Journal of Ethnopharmacology 137 (2011) 77–84

Contents lists available at ScienceDirect

Journal of Ethnopharmacology

journal homepage: www.elsevier.com/locate/jethpharm

TCM grammar systems: An approach to aid the interpretation of the molecular interactions in Chinese herbal medicine

Jing Yan, Yun Wang*, Si-jun Luo, Yan-jiang Qiao

Beijing University of Chinese Medicine, School of Chinese Pharmacy, No. 6, Sourthern Rd., Beijing 100102, China

ARTICLE INFO

Article history:
Received 19 October 2010
Received in revised form 23 January 2011
Accepted 20 April 2011
Available online 28 April 2011

Keywords:
Entity grammar systems
Traditional Chinese medicines
Herba Ephedrae Decoction

ABSTRACT

Ethnopharmacological relevance: Interpreting the molecular interactions in Chinese herbal medicine will help to understand the molecular mechanisms of Traditional Chinese medicines (TCM) and predict the new pharmacological effects of TCM. Yet, we still lack a method which could integrate the concerned pieces of parsed knowledge about TCM.
Materials and methods: To solve the problem, a new method named TCM grammar systems was proposed in the present article. The possibility to study the interactions of TCM at the molecular level using TCM grammar systems was explored using Herba Ephedrae Decoction (HED) as an example.
Results: A platform was established based on the formalism of TCM grammar systems. The related molecular network of Herba Ephedrae Decoction (HED) can be extracted automatically. The molecular network indicates that Beta2 adrenergic receptor, Glucocorticoid receptor and Interleukin12 are the relatively important targets for the anti-anaphylaxis asthma function of HED. Moreover, the anti-anaphylaxis asthma function of HED is also related with suppressing inflammation process. The results show the feasibility using TCM grammar systems to interpret the molecular mechanism of TCM. Although the results obtained depend on the database absolutely, recombination of existing knowledge in this method provides new insight for interpreting the molecular mechanism of TCM.
Conclusions: TCM grammar systems could aid the interpretation of the molecular interactions in TCM to some extent. Moreover, it might be useful to predict the new pharmacological effects of TCM. The method is an in silico technology. In association with the experimental techniques, this method will play an important role in the understanding of the molecular mechanisms of TCM.

1. Introduction

Traditional Chinese medicines (TCM) usually contain complex active ingredients and exhibit broad pharmacologic effects. However, some of them also contain toxic ingredients. In order to apply TCM in a better way, it is necessary to know their mechanisms at molecular level.

Nowadays, there are many approaches to study the molecular mechanism of TCM, such as experimental approaches, particularly molecular pharmacological experiments, including high-throughput screening (Al-Sayah et al., 2008), high content screening (Nadanaciva et al., 2010), and metabonomics (Barton, 2011); and informatics approach, such as virtual screening, including molecular docking (Adane and Bharatam, 2010), and pharmacophore (Deng, 2008). These methods mainly focus on the identification of targets for chemical constituents of TCM. Because of the complexity of the chemical components of TCM and the sophisticated mechanisms *in vivo*, it is hard to elucidate the exact functional mechanisms of TCM and the interaction between the chemical components by these methods. In recent years, the proposed chemical systems biology approach offered a new avenue for the study of the molecular mechanisms of TCM (Van et al., 2009). However, due to the complex nature of TCM, which include combination of Chinese Herbal Formula apart from the herb itself, there is currently no systematic method to study the interactions of TCM at the molecular level. We still lack a method which could integrate the concerned pieces of parsed knowledge about TCM. In this paper, we present a new method based on the entity grammar systems to solve the problem and explore the possibilities to study the interactions of TCM at the molecular level using this method. Because this method is proposed to study TCM, we call it TCM grammar system for convenient.

