Food and Chemical Toxicology 112 (2018) 526–534

Contents lists available at ScienceDirect

Food and Chemical Toxicology

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

Chemical toxicity prediction for major classes of industrial chemicals: Is it possible to develop universal models covering cosmetics, drugs, and pesticides?

Vinicius M. Alves [a, b], Eugene N. Muratov [a, c], Alexey Zakharov [d], Nail N. Muratov [c], Carolina H. Andrade [b], Alexander Tropsha [a, *]

[a] Laboratory for Molecular Modeling, Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, 27599, USA
[b] Laboratory of Molecular Modeling and Drug Design, Faculty of Pharmacy, Federal University of Goiás, Goiânia, GO, 74605-170, Brazil
[c] Department of Chemical Technology, Odessa National Polytechnic University, Odessa, 65000, Ukraine
[d] National Center for Advancing Translational Sciences (NCATS), National Institutes of Health, Rockville, MD, 20850, USA

ARTICLE INFO

Article history:
Received 31 December 2016
Received in revised form
16 March 2017
Accepted 10 April 2017
Available online 12 April 2017

Keywords:
Cosmetics
Drugs
Pesticides
Chemical space
QSAR models
Prediction

ABSTRACT

Computational models have earned broad acceptance for assessing chemical toxicity during early stages of drug discovery or environmental safety assessment. The majority of publicly available QSAR toxicity models have been developed for datasets including mostly drugs or drug-like compounds. We have evaluated and compared chemical spaces occupied by cosmetics, drugs, and pesticides, and explored whether current computational models of toxicity endpoints can be universally applied to all these chemicals. Our analysis of the chemical space overlap and applicability domain (AD) of models built previously for twenty different toxicity endpoints showed that most of these models afforded high coverage (>90%) for all three classes of compounds analyzed herein. Only *T. pyriformis* models demonstrated lower coverage for drugs and pesticides (38% and 54%, respectively). These results show that, for the most part, historical QSAR models built with data available for different toxicity endpoints can be used for toxicity assessment of novel chemicals irrespective of the intended commercial use; however, the AD restriction is necessary to assure the expected prediction accuracy. Local models may need to be developed to capture chemicals that appear as outliers with respect to global models.

1. Introduction

Chemical toxicity assessment is a critical point in regulatory decision making that concerns the release of drugs or industrial chemicals into production, which enables their human or environmental exposure (Parasuraman, 2011). There exists also a variety of natural and synthetic substances that are exposed to humans and/or the environment that have never been evaluated in any toxicity testing protocol (Chuprina et al., 2010; Egeghy et al., 2012). Over the years, the society has tolerated the use of animals in laboratory toxicity testing. However, in recent years, there has been an increased pressure on scientists and regulatory agencies to replace potentially hazardous chemicals by safer alternatives (Collins, 2003; Schulte et al., 2013). In addition, there has been a strong push on the part of both regulatory agencies such as FDA and EPA in the United States and their counterparts around the world to avoid animal testing of every chemical as such testing has become increasingly unsustainable in terms of both cost and time needed to conduct animal trials (Burden et al., 2015).

The development of the alternative *in vitro* and *in silico* approaches has been encouraged and supported by both NIH and EPA through large-scale programs such as ToxCast project (Dix et al., 2007) and the Tox21 consortium (Tice et al., 2013). Similar programs such as Endocrine Disruptors Prioritization List (http://ec.europa.eu/environment/chemicals/endocrine/index_en.htm) and the priority substances for water safety (European Union, 2013) have been funded by the European Union. Since the acceptance of Registration, Evaluation, Authorization, and Restriction of Chemicals (REACH) legislation in 2006 by the European Union (European

* Corresponding author. 100K Beard Hall, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, 27599, USA.
E-mail address: alex_tropsha@unc.edu (A. Tropsha).

http://dx.doi.org/10.1016/j.fct.2017.04.008
0278-6915/

Union, 2007; Nicolotti et al., 2014), the use of structural alerts and statistical QSAR models (often collectively referred to as (Q)SAR) have become a major computational approach to chemical safety assessment and regulatory decision support.

The majority of publicly available models for toxicity prediction have been built for drugs or drug candidates (Benfenati et al., 2009; Melnikov et al., 2016) or environmental chemicals (Naven and Louise-May 2015). In contrast, computational toxicity models for another large group of industrial chemicals, namely cosmetics products have been developed to a much lesser extent as the animal testing has been used as a preferred approach. However, with recent regulations banning the use of animals for testing of the cosmetics products (European Commission, 2013), there has been a resurgence of interest in employing computational models for their toxicity assessment (Bois et al., 2016; Cronin et al., 2012).

