supra*). The need to apply scientific testing to the preparation and administration of herbal medicines and food supplements has been highlighted by several recent reports of toxicity resulting from ingesting herb-based formulations. For example, one patient who took an herbal-based dietary supplement experienced digitalis toxicity (Slifman *et al.*, 1998, N. Engl. J. Med. 339:806-811). It was subsequently determined that the herb ingredient labeled as plantain in the

supplement was actually contaminated with *Digitalis lanata*, an herb known to contain at least 60 cardiac glycosides. In another instance, an herbal preparation was found to be the cause of chronic lead intoxication in a patient (Beigel *et al.*, 1998, N. Engl. J. Med. 339:827-830) This is not a completely unexpected occurrence since contamination of traditional Asian herbal remedies by lead and other heavy metals is well documented (Woolf *et al.*, 1994, Ann. Intern. Med. 121:729-735).

Characterization of Botanicals. It is well known that the genetic identity (*e.g.*, genera, species, cultivar, variety, clone), age of herbal growth, harvest time, the specific plant part utilized, processing method, geographical origin, soil type, weather patterns, type and rate of fertilizer, and other growth factors have a great impact on the particular chemical composition of any particular herb "harvested" from any particular area.

Increasing numbers of various types of tests have been instituted to assure the consistent quality of herbs used in medicine and as dietary supplements; including inspections at the macro- and microscopic levels as well as a variety of chemical analyses. Recently, high performance liquid chromatography (HPLC) profile of marker molecules in an herbal extract has become one reference standard. However, there are problems with this approach, including that some of the bioactive molecules may not adsorb UV or the visible lights for HPLC detection, and the amount of a chemical is not necessarily proportional to its biological potency. For these reasons, herbal manufacturers resort to a practice of mixing raw herbs from different sources to minimize chemical variations.

Mass spectrometry (MS) is an analytical method for determining the relative masses and relative abundances of components of a beam of ionized molecules or molecular fragments produced from a sample in a high vacuum. MS, unlike HPLC, is not optical density-dependent. In practice it is used in conjunction with HPLC or capillary electrophoresis (CE): the HPLC separates the chemicals and the MS then can be used to identify what they are. Commercial systems are available which integrate

MS and HPLC for biological uses. Mass spectrometry is limited to samples that are gaseous or volatile at low pressure, or that can be so rendered by derivatization.

These steps are no longer adequate. Recent publications report a greater variation in the quality of herbs by specific suppliers, and the difficulty of providing biological equivalence of herbal extracts. Furthermore, the correlation between safety and efficacy and chemicals in an herb is not well defined in most cases. Recently, in response to complaints from consumer groups and regulatory agencies (Federal Register, February 6, 1997, Volume 62, No. 25, Docket No. 96M-0417, cGMP in Manufacturing, Packing or Holding Dietary Supplements, Proposed Rules), some herbal manufacturers have begun to implement Good Manufacturing Practice (GMP) which requires stringent controls at all levels.

Chemical and spectroscopic methods have been used to characterize the components of herbal medicines and food supplements. For example, three new hederagenin-based acetylated saponins were isolated from the fruits of *Gliricidia sepium* using these two methods (Kojima *et al.*, 1998, Phytochemistry 48(5):885-888). The botanical sources of Chinese herbal drugs in a number of commercial samples were inferred by comparing the contents of some characteristic constituents which were analyzed with high-performance chromatography (HPLC) or capillary electrophoresis (CE) (Shuenn-Jyi Sheu, 1997, Journal of Food and Drug Analysis 5(4):285-294). For example, the ratio of ephedrine/pseudoephedrine was used as a marker to differentiate *Ephedra intermedia* from other species; total alkaloid contents were used to distinguish between species of *Phellodendron*; and the contents of ginsenosides were used to differentiate between species of *Panax*. However, these methods do not provide a direct measurement of the effect of the various herbs on the molecular, physiological or morphological responses following human treatment with the herbs.

Using gas chromatography-mass spectrometry and atomic-absorption methods, the California Department of Health Sciences, Food and Drug Branch, recently tested Asian medicines obtained from herbal stores for contaminants (R. J. Ko, 1998, N.

Engl. J. Med. 339:847). Of the 260 products they tested, at least 83 (32 percent) contained undeclared pharmaceuticals or heavy metals, and 23 had more than one adulterant. Using high-performance liquid chromatography, gas chromatography, and mass spectrometry, a commercially available combination of eight herbs (PC-SPES), was found to contain estrogenic organic compounds (DiPaola *et al.*, 1998, N. Engl. J. Med. 339:785-791). The researchers concluded that PC-SPES has potent estrogenic activity and that prostate cancer patients that took PC-SPES could confound the results of standard therapies and may experience clinically significant adverse effects. Gas chromatography data was also collected for different samples of the traditional Chinese medicine 'wei ling xian' and correlated to the antiinflammatory activity of the samples (Wei et al., Study of chemical pattern recognition as applied to quality assessment of the traditional Chinese medicine "wei ling xian," Yao Hsueh Hsueh Pao 26(10):772-772 (1991)). This study did not provide relevant HBR Array data, such as time course, dose dependent response, control samples to substantiate the differential power of the biomarkers, nor it utilize a reiterative type of data construction process to establish a comprehensive database for characterizing effects of the herbal composition.

Changes in protein levels have also been used to characterize the effects of herbal compositions or specific components of herbs. For example, the production of granulocyte colony-stimulating factor (G-CSF) from peripheral blood mononuclear cells was found to vary depending on which specific Chinese herb was added to the culture (Yamashiki *et al.*, 1992, J. Clin. Lab. Immunol. 37(2):83-90). Expression of interleukin-1 alpha receptors was markedly up regulated in cultured human epidermal keratinocytes treated with Sho-saiko-to, the most commonly used herbal medicine in Japan (Matsumoto *et al.*, 1997, Jpn. J. Pharmacol. 73(4):333-336). The expression of Fc gamma 11/111 receptors and complement receptor 3 of macrophages were increased by treatment with Toki-shakuyakusan (TSS) (J. C. Cyong, 1997, Nippon Yakurigaku Zasshi 110(Suppl. 1):87-92). Tetrandrine, an alkaloid isolated from a natural Chinese herbal medicine, inhibited signal-induced NF-kappa B activation in rat alveolar

macrophages (Chen *et al.*, 1997, Biochem. Biophys. Res. Commun. 231(1):99-102). The herbs Sairei-to, alismatis rhizoma (Japanese name "Takusha") and hoelen (Japanese name "Bukuryou") inhibited the synthesis and expression of endothelin-1 in rats with anti-glomerular basement membrane nephritis (Hattori et al., 1997, Nippon Jinzo Gakkai Shi 39(2):121-128).

The increase or decrease in mRNA levels has also been used as an indicator of the effect of various herbs and herbal components. Intraperitoneal injection of Qingyangshen (QYS), a traditional Chinese medicine with antiepileptic properties, and diphenylhydantoin sodium reduced alpha- and betta-tublin mRNAs and hippocampal c-fos mRNA induction during kainic acid-induced chronic seizures in rats (Guo et al., 1993, J. Tradit. Chin. Med. 13(4):281-286; Guo *et al.*, 1995, J. Tradit. Chin. Med. 15(4):292-296; Guo *et al.*, 1996, J. Tradit. Chin. Med. 16(1):48-51). Treatment of cultured human umbilical vein endothelial cells (HUVECs) with the saponin astragaloside IV, a component purified from *Astragalus membranaceus*, decreased plasminogen activator inhibitor type I (PAI-1) specific mRNA expression and increased tissue-type plasminogen activator (t-PA) specific mRNA (Zhang *et al.*, 1997, J. Vasc. Res. 34(4):273-280). One component isolated from the root of *Panax ginseng* was found to be a potent inducer of interleukin-8 (IL-8) production by human monocytes and by the human monocytic cell line THP-1, with this induction being accompanied by increased IL-8 mRNA expression (Sonoda *et al.*, 1998, Immunopharmacology 38:287-294).

Recent advances in cDNA microarray technology enable massive parallel mining of information on gene expression. This process has been used to study cell cycles, biochemical pathways, genome-wide expression in yeast, cell growth, cellular differentiation, cellular responses to a single chemical compound, and genetic diseases, including the onset and progression of the diseases (M. Schena *et al.*, 1998, TIBTECH 16:301). No researchers to date, if any, have attempted to apply these new technologies to study the molecular effects of whole herbal treatments and

supplements.

Some researchers have attempted to characterize the effects of the major active constituents isolated from selected herbs. For example, treatment of HUVECs with notoginsenoside R1 (NR1), purified from *Panax notoginseng*, resulted in a dose- and time-dependent increase in TPA synthesis (Zhang *et al.*, 1994, Arteriosclerosis and Thromobosis 14(7):1040-1046). Treatment with NR1 did not change urokinase-type plasminogen activator and PAI-1 antigen synthesis, nor did it effect the deposition of PAI-1 in the extracellular matrix. TPA mRNA increased as much as twofold when HUVECs were treated with NR1, whereas expression of PAI-1-specific mRNA was not significantly affected by NR1. Since most studies on *P. notoginseng* have involved its mixture with other herbs, the researchers noted that it was difficult to assess how their results relate to the situation *in vivo* when is used therapeutically in humans (Id., at 1045, second column, first paragraph). In addition, since the researchers only studied one major component of the herb, it is not possible to ascertain the molecular effect of the whole herb or the interactions among components of the herb from this study.

Dobashi *et al.* (1995, Neuroscience Letters 197:235-238) studied the effect of two of the main components of saiko agents, a Chinese herbal drug used to treat nephrotic syndrome, bronchial asthma and chronic rheumatoid arthritis. Administration of SS-d increased plasma adrenocorticotropin (ACTH) levels, proopiomelanocortin mRNA levels in the anterior pituitary and the CRF mRNA level in the rat hypothalamus in a dose dependent manner. In contrast, treatment with SS-a failed to affect the levels of these molecular markers. While this study indicates that administration of SS-d may have an important role in saiko agents-induced CRF release and CRF gene expression in rat hypothalamus, it fails to address the molecular effect of the herbal medication as a whole.

Kojima *et al.* (1998, Biol. Pharm. Bull. 4:426-428) describe the utilization of differential display of mRNA to isolate and identify genes transcriptionally regulated in

mouse liver by sho-saiko-to, an herbal medicine used for treating various inflammatory diseases in Japan. These researchers limited their study to the use of mRNA differential display techniques in investigating the molecular mechanisms of herbal medicine. It also failed to address effects in multiple organs of treated animals and did not provide any guidance for quality control, new use, and standardization of effects. In addition, the study failed to analyze the individual components of the herb and compare the individual results with the results obtained using the whole herbal mixture.

Ma Ji *et al.* (1998, Chinese Medical Journal 111(1):17-23) investigated the therapeutic effect of the herb *Astragali membranaceus* on sodium and water retention in rats experiencing aortocaval fistula-caused experimental congestive heart failure. Chronic heart failure rats with and without *Astraglia* treatment were compared for changes in various morphological characteristics (*e.g.*, body weight, serum sodium concentration); physiological characteristics (*e.g.*, mean arterial pressure, heart rate, hematocrit and plasma osmolality); mRNA expression levels (*e.g.*, hypothalamic arginine vasopressin (AVP), AVP V_1a receptor, renal AVP V_2 receptor, aquaporin-2 (AXP2)) and protein excretion (e.g., plasma atrial monophosphate peptide (ANP) and urinary cyclic guanidino monophosphate (cGMP)). The researchers found that treatment with *Astraglia* improved cardiac and renal functions, partially corrected abnormal mRNA expressions of the AVP system and AQP2, and improved the renal reaction to ANP. This study did not address using the collected data to guide the development of new formulations or for elucidating the synergistic or other interactions among various herbs in a formula, or validate the differential power of the effects for quality control purposes.

As shown by the above review of relevant scientific articles, molecular-based technology has not been used to explore and validate cellular and molecular responses in biological systems that are treated or challenged with multiple chemicals at the same time, such as herbal medicines and TCM. Furthermore, these recent advances have not been integrated with other technologies and methods to produce a process for the

systematic exploration of biological effects of herbal medicines and TCM.

SUMMARY OF THE INVENTION

This invention provides the tools and methodologies for creating, maintaining, improving and utilizing Herbal BioResponse Arrays (HBR Arrays), wherein the HBR Arrays constitute data sets associated with particular herbal compositions. The HBR Arrays of the present invention may include information on the plant-related parameters of the herbal constituents, marker information collected following the exposure of a biosystem to the herbal composition, and biological response information collected following the exposure of a biosystem to the herbal composition.

The present invention provides the tools and methodologies necessary for establishing standardized HBR Arrays for particular herbal compositions, wherein the standardized HBR Arrays are used as benchmarks by which to evaluate batches of similar or different herbal compositions. The present invention further provides the tools and methodologies necessary to update and maintain the standardized HBR Arrays. Particular embodiments of the present invention involve iterative processes whereby data for additional batches of the herbal composition, additional plant-related data, additional marker information, and/or additional BioResponse information is periodically added to the standardized HBR Arrays. Thus, the present invention provides the tools and methodologies for creating, maintaining, updating and using HBR Arrays on an ongoing basis.

The present invention provides the tools and methodologies necessary to guide the standardization of herbal compositions, to determine which specific components of herbal compositions are responsible for particular biological activities, to predict the biological activities of herbal compositions, for the development of improved herbal therapeutics; for adjusting or modifying an herbal composition; for identifying specific molecules in the batch herbal composition which retain the desired biological activity;

for determining which herbal components of a known herbal composition can be eliminated from the known herbal composition while maintaining or improving the desired biological activity of the known herbal composition; for identifying new uses and previously unknown biological activities for the batch herbal composition; and for using the predicted biological activity of the batch herbal composition to aid in the design of therapeutics which include herbal components and synthetic chemical drugs, including the design of therapeutics using the methods of combinatorial chemistry.

More specifically, the present invention provides methods of establishing standardized Herbal BioResponse Arrays (HBR Arrays) for herbal compositions, wherein the methods comprise:

1). selecting an herbal composition with at least one known BioResponse;

2). exposing a biosystem to a batch of the herbal composition and collecting data on two or more markers;

3). storing the marker data of step 2) as an HBR Array;

4). repeating steps 2) and 3) for one or more additional batches of the herbal composition using two or more of the same or different markers than used in step 2);

5). combining the HBR Arrays obtained in steps 3) and 4); and,

6). analyzing the combined HBR Array of step 5) to generate a standardized HBR Array for the herbal composition, wherein the standardized HBR Array has data for two or more markers which are correlated with at least one known BioResponse of the herbal composition. The present invention further provides such methods which further comprise exposing a biosystem to one or more batches of the herbal composition, collecting the data on one or more BioResponses, and adding the collected BioResponse data to the standardized HBR Array for that herbal composition.

The present invention provides methods of evaluating herbal compositions, wherein the methods comprise exposing a biosystem to a batch of the herbal composition and collecting data on two or more markers; and comparing the collected

marker data with a standardized HBR Array for the same or a substantially same herbal composition as that of the batch herbal compositions.

The present invention provides a system for predicting the biological activity of an herbal composition comprising:

1). a biosystem comprising one or more different types of cells, tissues, organs or *in vitro* assays;

2). a batch herbal composition;

3). two or more molecular markers;

4). a means for exposing the biosystem to the batch herbal composition and measuring the differential responses of the molecular markers;

5). a computer processor, including memory, for analyzing and storing the differential response measurements of the molecular markers so as to create an Herbal BioResponse Array (HBR Array) data set for the batch herbal composition;

6). a computer processor, including memory, for comparing the HBR Array of the batch herbal composition to one or more previously-stored HBR Arrays so as to predict the biological activity of the batch herbal composition, wherein the biological activities of the herbal compositions used to generate the one or more previously-stored HBR Arrays are known.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Figure 1 provides a schematic of the basic method steps for constructing a Standardized Herbal BioResponse Array (HBR Array) for any selected herbal composition. The figure is shown in its most basic form for ease of understanding. As discussed herein, each of the pathways of the schematic can be done iteratively. Furthermore, any information contained in one box can be used to guide decisions regarding gathering information for any other box. In this way, numerous feedback loops are also possible throughout the scheme.

Figure 2. Figure 2 provides a schematic of the basic method steps for constructing a an Herbal BioResponse Array (HBR Array) for any batch herbal composition and for comparing this batch HRB Array to a selected subset of information from the Standardized HBR Array. The figure is shown in its most basic form for ease of understanding. As discussed herein, each of the pathways of the schematic can be done iteratively. Furthermore, any information contained in one box can be used to guide decisions regarding gathering information for any other box. In this way, numerous feedback loops are possible throughout the scheme.

Figure 3. Figure 3 provides a schematic of the basic method steps for establishing and using a major data set. The figure is shown in its most basic form for ease of understanding. As discussed herein, each of the pathways of the schematic can be done iteratively. Furthermore, any information contained in one box can be used to guide decisions regarding gathering information for any other box. In this way, numerous feedback loops are possible throughout the scheme.

Figure 4. Western blot for various herbal compositions.

A. No herbal composition.
B. Huang Qing Tang A (HQT A) (0.2 mg/ml).
C. HQT A (4 mg/ml).
D. HQT B (0.2 mg/ml).
E. HQT B (4 mg/ml).
F. Scute (0.2 mg/ml).
G. Scute (4 mg/ml).

Figure 5. HPLC for *Paeonie lactiflora pallus*.

Figure 6. HPLC for *Ziziphi fructus*.

DETAILED DESCRIPTION OF THE INVENTION

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described.

Overview of the Invention

As set forth above, the present invention is directed to tools and methods useful for predicting the biological response of an herbal composition. More particularly, this invention provides methods of creating Herbal BioResponse Array (HBR Array) databases as well as methods for using such databases to improve the design of effective herbal-based therapeutics. The goal of the present invention is the overall design, creation, improvement and use of HBR Arrays for the preparation, testing and administration of herbal compositions, and guide development of new herbal compositions and novel uses of existing herbal compositions.

Phytomics. As used herein, depending on the context in which it is used, "phytomics" refers to using bioinformatics and statistical approaches to address the qualitative and quantitative aspects of the components of herbal compositions or to the actual data bases which are developed for addressing such aspects.