Abbreviations: TCM, Traditional Chinese medicines; HED, Herba Ephedrae Decoction; EGS, entity grammar systems; ASP, answer set programming; IL1, Interleukin 1; IL1R1, Interleukin 1 receptor 1; IRAK1, Interleukin-1 receptor-associated kinase; TNFR, Tumor Necrosis Factor Receptor; TRAF6, Tumor Necrosis Factor Receptor Associated Factor 6; IKK, IkB kinase; TNF-alpha, Tumor Necrosis Factor alpha; NO, nitric oxide; PKA, Protein Kinase A; AA, arachidonic acid; IL12, Interleukin12; EOS, eosinophils; BHR, bronchial hyperresponsiveness.

* Corresponding author. Tel.: +86 10 84738620.
E-mail address: wangyun@bucm.edu.cn (Y. Wang).

0378-8741/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.jep.2011.04.057

Entity grammar systems (EGS) are a formal language proposed for modeling the complex hierarchies in biological systems. The definition of EGS has been introduced by Wang (2004). The relationship between EGS and Chomsky formal language has been studied in the same paper. Using the EGS as the framework, the formal model of complex systems can be easily constructed using the method mentioned in the paper. Therefore, EGS have been successfully used in the modeling, learning and simulation of biological cells. A recently published paper has discussed the emergence in entity grammar systems (Wang et al., 2007). The formal definition of emergence based on entity grammar systems provide formal tool to study the entangled hierarchies in biological systems and to control the generation of emergence in certain conditions. In this paper, we use EGS as the framework to construct the TCM grammar system.

2. TCM grammar systems

2.1. Brief introduction of entity grammar systems

In this section, we will recall the basic definitions of entities and entity grammar systems.

An alphabet V is a set of symbols. An organizer F is a set of operations (or functions). The set $E(V, F)$ of all F-entities over V is inductively defined as:

(1) $\lambda \in E(V, F)$, λ is the empty entity with no symbols;
(2) $V \subseteq E(V, F)$;
(3) For all $f \in F$, and all $\xi_1, \ldots, \xi_n \in E(V, F)$, we have $f(\xi_1, \ldots, \xi_n) \in E(V, F)$.

The set $E^+(V, F)$ of all non-empty F-entities over V is $E^+(V, F) = E(V, F) - \{\lambda\}$.

Suppose f is an n-ary operation in F, the set of positions of f is the set $Pos(f) = \{1, 2, \ldots, n\}$. The set of positions of an entity $\xi = f(\xi_1, \ldots, \xi_n) \in E(V, F)$, denoted by $Pos(\xi)$, is inductively defined as: (1) $Pos(\lambda) = \Phi$; (2) for $\xi \in V$, $Pos(\xi) = \{\lambda\}$; (3) for an entity $\xi = f(\xi_1, \ldots, \xi_n)$, $Pos(\xi) = \bigcup_{i=1}^{n} \{ip | p \in Pos(\xi_i)\}$. The size $|\xi|$ of an entity is the cardinality of $Pos(\xi)$.

For an entity with the form $f(\xi_1, \ldots, \xi_n)$, if at least one ξ_i is λ, the entity is called pseudo-f-entity. For all operations f in F, the set of the all pseudo-f-entities is called pseudo-F-entities, denoted by PFE. An operation $f : (E(V, F))^n \to E(V, F)$ in F is called pseudo-operation, if for any set $A \subseteq E(V, F)$, $f(A) = A$.

An entity grammar G is a quintuple, $G = (V_N, V_T, F, P, S)$, where, V_N is finite set of non-terminal symbols, V_T is finite set of terminal symbols, and $V_N \cap V_T = \Phi$, F is finite set of operations, $F = \{f_i | f_i : (E(V, F))^n \to E(V, F), 1 \le i \le m, m, n \in N\}$, where $V = V_N \cup V_T$, P is a finite set of productions $\alpha \to \beta$ with $\alpha \in E^+(V, F)$ and $\beta \in E(V, F)$, S is the start entities.