Naturally, a question can be posed as to whether toxicity prediction models built for environmental chemicals or drug molecules could be employed for the cosmetics products. The answer to this question depends on the overlap of the chemical spaces occupied by cosmetics, drugs, and environmental chemicals and the size of the applicability domain (AD) of the respective models. AD is commonly defined as the threshold of similarity between a new chemical and molecules in the training set used to develop the respective QSAR model (Netzeva et al., 2005; Tropsha, 2010; Tropsha and Golbraikh, 2007); only predictions for new molecules within the AD of QSAR models, *i.e.*, relatively similar to the modeling set are considered reliable. Importantly, the size of the AD is fully defined by the size and diversity of the modeling set and the computational method used to develop QSAR models. For instance, it is known that the chemical space of drugs has been changing over the past few decades (Deng et al., 2013) creating a challenge for "old" models' ability to evaluate new compounds. The applicability of current models to many new compounds was also questioned due to limited size and diversity of data available publicly for model building (Kulkarni et al., 2016).

The considerations above capture both significant advantages and challenges associated with the idea of using models developed with one group of industrial chemicals to evaluate toxicity of another group. Obvious advantages deal with significant savings in time and effort afforded by the opportunity to use previously developed models of multiple toxicity endpoints relevant to drugs and/or environmental chemicals (*e.g.*, pesticides) to evaluate toxicity of cosmetic products. However, since chemicals used in different areas of commerce such as drug, chemical, or cosmetic industries are developed with very different applications in mind, there is no *a priori* reason to expect that their respective chemical spaces overlap. Taking the issue of the AD into account, investigations into studying the degree of such overlap and the applicability of models developed for one group of chemicals to predict toxicity of another group are potentially highly impactful for the respective industries, especially, cosmetics. To the best of our knowledge, such investigations have not been conducted in the public domain with large groups of industrial chemicals.

Herein, we have aimed to compare chemical spaces occupied by cosmetics, drugs, and pesticides, and analyze whether current computational models of different toxicity endpoints can be universally applied to all chemicals. To achieve these aims, we have (i) compiled, curated, and integrated chemical structures of known cosmetics, drugs, and pesticides; (ii) analyzed the distribution of these compounds in chemical space and estimated the structural similarity between the datasets; (iii) performed cluster analysis followed by toxicity annotation comparison for structurally similar compounds in the same clusters; (iv) predicted toxicities of investigated compounds with QSAR models for endpoints developed earlier by us; (v) and analyzed the coverage of these models separately for drugs, cosmetics, and pesticides. We observed that, with some exceptions, the majority of compounds in all three groups of industrial chemicals were found within the AD of QSAR models built previously for twenty different toxicity endpoints. These findings open the door for the development and employment of global toxicity models applicable to the majority of chemicals in commerce while suggesting the need to develop local models that could capture AD outliers of the global models.

2. Materials and methods

2.1. Datasets

2.1.1. Cosmetic ingredients (Dataset A)

The cosmetics ingredients were retrieved from the CosIng, the European Commission database for information on cosmetic substances and ingredients (https://ec.europa.eu/growth/sectors/cosmetics/cosing_en). This dataset included 5166 chemical records with a defined chemical structure. After curation (*vide infra*), 3930 unique chemical substances were kept for this study.

2.1.2. Drugs (Dataset B)

We retrieved 7000 chemical records from the 2014 Leadscope Marketed Drugs Database (http://www.leadscope.com/marketed_drugs_database/). After curation, 4671 unique chemical substances were kept for this study.

2.1.3. Pesticides (Dataset C)

We retrieved 3001 chemical records from the EPA's Pesticide Product Information System Database (https://www.epa.gov/ingredients-used-pesticide-products/ppis-download-product-information-data). After curation, 2044 unique chemical substances were kept for this study.

2.2. Data curation

The datasets were thoroughly curated using the workflows proposed earlier by our group (Fourches et al., 2016, 2015, 2010). Briefly, specific chemotypes such as aromatic and nitro groups as well as double bonds were normalized, and absolute stereo configurations removed using the ChemAxon Standardizer (v.16.10.24.0, ChemAxon, Budapest, Hungary, http://www.chemaxon.com). Polymers, substances with undefined chemical substructure, and substances with molecular weight above 1000 DA were removed. Counterions, inorganic salts, organometallic compounds, and mixtures were removed. After structural standardization, the duplicates were identified with HiT QSAR software (Kuz'min et al., 2008) and carefully analyzed. Within the same dataset, only one record was kept and all duplicates were eliminated. The entire collection (datasets A, B, and C) comprised 9785 unique chemical compounds. As one can see in Fig. 1, 99 compounds were simultaneously labeled as cosmetics, drug, and pesticide; 220 were labeled as cosmetics and drugs; 270 were labeled as cosmetics and pesticides; 172 were labeled as drugs and pesticides; 3341 compounds were labeled only as cosmetics; 4180 were labeled only as drugs; and 1503 were labeled only as pesticides.

2.3. Molecular descriptors

We have calculated the same molecular descriptors as in our previously built QSAR models of toxicity endpoints used in this study (see Table 1 for more detailed information about descriptors, models, and respective references). Majority of the models were built using DRAGON descriptors (Talete SRL, 2007). hERG models