Herbal BioResponse Array. As used herein, an HBR Array constitutes a data set of two or more observations or measurements associated with an herbal composition. The HBR Array may include qualitative and quantitative data on the plants in the composition (plant-related data), marker information obtained after exposure of a biosystem to the herbal composition including a dose dependent study, and BioResponse data obtained after exposure of a biosystem to the herbal composition. The data in any particular HBR Array can be statistically analyzed in either 2- or 3-dimensional space.

HBR Arrays may be designated as batch HBR Arrays and standardized HBR Arrays. Batch HBR Arrays are arrays of data associated with specific batches of an herbal composition. Standardized HBR Arrays are arrays of data associated with a standardized herbal composition.

Major Data Set. As used herein, the term "major data set" refers to the data set which acts as the baseline set of data by which various other sets of data are compared or otherwise analyzed for the same or different herbal compositions. Generally, the major data set is created using biotechnological techniques to ascertain some genetic or protein aspect of the herbal compositions. Thus, the major data set will usually, but not always, be based on a genomic or proteomic set of data. For example, DNA microarray results could be the major data set which is used to compare to other, dependent or minor data sets.

Minor or Dependent Data Set. As used herein, the "minor data set" or "dependent data set" refers to one or more data sets which are used for comparing to the major data set. Generally, but not always, the minor data set will consist of information on an herbal composition which are collected by more traditional methods. For example, the minor, or dependent, data set may consist of a collection of plant-related data obtained by more conventional means. Examples of plant-related data include, but are not limited to, the genus/species of the herb(s) in the herbal composition, the particular plant parts of the herb(s) in the composition and the geographic location where the herb(s) were located. Another example of a minor data set might consist of a set of biological responses of a cell, tissue, organ or organism after treatment with one or more different amounts of the herbal composition. Examples of such biological data or a whole organism may include, but are not limited to, cell toxicity studies, enzyme treatment studies, growth rates, weight gain or loss, changes in motor skills and changes in mental abilities.

Herb. Technically speaking an herb is a small, non-woody (i.e., fleshy stemmed), annual or perennial seed-bearing plant in which all the aerial parts die back

at the end of each growing season. Herbs are valued for their medicinal, savory or aromatic qualities. As the word is more generally used and as the word is used herein, an "herb" refers to any plant or plant part which has a food supplement, medicinal, drug, therapeutic or life-enhancing use. Thus, as used herein, an herb is not limited to the botanical definition of an herb but rather to any botanical, plant or plant part used for such purposes, including any plant or plant part of any plant species or subspecies of the Metaphyta kingdom, including herbs, shrubs, subshrubs, and trees. Plant parts used in herbal compositions include, but are not limited to, seeds, leaves, stems, twigs, branches, buds, flowers, bulbs, corms, tubers, rhizomes, runners, roots, fruits, cones, berries, cambium and bark.

Herbal Composition. As used herein, an "herbal composition" refers to any composition which includes herbs, herbal plants or herbal plant parts. Thus, as used herein, an herbal composition is any herbal preparation, including herbal food supplements, herbal medicines, herbal drugs and medical foods. Examples of herbal compositions include, but are not limited to, the following components: a whole plant or a plant part of a single plant species; whole plants or plant parts of multiple plant species; multiple components derived from a single plant species; multiple components derived from multiple plant species; or any combination of these various components. For a thorough review of various herbal compositions, see, for example, Kee Chang Huang, The Pharmacology of Chinese Herbs, CRC Press (1993), herein incorporated in its entirety. Representative examples of various herbal compositions are provided in the following paragraphs.

Herbal compositions which include the bark of the willow tree have been used to treat fever since the mid-eighteenth century in England. The active ingredient in willow bark is a bitter glycoside called salicin, which on hydrolysis yields glucose and salicylic alcohol. Aspirin (acetylsalicylic acid) and aspirin-like drugs (e.g., ibuprofen), all of which are often called nonsteroidal antiinflammatory drugs (NSAIDs), are frequently used to treat pain, fever, and inflammation. Meadowsweet is another herb

that contains salicylates. Treatment of arthritic and arthritic-like symptoms with willow bark or meadowsweet requires the consumption of large quantities of herbal teas made from these plants. The entire *Populus* species (i.e., poplar trees and shrubs) also contains salicylate precursors and poplar-buds have been used in antiinflammatory, antipyretic and analgesic medications.

U.S. Patents have been issued for herbal compositions used for the treatment of various diseases and other health-related problems afflicting humans and animals. For example, U.S. Patent No. 5,417,979 discloses a composition comprising a mixture of herbs, including species of *Stephania* and *Glycyrrhiza*, as well as their extracts, which is used as an appetite stimulant and for the treatment of pain. Herbal compositions which include *Glycyrrhiza uralensis* have been found useful for treating eczema, psoriasis, pruritis and inflammatory reactions of the skin (U.S. Patent No. 5,466,452). U.S. Patent No. 5,595,743 discloses various herbal compositions which include licorice extract (*Glycyrrhiza*) and siegesbeckia, sophora, stemona and tetrandra herbs used for the treatment of various mammalian diseases, including inflammation and rheumatoid arthritis. Ocular inflammation can be treated with a pharmaceutical composition containing the plant alkaloid tetrandrine (U.S. Patent No. 5,627,195). U.S. Patent No. 5,683,697 discloses a pharmaceutical composition having anti-inflammatory, anti-fever, expectorant or anti-tussive action, wherein the composition includes plant parts from the species *Melia, Angepica, Dendrobium, Impatiens, Citrus, Loranthus, Celosia, Cynanchum* and *Glehnia*. An herbal composition which includes extracts of the roots, rhizomes, and/or vegetation of *Alphinia, Smilax, Tinospora, Tribulus, Withania* and *Zingiber* has been found to reduce or alleviate the symptoms associated with rheumatoid arthritis, osteoarthritis, reactive arthritis and for reducing the production of proinflammatory cytokines (U.S. Patent No. 5,683,698).

Herbal compositions are available in many forms, including capsules, tablets, or coated tablets; pellets; extracts or tinctures; powders; fresh or dried plants or plant parts; prepared teas; juices; creams and ointments; essential oils; or, as combinations of

any of these forms. Herbal medicines are administered by any one of various methods, including orally, rectally, parenterally, enterally, transdermally, intravenously, via feeding tubes, and topically.

Herbal compositions encompassed by the present invention include herbal compositions which also contain non-herbal components. Examples of such non-herbal components include, but are not limited to, whole insects and insect parts, worms, animal or insect feces, natural or petroleum oils, carbonate of ammonia, salt of tartar, liquor, water, glycerin, steroids, pharmaceuticals, vitamins, nutrient extracts, whey, salts, and gelatin.

For oral administration, the herbal compositions disclosed may take the form of, for example, tablets or capsules prepared by conventional means in admixture with generally acceptable excipients such as binding agents (*e.g.*, pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (*e.g.*, lactose, microcrystalline cellulose or calcium phosphate); lubricants (*e.g.*, magnesium stearate, talc or silica); disintegrants (*e.g.*, potato starch or sodium starch glycolate); or wetting agents (*e.g.*, sodium lauryl sulphate); glidants; artificial and natural flavors and sweeteners; artificial or natural colors and dyes; and solubilizers. The herbal compositions may be additionally formulated to release the active agents in a time-release manner as is known in the art and as discussed in U.S. Patent Nos. 4,690,825 and 5,055,300. The tablets may be coated by methods well known in the art.

Liquid preparations for oral administration may take the form of, for example, solutions, syrups, suspensions, or slurries (such as the liquid nutritional supplements described in Mulchandani *et al.*, 1992 U.S. Patent No. 5,108,767), or they may be presented as a dry product for reconstitution with water or other suitable vehicles before use. Liquid preparations of folic acid, and other vitamins and minerals may come in the form of a liquid nutritional supplement specifically designed for ESRD patients. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (*e.g.*, sorbitol syrup,

methyl cellulose or hydrogenated edible fats); emulsifying agents (*e.g.*, lecithin or acacia); non-aqueous vehicles (*e.g.*, almond oil, oily esters or ethyl alcohol); preservatives (*e.g.*, methyl or propyl p-hydroxybenzoates or sorbic acid); and artificial or natural colors and/or sweeteners.

For topical administration, herbal components may be combined in admixture with at least one other ingredient constituting an acceptable carrier, diluent or excipient in order to provide a composition, such as a cream, gel, solid, paste, salve, powder, lotion, liquid, aerosol treatment, or the like, which is most suitable for topical application. Sterile distilled water alone and simple cream, ointment and gel bases may be employed as carriers of the herbal components. Preservatives and buffers may also be added. The formulation may be applied to a sterile dressing, biodegradable, absorbable patches or dressings for topical application, or to slow release implant systems with a high initial release decaying to slow release.

For a more complete overview and discussion of herbal-based compositions see Earl Mindell, Earl Mindell's Herb Bible, Simon & Schuster (1992); Culpeper's Complete Herbal, W. Foulsham & Co., Ltd. (originally published in the mid 1600's); and, Rodale's Illustrated Encyclopedia of Herbs, Rodale Press (1987).

Standardized Herbal Composition. As used herein, a "standardized herbal composition" refers to a particular herbal composition which is chosen as the standard herbal composition for evaluating batch herbal compositions which have the same, similar or different components as the components of the standardized herbal composition. Sometimes herein also referred to as the "master herbal composition." Standardized herbal compositions are generally herbal compositions which have been well characterized and which demonstrate the desired biological responses in a particular biosystem. Standardized herbal compositions are usually standardized by chemical tests well known to one skilled in the art and are properly stored for long term usage and reference. The standardized herbal composition is used to establish a standardized HBR Array based on observations and measurements for the plants (i.e.,

plant-related data), markers and BioResponses so as to characterize the herbal composition.

Batch Herbal Composition. As used herein, a "batch herbal composition" refers to any test herbal composition which is used to establish a HBR Array based on observations and measurements for the plants and markers so as to characterize the herbal composition. Sometimes herein also referred to as a "test" or "batch" herbal composition. Observations and measurements of BioResponses may or may not be included. The herbal compositions used to establish the standardized herbal composition may also be referred to as "batch herbal compositions" until designated as "standardized herbal compositions."

Batch. As used herein, a "batch" refers to a particular quantity of an herbal composition which can be identified as to some particular attribute so as to distinguish it from any other particular quantity of that same herbal composition. For example, one batch of an herbal composition may differ from another batch of that same herbal composition in that one of the batches was harvested at a different time or in a different geographical location than the other batch. Other differences that distinguish particular batches may include, but are not limited to, the following: 1) the particular plant part used (*e.g.*, the root of an herb was used in one batch while the leaves of that same herb were used in a different batch); 2) the post-harvest treatment of the individual herbs or herbal composition (*e.g.*, one batch may be processed with distilled water while a different batch may be processed with Hydrogen Chloride to simulate the acidity of the human stomach); and, 3) the relative proportions of the individual herbs in an herbal composition (*e.g.*, one batch may have equal parts by weight or volume of three different herbs while another batch has proportionally more of one herb than the other two).

Biosystem. As used herein, a "biosystem" refers to any biological entity for which biological responses may be observed or measured. Thus, a biosystem includes, but is not limited to, any cell, tissue, organ, whole organism or *in vitro* assay.

Biological Activity. As used herein, the "biological activity" of an herb refers to the specific biological effect peculiar to an herbal composition on a given biosystem.

Plant-Related Data. As used herein, "Plant-related data" refers to the data collected on the herbal composition including, but not limited to, data about the plants, their growing conditions and the handling of the plants during and after harvesting. The plant-related data also includes the relative proportions of the components in an herbal compositions, wherein the components may be different plant parts, different plant species, other non-plant ingredients (*e.g.*, insect parts, chemical drugs) or any combinations of these variables.

Plant-related data which may be gathered for an herbal composition includes, but is not limited to, the following: 1) the plant species (and, if available, the specific plant variety, cultivar, clone, line, etc.) and specific plant parts used in the composition; 2) the geographic origin of the herbs, including the longitude/latitude and elevation; 3) the growth conditions of the herbs, including fertilizer types and amounts, amounts and times of rainfall and irrigation, average microEinsteins received per day, pesticide usage, including herbicides, insecticides, miticides and fungicides, and tillage methods; 4) methods and conditions used for processing the herbs, including age/maturity of the herbs, soaking times, drying times, extraction methods and grinding methods; and 5) storing methods and conditions for the herbal components and the final herbal composition.

Additionally, the standardized herbal composition may be analyzed chemically. Chemical characterization may be accomplished by any chemical analysis method generally known by one skilled in the art. Examples of applicable chemical analyses include, but are not limited to, HPLC, TLC, chemical fingerprinting, mass spectrophotometer analyses and gas chromatography.

Bioinformatics. As used herein, "bioinformatics" refers to the use and organization of information of biological interest. Bioinformatics covers, among other things, the following: (1) data acquisition and analysis; (2) database development; (3)

integration and links; and (4) further analysis of the resulting database. Nearly all bioinformatics resources were developed as public domain freeware until the early 1990s, and much is still available free over the Internet. Some companies have developed proprietary databases or analytical software.

Genomic or Genomics. As used herein, the term "genomics" refers to the study of genes and their function. Genomics emphasizes the integration of basic and applied research in comparative gene mapping, molecular cloning, large-scale restriction mapping, and DNA sequencing and computational analysis. Genetic information is extracted using fundamental techniques, such as DNA sequencing, protein sequencing and PCR.

Gene function is determined (1) by analyzing the effects of DNA mutations in genes on normal development and health of the cell, tissue, organ or organism; (2) by analyzing a variety of signals encoded in the DNA sequence; and (3) by studying the proteins produced by a gene or system of related genes.

Proteomic or Proteomics. As used herein, the term "proteomics", also called "proteome research" or "phenome", refers to the quantitative protein expression pattern of a genome under defined conditions. As used generally, proteomics refers to methods of high throughput, automated analysis using protein biochemistry.

Conducting proteome research in addition to genome research is necessary for a number of reasons. First, the level of gene expression does not necessarily represent the amount of active protein in a cell. Also, the gene sequence does not describe post-tranlsational modifications which are essential for the function and activity of a protein. In addition, the genome itself does not describe the dynamic cell processes which alter the protein level either up or down.

Proteome programs seek to characterize all the proteins in a cell, identifying at least part of their amino acid sequence of a isolated protein. In general, the proteins are first separated using 2D gels or HPLC and then the peptides or proteins are sequenced using high throughput mass spectrometry. Using a computer, the output of the mass

spectrometry can be analyzed so as to link a gene and the particular protein for which it codes. This overall process is sometimes referred to as "functional genomics". A number of commercial ventures now offer proteomic services (*e.g.*, Pharmaceutical Proteomics™, The ProteinChip™ System from Ciphergen Biosystem; PerSeptive Biosystems).

For general information on proteome research, see, for example, J.S. Fruton, 1999, Proteins, Enzymes, Genes: The Interplay of Chemistry and Biology, Yale Univ. Pr.; Wilkins et al., 1997, Proteome Research: New Frontiers in Functional Genomics (Principles and Practice), Springer Verlag; A.J. Link, 1999, 2-D Proteome Analysis Protocals (Methods in Molecular Biology, *112*, Humana Pr.; Kamp et al., 1999, Proteome and Protein Analysis, Springer Verlag.

Signal Transduction. As used herein, "signal transduction", also known as cellular signal transduction, refers to the pathways through which cells receive external signals and transmit, amplify and direct them internally. Signaling pathways require intercommunicating chains of proteins that transmit the signal in a step wise fashion. Protein kinases often participate in this cascade of reactions, since many signal transductions involve receiving an extracellular chemical signal, which triggers the phosphorylation of cytoplasmic proteins to amplify the signal.

Post-translational Modification. As used herein, "post-translational modification" is a blanket term used to cover the alterations that happen to a protein after it has been synthesized as a primary polypeptide. Such post-translational modifications include, but are not limited to, glycosylation, removal of the N-terminal methionine (or N-formyl methionine), signal peptide removal, acetylation, formylation, amino acid modifications, internal cleavage of peptide chains to release smaller proteins or peptides, phosphorylation, and modification of methionine.

Array or Microarray. As used herein, an "array" or "microarray" refers to a grid system which has each position or probe cell occupied by a defined nucleic acid fragment. The arrays themselves are sometimes referred to as "chips", "biochips",

"DNA chips" or "gene chips". High-density DNA microarrays often have thousands of probe cells in a variety of grid styles.

Once the array is fabricated, a batch is added and some form of chemistry occurs between the batch and the array to give some recognition pattern which particular to that array and batch. Autoradiography of radiolabled batches is a traditional detection strategy, but other options are available, including electronic signal transduction.

Markers. As used herein, the term "markers", refers to any biological-based measurement or observation for a particular herbal composition that is characteristic of a particular biosystem which is being exposed to a particular batch of an herbal composition. The term "marker" encompasses both qualitative and qualitative measurements and observations of a biosystem. The marker database constitutes a data set which characterizes gene expression patterns in response to herbal therapies, wherein the patterns show which genes are turned on, off, up or down in response to specific herbal compositions. Thus, "markers" refers to any biologically-based measurement or observation whose up- and down- or temporal regulations, or qualitative or quantitative changes of expression levels in a biosystem are used to characterize differential biological responses of a biosystem to an herbal composition.

The particular batch of an herbal composition to which the biosystem is exposed may be an unknown herbal composition, a known herbal composition, or a standardized herbal composition. Examples of markers useful in accomplishing the present invention include, but are not limited to, molecular markers, cytogenetic markers, biochemical markers or macromolecular markers. Macromolecular markers include, but are not limited to, enzymes, polypeptides, peptides, sugars, antibodies, DNA, RNA, proteins (both translational proteins and post-translational proteins), nucleic acids, polysaccarides.

Any marker that satisfies the definition of "marker" herein is appropriate for conducting the present invention. The term "markers" includes related, alternative

terms, such as "biomarker" or "genetic marker" or "gene marker." There may be one or more primary markers along with secondary markers, or a hierarchy of markers for achieving the purposes of increasing the discriminating power of a HBR array. Thus, selected molecular markers may be combined with various other molecular, cytogenetic, biochemical or macromolecular markers to enable an even more accurate, extended HBR Array.

A molecular marker comprises one or more microscopic molecules from one or more classes of molecular compounds, such as DNA, RNA, cDNA, nucleic acid fragments, proteins, protein fragments, lipids, fatty acids, carbohydrates, and glycoproteins.

