

Banco de DAdos de PROdutos Naturais (BADAPRON) e inteligência artificial na desreplicação de matrizes vegetais

Tiago Branquinho Oliveira¹ (PG)*, Fernando Batista da Costa¹ (PQ)

¹Universidade de São Paulo (USP), Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Laboratório de Farmacognosia, Ribeirão Preto, SP. * tiagotbo@fcrp.usp.br

Palavras Chave: produtos naturais, inteligência artificial, quimioinformática, quantitative structure-retention relationships, redes neurais artificiais.

Introdução

Uma alternativa para reduzir o tempo e custos no processo de desreplicação de matrizes vegetais é o uso de bibliotecas de produtos naturais. Estas bibliotecas, sejam elas acadêmicas ou comerciais, geralmente são constituídas de várias substâncias e seus dados associados, como espectrométricos ou cromatográficos^{1,2}.

No entanto, a simples elaboração de uma biblioteca de produtos naturais não garante o aumento de produtividade. É recomendável que parte do processo seja automatizada, como no caso da desreplicação. Neste contexto, destacam-se o emprego de inteligência artificial e de algumas ferramentas de quimioinformática, como a elaboração de bancos de dados de estruturas químicas e ferramentas de predição de tempos de retenção (t_R). Neste sentido, propomos o desenvolvimento do Banco de DAdos de PROdutos Naturais (BADAPRON).

Resultados e Discussão

O BADAPRON foi elaborado a partir de uma biblioteca de substâncias puras isoladas basicamente de espécies da família Asteraceae associadas a seus respectivos dados de cromatografia líquida de alta eficiência (CLAE), espectrométricos, informações taxonômicas e bibliográficas. A biblioteca é composta por diferentes classes de metabólitos secundários, perfazendo atualmente cerca de 110 substâncias. Seu gerenciamento é realizado através do software Symyx Isentris 3.3 e os dados podem ser filtrados e exportados para o Microsoft Excel™ ou em arquivos SDF.

Um estudo de caso aplicando a técnica *Quantitative Structure-Retention Relationships* (QSRR) com as estruturas químicas do BADAPRON, especialmente as lactonas sesquiterpênicas (LST), foi empregado para construir modelos para a predição de seus t_R em CLAE.

A obtenção dos t_R experimentais envolveu o uso de coluna de fase reversa C-18, eluição isocrática em dois eluentes (MeOH-H₂O 1:1 e MeCN-H₂O 3:7) e detecção no UV. Os t_R são relativos a substância de referência 2,3-dimetilfenol (DMP).

As estruturas em 2D foram desenhadas no software BKChem 0.130 e as coordenadas em 3D foram obtidas pelo software Corina 3.4.8.

Foram calculados descritores constitucionais, propriedades moleculares, grupos funcionais e descritores 3D utilizando o software DRAGON Professional 5.5.

Para se estimar os t_R , nos dois eluentes, foram elaborados modelos empregando-se regressão linear múltipla (RLM) e redes neurais artificiais (RNA). As estruturas químicas codificadas compuseram a entrada de dados nos modelos e o t_R foi a resposta. Os melhores descritores foram selecionados por serem relacionados ao mecanismo de retenção de analitos em fase reversa.

Em um primeiro estudo, os modelos foram feitos com 51 estruturas de LST e a validação foi através de validação cruzada 10-fold. No segundo estudo, 19 estruturas foram selecionadas para treino dos modelos utilizando-se SOMs (*Self Organizing Maps*), restando 32 para teste (validação externa).

Verificou-se, também, que o desempenho da RNA foi melhor ao dos modelos de RLM e que as melhores predições foram obtidas para os modelos utilizando MeCN-H₂O como fase móvel.

Em uma próxima etapa serão empregados dados em sistema isocrático.

Conclusões

O BADAPRON foi capaz de produzir uma fonte conveniente de informações e a utilidade de seus dados foi ilustrada pelo estudo de caso envolvendo QSRR para se investigar o mecanismo de retenção de LST.

Agradecimentos

Alexander von Humboldt Stiftung (Alemanha), CAPES, FAPESP, CNPq, Molecular Networks GmbH (Alemanha), Computer-Chemie Centrum (Universidade de Erlangen, Nuremberg, Alemanha) e Centro Universitário de Anápolis-GO.

¹ Quinn, R.J.; Carroll, A.R.; Pham, N.B.; Baron P.; Palframan, M.E.; Suraweera, L.; Pierens, G.K.; Muresan S. J. Nat. Prod. **2008**, 71, 464-468.

² Bindseil, K.U.; Jakupovic, J.; Wolf, D.; Lavayre, J.; Leboul, J.; van der Pyl, D. DDT, **2001**, 6, 840-847.