Let $G = (V_N, V_T, F, P, S)$ be an entity grammar. Then the set

$$L(G) = \{\xi \in E(V_T, F) | S \Rightarrow_G^* \xi\}$$

is the language generated by G, where $S \Rightarrow_G^* \xi$ represents an unspecified number of derivations (including zero) that can be taken from S to ξ. For $x, y \in E(V, F)$, $f \in F$ we say that y is directly derivable from x in G, denoted by $x \Rightarrow_G y$, if and only if for some $\alpha \to \beta$ in P and $u, v \in E(V, F)$, we get $x = f(u, \alpha, v)$ and $y = f(u, \beta, v)$.

If the distinction between non-terminals and terminals are not concerned about, an entity grammar can be expressed as $G = (V, F, P, S)$, where V is the alphabet of the system.

Using the following several steps, a new grammar system can be constructed (Wang, 2004).

(1) Define the structure of a kind of entities and the alphabet of the grammar. From the definition of entity, any concrete entity has the general form $f(\xi_1, \ldots, \xi_n)$, which is composed of a set of components $\{\xi_1, \ldots, \xi_n\}$ and an operation f on the set. The structure of the entity or the relationship of components is determined by the operation. A concrete form of the operation is corresponding to a concrete kind of entities. The operation f might be composed of many different operations, which operate on different subset of $\{\xi_1, \ldots, \xi_n\}$ or define the different relationship of components.
(2) Define the elements of the set of operations. The set of operations is an important part and the basis of entity grammar systems. A successful definition of the set of operations is the prerequisite of a useful entity grammar system. The elements of the set of operations can be divided into two categories: (1) the operations to define the basic structure of entities, which has been defined in the first step, (2) the operations to define the interaction of entities. As defined in Definition 3.1, all the operations are closed in $E(V, F)$, so all of the results of the operations are also entities. Suppose $f(x_1, \ldots, x_n)$ is a general operation in F, if it is the first type operation, it could be used for all $x_1, \ldots, x_n \in E(V, F)$. If it is the second type operations, it satisfies at least one of $x_i \notin V$.
(3) Define the production rules of grammar. The basic type of production rules is recursive productions, context-sensitive productions, context-free productions and regular productions. For a successful grammar, the productions should reflect the changing law of the systems being studied.

Usually, if the entities we defined in step 1 is called "X", the corresponding grammar is called "X Grammar" and the system described by this grammar is called "X Grammar System". From the definition of entities grammar system, TCM grammar systems can be derived from it.

2.2. Definition of TCM grammar systems

Definition 1. An entity grammar system $G = (V, F, P, S)$ is called an TCM grammar system, if

(1) $V = V_1 \cup V_2 \cup V_3 \cup V_4$, where V_1 is a set of TCM formula, V_2 is a set of TCM herb medicine, V_3 is a set of components of TCM medicine and V_4 is a set of endogenous molecules.
(2) $F = F_1 \cup F_2 \cup F_3 \cup F_4$;

$$F_1 = \{f_1 | f_1 : (V_i, V_j) \to E(V, F), 1 \le i \le 3, 2 \le j \le 3\},$$
$$F_2 = \{f_2 | f_2 : (V_3, V_4) \to E(V, F)\},$$
$$F_3 = \{f_3 | f_3 : (V_4, V_4) \to E(V, F)\},$$
$$F_4 = \{f_4 | f_4 : V \to E(V, F)\}.$$

(3) $P = P_1 \cup P_2$,

$$P_1 = \{p_1 | p_1 : \alpha \Rightarrow \beta, \alpha \in E^+(V_1 \cup V_2 \cup V_3, F_1), \beta \in E^+(V_1 \cup V_2 \cup V_3, F_1)\},$$
$$P_2 = \{p_2 | p_2 : \alpha \Rightarrow \beta, \alpha \in E^+(V_3 \cup V_4, F_2 \cup F_3, F_4), \beta \in E^+(V_3 \cup V_4, F_2 \cup F_3 \cup F_4)\}$$

In this definition, F_1 is the set of the relationships between TCM formula and its ingredients or herbal medicine and its components. F_2 is the set of the interaction between the components of herbal medicine and the biological targets. F_3 is the set of interaction between endogens molecules. F_4 is the set of properties of the entities. Using the rules in P_1, the relationship chain from formulas,