The establishment, generation and use of applicable molecular markers are well known to one skilled in the art. Examples of particularly useful technologies for the characterization of molecular markers include differential display, reverse transcriptase polymerase chain reactions (PCR), large-scale sequencing of expressed sequence tags (ESTs), serial analysis of gene expression (SAGE), Western immunoblot or 2D, 3D study of proteins, and microarray technology. One skilled in the art of molecular marker technology is familiar with the methods and uses of such technology (see, *e.g.*, Bernard R. Glick and Jack J. Pasternak, Molecular Biotechnology, Principles and Applications of Recombinant DNA, Second Edition, ASM Press (1998); Mathew R. Walker and Ralph Rapley, Route Maps in Gene Technology, Blackwell Science (1997); Roe *et al.*, DNA Isolation and Sequencing, John Wiley & Sons (1996)James D. Watson *et al.*, Recombinant DNA, Second Edition, Scientific American Books (1992)).

DNA, RNA and protein isolation and sequencing methods are well known to those skilled in the art. Examples of such well known techniques can be found in Molecular Cloning: A Laboratory Manual 2nd Edition, Sambrook et al, Cold Spring Harbor, N.Y. (1989); Hanspeter Saluz and J. P. Jost, A Laboratory Guide to Genomic Sequencing: The Direct Sequencing of Native Uncloned DNA (Biomethods Vol 1), Birkhauser (1988); and B. Roe et al., DNA Isolation and Sequencing, Wiley (1996).

Examples of conventional molecular biology techniques include, but are not limited to, *in vitro* ligation, restriction endonuclease digestion, PCR, cellular transformation, hybridization, electrophoresis, DNA sequencing, cell culture, and the like. Specific kits and tools available commercially for use in the present invention include, but are not limited to, those useful for RNA isolation, PCR cDNA library construction, retroviral expression libraries, vectors, gene expression analyses, protein antibody purification, cytotoxicity assays, protein expression and purification, and high-throughput plasmid purification (see, *e.g.*, CLONTECHniques product catalog, XIII(3), 1-32 (1998) or www.clontech.com; Atlas™ cDNA Expression Assays product catalog (1998); SIGMA® product catalog (1997)).

For discussions, methodologies and applications of oligonucleotide arrays, microarrays, DNA chips or biochips, see, for example, U.S. Patent Numbers 5,445,934, 5,605,662, 5,631,134, 5,736,257, 5,741,644, 5,744,305, 5,795,714; Schena et al., Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes, Proc. Natl. Acad. Sci. USA 93, 10614-10619 (1996); DeRisi et al., Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale, Science 278, 680-686 (1997); Wodicka, et al., Genome-wide Expression Monitoring in *Saccharomyces cerevisiae*, Nature Biotechnology 15, 1359-1367 (1997); Pardee, Complete Genome Expression Monitoring: The Human Race, Nature Biotechnology 15, 1343-1344 (1997); Schafer et al., DNA Variation and the Future of Human Genetics, Nature Biotechnology 16, 33-39 (1998); DeRisi et al., Use of a cDNA Microarray to Analyze Gene Expression Patterns in Human Cancer, Nature Genetics 14, 457-460 (1996); Heller et al., Discovery and Analysis of Inflammatory Disease-Related Genes Using cDNA Microarrays, Proc. Natl. Acad. Sci. USA 94, 2150-2155 (1997); Marshall et al., DNA Chips: An Array of Possibilities, Nature Biotechnology 16, 27-31 (1998); Schena et al., Microarrays: Biotechnology's Discovery Platform for Functional Genomics, Tibtech 16, 301-306 (1998); Ramsay, DNA Chips: State-of-the-art, Nature Biotechnology 16, 40-44 (1998); Chee et al., Accessing Genetic

Information with High-Density DNA Arrays, Science 274, 610-614 (1996); and Chen et al., Profiling Expression Patterns and Isolating Differentially Expressed Genes by cDNA Microarray System with Colorimetry Detection, Genomics 50, 1-12 (1998); P. Andrew Outinen et al., Characterization of the stress-inducing effects of homocysteine, Biochem. J. 332, 213-221 (1998); and Gelbert et al., Will genetics really revolutionize the drug discovery process, Curr Opin Biotechnol 8(6), 669-674 (1997).

Other, more specific, references applicable to the instant invention include, but are not limited to, those addressing the expression technologies, such as ESTs (see, *e.g.*, Michael R. Fannon, Gene expression in normal and disease states – identification of therapeutic targets, TIBTECH 14, 294-298 (1996)); the generation of protein profiles (see, *e.g.*, Robinson et al., A Tyrosine Kinase Profile of Prostate Carcinoma, Proc. Natl. Acad. Sci. USA 93, 5958-5962 (1996)); chemical and spectroscopic methods for identifying components of herbal compositions (Kojima et al., Saponins from Gliricidia sepium, Phytochemistry 48(5), 885-888 (1998)); the determination of functional antigens (see, *e.g.*, Aris Persidis, Functional antigenics, Nature Biotechnology 16, 305-307 (1998)); HPLCs (see, *e.g.*, Milton T. W. Hearn (Editor), HPLC of Proteins, Pepties, and Polynucleotides: Contemporary Topics and Applications (Analytical Techniques in Clinical Chemistry and Laboratory Manual), VCH Pub. (1991); electrophoresis (see, *e.g.*, Westermeier *et al.*, Electrophoresis in Practice: A Guide to Methods and Applications of DNA and Protein Separations, John Wiley & Sons (1997)); and cross-reactivity marker assays (see, *e.g.*, Irving Millman *et al.*, Woodchuck Hepatitis Virus: Experimental Infection and Natural Occurrence, Hepatology 4(5):817-823 (1984)). The use of structural genomics for solving the structures of all the proteins encoded for in completed genomes, wherein the methodology includes high-throughput direct structure determinations and computational methods, is discussed by Terry Gaasterland, Structural genomics: Bioinformatics in the driver's seat, Nature Biotechnology 16, 625-627. For bioinformatics methodologies, see, for example, Andreas Baxevanis (Editor),

Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, John Wiley & Sons (1998) and Luke Alphey, DNA Sequencing: From Experimental Methods to Bioinformatics (Introduction to Biotechniques Series), Springer Verlag (1997).

Cytogenetic parameters include, but are not limited to, karyotype analyses (*e.g.*, relative chromosome lengths, centromere positions, presence or absence of secondary constrictions), ideograms (*i.e.*, a diagrammatic representation of the karyotype of an organism), the behavior of chromosomes during mitosis and meiosis, chromosome staining and banding patterns, DNA-protein interactions (also known as nuclease protection assays), neutron scattering studies, rolling circles (A.M. Diegelman and E.T. Kool, Nucleic Acids Res 26(13):3235-3241 (1998); Backert *et al.*, Mol. Cell. Biol. 16(11):6285-6294 (1996); Skaliter *et al.*, J. Viol. 70(2):1132-1136 (1996); A. Fire and S.Q. Xu, Proc. Natl. Acad. Sci. USA 92(10):4641-4645 (1995)), and autoradiography of whole nuclei following incubation with radiolabelled ribonucleotides.

Biochemical parameters include, but are not limited to, specific pathway analyses, such as signal transduction, protein synthesis and transport, RNA transcription, cholesterol synthesis and degradation, glucogenesis and glycolysis.

Fingerprinting. As used herein, the term "fingerprinting" as used herein refers to the means of making a characteristic profile of a substance, particularly an herb, in order to identify it. The term "fingerprint" as used herein refers to the display of the result of the particular means employed for the fingerprinting.

Examples of the various types of fingerprinting means include, but are not limited to, DNA fingerprinting, protein fingerprinting, chemical fingerprinting and footprinting.

DNA fingerprinting, or profiling, refers to a way of making a unique pattern from the DNA of particular biological source (*e.g.*, a particular plant, plant species, genus of plant, plant part or plant tissue). The DNA fingerprint, or profile, can be used to distinguish that particular biological source from a different biological source. The pattern obtained by analyzing a batch using microarrays, oligonucletide arrays, DNA

chips or biochips are also referred to as "fingerprints".

Protein fingerprinting refers to generating a pattern of proteins in a cell, tissue, organ or organism, such as a plant, which provides a completely characteristic "fingerprint" of that cell, tissue, organ or organism at that time.

Chemical fingerprinting refers to the analysis of the low molecular weight chemicals in a cell and the resulting pattern used to identify a cell, tissue, organ or organism, such as a plant. The analysis is usually done using Gas Chromotography (GC), HPLC or mass spectrometry.

Footprinting refers to a method of finding how two molecules stick together. In the case of DNA, a protein is bound to a labeled piece of DNA, and then the DNA is broken down, by enzymes or by chemical attack. This process produces a "ladder" of fragments of all sizes. Where the DNA is protected by the bound protein it is degraded less, and so the "ladder" appears fainter. Footprinting is a common technique for homing in on where the proteins that regulate gene activity actually bind to the DNA.

The means, or methods, used to accomplish each type of fingerprinting are described in detail elsewhere herein.

BioResponses. As used herein, a "BioResponse" refers to any observation or measurement of a biological response of a biosystem following exposure to an herbal composition. Sometimes herein also referred to as a "biological effect." A BioResponse is a qualitative or quantitative data point for the biological activity of a particular herbal composition. BioResponse data includes both dosage and temporal information, wherein such information is well known to one skilled in the art of measuring responses of biosystems to various treatments. Thus, BioResponse data includes information on the specific biological response of a specific biosystem to a specific dosage of herbal composition administered in a particular manner for a specific period of time.

BioResponses include, but are not limited to, physiological responses, morphological responses, cognitive responses, motivational responses, autonomic

responses and post-translational modifications, such as signal transduction measurements. Many herbal compositions demonstrate more than one BioResponse (see, *e.g.*, Kee Chang Huang, The Pharmacology of Chinese Herbs, CRC Press (1993)). Some particular BioResponses may be included in more than one of the delineated groups or have aspects or components of the response that encompass more than one group. BioResponses applicable to the instant invention are well known to one skilled in the art. The following references are representative of the state of art in the field: Kee Chang Huang, The Pharmacology of Chinese Herbs, CRC Press (1993); Earl Mindell, Earl Mindell's Herb Bible, Simon & Schuster (1992); Goodman & Gilman's The Pharmacological Basis of Therapeutics, Ninth Edition, Joel G. Hardman, *et. al.* (eds.), McGraw Hill, Health Professions Division (1996); P. J. Bentley, Elements of pharmacology. A primer on drug action, Cambridge University Press (1981); P. T. Marshall and G. M. Hughes, Physiology of mammals and other vertebrates, Second Edition, Cambridge University Press (1980); Report of the Committee on Infectious Diseases, American Academy of Pediatrics (1991); Knut Schmidt-Nielsen, Animal Physiology: Adaptation and Environment, 5th Edition, Cambridge University Press (1997); Elain N. Marieb, Human Anatomy & Physiology, Addison-Wessley Pub. Co. (1997); William F. Ganong, Review of Medical Physiology (18th Ed), Appleton & Lange (1997); Arthur C. Guyton and John E. Hall, Textbook of Medical Physiology, W. B. Saunders Co. (1995).

A "physiological response" refers to any characteristic related to the physiology, or functioning, of a biosystem. Physiological responses on a cellular, tissue or organ level include, but are not limited to, temperature, blood flow rate, pulse rate, oxygen concentration, bioelectric potential, pH value, cholesterol levels, infection state (*e.g.*, viral, bacterial) and ion flux. Physiological responses on a whole organism basis include gastrointestinal functioning (*e.g.*, ulcers, upset stomach, indigestion, heartburn), reproductive tract functioning (*e.g.*, physiologically-based impotence, uterine cramping, menstrual cramps), excretory functions (*e.g.*, urinary tract problems,

kidney ailments, diarrhea, constipation), blood circulation (*e.g.*, hypertension, heart disorders), oxygen consumption, skeletal health (*e.g.*, osteoporosis), condition of the cartilage and connective tissues (*e.g.*, joint pain and inflammation), locomotion, eyesight (*e.g.*, myopia, blindness), muscle tone (*e.g.* wasting syndrome, muscle strains), presence or absence of pain, epidermal and dermal health (*e.g.*, skin irritation, itching, skin wounds), functioning of the endocrine system, cardiac functioning, nervous coordination, head-related health (*e.g.*, headaches, dizziness), age (*e.g.*, life span, longevity) and respiration (*e.g.*, congestion, respiratory ailments).

A "morphological response" refers to any characteristic related to the morphology, or the form and structure, of a biosystem following exposure to an herbal composition. Morphological responses, regardless of the type of biosystem, include, but are not limited to, size, weight, height, width, color, degree of inflammation, general appearance (*e.g.*, opaqueness, transparency, paleness), degree of wetness or dryness, presence or absence of cancerous growths, and the presence or lack of parasites or pests (*e.g.*, mice, lice, fleas). Morphological responses on a whole organism basis include, but are not limited to, the amount and location of hair growth (*e.g.*, hirsutism, baldness), presence or absence of wrinkles, type and degree of nail and skin growth, degree of blot clotting, presence or absence of sores or wounds, and presence or absence of hemorrhoids.

A "cognitive response" refers to any characteristic related to the cognitive, or mental state, of a biosystem following exposure to an herbal composition. Cognitive responses include, but are not limited to, perceiving, recognizing, conceiving, judging, memory, reasoning and imagining.

A "motivational response" refers to any characteristic related to the motivation, or induces action, of a biosystem following exposure to an herbal composition. Motivational responses include, but are not limited to, emotion (*e.g.*, cheerfulness), desire, learned drive, particular physiological needs (*e.g.*, appetite, sexual drive) or similar impulses that act as incitements to action (*e.g.*, stamina, sex drive).

An "autonomic response" refers to any characteristic related to autonomic responses of a biosystem following exposure to an herbal composition. Autonomic responses are related to the autonomic nervous system of the biosystem. Examples of autonomic responses include, but art not limited to, involuntary functioning (*e.g.*, nervousness, panic attacks), or physiological needs (*e.g.*, respiration, cardiac rhythm, hormone release, immune responses, insomnia, narcolepsy).

BioResponses of cells, tissues, organs and whole organisms treated with various herbal compositions or herbal components are well known in the herbal arts. For example, the herbal compositions Sairei-to (TJ-114), alismatis rhizoma (Japanese name 'Takusha') and hoelen (Japanese name 'Bukuryou') were each found to inhibit the synthesis and expression of endothelin-1 in rats (Hattori *et al.*, Sairei-to may inhibit the synthesis of endothelin-1 in nephritic glomeruli, Nippon Jinzo Gakkai Shi 39(2), 121-128 (1997)). Interleukin (IL)-1 alpha production was significantly promoted by treatment of cultured human epidermal keratinocytes with the herbal medicine Sho-saiko-to (Matsumoto *et al.*, Enhancement of interleukin-1 alpha mediated autocrine growth of cultured human keratinocytes by sho-saiko-to, Jpn J. Pharmacol 73(4), 333-336 (1997). Adding Sho-saiko-to to a culture of peripheral blood mononuclear cells obtained from healthy volunteers resulted in a dose-dependent increase in the production of granulocyte colony-stimulating factor (G-CSF) (Yamashiki *et al.*, Herbal medicine "sho-saiko-to" induces in vitro granulocyte colony-stimulating factor production on peripheral blood mononuclear cells, J Clin Lab Immunol 37(2), 83-90 (1992)). These researchers concluded that the administration of Sho-saiko-to may be useful for the treatment of chronic liver disease, malignant diseases and acute infectious diseases where G-CSF is efficacious. Plasminogen activator inhibitor type 1 (PAI-1)-specific mRNA expression decreased and tissue-type plasminogen activator (t-PA)-specific mRNA increased after treatment of human umbilical vein endothelial cells (HUVECs) with the saponin astragaloside IV (AS-IV) purified from the Chinese herb Astragalus membranaceus (Zhang *et al.*, Regulation of the fibrinolytic potential of

cultured human umbilical vein endothelial cells: astragaloside IV down regulates plasminogen activator inhibitor-1 and up regulates tissue-type plasminogen activator expression, J Vasc Res 34(4), 273-280 (1997)). One component out of four components isolated from the roots of Panax ginseng was found to be a potent inducer of IL-8 production by human monocytes and THP-1 cells, and this induction was accompanied by increased IL-8 mRNA expression (Sonoda *et al.*, Stimulation of interleukin-8 production by acidic polysaccharides from the root of panax ginseng, Immunopharmacology 38(3), 287-294 (1998)). By flow cytometric analysis, the expression of Fc gamma 11/111 receptors and complement receptor 3 (CR3) on macrophages were found to be increased by treatment with the Kampo-herbal medicine Toki-shakuyakusan (TSS) (Cyong, New BRM from kampo-herbal medicine, Nippon Yakurigaku Zasshi 110 Suppl 1, 87P-92P (1997)). Using computer image analysis, Chen *et al.* (Image analysis for intercellular adhesion molecule-1 expression in MRI/1pr mice: effects of Chinese herb medicine, Chung Hua I Hsueh Tsa Chih 75(4), 204-206 (1995)) found that the distribution intensity of intercellular adhesion molecule-1 (ICAM-1), immunoglobulins and C3 were significantly decreased in MRL/lpr mice after treatment with the Chinese herb stragalin. Western blot analysis showed that tetradrine, isolated from a natural Chinese herbal medicine, inhibited signal-induced NF-kappa B activation in rat alveolar macrophages (Chen *et al.*, Tetrandrine inhibits signal-induced NF-kappa B activation in rat alveolar macrophages, Biochem Biophys Res Commun 231(1), 99-102 (1997)).

Algorithm. As used herein, an "algorithm" refers to a step-by-step problem-solving procedure, especially an established, recursive computational procedure with a finite number of steps. Appropriate algorithms for two- and three-dimensional analyses of the plant-related, marker and BioResponse data sets are well known to one skilled in the computational arts. Such algorithms are useful in constructing the Herbal BioResponse Arrays of the present invention. For general information on algorithms, see, for example, Jerrod H. Zar, Biostatistical Analysis, second edition, Prentice Hall

(1984); Robert A. Schowengerdt, Techniques for image processing and classification in remote sensing, Academic Press (1983); Steven Gold et al., New Algorithms for 2D and 3D Point Matching: Pose Estimation and Correspondence, Pattern Recognition, 31(8):1019-1031 (1998); Berc Rustem, Algorithms for Nonlinear Programming and Multiple-Objective Decisions, Wiley-Interscience Series in Systems and Optimization, John Wiley & Sons (1998); Jeffrey H. Kingston, Algorithms and Data Structures: Design, Correctness, Analysis, International Computer Science Series, Addison-Wesley Pub. Co. (1997); Steven S. Skiena, The Algorithm Design Manual, Springer Verlag (1997); and Marcel F. Neuts, Algorithm Probability: A Collection of Problems (Stochastic Modeling), Chapman & Hall (1995). For information more specific to the application of algorithms to genetic-based data, see, for example, Dan Gusfield, Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology, Cambridge University Press (1997); Melanie Mitchell, An Introduction to Genetic Algorithms (Complex Adaptive Systems), MIT Press (1996); David E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wessley Pub. Co. (1989); Zbigniew Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, Springer Verlag (1996); Andre g. Uitterlinden and Jan Vijg, Two-Dimensional DNA Typing: A Parallel Approach to Genome Analysis, Ellis Horwood Series in Molecular Biology, Ellis Horwood Ltd. (1994); and Pierre Baldi and Soren Brunak, Bioinformatics: The Machine Learning Approach (Adaptive Computation and Machine Learning), MIT Press (1998).

Combinatorial Chemistry. As used herein, "combinatorial chemistry" refers to the numerous technologies used to create hundreds or thousands of chemical compounds, wherein each of the chemical compounds differ for one or more features, such as their shape, charge, and/or hydrophobic characteristics. Combinatorial chemistry can be utilized to generate compounds which are chemical variations of herbs or herbal components. Such compounds can be evaluated using the methods of the present invention.

Basic combinatorial chemistry concepts are well known to one of ordinary skill in the chemical arts and can also be found in Nicholas K. Terrett, Combinatorial Chemistry (Oxford Chemistry, Masters), Oxford Univ. Press (1998); Anthony W. Czarnik and Sheila Hobbs Dewitt (Editors), A Practical Guide to Combinatorial Chemistry, Amer. Chemical Society (1997); Stephen R. Wilson (Editor) and Anthony W. Czarnik (Contributor), Combinatorial Chemistry: Synthesis and Application, John Wiley & Sons (1997); Eric M. Gordon and James F. Kerwin (Editors), Combinatorial Chemistry and Molecular Diversity in Drug Discovery, Wiley-Liss (1998); Shmuel Cabilly (Editor), Combinatorial Peptide Library Protocols (Methods in Molecular Biology), Human Press (1997); John P. Devlin, High Throughput Screening, Marcel Dekker (1998); Larry Gold and Joseph Alper, Keeping pace with genomics through combinatorial chemistry, Nature Biotechnology 15, 297 (1997); Aris Persidis, Combinatorial chemistry, Nature Biotechnology 16, 691-693 (1998).

EXAMPLES

Example 1. Establishing a Standardized HBR Array for Selected Herbal Compositions.

The basic scheme for establishing a Standardized HBR Array is provided in Figure 1. Definitions of each component of the schematic are provided above.

Following selection of an herbal composition of interest, data is collected for various traits associated with the herbal composition, including, but not limited to plant-related characteristics and marker and BioResponse information.

Plant-related data includes, but is not limited to, the plant species, specific plant parts, geographic origin of the plants in the herbal composition, the growth conditions of the plants, the processing methods used to prepare the herbal components, storage methods and conditions, and various chemical analyses of the herbal composition. Marker information includes qualitative and quantitative data for markers collected after exposure of a biosystem to the herbal compost. Applicable markers include, but

are not limited to, molecular markers, cytogenetic markers, biochemical markers and macromolecular markers. BioResponse information includes qualitative and quantitative data for biological responses collected after exposure of a biosystem to the herbal composition.

Each type of data (*e.g.*, chemical, marker, BioResponse) can be obtained using one or more assays on the same, similar, substantially similar, or different batches of the herbal composition of interest. Such different assays can be conducted at the same or different times. In addition, data can be collected for the same or different markers at the same or different times. Similarly, BioResponse data can be collected for the same or different biological responses at the same or different times. Thus, collection of the data for the HBR Array is either collected at one time or collected on an on-going basis. Where a biosystem is exposed to an herbal composition so as to collect data. information is recorded on the administered dosages of the herbal composition as well as treatment times. BioResponse data may also consist of post-translational modifications, such as measurements of signal transduction.

After collection of two or more types of data (*e.g.*, data for two or more markers and a BioResponse; data for plant-related traits and data for a BioResponse), the data is analyzed using algorithms so as to create 2- and/or 3-dimensional Herbal BioResponse Arrays.

Various statistical parameters may be calculated for the HBR Array and may become part of the HBR Array data set. These statistical parameters may include, but are not limited to, means, standard deviations, correlation or match (or mismatch) matrices, ratios, regression coefficients, and transformed values (*e.g.*, arcsin percentage transformations of the raw data). Thus, the HBR Array may consist of the raw data as well as certain calculations, distributions, graphical presentations and other data manipulations associated with the raw data. Particular examples of such information include, but are not limited to, digital images, scatter graphs, cluster analyses and large scale gene expression profiles for marker data.

The total accumulated data and resultant analyses constitute a standardized HBR Array for the particular herbal composition used to establish the HBR Array data set. Due to the iterative nature of the process used to establish and maintain an HBR Array for an herbal composition, such arrays can be viewed as either static at any one point in time or dynamic over time.

The resulting analyses can identify subsets of the standardized HBR Arrays which are correlated (positively or negatively) or associated (*i.e.*, showing a general trend) with one or more specific biological activities of any particular herbal composition.

Example 2. Establishing a Batch HBR Array for Batch Herbal Compositions.

The basic scheme for establishing a HBR Array for a batch of an herbal composition is provided in Figure 2. Definitions of each component of the schematic are provided above. The procedure for establishing such an array is the same as that set forth immediately above for the standardized HBR Array.

Generally, the amount of data collected for a batch HBR Array will be less than that collected to establish a standardized HBR Array. However, data collected for a batch herbal composition may be added to an established HBR Array or used to establish a new standardized HBR Array.

Generally, the only data collected for a batch herbal composition is that data which has been found to be highly correlated or associated with the desired biological activities of the herbal composition being tested. For example, if it has been determined that a particular subset of plant-related and marker data is highly correlated to a desired biological activity of a particular herbal composition (based on the standardized HBR Array data and analyses discussed above), it is only necessary to test the batch herbal composition for that subset of traits in order to determine whether or not the batch has the desired biological activity. By comparing the data obtained for that subset of traits obtained from the batch (*i.e.*, the batch HBR Array) with the

standardized HBR Array for that particular herbal composition, one skilled in the art can determine whether or not that particular batch has the desired biological activity.

Example 3. Establishing and Using a Major Data Set.

The basic scheme for establishing and using a major data set for an herbal composition is provided in Figure 3. Definitions of each component of the schematic are provided above.

The first step is the establishment of a major data set for a selected herbal composition or batch herbal composition. This is accomplished by exposing a biosystem to the herbal composition and collecting the resultant marker information which will constitute the major data set. In most, but not instances, the major data set will consist of genomics and/or proteomics data in the form of an array, such as an array obtained with a DNA biochip.

Next, the major data set is analyzed to see if differential expression/results have been obtained for the tested herbal composition. Differential expression/results are necessary in order to generate meaningful algorithms in the next step. Examples of such differential expression/results include, but are not limited to, indications that certain genes are up- or down-regulated in response to exposure to the herbal composition or that the levels of certain proteins have been increased or decreased in response to the exposure.

If no meaningful or useful differential expression/results are obtained, then it is necessary to repeat the exposure and marker collection step. If it is believed that experimental error lead to the lack of a adequate result the first time then the exposure/data collection step can be repeated with all of the variables the same as the first time (*e.g.*, same biosystem, same marker set, same experimental protocol, etc.). However, it may be necessary to vary the biosystem sampling (*e.g.*, type of cell utilized, stage of cell growth), use a different marker set and/or change the experimental protocol in order to get differential expression/result..

Example 4. Using HBR Array Information.

The HBR Array information discussed herein can be used for many different purposes including, but not limited to, the following: 1) evaluating the components of an herbal composition; 2) predicting the BioResponse of an herbal composition; 3) determining which marker information is most highly correlated with a particular BioResponse of an herbal composition; 3) determining what data set of information (*i.e.*, plant-related data, marker data, and BioResponse data) is most correlated with a particular BioResponse of an herbal compost; 4) determining which type of biosystem is best for evaluating the biological activity of an herbal composition; 5) adjusting or changing the components of a herbal composition so that the HBR Array of that herbal composition corresponds to a standardized HBR Array for the same or substantially the same herbal composition; 6) adjusting or changing the components of an herbal composition so that the herbal composition will have the desired biological activity; 7) creating and updating standardized HBR Arrays; 8) identifying specific components (*e.g.*, plant parts, proteins, molecules) which retain the desired biological activity of an herbal composition; 9) determining which components of an herbal composition can be eliminated while maintaining or improving the desired biological activity of the herbal composition; 10) identifying one or more previously unknown biological activities for an herbal composition; 11) aiding in the design of therapeutics which include herbal and non-herbal components, such as chemically-synthesized drugs or pharmaceuticals and 12) utilizing the HBR Array information to complement combinatorial chemistry methods of designing therapeutics. Each of these embodiments of the present invention can be accomplished by one skilled in the applicable art using the methods and tools provided herein.

Example 5. Quality Control.

The HBR Array technology of the present invention is used to correlate or to determine a substantial equivalence of a specific batch of an herbal composition (single herb or multiple herbs of a formula) to a standardized, or master, batch of a same or substantial similar herbal composition. The HBR Arrays utilized in this process

include the acceptable range of quantitative variation for each of the biological effects (*i.e.*, BioResponse), and possibly a global score composed of weighted values assigned to each of the biological effects, which may consist of markers from multiple biochemical pathways of a biosystem.

"Data mining" refers to a process used to determine or select which subset of biological effects is the minimum number of biological effects required in any specific HBR Array. The information for data mining results from exposing a biosystem (*e.g.*, a cell line) in a dose dependent manner to a standardized herbal composition to establish a standardized HBR Array. This standardized HBR Array can then be compared to various HBR Arrays established for test herbal compositions. These test herbal compositions include, but are not limited to, different batches prepared at different dates; different batches prepared from raw herbs collected at different times; and different batches prepared from raw herbs collected at different locations.

Example 6. Improving an Herbal Composition or Identifying New Uses for an Herbal Composition.

HBR Arrays are generated by exposing biosystems to either extracts from individual herbs of a formula, or to extracts from the whole formula, and examining the biological effects of the extracts. The observed biological effects can be from multiple biochemical pathways of a biosystem and/or from multiple tissues of an animal, wherein various markers are evaluated for their corresponding qualitative and/or quantitative changes. The resulting HBR Arrays can be compared to novel HBR Arrays or to similar HBR Arrays from different herbal compositions or herbal compositions prepared by different processes. This procedure is useful for selecting a given set of biological effects and the minimum number of markers required to predict that a given batch herbal composition has the given set of biological effects.

In order to construct HBR Arrays, one skilled in the art utilizes various data mining tools including, but are not limited to, statistical analyses, artificial intelligence, and database research on neural work. The statistical methods of choice include, but

are not limited to, basic exploratory data analysis (EDA), graphic EDA (such as bushing) and multivariate exploratory techniques (*e.g.*, cluster analysis, discriminating factor analyses, stepwise linear on non-linear regression, classification tree) (see, *e.g.*, STATISTICA™, software packages from StatSoft, Tulsa, OK 74104; Tel: 918-749-1119; Fax: 918-749-2217; www.statsoft.com).

Data mining tools are used to explore large amounts of HBR Array data in search of constructing an HBR Array and consistent pattern within, between or among various HBR Arrays. The procedure consists of exploration, construction of an HBR array, and validation. This procedure is typically repeated iteratively until a robust HBR Array, or standardized HBR Array, is identified.

The foregoing detailed description has been given for clearness of understanding only and no unnecessary limitations should be understood therefrom as modifications will be obvious to those skilled in the art.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth and as follows in the scope of the appended claims.

Example 7. Establishing a Standardized HBR Array for Ginseng Recipes.

For the purposes of this example, standard ginseng is chosen to be *Panax Ginseng* C.A. Meyer G115 grown either in Manchuria or in Korea. The climate for growth is between -10 to +10°C with an annual rainfall of 50-100 cm (see Huang in The Pharmacology of Chinese Herbs, (1993) pp 21-45, CRC Press, Boca Raton, FL, fully incorporated by reference). Ginseng batches will first be characterized by geographic origin, species, plant part (e.g., rhizome, root, leaf skin, seed, bud and flower); growth conditions, processing methods and storage conditions both before and

after processing. Verification of chemical content for these batches will be performed by qualitative HPLC analysis for determination of ginsenoside saponins *(e.g.*, Ro, Ra1, Ra2, Rb1, Rb2, Rb3, Rc, Rg1, Rg2, Rd, Re, Rf, Rh1, Rh2, NG-R2 and Z-R1), including TLC qualitative analysis for lipophilic constituents (see, Elkin *et al.*, Chumg Kuo Yao Li Hsueh Pao (1993) 14: 97-100 and Yoshikawa *et al.*, Yakugaku Zasshi (1993) 113: 460-467). The saponin content of different herbs should be between 2.1 and 20.6% (by weight) depending on the species (see Table 1). These data will then be stored, preferably in the memory of a computer processor, for further manipulation.

Table 1. Saponin Content of Different Ginseng Herbs.*

Species	Total saponins (% by weight)
Panax ginseng C.A. Meyer	2.1-4.4%
Panax quiquefolius	4.9%
Panax notoginseng and *Panax japonica*	13.6-20.6%
Panax japonica var. major	9.34%

*from Huang in The Pharmacology of Chinese Herbs, (1993) page 29, CRC Press, Boca Raton, FL.

Expression biomarkers for standard ginseng (*i.e.*, G115) include the following: IL-8, IL-2, GM-CSF, NfkB, ICAM-1, interferon gamma, choline acetyl transferase, trk A, nerve growth factor (Kim *et al.*, Planta Med (1998) 64: 110-115; Sonoda *et al.*, Immunopharmacology (1998) 38: 287-294; Baum *et al.*, Eur J Appl Physiol (1997) 76: 165-169; Iwangawa *et al.*, Free Radic Biol Med (1998) 24: 1256-1268; Rhind *et al.*, Eur J Appl Physiol (1996) 74: 348-360). Alternatively, for a broader batch size, the 400,000 oligonucleotide group/1.6 cm^2 chip of Affimatrix can be used (U.S. Pat. No.5,556,752). The expression biomarkers for standard ginseng will be prepared by microarray technology of cDNA using either photolithography, mechanical microspotting or ink jet application (see Schena *et al.*, TIBTECH (1998) 16: 301-306). Selected sets of cells will be contacted with standard ginseng for varying periods of

time, under varying conditions to generate multiple microarray sets. The microarray sets will then be analyzed by hybridization-based expression monitoring of biochemical extracts via deduction of steady state mRNA levels from fluorescence intensity at each position on the microarrays (Schena *et al.*, Science (1995) 270: 467-470; Schena *et al.*, Proc Natl Acad Sci *USA* (1996) 93: 10614-10619; Lockhart *et al.* Nat Biotechnol (1996) 14: 1675-1680; DeRisi *et al.*, Nat Genet (1996) 14: 457-460; Heller *et al.*, Proc Natl Acad Sci USA (1997) 94: 2150-2155). The array data sets are then input into algorithms to generate statistical expression biomarker values for standard ginseng. Biochemical biomarkers for standard ginseng include quantitative analysis for increases in cycloheximide sensitive [^{3}H]-leucine incorporation proportional to protein synthesis and [^{3}H]-thymidine incorporation reflective of mitosis. (see Yamamoto *et al.*, Arzneimittelforschung (1977) 27: 1169-1173). For biochemical biomarkers, bone marrow cells will be contacted with standard ginseng for varying time periods under varying conditions in the presence of [^{3}H]-thymidine (for DNA synthesis) or in the presence and absence of cycloheximide and [^{3}H]-leucine (for protein synthesis) to perform multiple quantitative analysis of biochemical biomarkers (i.e., BBM sets). The BBM sets are then input into algorithms to generate statistical biochemical biomarker values for standard ginseng. Statistical data will then be stored, preferably in the memory of a computer processor, for further manipulation.

Biological response of a biosystem *(i.e.*, BioResponses) will be determined using cells and whole animals. For cells, ginseng batches will be exposed to specific cell types, including, but not limited to, fibroblasts, macrophages, monocytes, PMNL, LAK cells, B16-F10 melanoma cells, THP-1 cells and hippocampal neurons at a concentration of 0.5 mg/ml to 100 mg/ml. For animal treatments, 0.5-100 mg/kg of ginseng herbal extract will be administered orally, by intraperitoneal injection or subcutaneous injection.

To determine a biological response of a biosystem to standardized ginseng, human ovarian cancer cells will be inoculated into nude mice, which results in the

formation of palpable tumors. After tumor formation the mice will be treated by co-administration of cis-diamminecichloroplatinum and standard ginseng. Mice will be examined for tumor growth inhibition, increase in survival time and lowered adverse side-effects on hematocrit values and body weight (Nakata *et al.*, *Jpn J Cancer Res* (1998) 89:733-740). The assay will be repeated using various concentrations of standard ginseng to generate measures of central tendency, dispersion and variability for each variable.

The data collected will then be subjected to multidimensional analysis to generate multivariant normal distribution sets as a means of determining a baseline correlation between biological activity and standard ginseng (see Zar, J. H., in Biostatistical Analysis, 2nd ed. (1984), pp 328-360, Prentice Hall, Englewood Cliffs, NJ). A second independent determination of a biological response of a biosystem to standard ginseng will be the effect of standard ginseng on physical performance during exercise. Rats will be treated for 4 days with standard ginseng at various concentrations (between 0.5-100 mg/kg/day) and animals will be tested for increased plasma free fatty acid level and maintenance of glucose level during exercise at approximately 70% VO2max (see Wang *et al.*, Planta Med (1998) 64130-133). The data generated will be collected and then subjected to multidimensional analysis to generate multivariant normal distribution sets as a means of determining a baseline correlation between biological activity and standard ginseng (see Zar, J. H., in Biostatistical Analysis, 2nd ed. (1984), pp 328-360, Prentice Hall, Englewood Cliffs, NJ, Herein, fully incorporated by reference). The distribution sets for each BioResponse are then put into algorithms to generate statistical values for standard ginseng. Statistical data will then be stored, preferably in a memory of a computer processor, for further manipulation.

Each of these steps (*i.e.*, chemical analysis, generation of biomarker information and determination of responses of a biosystem) is reiterated to generate a large database of statistical values. These values are compiled and input into an algorithm to generate 2- and 3-dimensional Herbal Response Arrays (HBR Array) for

standardized ginseng. Through reiteration, the resulting arrays (*i.e.*, Standardized Arrays) display the highest correlation between composition (including growth conditions), biomarker information and biological response for standardized ginseng. By determining two or more known associated variables for composition and biomarker information values via display on an HBR Array for a test batch, the values for biological response variables can be predicted for the test batch by comparing test values against Standardized HBR Array values for standardized ginseng. The resulting prediction will be used to evaluate the quality of a given ginseng batch without necessitating the use of an observed biological response of a biosystem (see Example 2).

Example 8. Evaluation of a Selected Herbal Composition of Ginseng Using a Subset of Variables Correlated with a Specific Biological Response.

To evaluate the quality of a test batch herbal composition, data is first collected concerning the plant-related parameters for the herbs in the selected herbal composition (*e.g.*, plant species, plant parts, geographic origin, growth conditions, processing methods and storage conditions). The selected herbal composition is then manipulated such that chemical analysis can be performed to determine the chemical content of the herb (see Elkin *et al.*, Chumg Kuo Yao Li Hsueh Pao (1993) 14: 97-100 and Yoshikawa *et al.*, Yakugaku Zasshi (1993) 113: 460-467). Previously obtained ginseng data has demonstrated a strong correlation between oxygen consumption during aerobic exercise performance and the presence of a subset of saponin components, especially Rg1 and Rb1 (Wang *et al.*, Planta Med (1998) 64: 130-133).

The test batch is then exposed to test cells including, but not limited to, fibroblasts, macrophages, monocytes, PMNL, LAK cells, B16-F10 melanoma cells, THP-1 cells and hippocampal neurons at a concentration of 0.5 mg/ml to 100 mg/ml to determine expression biomarker values. mRNA is isolated from exposed cells which is subsequently manipulated to serve as a substrate for hybridization-based expression monitoring of biochemical extracts using microarrays comprising IL-8, IL-2 and

Interferon gamma cDNA (Schena *et al.*, Science (1995) 270: 467-470; Schena *et al.*, Proc Natl Acad Sci USA (1996) 93: 10614-10619; Lockhart *et al.*, Nat Biotechnol (1996) 14: 1675-1680; DeRisi *et al.*, Nat Genet (1996) 14: 457-460; Heller *et al.*, Proc Natl Acad Sci USA (1997) 94: 2150-2155). Previously obtained ginseng data has demonstrated a strong correlation between oxygen consumption during aerobic exercise performance and the induction of the expression biomarkers IL-8, IL-2 and Interferon gamma in test cells (Venkatraman *et al.*, Med Sci Sports Exerc (1997) 29: 333-344 and Wang *et al.*, Planta Med (1998) 64: 130-133). For biochemical biomarkers, rat bone marrow cells will then be exposed to the test batch and assayed for [^{3}H]-thymidine incorporation reflective of mitosis. Previously obtained ginseng data has demonstrated that Rb1 and Rg1 show a strong correlation with DNA synthesis in rat bone marrow cells (Yamamoto *et al.*, Arzneimittelforschung (1978) 28: 2238-2241).

After reiterative analysis, data from each assay will be input into an algorithm to generate a test HBR array for the selected herbal composition based on the enumerated plant-related data, including chemical analyses, and data concerning the subset of biomarkers. The quality of a test batch will be determined by comparing test HBR and standard ginseng Standardized HBR Array variables directed toward analysis of the above observations and subsets, wherein the demonstration of the induction of IL-2, IL-8 and INF gamma mRNA *in vitro* and an increase in [^{3}H]-thymidine incorporation in rat bone marrow cells (including data collected on growth conditions, origin, and verification of the saponins Rg1 and Rb1) is predictive of an equivalent BioResponse effect of the test batch on oxygen consumption as that exhibited by standard ginseng. Based on this procedure it can be determined whether or not the test batch is of a similar or different quality than that of the standard for the given biological response or biological response of interest.

Example 9. Establishing a Standardized HBR Array for Huang Ling (HL)

Recipes.

For the purposes of this example, standard huang ling (HL) is chosen to be *Coptis chinesis* France, from southwest Asia, wherein growth conditions are well known to one skilled in the art (see Huang in The Pharmacology of Chinese Herbs, (1993), pp 69 and 287-288, CRC Press, Boca Raton, FL). Dried rhizomes of *Coptis chinesis* France will be verified for chemical content by quantitative chemical analysis for determination arsenic, berberine, caeruleic acid, columbamine, copsine, coptine, coptiside-I, coptiside-II, coptisine, coreximine, epiberberine, ferulic acid, greenlandicine, isocoptisine, lumicaerulic acid, magnoflorine, oxyberine, thalifendine, umbellatine, urbenine, worenine, palmatine, jatrorrhizine and colubamine (see also Zhu M., Chung Yao Tung Pao (1984) 9: 63-64). Content of the alkaloid berberine of different herbs should be between 7-9% (by weight). These data will be stored, preferably in the memory of a computer processor, for further manipulation.

Expression biomarkers for standard HL include the following: NfκB; bcl-2 analog, A1; zinc finger protein, A20; IL-2 receptor; cell cycle probes; c-Ki-ras2; growth regulators probes and glucocorticoid receptor dependent apoptosis probes (see Chi *et al.*, Life Sci (1994) 54: 2099-2107; Yang *et al.*, Naunyn Schmiedebergs Arch Pharmacol (1996) 354: 102-108; Miura et al., Biochem Pharmacol (1997) 53; Chang K.S., J Formos Med Assoc (1991) 90: 10-14). Alternatively, for a broader batch size, the 400,000 oligonucleotide group/1.6 cm^2 chip of Affimatrix can be used (U.S. Pat. No.5,556,752). The expression biomarkers for standard HL will be prepared by microarray technology as described in Example I, including analysis and statistical data generation. Biochemical biomarkers for standard HL include increase in glucocorticoid receptor and inhibition of alpha-fetoprotein secretion in HL exposed HepG2 cells (see Chi *et al.*, Life Sci (1994) 54: 2099-2107). BBM sets are generated and analyzed as described in Example 1. Statistical data will then be stored, preferably in the memory of a computer processor, for further manipulation.

Biological response of a biosystem will be determined using cells and whole

animals. Batches of the selected herbal composition will be exposed to specific cell types, including but not limited to, human HepG2 hepatoma cells, human embryonal carcinoma cells and thymocytes at concentrations from 0.1-100mg/ml. For animal treatments 0.1mg-2g/kg of coptic herbal composition (i.e., HL) will be administered orally, by intraperitoneal injection or subcutaneous injection. To determine a biological response of a biosystem to standardized HL, human embryonal carcinoma clone, NT2/D1 is exposed to various concentrations of standard HL and cells will be examined for differentiation into cells with neuronal-like cell morphology (Chang K.S., J Formos Med Assoc (1991) 90: 10-14). The assay will be repeated to generate measures and analysis will be performed as described for ginseng in Example 1. A second independent determination of a biological response of a biosystem to standard HL will be the effect of standard HL on diarrhea due to enterotoxigenic *Escherichia coli* (ETEC). Patients with active diarrhea due to ETEC will be treated with various concentrations of HL (e.g., 2g/kg) and stool volumes will be determined (see, *e.g.*, Rabbani G.H., Dan Med Bull (1996) 43: 173-185). The assay will be repeated to generate measures and analysis will be performed as described for ginseng in Example 1. The distribution sets for each biological system are then put into algorithms to generate statistical values for standard HL. Statistical data will then be stored, preferably in the memory of a computer processor, for further manipulation.

Lastly, as in Example 1, the steps are reiterated to generate HBR arrays for standardized HL, wherein the resulting HBR arrays will then be used to predict biological activity and evaluate batch quality. Using this method, a Standardized HBR Array can be generated and updated periodically.

Example 10. Evaluation of a Selected Herbal Composition of Huang Ling Using a Subset of Variables Correlated with a Specific Biological Response.

To evaluate the quality of a selected test batch of an herbal composition of Huang Ling, data is first collected concerning the plant-related characteristics (e.g., plant species, plant parts, geographic origin, growth conditions, processing methods

and storage conditions). The herbal composition is then manipulated such that chemical analysis can be performed to determine the chemical content of the composition (see also Zhu M., Chung Yao Tung Pao (1984) 9: 63-64).

Previously obtained HL data has demonstrated terminal differentiation of human embryonal carcinoma clones into neuronal-like cells is strongly correlated with the presence of berberine (see Chang K.S., J Formos Med Assoc (1991) 90: 10-14). The test batch is then exposed to test cells including human embryonal carcinoma clone, NT2/D1 at a concentration starting at a non-toxic concentration (determination of which is within the skill of the ordinary artisan). mRNA is isolated from exposed cells which is subsequently manipulated to serve as substrate for hybridization based expression monitoring of biochemical extracts using microarrays comprising IL-2 receptor and NfκB; (see Chi *et al.*, Life Sci (1994) 54: 2099-2107; Yang *et al.*, Naunyn Schmiedebergs Arch Pharmacol (1996) 354: 102-108; Miura et al., Biochem Pharmacol (1997) 53; Chang K.S., J Formos Med Assoc (1991) 90: 10-14; U.S. Pat. No.5,556,752), and which can be used to determine down regulation of c-Ki-ras2 gene expression in said cells. Previously obtained HL data has demonstrated terminal differentiation of human embryonal carcinoma clones into neuronal-like cells is strongly correlated with induction of mitogen probes and down regulation of c-Ki-ras2 gene expression (see Chang K.S., J Formos Med Assoc (1991) 90: 10-14).

For biochemical markers, HepG2 cells are exposed to the test composition and cells are assayed for increase in glucocorticoid receptor and inhibition of alpha-fetoprotein secretion (see Chi *et al.*, Life Sci (1994) 54: 2099-2107). Previously obtained HL data has demonstrated that inhibition of glucocorticoid induced apoptosis is strongly correlated with berberine-type alkaloids (see Miura *et al.*, Biochem Pharmacol (1997) 53: 1315-1322). After reiterative analysis, data from each assay will be input into an algorithm to generate a test HBR array based on the enumerated observational data, chemical data and data concerning the subset of biomarkers.

The quality of a test batch will be determined by comparing test HBR and

standard HL HBR Array variables directed toward analysis of the above observations and subsets, wherein the demonstration of the induction of IL-2 receptor and NfκB, the down regulation of c-Ki-ras2 gene expression, an increase in glucocorticoid receptor and inhibition of alpha-fetoprotein secretion for HepG2 cells (to including data collected on growth conditions, origin, and verification of berberine alkaloid) is predictive of an equivalent BioResponse effect of the test batch on terminal differentiation of human embryonal carcinoma clones into neuronal-like cells and inhibition of dexamethasone induced apoptosis as that exhibited by standard HL. Based on this procedure it can be determined whether or not the test batch is of a similar or different quality than that of the standard.

Example 11. Evaluation of Xiao Chai Hu Tang (sho-saiko-to) Using Two Bioassays.

To evaluate the quality of three sources of Xiao Chai Hu Tang, two bioassays were used: 1) cell growth inhibition and 2) hepatitis B virus secretion from infected cells. The Xiao Chai Hu Tang composition is made from a mixture of 6-7 herbal plants (*Radix Bupeuri, Rhizoma Pinelliae, Rhizoma Zingiberis, Radix Scutellariae, Fructus Ziziphi, Codonopsis Pilosula, Radix Ginseng and Radix Glycyrrhizae*, see Table 2 for relative amounts, by weight).

Table 2. Composition of Xiao Chai Hu Tang.

Source	Plant Species							
	Radix bupleuri	*Rhizoma pinelliae*	*Rhizoma zingiberis*	*Radix scutellariae*	*Fructus ziziphi*	*Codonopsis pilosula*	*Radix ginseng*	*Radix glycyrrhizae*
	Relative Amount by Weight							
Singapore	1	1	0.375	0.375	0.375	-----	0.375	0.375
Korea	1	0.717	0.571	0.492	----	0.429	----	0.288
Taiwan	1	0.25	0.375	0.375	0.25	----	0.375	0.375

The three "recipes" originate in either Singapore, Korea or Taiwan. Batches were evaluated for toxicity and for the ability to inhibit hepatitis B virus as detected by

DNA quantitation or detection of hepatitis B surface antigen (HbsAg) (see Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499).

Briefly, one gram of preparation was added with 10 ml of water. The mixture was boiled for 30 minutes. The supernatant was collected after centrifugation and filtered through a 0.22 μm filter. Two cell types were used: a) 2.2.15 cells which secrete hepatitis B virons (kindly provided by Professor G. Ace; see Ace et al. Proc Natl Acad Sci USA (1987) 84: 1005-1009) and b) HepG2 cells (ATCC cat # HB-8065). One to fifty dilutions were used for each assay. The cell growth inhibition assay was performed for 72 hours. All other procedures were performed as described by Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499. The results of the assays using the three batches is displayed in Table 3. Based on these data, the Taiwan source would be selected as a standard herbal composition because of its low toxicity combined with its effectiveness in reducing secretion HbsAG (which is proportional to viral release) by more than half.

Table 3. Bioassay of Xiao Chai Hu Tang (Sho-saiko-to).

Source	Cell Growth Inhibition (%)		Hepatitis B Virus (secreted) % Inhibition	
	HepG2	2.2.15	DNA	HbsAG
Singapore	73	100	65	38
Korea	13	60	20	42
Taiwan	0	42	0	47

The data presented in Tables 2 and 3 for the Taiwan herbal composition constitute the initial data for the standardized HBR Array for this herbal composition. Therefore, this data set would initially include the source of the herbal composition, the plant species and relative amounts of each the herbal compostion, and two BioResponses (*i.e.*, cell growth inhibition and hepatitis B virus secretion from infected cells).

Using the procedures set forth in the schematic of Figure 1 and in Examples 1 and 3, additional data can be collected on plant-related data, markers and BioResponses for the standard herbal composition. This additional data is added to the initial standardized HBR Array to generate an expanded standardized HBR Array. Appropriate analyses of the resulting database can be conducted as set forth in the detailed description and the examples in order to ascertain the subset of variables which is most highly correlated or associated with the BioResponse of interest. Batch HBR Arrays may be determined using the methods depicted in Figure 2 and in the procedures of Examples 2 and 4.

The resultant batch HBR Array can be compared to the standardized HBR Array so as to predict the BioResponse of the batch herbal compositions.

Example 12. Herbal Preparation

The standardized protocol for the herbal extract preparation was as follows: The ingredients of herbal raw materials with proper ratios were placed in a jacketed reactor and extracted with water at an elevated constant temperature with mixing. The solid was separated from the liquid with a 120-mesh screen. The resultant filtrate was collected and then concentrated by evaporating the water under reduced pressure. The concentrated liquor was spray dried at elevated temperature to yield granulated powder. This bulk substance was then formulated into the desired dosage form.

Example 13. Evaluation of Huang Qing Tang

Huang Qing Tang (HQT) is an ancient Chinese botanical formula composed of four distinct herbs: *Scutellariae (scute), Glycyrrhizae (licorice), Paeonie lactiflora pallus (white peony root), and Fructus zizipho (date).* (Table 4). This herbal formula has been long used in Asia to treat a variety of gastrointestinal ailments since 300 AD.

Table 4. Herbal Ingredients of TCM Formula HQT

Scientific Name	Common Name	Traditional Use
Scutellariae Radix	Scute Baical Skullcap root	Used to reduce capillary permeability: to reduce inflammation: to treat enteritis and dysentery: increase the secretion of bile to treat jaundice: to relieve muscle spasms to treat coughing: to expel parasites.
Glycyrrhizae Radix (Gancao)	Licorice Root	Used to moisten the lungs and stop coughs: to relax spasm and stop pain: to moderate the action of herbs; to reduce fire and release toxins.
Fructus Ziziphi	Date	Has diuretic and strengthening effects.
Paeonie lactiflora pallus radix	White Peony Root	Used to suppress and soothe pain; to soothe ligaments and purify the blood.

Biological and Enzyme Assays

Table 5. Batch Properties (HQT)

Property	Batch A	Batch B	Batch C
Origin	Taiwan, Sun-Ten	Taiwan, Sun-Ten	Taiwan, Sun-Ten
Preparation method	Standard	Standard	Boiled 30 min.
Plant part	Root	Root	-

Briefly, one gram of each batch of Huang Qing Tang (HQT) was added with 10 ml of water (1 mg/ml). The mixture was treated as outlined in Table 5. The supernatant was collected after centrifugation and filtered through a 0.22 μm filter. Two cell types were used to test for biological effects of each batch of HQT: a) Jurkat T cells (ATCC cat #TIB-152) and b) HepG2 cells (ATCC cat # HB-8065). One to fifty dilutions were used for each assay. Frozen cells (10^7/ml) were quickly thawed in a water bath at 37 °C. The cells were then diluted in 10 ml of pre-warmed media (see Life Technologies, Inc., Catalogue and Reference Guide, 1998-1999, Cell Culture section) followed by centrifugation at 1500 rpm for 5 min. The supernatant was then

discarded and the cells were cultured in 100 ml media at 37 °C, 5% CO_2. After 2 days, the cells were counted (approximately 8 x 10^5/ml, total 100 ml).

Batches were also evaluated for the ability to inhibit hepatitis B virus as detected by DNA quantitation (see Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499). Briefly, one gram of preparation was added with 10 ml of water. The mixture was boiled for 30 minutes. The supernatant was collected after centrifugation and filtered through a 0.22 μm filter. HepG2.2.15 cells which secrete hepatitis B virons (kindly provided by Professor G. Ace; see Ace et al. Proc Natl Acad Sci USA (1987) 84: 1005-1009) were used in this assay. One to fifty dilutions were used for each assay. The cell growth inhibition assay was performed for 72 hours. All other procedures were performed as described by Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499.

β-glucuronidase was assayed as HQT is known for its anti-diarrhea properties. Different HQT extracts were added to triplicate wells of a 96-well plate which contained 0.1mM phenolphthalein glucuronidate, 70 mM Tris-HCl (pH 6.8) and 0.8 ng of dialyzed β-glucuronidase (from *E.Coli*, purchased from Sigma™) to a final volume of 80 μl. After 2 hr incubation at 37°C, the reactions were terminated with 200 μl of stopping solution which contained 0.2 M Glycine and 0.2 M NaCl (pH 10.4), and the OD was monitored with a kinetic microplate reader at 540 nm.

The results of the assays using the three batches are displayed in Table 6. Based on these data, HQT sources A and B have relatively low toxicities combined with higher inhibitory activity relative to batch HQT C (*i.e.*, approximately 5 fold greater toxicity toward HepG2 cells and 3.3 fold less inhibitory activity against β-glucuronidase than either HQT A or B, see Tabl 6).

Table 6. Biological Assay of Three Preparations of HQT*

	E. Coli β-Glucuronidase	HepG2	Jurkat	HBV‡ DNA
HQT A	0.6	1.50	0.76	None
HQT B	0.7	1.6	0.81	ND
HQT C	2.2	0.32	ND	ND

*Values represent IC_{50} values. ‡, % of Control

ND, not determined.

Evaluation of HQT Effects on Protein Expression

HepG2 cells (1 x 10^6) were seeded in 25 cm^2 flasks in 3.0 ml of RPMI-1640 medium (see Life Technologies, Inc., Catalogue and Reference Guide, 1998-1999, Cell Culture section) 24 hr before the drug addition. The cells were treated with or without herbal medicine, where the former is added at two final concentrations of 0.2 mg/ml or 4 mg/ml, respectively, and incubated at 37°C for 24 hours. The medium was removed and the cells were washed twice with cold PBS. The cells were harvested into 1 ml of PBS and centrifuged at 10,000 rpm for 2 minutes, extracted on ice with a buffer containing 50 mM Tris-Cl (pH 7.5), 0.2 mM PMSF and 10% glycerol, followed by three freeze-thaw cycles. Potassium chloride was added to the cell lysate at a final concentration of 0.15 M prior to centrifugation. The protein concentration was determined and the cell extract was electrophoresed according to the method of Laemmli (*Nature* (1970) 227:680-685). Western blots were performed by standard techniques known in the art, see for example Sambrook, *et al* (1989). The antibodies used were directed to the following proteins: Topo I; Stat (20707); Cyclin B1; MAPK (Ab2) and Nm 23 H1.

Figure 4 demonstrates that the higher concentrations of HQT A or HQT B differentially effects the expression of cyclin B1 polypeptide.

HPLC Analysis

The herbal batches were analyzed by HPLC with a Beckman ODS Ultrasphere™ column (5 micron particles, 4.6 mm X 25 cm) and detected with a UV spectrophotometer (Perkin Elmer). The wavelengths for UV detection were monitored at 280 nm and 340 nm. The mobile phase was pumped at 1 ml/min and consisted of Solvent A: H_2O and Solvent B: 20% MeOH with the following gradient: 1) the solvent was 100% solvent A for the first 5 minutes; 2) the solvent composition was changed to 10% solvent A / 90% solvent B for the next 10 minutes; and 3) the solvent was changed to 10% solvent A / 90% solvent B for the next 40 minutes. This was followed by the addition of 100 % solvent A for 5 minutes. The HPLC markers are baicalin and baicalein.

Mass Spectrometry

The herbal extract was analyzed by Mariner™ ESI-TOF Mass Spectrometry (MS) from PE Biosystems. Control tracings were generated using baicalein and baicalin, two known active ingredients in HQT.

HQT samples in water and acid treated batches were been analyzed by HPLC and Mass Spectrometry. While water treated HQT batches A and B had distinct HPLC and MS tracings, acid treated batches gave almost identical patterns (data not shown).

Algorithm

The data collected form part of the multidimensional analysis used to generate multivariant normal distribution sets as a means of determining a baseline correlation between biological activity and standard HQT chemical (HPLC and Mass Spec), and origin/growth characteristics.

Example 14. Individual components

A. Licorice.

Evaluation of *Glycyrrhizae Radix* (licorice)

Licorice is useful for moistening the lungs and reducing coughs, helps to relax spasm and pain. The properties of the licorice batches used in this example are

presented in Table 7.

Table 7. Batch Properties (licorice)

Property	Batch A	Batch B	Batch C	Batch D
Plant Name	*Glycyrrhizae Radix*	*Glycyrrhizae Radix*	*Glycyrrhizae Radix*	*Glycyrrhizae Radix*
Origin	Inner Mongolia	Inner Mongolia	U.S., Kin Man Herb Center	U.S., Kin Man Herb Center
Preparation method	Standard	Standard	Boiled 30 min.	Warm H_2O, 30 min.
Plant part	Root	Root	-	-

Biological and Enzyme Assays

To assay the quality of herbal sources, each herbal extract supernatant was assayed and the analysis was repeated three times. For a given sample to be assayed, 1 gram of herbal powder was dissolved in 10 ml of 80° C deionized water (neutral pH) in a polypropylene tube. The tube was then incubated as outlined in Table 7, then centrifuged to obtain the supernatant. Batches of licorice were tested against either HepG2 cells (ATCC cat # HB-8065) or Jurkat T cells (ATCC cat #TIB-152) or both. Cells were cultured for 24 hours as described above.

Batches were also evaluated for the ability to inhibit hepatitis B virus as detected by DNA quantitation (see Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499). Briefly, one gram of preparation was added with 10 ml of water. The mixture was boiled for 30 minutes. The supernatant was collected after centrifugation and filtered through a 0.22 μm filter.2.2.15 cells which secrete hepatitis B virons (kindly provided by Professor G. Ace; see Ace et al. Proc Natl Acad Sci USA (1987) 84: 1005-1009) were used in this assay. One to fifty dilutions were used for each assay. The cell growth inhibition assay was performed for 72 hours. All other procedures were performed as described by Dong *et al.*, Proc Natl Acad Sci USA

(1991) 88: 8495-8499.

Again, β-glucuronidase was assayed. Different licorice extracts were added to triplicate wells of a 96-well plate which contained 0.1mM phenolphthalein glucuronidate, 70 mM Tris-HCl (pH 6.8) and 0.8 ng of dialyzed beta-glucuronidase (from *E.Coli*, purchased from Sigma) to a final volume of 80 µl and assayed as above.

The results of the assays using the two batches is displayed in Table 8. Based on these data, licorice batch A was much more toxic to Jurkat cells than batches B (approximately 9 fold) and a more effective inhibitor of th β-glucuronidase (see Table 8).

Table 8. Biological Assay of Four Preparations of Licorice*

	E. Coli β-Glucuronidase	HepG2	Jurkat	HBV‡ DNA
Licorice A	1.1	1.07	0.41	None
Licorice B	ND	ND	3.6	ND
Licorice C	2.1	ND	ND	ND
Licorice D	ND	ND	>2.0	53.8

*Values represent IC_{50} values. ‡, % of Control

ND, not determined.

Expression Assay

In order to assay gene expression, Jurkat T cells were treated with herbal extract as follows: Jurkat cells (10^7/ml) were quickly thawed in a water bath at 37 °C. The cells were then diluted in 10 ml of pre-warmed media (see Life Technologies, Inc., Catalogue and Reference Guide, 1998-1999, Cell Culture section) followed by centrifugation at 1500 rpm for 5 min. The supernatant was then discarded and the cells were cultured in 100 ml media at 37 °C, 5% CO_2. After 2 days, the cells were counted (approximately 8 x 10^5/ml, total 100 ml).

The herbal extract solution was prepared as outlined above (e.g., 2 g of an

herbal powder to obtain 20 ml of sterile solution (0.1 g/ml). The cells were divided into 3 flasks at a density of 2.5 x 10^5/ml, 100 ml/each flask. Assays were carried out with control (no extract), and 10 ml of extract at 10 mg/ml, and 1 mg/ml. Again, toxicity results were used to determine the "high" and "low" concentrations for any given extract. After extract addition, cell cultures were incubated for 24 hours under conditions as outlined above. The cells were counted and subsequently collected in 50 ml centrifuge tubes. The resulting cell pellet was treated with an RNA isolation means to extract mRNA (see, for example, Sambrook *et al.*, 1989 at pages 7.3-7.39).

Microarray

Microarray printing was carried out as follows:

Gene clones were obtained from the IMAGE Consortium libraries through Research Genetics (Huntsville, AL) and comprise genes from various tissues. Most clones have been partially sequenced and were available as expressed sequence tags form the dbEST database of GenBank. Clones comprising pBluescript plasmids were separately cultured and amplified using commercially available primers prior to application on nylon membranes (Chen *et al.*, *Genomics* (1998) 51:313-324). Approximately 10 ng of each amplified target was applied on a positively charged nylon membrane using a PC (personal computer) controlled arraying system. Roughly 85,000 spots can be placed on a piece of nylon membrane measuring 35 by 55 mm using a 24-pin arraying tool.

cDNA probe and Membrane Hybridization

One microgram of each mRNA sample (mRNA was isolated as outlined above) was labeled with biotin and/or digoxigenin using random primed reverse transcription. The labeled samples were treated with alkali and the resulting labeled nucleic acids were precipitated prior to use in hybridization. Membrane hybridization and washing were carried out using the labeled probes as disclosed in Chen *et al.* (1998). To detect the spots on the membrane in dual color mode (*i.e.*, both biotin and digoxigenin), β-galactosidase-conjugated streptavidin (Strept-Gal) and alkaline phosphatase-

conjugated digoxigenin antibody (anti-Dig-AP) were employed. After color development, image digitization using a imaging means was employed (*e.g.*, a flatbed scanner or digital camera). Quantitative measurements were determined by computer analysis which uses a program that measures the integrated density of the primary color components of each spot, performs regression analysis of the integrated density data and locates statistical outliers as differentially expressed genes.

Genes expression data for samples 1, 2 and licorice (ST117)

Extract 1, 2 and 6 corresponding to extract of Cordyceps sinensis, *Poria cocos* (ST 027) and licorice, respectively, were assayed by the following method: Batches were evaluated for toxicity using Jurkat T cells.

The extracts were prepared as outlined in Example 6. The cells were divided into 24 well culture plates by adding 1 ml of Jurkat cells at a density of 5 x 10^5/ml. Assays were carried out with control (no extract), and 5 concentrations of extracts as described (see Table 9). The high and low concentrations for the cell culture assays were varied between 10 mg/ml and 0.05 mg/ml (*i.e.*, mg dry weight of herbal extract per ml) depending on the toxicity of the extract to cells. For certain samples the toxicities at 10 mg/ml were such that "high" and "low" concentrations were adjusted downward, nevertheless, at least one order of magnitude between extremes was maintained. For example, for licorice (ST117) the "high" was 0.5 mg/ml and the "low" was 0.05 mg/ml (see Table 9). After extract addition, cell cultures were incubated for 24 hours under conditions as outlined in Example 6. The cells were counted and the resulting data tabulated to demonstrate extract toxicity. The resulting data is shown in Table 9.

Table 9. Survival Cell Number at Different Concentration of Herbal Extract Solution

							experiment concentration	
no.x10^{5}ml	10 mg/ml	5 mg/ml	2 mg/ml	1 mg/ml	0.5 mg/ml	No drug	High conc	Low conc.
1 Cordyceps sinensis	8.4	11.9	11.5	9.2	9.0	12.4	10	1
2 ST027	4.2	8.7	10	7.5	10	10	10	1
3 ST0-44	-	5.9	8.4	9.9	9.4		5	0.5
4 ST051	-	-	-	1.7	5.4		0.5	0.05
5 ST093	-	-	1.9	3.8	4.4		0.5	0.05
6 ST117	-	1.6	3.6	4.6	5.8		0.5	0.05
7 ST123	3.4	6.4	8	9.3	7.8		5	0.5
8 ST128	3.5	7.7	7.9	7.7	8.3		5	0.5
9 ST134	2.9	6.1	11.2	9.6	9.8		5	0.5
10 ST237	-	-	2.5	6.6	8.7		1	0.1

Note: original cell number is 5x10^5/ml and the number to 10x10^5/ml after 24h incubation. "-" describes all dead cells.

Protocol:
1. Add 1 ml of 5x10^5/ml Jurkat cells into 24 well culture plates.
2. Prepare 12 kinds of herbal extract solutions and sterilize.
3. Test 5 concentration per sample. 10 mg/ml, 5 mg/ml, 2 mg/ml, 1 mg/ml, 0.5 mg/ml
4. Culture the cells for 24 h in 37C with 5% CO_2 incubator.

In each analysis, 144X96 genes (i.e., 13,824 genes) were analyzed (data not shown) and about 100 genes showed significant differences in comparison with that of control (Table 10). Some of the genes were up-regulated and others were down-regulated. The magnitude of the difference with the control sometimes varied depending on the relative amount of the herbal composition to which the particular cells were exposed. Numbers under C1 (control treatment) and H or L (herbs) represent intensities of mRNA expressed after subtraction of background (Table 10). The gene designation is encoded in Array AD, which can be traced to a specific GenBank clone. The level of expression was determined by H or L divided by C.

Only a fraction of 13,824 genes in each herb treated samples showed significant changes, namely, up, down or unchanged (see Table 10).

TABLE 10.

rem	The value show in colume K,L,M,N,O,P represent the ratio of mRNA express level for #1, #2, #6 herbs treatment(H, L for high conc. And low conc) compared with untreated cell (C1)													
	Title	C1	1H	1L	2h	2L	6H	6L	1H	1L	2h	2L	6H	6L
ArrayAD	TRANSDUCIN-LIKE ENHANCER PROTEIN 1	Sg-Bk	Sg-Bk	Sg-Bk	Sg-Bk	Sg-Bk	Sg-Bk	Sg-Bk	COMPARE TO CONTROL					
<94,030>	Adenylosuccinate lyase	55.29	8.55	7.15	46.74	38.07	29.21	30.82	0.154639	0.129318	0.845361	0.688551	0.528305	0.557424
<91,097>	Neurotrophic tyrosine kinase, receptor, type 1	89.55	98.29	70.26	84.83	93.83	36.02	89.63	1.097599	0.78459	0.947292	1.047795	0.402233	1.000893
<89,083>	MHC class I protein HLA-A (HLA-A28,-B40, -Cw3)	80.68	61.91	46.13	59.86	29.64	46.16	32.35	0.767353	0.571765	0.741943	0.367377	0.572137	0.400967
<87,083>		85.33	50.96	37.75	53.25	25.54	37.5	49.82	0.597211	0.4424	0.624048	0.299309	0.43947	0.583851
<86,131>	ESTs, Highly similar to HELIX-LOOP-HELIX PROTEI	99.48	109.46	105.62	56.68	67.54	35.65	75.78	1.100322	1.061721	0.569763	0.67893	0.358363	0.761761
<85,129>	INTEGRAL MEMBRANE PROTEIN E16	90.64	72.12	92.75	37.97	70.08	48.47	59.36	0.795675	1.023279	0.41891	0.773169	0.534753	0.654898
<85,112>		35.23	60.82	48.87	44.04	29.71	28.25	59.33	1.72637	1.38717	1.250071	0.843315	0.801873	1.684076
<84,075>		89.56	77.54	77.9	66.1	59.44	43.37	42.97	0.865788	0.869808	0.738053	0.663689	0.484256	0.47979
<83,129>		118.83	108.15	118.05	75.84	64.09	55.8	66.27	0.910124	0.993436	0.638223	0.539342	0.469578	0.557687
<82,082>	Homo sapiens androgen receptor associated protein 24 (A	60.23	47.87	34.96	37.41	9.84	44.8	46.5	0.794787	0.580442	0.621119	0.163374	0.743815	0.772041
<82,027>		65.11	90.01	66.74	66.3	64.44	39.3	78.37	1.38243	1.025035	1.018277	0.98971	0.603594	1.203655
<80,129>		96.27	99.96	109.3	81.88	87.66	48.32	53.87	1.03833	1.135348	0.850525	0.910564	0.501922	0.559572
<80,035>	Human small GTP binding protein Rab7 mRNA, complet	44.91	46.23	25.19	78.21	40.75	37.48	42.49	1.029392	0.5609	1.741483	0.90737	0.834558	0.946114
<74,108>		65.2	77.25	68.09	60.2	66.21	51.52	70.19	1.184816	1.044325	0.923313	1.015491	0.790184	1.076534
<74,012>		34.68	60.92	57.73	24.61	48.93	42.34	65.02	1.756632	1.664648	0.709631	1.4109	1.220877	1.874856
<73,132>	Human mRNA for KIAA0078 gene, complete cds	109.86	103.38	82.35	85.97	69.15	61.91	32.44	0.941016	0.74959	0.782541	0.629437	0.563535	0.295285
<72,101>	Human trans-Golgi p230 mRNA, complete cds	46.96	25.76	18.35	10.54	17.87	1.27	17.51	0.548552	0.390758	0.224446	0.380537	0.027044	0.372871
<72,014>	Chlordecone reductase	4.37	15.08	13.31	42.87	22.76	20.64	30.11	3.450801	3.045767	9.810069	5.208238	4.723112	6.89016
<70,101>		60.86	13.55	35.23	33.39	40.65	16.36	18.7	0.222642	0.57887	0.548636	0.667926	0.268814	0.307263
<69,107>		37.92	31.41	26.12	82.82	28.64	36.95	50.55	0.828323	0.688819	2.184072	0.755274	0.97442	1.33307
<68,064>	Human mRNA for KIAA0034 gene, complete cds	2.24	1.27	5.31	0.21	2.5	4.2	41.64	0.566964	2.370536	0.09375	1.116071	1.875	18.58929
<62,117>		106.71	94.67	83.98	81.39	61.45	77	41.8	0.887171	0.786993	0.762721	0.57586	0.721582	0.391716
<62,084>	NA	1.13	0	0	13.55	0.58	32.81	17.71	0	0	11.99115	0.513274	29.0354	15.67257
<62,001>	Heterogeneous nuclear ribonucleoprotein A1	97.11	71.92	73.16	42.01	78.94	83.94	96.87	0.740603	0.753372	0.432602	0.812893	0.864381	0.997529
<61,118>	NA	89.52	59.97	54.11	66.86	43.05	65.65	34.92	0.669906	0.604446	0.746872	0.480898	0.733356	0.39008
<59,142>	ATPase, Na+/K+ transporting, alpha 1 polypeptide	13.03	15.48	47.75	27.71	27.86	23.22	16.58	1.188028	3.66462	2.126631	2.138143	1.782041	1.272448
<59,135>		60.85	90.23	39.83	93.01	38.17	86.39	31.2	1.482827	0.65456	1.528513	0.62728	1.419721	0.512736
<58,087>	ESTs, Weakly similar to KIAA0062 [H.sapiens]	77.72	46.86	76.33	64.43	22.52	56.78	63.4	0.602934	0.982115	0.829002	0.289758	0.730571	0.815749
<57,016>	Neuroblastoma RAS viral (v-ras) oncogene homolog	71.71	16.47	44.61	39.51	66.89	50.43	35.52	0.229675	0.622089	0.550969	0.932785	0.703249	0.495328
<56,088>	3-HYDROXY-3-METHYLGLUTARYL-COENZYME A	56.77	52.66	59.53	59.6	16.89	72.87	60.28	0.927603	1.048617	1.04985	0.297516	1.2836	1.061828
<56,060>	Translation elongation factor 1-alpha-1	81.26	92.28	92.08	36.86	79.09	71.74	70.47	1.135614	1.133153	0.453606	0.973296	0.882845	0.867216
<55,040>		0	43.58	39.47	53.67	46.42	40	65.93						
<54,134>	CARBONIC ANHYDRASE III	0	0	4.38	43.47	8.23	32.61	9.74						
<53,048>	H.sapiens mRNA for interferon regulatory factor 3	0.25	0	0	33.47	0	0	0.53	0	0	133.88	0	0	2.12
<53,023>	Human mRNA for proteasome subunit HsC7-I, complete	67.76	14.98	42.33	38.48	66.52	27.24	37.24	0.221074	0.624705	0.567887	0.9817	0.402007	0.549587
<52,135>		49.26	46.81	44.4	72.95	55.5	63.21	30.35	0.950264	0.90134	1.480918	1.126675	1.283191	0.616119
<51,131>	ATP citrate lyase	2.02	0	52.06	39.33	6.03	22.36	5.6	0	25.77228	19.4703	2.985149	11.06931	2.772277
<51,101>	POLYADENYLATE-BINDING PROTEIN	8.65	0.17	0.45	38.99	14.31	38.25	11.64	0.019653	0.052023	4.507514	1.654335	4.421965	1.345665

CloneID
504540
1013392
510048
1188706
72215
118548
115134
172767
1184183
118235
171864
1172135
1172266
37502
45672
79342
563992
21759
564469
29009
51927
241351
28012
236388
240018
121270
510863
645239
47526
509520
31027
328351
287006
116915
221285
286222
624420
529138

WO 00/24934 PCT/US99/24851

TABLE 10 (CONT.)

<50,035>	SERUM ALBUMIN PRECURSOR	0.88	0	10.21	44.14	0	0.79	0.01	0	11.60227	50.15909	0	0.897727	0.011364	510245
<47,055>	ESTs	38.53	31.69	19.45	43.8	56.1	59.36	53.2	0.822476	0.504801	1.136777	1.456008	1.540618	1.380742	26801
<46,097>		34.26	35.95	32.49	39.81	63.19	30.86	51.76	1.049329	0.948336	1.161996	1.844425	0.900759	1.5108	511591
<45,100>	Ribosomal protein L5	26.15	33.6	26.3	55.86	48.59	31.26	43.29	1.284895	1.005736	2.136138	1.858126	1.195411	1.655449	511410
<44,103>	NA	72.01	54.36	70.27	75.43	91.53	70.5	80.86	0.754895	0.975837	1.047493	1.271073	0.979031	1.1229	26099
<43,039>		55.86	18.02	20.83	11.68	17.8	11.82	41.06	0.322592	0.372897	0.209094	0.318654	0.2116	0.735052	592919
<43,035>	14-3-3 PROTEIN TAU	86.86	23.18	36.23	18.14	12.76	12.99	42.28	0.266866	0.417108	0.208842	0.146903	0.149551	0.48676	120011
<43,019>		72.48	10.4	23.38	27.59	12.83	3.47	15.99	0.143488	0.322572	0.380657	0.177014	0.047875	0.220613	488154
<42,099>		30.25	33.73	36.83	61	51.56	40.36	48.43	1.115041	1.217521	2.016529	1.704463	1.334215	1.600992	86102
<40,100>	ESTs	0	0	0.2	38.37	12.78	2.59	7.47							84335
<39,118>	Thioredoxin	4.26	0.34	2.83	37.28	8.03	2.25	9.67	0.079812	0.664319	8.751174	1.884977	0.528169	2.269953	123764
<39,007>	Pre-alpha (globulin) inhibitor, H3 polypeptide	9.46	24.06	15.82	26.08	25.48	22.8	44.36	2.54334	1.672304	2.756871	2.693446	2.410148	4.689218	23643
<38,097>	NA	11.81	6.94	23.82	80.14	16.65	6.13	18.51	0.587638	2.016935	6.785775	1.409822	0.519052	1.567316	81665
<38,081>	Human GP36b glycoprotein mRNA, complete cds	101.51	82.61	89.86	22.88	71.4	89.93	79.82	0.813811	0.885233	0.225397	0.703379	0.885923	0.786326	1185901
<35,096>		47.35	24.52	0	23.2	32.88	37.19	35.29	0.517846	0	0.489968	0.694403	0.785428	0.745301	723779
<34,143>		61.23	2.99	25.05	30.61	16.93	27.03	40.4	0.048832	0.409113	0.499918	0.276498	0.44145	0.659807	712549
<32,118>	ESTs, Weakly similar to CASEIN KINASE I HOMOLOG	21.98	8.91	14.11	53.54	24.01	17.46	22.99	0.405369	0.641947	2.435851	1.092357	0.794359	1.045951	265480
<32,034>	Pyruvate kinase, muscle	78.41	54.15	79.36	29.56	40.5	28.52	61.59	0.690601	1.012116	0.376993	0.516516	0.363729	0.785487	1170289
<30,102>	RETINOBLASTOMA BINDING PROTEIN P48	27.3	23.1	16.06	62.35	34.39	16.52	47.49	0.846154	0.588278	2.283883	1.259707	0.605128	1.73956	612365
<30,033>		70.49	30.32	65.13	17.18	33.65	49.07	36.52	0.430132	0.923961	0.243723	0.477373	0.696127	0.518088	82236
<29,058>	NA	25.77	16.66	0	59.85	33.24	27.05	18.54	0.646488	0	2.322468	1.289872	1.04967	0.719441	38798
<29,035>		60.19	57.12	86.66	46.47	56.59	61.01	58.47	0.948995	1.439774	0.772055	0.940189	1.013624	0.971424	81491
<28,104>	Ribosomal protein S13	0	0	0	37.19	1.37	0	5.97							595769
<28,035>		71.46	35.93	51.86	11.27	30.75	40.12	36.54	0.502799	0.725721	0.157711	0.430311	0.561433	0.511335	994662
<27,119>		0	0	0	48.27	1.66	0	1.1							114662
<27,097>	ESTs, Moderately similar to Etr-3 [H.sapiens]	16.06	24.53	9.26	76.01	24.34	24.16	30.01	1.527397	0.576588	4.732877	1.515567	1.504359	1.868618	68744
<25,126>	ESTs	42.93	22.03	29.88	67.76	40.39	24.86	30.53	0.513161	0.696017	1.578383	0.940834	0.579082	0.711158	427843
<25,106>		61.9	48.38	43.41	91.26	58.89	39.88	51.83	0.781583	0.701292	1.474313	0.951373	0.644265	0.837318	42714
<24,098>	Human splicesomal protein (SAP 61) mRNA, complete c	24.26	25.43	21.8	56.18	37.28	30.59	46.77	1.048228	0.898599	2.315746	1.536686	1.260923	1.927865	274769
<24,071>	HEAT SHOCK 70 KD PROTEIN 1	63.92	58.19	61.1	50.86	66.27	80.57	68.86	0.910357	0.955882	0.795682	1.036765	1.260482	1.077284	255134
<23,126>		5.18	1.1	10.08	54.71	12.11	2.9	10.53	0.212355	1.945946	10.56178	2.337838	0.559846	2.032819	416946
<23,099>	Ribosomal protein S3A	5.13	1.93	1.25	38.99	14.35	4.51	18.73	0.376218	0.243665	7.60039	2.797271	0.879142	3.651072	298187
<22,013>	EUKARYOTIC INITIATION FACTOR 4B	61.8	49.95	41.52	22.99	49.11	42.7	20.04	0.808252	0.671845	0.372006	0.79466	0.690939	0.324272	51894
<22,006>	3-HYDROXY-3-METHYLGLUTARYL-COENZYME A	83.17	57.84	39.8	40.04	44.33	46.33	35.49	0.695443	0.478538	0.481424	0.533005	0.557052	0.426716	563866
<20,126>	MITOCHONDRIAL STRESS-70 PROTEIN PRECURSO	27.02	22.19	33.86	56.87	32.97	27.24	33.86	0.821244	1.253146	2.104737	1.220207	1.008142	1.253146	382766
<20,104>	Human eukaryotic translation initiation factor (eIF3) mR	22.09	15.17	17.33	74.98	19.17	8.3	35.84	0.686736	0.78[illegible]518	3.394296	0.867813	0.375736	1.622454	562992
<20,103>	Laminin receptor (2H5 epitope)	1.45	1.26	0	39.66	8.67	0.73	9.03	0.868966	0	27.35172	5.97931	0.503448	6.227586	562928
<19,098>	Human semaphorin (CD100) mRNA, complete cds	17.79	14.38	10.79	52.94	28.97	12.5	31.25	0.808319	0.606521	2.975829	1.628443	0.702642	1.756605	248687
<18,103>	Ribosomal protein L5	31.47	21.13	30.07	78.6	27.18	27.47	43.29	0.671433	0.955513	2.497617	0.86368	0.872895	1.375596	549382
<18,084>	ESTs	54.61	75.17	46.16	79.5	60.01	74.37	56.43	1.376488	0.845266	1.455777	1.098883	1.361838	1.033327	512336
<17,103>	NA	0	0	0	33.96	2.87	3.64	9.87							549224

TABLE 10. (CONT.)

<17,097>	NA	1.52	2.1	1.21	37.77	11.59	3.43	14.35	1.381579	0.796053	24.84868	7.625	2.256579	9.440789	327474
<17,062>		92.72	73.27	103.64	69.79	69.35	68.85	83.14	0.790229	1.117774	0.752696	0.747951	0.742558	0.896678	50753
<16,103>	Integrin, beta 1 (fibronectin receptor, beta polypeptide, an	0	0.92	0	41.3	1.84	6.67	11.18							547624
<16,100>	Homo sapiens E1B 19K/Bcl-2-binding protein Nip3 mRN	0.79	0	0.7	38.06	8.28	0.89	16.63	0	0.886076	48.17722	10.48101	1.126582	21.05063	308754
<15,112>		16.47	11.65	11.87	46.52	22.98	28.22	22.43	0.707347	0.720704	2.824529	1.395264	1.713418	1.36187	49967
<14,103>	NA	27.61	34.26	30.18	55.67	38.45	35.52	50.43	1.240855	1.093082	2.016298	1.392611	1.28649	1.826512	546545
<14,011>	NA	47.42	34.87	68.28	38.71	53.6	50.03	52.24	0.735344	1.439899	0.816322	1.130325	1.05504	1.101645	32672
<13,114>	Homo sapiens clone 24689 mRNA sequence	20.64	13.87	19.2	51.26	32.34	20.27	27.7	0.671996	0.930233	2.483527	1.56686	0.982074	1.342054	144001
<12,144>	Homo sapiens NADH:ubiquinone oxidoreductase NDUFS	76.78	67.2	76.11	64.5	50.04	53.7	6.92	0.875228	0.991274	0.840063	0.651732	0.699401	0.090128	325580
<12,111>		40.04	55.04	46.34	77.78	46.07	69.51	55.69	1.374625	1.157343	1.942557	1.150599	1.736014	1.390859	49373
<11,135>		90.34	114.29	87.88	80.49	87.22	60.79	87.28	1.26511	0.97277	0.890967	0.965464	0.672902	0.966128	510620
<11,105>	Ribosomal protein L3	8.27	14.41	15.59	36.82	25.87	31.97	47.1	1.742443	1.885127	4.452237	3.128174	3.86578	5.695284	31866
<11,053>	Calpain, large polypeptide L2	10.48	23.58	14.8	15.69	8.87	47.63	28.91	2.25	1.412214	1.497137	0.846374	4.544847	2.758588	544957
<10,136>	Malic enzyme 2, mitochondrial	74.34	81.63	90.94	83.18	74.85	54.35	65.41	1.098063	1.223298	1.118913	1.00686	0.7311	0.879876	510323
<10,101>	ESTs	13.65	16.34	13.97	47.57	26.68	36.05	40.14	1.19707	1.023443	3.484982	1.954579	2.641026	2.940659	613442
<09,136>	NA	59.82	95	87.16	77.34	83.51	69.24	71.68	1.588098	1.457038	1.292879	1.396021	1.157472	1.198261	509836
<09,110>	ER LUMEN PROTEIN RETAINING RECEPTOR 1	17.3	41.73	48.92	55.32	35.56	45.73	60.03	2.412139	2.827746	3.197688	2.055491	2.643353	3.469942	47800
<09,085>	Human phospholipid transfer protein mRNA, complete cd	80.14	88.83	96.22	80.4	77.44	76.05	82.81	1.108435	1.200649	1.003244	0.966309	0.948964	1.033317	613511
<08,138>		41.57	60.4	69.53	73.27	81.26	59.49	66.73	1.452971	1.6726	1.762569	1.954775	1.43108	1.605244	632480
<08,132>		67.22	102.84	103.65	90.56	88.43	76.85	81.15	1.529902	1.541952	1.347218	1.315531	1.143261	1.20723	66427
<08,122>	Homo sapiens mRNA for DRAK1, complete cds	32.01	23.21	47.74	63.02	35.64	32.48	46.85	0.725086	1.491409	1.96876	1.113402	1.014683	1.463605	308580
<08,105>		69.58	58.7	84.49	90.5	68.91	89.98	94.96	0.843633	1.214286	1.300661	0.990371	1.293188	1.36476	30607
<08,055>		19.5	20.01	22.83	27.87	34.66	48.86	50.43	1.026154	1.170769	1.429231	1.777436	2.505641	2.586154	611114
<07,111>		30.8	44.67	48.75	73.84	51.31	59.76	68.29	1.450325	1.582792	2.397403	1.665909	1.94026	2.217208	46880
<07,107>	ESTs	4.57	3.23	9.01	41.32	8.83	22.9	23.56	0.706783	1.971554	9.041575	1.932166	5.010941	5.155361	29970
<06,135>		71.42	112.68	128.81	102.8	101.48	93.72	94.2	1.577709	1.803556	1.439373	1.420891	1.312237	1.318958	286790
<06,102>	ESTs	39.82	36.03	32.57	67.74	46.04	59.07	57.13	0.904822	0.817931	1.701155	1.156203	1.483425	1.434706	530700
<06,101>	Intercellular adhesion molecule 2	4.43	2.95	4.21	38.95	16.6	15.85	32.59	0.665914	0.950339	8.792325	3.747178	3.577878	7.356659	530665
<06,034>	Ribosomal protein L5	0	0	0	0	33.93	0	0							666094
<06,006>	Human fragile X mental retardation syndrome related prot	0	31.71	0	0	0	0	0							530672
<04,134>		39.03	54.62	46.61	71.8	58.82	64.83	64.43	1.399436	1.19421	1.839611	1.507046	1.66103	1.650781	285979
<04,106>	Ferritin, light polypeptide	15.74	23.81	21.94	46.74	21.94	56.63	55.12	1.512706	1.393901	2.969504	1.393901	3.59784	3.501906	31309
<03,140>		56.08	88.33	61.29	93.1	74.04	74.42	81.16	1.575071	1.092903	1.660128	1.320257	1.327033	1.447218	624595
<03,099>	ESTs	22.97	18.57	12.1	52.83	23.48	53.26	41.55	0.808446	0.526774	2.299956	1.022203	2.318677	1.808881	42281
<02,111>	ESTs, Highly similar to HYPOTHETICAL 64.5 KD PRO	34.89	23.63	25	77.41	28.84	56.6	60.36	0.677271	0.716538	2.218687	0.826598	1.622241	1.730009	45327
<01,136>	Homo sapiens mRNA for 5-aminoimidazole-4-carboxami	21.36	34.84	23.95	53.14	45.66	40.67	51.53	1.631086	1.121255	2.487828	2.13764	1.904026	2.412453	567287
<01,058>	Acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight ch	60.68	53.35	73.26	62.56	56.58	72.81	91.71	0.879202	1.207317	1.030982	0.932432	1.199901	1.511371	30921

Total 116

In this manner, we are able to correlate specific gene expression with the exposure of a cell to no, low (L) or high (H) amounts of an herbal composition. Many of the genes identified in this way code for proteins important in known metabolic or biochemical pathways. Many of these proteins have direct and indirect effects on certain physiological, morphological and psychological parameters. Thus, this method permits the association of a particular genetic fingerprint of an herbal composition with its array biological effects. Such associations can be used to profile or characterize an herbal composition for the purposes of Quality Control and Quality Assurance and evaluating pharmacological or toxicological properties. The role of primary and secondary herbs in an herbal formula can also be assessed by this approach.

HPLC Analysis

The herbal batches were analyzed by HPLC with a Beckman ODS Ultrasphere™ column (5 micron particles, 4.6 mm X 25 cm) and detected with a UV spectrophotometer (Perkin Elmer). The wavelengths for UV detection were monitored at 280 nm and 340 nm. The mobile phase was pumped at 1 ml/min and consisted of Solvent A: H_2O and Solvent B: 20% MeOH with the following gradient: 1) the solvent was 100% solvent A for the first 5 minutes; 2) the solvent composition was changed to 10% solvent A / 90% solvent B for the next 10 minutes; and 3) the solvent was changed to 10% solvent A / 90% solvent B for the next 40 minutes. This was followed by the addition of 100 % solvent A for 5 minutes. The HPLC marker is glycyrrhizin.

Algorithm

The data collected form part of the multidimensional analysis used to generate multivariant normal distribution sets as a means of determining a baseline correlation between biological activity and standard licorice molecular, chemical (HPLC and Mass Spec), and origin/growth characteristics.

B. Scute

Evaluation of *Radix Scutellariae* (Scute)

Scute has been found to be useful in reducing capillary permeability and inflammation. It can also be used treat enteritis and dysentery, increases the secretion of bile to treat jaundice; to relieve muscle spasms; to treat coughing and to expel parasites. The properties of the scute batches used in this example are presented in Table 11.

Table 11. Batch Properties (Scute)

Property	Batch A	Batch B	Batch C	Batch D
Plant Name	*Scutellariae radix*	*Scutellariae radix*	*Scutellariae radix*	*Scutellariae radix*
Origin	Sanxi Province.	U.S., Kin Man Herb Center	U.S., Kin Man Herb Center	U.S., Boston
Preparation method	Standard	Boiled, 30 min	Warm H_2O, 30 min.	Boiled , 2 hours
Plant part	Root	-	-	-

Biological and Enzyme Assays

Briefly, one gram of each preparation of scute extract was added with 10 ml of water (1 mg/ml). The mixture was treated as Outlined in Table 11. The supernatant was collected after centrifugation and filtered through a 0.22 μm filter. Batches of scute were tested against either HepG2 cells (ATCC cat # HB-8065) or Jurkat T cells (ATCC cat #TIB-152) or both. One to fifty dilutions were used for each assay. Cells were cultured for 24 hours as described above.

Batches were also evaluated for the ability to inhibit hepatitis B virus as detected by DNA quantitation (see Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499). Briefly, one gram of preparation was added with 10 ml of water. The mixture was treated as Outlined in Table 11. The supernatant was collected after

centrifugation and filtered through a 0.22 μm filter.2.2.15 cells which secrete hepatitis B virons (kindly provided by Professor G. Ace; see Ace et al. Proc Natl Acad Sci USA (1987) 84: 1005-1009) were used in this assay. One to fifty dilutions were used for each assay. The cell growth inhibition assay was performed for 72 hours. All other procedures were performed as described by Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499.

For β-glucuronidase, different scute extracts were added to triplicate wells of a 96-well plate which contained 0.1mM phenolphthalein glucuronidate, 70 mM Tris-HCl (pH 6.8) and 0.8 ng of dialyzed β-glucuronidase (from *E.Coli*, purchased from Sigma) to a final volume of 80 μl. After 2 hr incubation at 37°C, the reactions were terminated with 200 μl of stopping solution which contained 0.2 M Glycine and 0.2 M NaCl (pH 10.4), and the OD was monitored with a kinetic microplate reader at 540 nm.

The results of the assays using the three batches is displayed in Table 12.

Table 12. Biological Assay of Four Preparations of Scute*

	E. Coli β-Glucuronidase	HepG2	Jurkat	HBV‡ DNA
Scute A	1.5	0.33	0.45	None
Scute B	1.8	ND	ND	ND
Scute C	0.3	ND	ND	ND
Scute D	ND	0.65	ND	27.5

*Values represent IC_{50} values.
‡, % of Control
ND, not determined.

Evaluation of Scute Effects on Protein Expression

HepG2 cells (1 x 10^6) were seeded in 25 cm^2 flasks in 3.0 ml of RPMI-1640 medium (see Life Technologies, Inc., Catalogue and Reference Guide, 1998-1999, Cell Culture section) 24 hr before the extract addition. The cells were treated with or

without herbal medicine, where the former is added at two final concentrations of 0.2 mg/ml or 4 mg/ml, respectively, and incubated at 37°C for 24 hours. The medium was removed and the cells were washed twice with cold PBS. The cells were harvested into 1 ml of PBS and centrifuged at 10,000 rpm for 2 minutes, extracted on ice with a buffer containing 50 mM Tris-Cl (pH 7.5), 0.2 mM PMSF and 10% glycerol, followed by three freeze-thaw cycles. Potassium chloride was added to the cell lysate at a final concentration of 0.15 M prior to centrifugation. The protein concentration was determined and the cell extract was electrophoresed according to the method of Laemmli U.K. (*Nature* (1970) 227:680-685). Western blots were performed by standard techniques known in the art, see for example Sambrook, *et al* (1989). The antibodies used were directed to the following proteins: Topo I; Stat (20707); Cyclin B1; MAPK (Ab2) and Nm 23 H1.

Figure 4 demonstrates that scute batches A and B do not differentially affect the expression of the polypeptides resolved on Western blots.

HPLC Analysis

The herbal batches were analyzed by HPLC with a Beckman ODS Ultrasphere™ column (5 micron particles, 4.6 mm X 25 cm) and detected with a UV spectrophotometer (Perkin Elmer). The wavelengths for UV detection were monitored at 280 nm and 340 nm. The mobile phase was pumped at 1 ml/min and consisted of Solvent A: H_2O and Solvent B: 20% MeOH with the following gradient: 1) the solvent was 100% solvent A for the first 5 minutes; 2) the solvent composition was changed to 10% solvent A / 90% solvent B for the next 10 minutes; and 3) the solvent was changed to 10% solvent A / 90% solvent B for the next 40 minutes. This was followed by the addition of 100 % solvent A for 5 minutes. The HPLC markers are baicalin and baicalein.

Scute batches in water and acid treated samples were analyzed by HPLC. Water and acid treated batches were virtually indistinguishable.

Algorithm

The data collected form part of the multidimensional analysis used to generate multivariant normal distribution sets as a means of determining a baseline correlation between biological activity and standard scute chemical (HPLC), and origin/growth characteristics.

C. White Peony Root

Evaluation of *Paeonie lactiflora pallus radix* (Peony)

Peony is used to suppress and soothe pain. It is also known to soothe ligaments and purify the blood. The properties of the peony batches used in this example are presented in Table 13.

Table 13. Batch Properties (Peony)

Property	Batch A	Batch B
Plant Name	*Paeonie lactiflora pallus*	*Paeonie lactiflora pallus*
Origin	Anwey Province	U.S., Boston
Preparation method	Standard	Boiled 2 hours.
Plant part	Root	Root

Biological and Enzyme Assays

Briefly, one gram of each preparation of scute extract was added with 10 ml of water (1 mg/ml). The mixture was treated as outlined in Table 13. The supernatant was collected after centrifugation and filtered through a 0.22 μm filter. Batches of peony were tested against either HepG2 cells (ATCC cat # HB-8065) or Jurkat T cells (ATCC cat #TIB-152) or both. One to fifty dilutions were used for each assay. Cells were cultured for 24 hours as described above.

Batches were also evaluated for the ability to inhibit hepatitis B virus as detected by DNA quantitation (see Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499).

Briefly, one gram of preparation was added with 10 ml of water. The mixture was treated as outlined in Table 13. The supernatant was collected after centrifugation

and filtered through a 0.22 μm filter.2.2.15 cells which secrete hepatitis B virons (kindly provided by Professor G. Ace; see Ace et al. Proc Natl Acad Sci USA (1987) 84: 1005-1009) were used in this assay. One to fifty dilutions were used for each assay. The cell growth inhibition assay was performed for 72 hours. All other procedures were performed as described by Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499.

Different peony extracts were added to triplicate wells of a 96-well plate which contained 0.1mM phenolphthalein glucuronidate, 70 mM Tris-HCl (pH 6.8) and 0.8 ng of dialyzed beta-glucuronidase (from *E.coli*, purchased from Sigma) to a final volume of 80 μl. After 2 hr incubation at 37°C, the reactions were terminated with 200 μl of stopping solution which contained 0.2 M Glycine and 0.2 M NaCl (pH 10.4), and the OD was monitored with a kinetic microplate reader at 540 nm. Results are shown in Table 14.

Table 14. Biological Assay of Two Preparations of Peony*

	E. Coli β-Glucuronidase	HepG2	Jurkat	HBV‡
Peony A	2.8	>1.5	1.1	None
Peony B	>2.5	ND	ND	ND

*Values represent IC_{50} values. ‡, % of Control

ND, not determined.

HPLC Analysis

The herbal batches were analyzed by HPLC with a Beckman ODS Ultrasphere column (5 micron particles, 4.6 mm X 25 cm) and detected with a UV spectrophotometer (Perkin Elmer). The wavelengths for UV detection were monitored at 280 nm and 340 nm. The mobile phase was pumped at 1 ml/min and consisted of Solvent A: H_2O and Solvent B: 20% MeOH with the following gradient: 1) the solvent was 100% solvent A for the first 5 minutes; 2) the solvent composition was changed to

10% solvent A / 90% solvent B for the next 10 minutes; and 3) the solvent was changed to 10% solvent A / 90% solvent B for the next 40 minutes. This was followed by the addition of 100 % solvent A for 5 minutes. HPLC marker is paeoniflorin.

Peony batches were analyzed by HPLC as shown in Figure 5.

Algorithm

The data collected form part of the multidimensional analysis used to generate multivariant normal distribution sets as a means of determining a baseline correlation between biological activity and standard peony chemical (HPLC), and origin/growth characteristics.

D. Date

Evaluation of *Ziziphi Fructus* (Date)

Date has been used for diuretic properties and strengthening effects. The properties of the date batches used in this example are presented in Table 15.

Table 15. Batch Properties (Date)

Property	Batch A	Batch B	Batch C
Plant Name	*Ziziphi Fructus*	*Ziziphi Fructus*	*Ziziphi Fructus*
Origin	Hebei Province.	U.S., Kin Man Herb Center	U.S., Kin Man Herb Center
Preparation method	Standard	Boiled, 30 min	Warm H_2O, 30 min.
Plant part	Fruit	-	-

Biological and Enzyme Assays

Briefly, one gram of each batch of scute extract was added with 10 ml of water (1 mg/ml). The mixture was treated as outlined in Table 15. The supernatant was collected after centrifugation and filtered through a 0.22 μm filter. Batches of date were tested against either HepG2 cells (ATCC cat # HB-8065) or Jurkat T cells (ATCC cat #TIB-152) or both. One to fifty dilutions were used for each assay. Cells

were cultured for 24 hours as described above.

Batches were also evaluated for the ability to inhibit hepatitis B virus as detected by DNA quantitation (see Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499). Briefly, one gram of preparation was added with 10 ml of water. The mixture was treated as outlined in Table 15. The supernatant was collected after centrifugation and filtered through a 0.22 μm filter. HepG2.2.15 cells which secrete hepatitis B virons (kindly provided by Professor G. Ace; see Ace et al. Proc Natl Acad Sci USA (1987) 84: 1005-1009) were used in this assay. One to fifty dilutions were used for each assay. The cell growth inhibition assay was performed for 72 hours. All other procedures were performed as described by Dong *et al.*, Proc Natl Acad Sci USA (1991) 88: 8495-8499.

Different peony extracts were added to triplicate wells of a 96-well plate which contained 0.1mM phenolphthalein glucuronidate, 70 mM Tris-HCl (pH 6.8) and 0.8 ng of dialyzed beta-glucuronidase (from *E.Coli*, purchased from Sigma) to a final volume of 80 μl. After 2 hr incubation at 37°C, the reactions were terminated with 200 μl of stopping solution which contained 0.2 M Glycine and 0.2 M NaCl (pH 10.4), and the OD was monitored with a kinetic microplate reader at 540 nm. Results are shown in Table 16.

Table 16. Biological Assay of Three Preparations of Date*

	E. Coli β-Glucuronidase	HepG2	Jurkat	HBV‡ DNA
Date A	1.2	1.5	5.1	None
Date B	ND	>2.0	ND	52.3
Date C	2.5	ND	ND	ND

*Values represent IC_{50} values.
‡, % of Control
ND, not determined.

HPLC Analysis

The herbal batches were analyzed by HPLC with a Beckman ODS Ultrasphere column (5 micron particles, 4.6 mm X 25 cm) and detected with a UV spectrophotometer (Perkin Elmer). The wavelengths for UV detection were monitored at 280 nm and 340 nm. The mobile phase was pumped at 1 ml/min and consist of Solvent A: H_2O and Solvent B: 20% MeOH with the following gradient: 1) the solvent was 100% solvent A for the first 5 minutes; 2) the solvent composition was changed to 10% solvent A / 90% solvent B for the next 10 minutes; and 3) the solvent was changed to 10% solvent A / 90% solvent B for the next 40 minutes. This was followed by the addition of 100 % solvent A for 5 minutes. HPLC markers for date are chelidonic acid and cAMP.

Date batches samples were analyzed by HPLC as shown in Figure 6.

Algorithm

The data collected form part of the multidimensional analysis used to generate multivariant normal distribution sets as a means of determining a baseline correlation between biological activity and standard peony chemical (HPLC), and origin/growth characteristics.

The foregoing detailed description has been given for clearness of understanding only and no unnecessary limitations should be understood therefrom as modifications will be obvious to those skilled in the art.

While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and this application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features hereinbefore set forth and as follows in the scope of the appended claims.



WHAT IS CLAIMED IS:

1. A method for predicting the biological activity of an herbal composition comprising:

a). exposing a biosystem to a batch herbal composition and measuring the differential responses of two or more molecular markers, wherein the set of differential response measurements constitute an Herbal BioResponse Array (HBR Array) data set;

b). comparing the HBR Array of the batch herbal composition to a least one previously-obtained HBR Array of a known herbal composition with at least one known biological activity; and,

c). predicting the biological activity of the batch herbal composition based on the HBR Array comparison made in step b).

2. The method of claim 1 further comprising repeating steps a) through c) for one or more additional batches of the herbal composition and selecting the batch herbal composition which has the desired biological activity.

3. The method of claim 1 wherein additional previously-obtained HBR Arrays are used for the HBR Array comparison of step b).

4. The method of claim 1 or 3 wherein the markers are selected from the group consisting of molecular markers, cytogenetic markers, biochemical markers and

macromolecular markers.



5. The method of claim 1 or 3 wherein the previously-stored HBR Arrays are standardized HBR Arrays for the same or substantially similar herbal compositions as that of the batch herbal composition.

6. The method of claim 5 further comprising adjusting or modifying the batch herbal composition to produce a HBR Array substantially similar to that of one or more of the standardized HBR Arrays.

7. The method of claim 1 or 3 further comprising using the results of the HBR Array comparisons to identify specific molecules in the batch herbal composition which retain the desired biological activity of the known herbal composition.

8. The method of claim 1 or 3 further comprising using the results of the HBR Array comparisons to determine which herbal components of a known herbal composition can be eliminated from the known herbal composition while maintaining or improving the desired biological activity of the known herbal composition.

9. The method of claim 1 or 3 wherein the results of the HBR Array comparisons identify one or more previously unknown biological activities for the batch herbal



composition.

10. The method of claim 1 or 3 further comprising using the predicted biological activity of the batch herbal composition to aid in the design of therapeutics which include herbal components and synthetic chemical drugs.

11. A method of establishing a standardized Herbal BioResponse Array (HBR Array) for an herbal composition, wherein the method comprises:

a). selecting an herbal composition with at least one known BioResponse;

b). exposing a biosystem to a batch of the herbal composition and collecting data on two or more markers;

c). storing the marker data of step b) as an HBR Array;

d). repeating steps b) and c) for one or more additional batches of the herbal composition using two or more of the same or different markers than used in step b);

e). combining the HBR Arrays obtained in steps c) and d); and,

f). analyzing the combined HBR Array of step e) to generate a standardized HBR Array for the herbal composition, wherein the standardized HBR Array has data for two or more markers which are correlated with at least one known BioResponse of the herbal composition.

12. The method of claim 11 further comprising exposing a biosystem to one or

more batches of the herbal composition, collecting data on one or more BioResponses, and adding the collected BioResponse data to the standardized HBR Array for that herbal composition.

13. A method of evaluating an herbal composition comprising:

a). exposing a biosystem to a batch of the herbal composition and collecting data on two or more markers; and

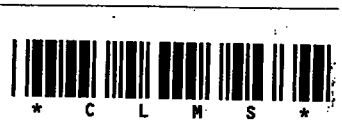
b). comparing the collected marker data with a standardized HBR Array for the same or a substantially same herbal composition as that of the batch herbal compositions.

14. The method of claim 13 further comprising predicting the BioResponse of the batch herbal composition.

15. The method of claim 11 or 13 wherein the markers are selected from the group consisting of molecular, cytogenetic, biochemical and macromolecular markers.

16. The method of claim 11 or 13 wherein the standardized HBR Array further comprises information on plant-related data.

17. A method of establishing a standardized Herbal BioResponse Array (HBR



Array) for an herbal composition with a known BioResponse, wherein the method comprises:

a). exposing a biosystem to a batch of the herbal composition and measuring the qualitative and quantitative changes of the resulting gene expressions to generate data as an HBR Array for that batch;

b). repeating step a) using different preparations of the herbal composition to generate data as additional HBR Arrays;

c). selecting a set of discriminating genetic markers by analyzing the HBR Arrays obtained in steps a) and b) to establish a standardized HBR Array for the herbal composition.

18. A method of evaluating an herbal composition comprising:

a). exposing a biosystem to a batch of the herbal composition and measuring the qualitative and quantitative changes of the resulting gene expressions to present as an Herbal BioResponse Array (HBR Array) for that batch; and,

b). comparing the HBR Array obtained in step a) with a standardized HBR Array for a substantially equivalent herbal composition.

19. The method of claim 17 or 18 wherein gene expression takes place transcriptionally.

* C L M S *

20. The method of claim 17 or 18 wherein gene expression takes place translationally.

21. The method of claim 1, 11, 12, 13, 17 or 18 wherein the biosystem is selected from the group consisting of cells, tissues, organs, whole organisms and *in vitro* assays.

22. The method of claim 11, 14 or 17 wherein the BioResponse is selected from the group consisting of physiological responses, morphological responses, cognitive responses, motivational responses and autonomic responses.

23. A system for predicting the biological activity of an herbal composition comprising:

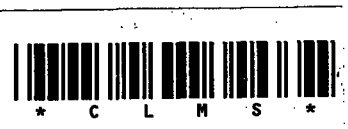
a. a biosystem comprising one or more different types of cells, tissues, organs or *in vitro* assays;

b. a batch herbal composition;

c. two or more molecular markers;

d. a means for exposing the biosystem to the batch herbal composition and measuring the differential responses of the molecular markers;

e. a computer processor, including memory, for analyzing and storing the differential response measurements of the molecular markers so as to create an Herbal BioResponse Array (HBR Array) data set for the batch herbal composition;



f. a computer processor, including memory, for comparing the HBR Array of the batch herbal composition to one or more previously-stored HBR Arrays so as to predict the biological activity of the batch herbal composition, wherein the biological activities of the herbal compositions used to generate the one or more previously-stored HBR Arrays are known.