Physicochemical space of synthetic and natural pesticides – a meta-analysis

Short Title : Physicochemical space of pesticides

Sylvain Dubey[1,2*], Florian Guignard[1], Alain Gaume[3], Katia Gindro[3], Sylvain Schnee[3], Mathis Joz-Roland[1], Olivier Broennimann[2+], Olga Dubey[1+]

[1] AgroSustain SA, CH-1260 Nyon, Switzerland

[2] Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland

[3]Agroscope, Swiss Federal Agricultural Research Station in Changins, CH-1260 Nyon, Switzerland

[*]Corresponding author

E-mail: sylvain.dubey@unil.ch

+ contributed equally

Abstract

The first commercial use of synthetic pesticides for crop protection dates back to the 1940s, followed by a fast spreading of their use and the development of a large number of compounds. In contrast to synthetic pesticides that are nowadays designed with the help of artificial intelligence that includes computational science and combinatorial chemistry, natural pesticides are the results of long evolutionary processes involving specific host-pathogens, predator-prey and competitor interactions. For these reasons, natural pesticides are often more specific and less harmful to the environment. Natural pesticides are very diverse and can be found in various living organisms. In the present study, we investigated differences in the physicochemical space occupied by synthetic and natural pesticides. In this respect, we measured the mean and breadth of synthetic and natural pesticides, as well as the overlap between these groups in a reduced physicochemical space derived from a set of 44 physicochemical variables. We showed that physicochemical space strongly differs between synthetic and natural pesticides and could be determined with 93-100% certainty, a result comparable to differences observed in drugs. Importantly, the physicochemical space occupied by synthetic pesticides was 2.6 fold smaller than the one of natural pesticides and toxicity potential was lower in the latter. In conclusion, our work showed that the design of commercialized synthetic pesticides is underexploiting the possible physicochemical space of known natural pesticides, likely due to specific constraints. Such limitations should trigger the development of efficient natural pesticides avoiding as much as possible detrimental effects on non-target organisms.

Keywords: pesticide, insecticide, fungicide, herbicide, synthetic, natural, marine organism, physicochemical space;

1. Introduction

The first commercial use of synthetic pesticides for crop protection dates to the late 1940s (Annand 1944), followed by a fast spreading of their use and the development of a large number of compounds that can be classified in three main categories, i.e., insecticides, fungicides and herbicides (Russel 2005; Duke and Powles 2008). Due to the observation of potentially dangerous effects (direct toxicity and bioaccumulation) of synthetic pesticides (not found in living organisms and hence synthetized) used to date, as well as the reinforcement of risk assessment criteria, the commercialized active ingredients are undergoing a risk re-evaluation. This situation, coupled with political and social concerns, is leading to a search for new substances with a cleaner ecotoxicological profile. The screening of bioactive natural products (natural pesticides; compounds found in living organisms and extracts or synthetize based on their molecular structure) is yet not very efficient (large number of candidates has to be evaluated for identification of only a few leads) and expensive, and their development is still mainly following traditional approach such as combinatorial chemistry and high-throughput screening (Lindell et al. 2009). Typical approaches include Quantitative Structure-Activity Relationship (QSAR) (Speck-Planche et al. 2011), Molecular docking (Cortes-Hernandez et al. 2020) or Structure-Based Design (SBD) (Walter 2002), following Lipinski's and derived rules (primarily developed for drug discoveries; Lipinski et al. 2001; Tice 2001; Hao et al. 2011; Avram et al. 2014). More recently, Machine Learning has also been used as a novel tool for the design of pesticides (Orsolic et al. 2021).

Natural pesticides are the results of long evolutionary processes involving specific host-pathogens, predators and competitor interactions (e.g., Bennett and Wallsgrove 1994; Künzler 2018; Erb and Kliebenstein 2020) rather than made from mostly toxic compounds with specific chemical patterns that are not present in nature and hence less biodegradable (such as e.g., nitroaromatic, halogenated hydrocarbons and organophosphate compounds; Chen and Cashman 2013; Jeschke 2017; Debarati 2020; see Figure S1, supplementary information, for some illustrated examples). Natural pesticides can be found in many living organisms from plants, bacteriae, fungi to marine organisms such as sponges and echinoderms and are hence very diverse (Copping and Duke 2007; Cheung et al. 2014; Singh et al.

2014; Dubey et al. 2020). For these reasons, natural pesticides are often more specific and less harmful to the environment and human health than synthetic ones (Reganold et al. 2001; Rattan et al. 2010; Nawaz et al. 2016; Neeraj et al. 2017; Marrone 2019), In addition, they are often highly biodegradable (Oguh et al. 2019; Duran-Lara et al. 2020; Lengai et al. 2020). Nevertheless, in some instances, natural pesticides can still have detrimental effects on the environment (e.g. Bahlai et al. 2010; Ndakidemi et al. 2016). Despite high value for a sustainable agriculture (Ndolo et al. 2019), the commercial use of biopesticides still represents a very small part of the overall pesticide market (~6%), partly due to the difficulty of identifying active ingredients from natural extracts and the challenges linked to upscale their production (Rijswijk et al. 2018; Suteu et al. 2020).

In an attempt to highlight physicochemical similarities/discrepancies between natural products and synthetic pesticides, scientists have developed libraries for pesticides based on natural products (Cordier et al. 2008). This approach is also called Biology Oriented Synthesis, which is based on the use of natural products as the core skeleton structure and has already led to the discovery of drugs (NorenMuller et al. 2006). Such an approach can be illustrated by nicotine, naturally found in tobacco plants that led to neonicotinoid insecticides (Seifert 2005) and strobilurins from Basidiomycete fungi that led to compounds such as azoxystrobin commonly used as fungicide (Rodrigues et al. 2013; Nofiani et al. 2018). However, it does not guarantee a lack of dramatic impacts on the environment (Yamamuro et al. 2019).

Several studies have attempted to pinpoint physicochemical differences of synthetic and natural compounds from various origins in order to design new drugs (Ertl and Schuffenhauer 2008; Muigg et al. 2013; Chen et al. 2018; Shang et al. 2018), but such studies focusing on pesticides are lacking. Studies focusing on drugs revealed that synthetic drug-like compounds commonly differ from natural ones by being e.g. small, flat, rigid with a high degree of aromatic characters compared to natural ones that are generally more complex and exhibit differences depending of their origin (e.g. from marine versus non marine organisms; Muigg et al. 2013; Stratton et al. 2015; Chen et al. 2020). A diversity that is limited in synthetic compounds by constraints related to their synthesis and conventional rules such

as Lipinski's Rules-of-five for orally absorbed drugs (Lipinski et al. 2001; Doak et al. 2014; Chen and Kirchmair 2020). However, the number of drugs with a chemical space differing from such a rule is increasing in the recent years, suggesting that the way most of the drugs are being designed should be adapted in consequence to lower potential lost opportunities (Doak et al. 2014; Poongavanam et al. 2018; Caron et al. 2021). In the pharmaceutical industry, studies related to the comparison of physicochemical properties between natural and synthetic compounds are quite abundant and well-known. However, we identified that to date there was only a limited number of research conducted on pesticides. Indeed, present studies are scarce and limited to a reduced number of compounds and/or analyzed parameters, precluding the drawing of general conclusions (e.g. Zaid et al. 2010; Akamatsu 2011; Chen et al. 2018; Smith and Perfetti 2020). The lack of in-depth studies related to the assessment of physicochemical properties in pesticides can be explained by different cost constraints on the synthesis or on the extraction and isolation of bioactive natural products (Rahman and Hosain 2003; Harvey et al. 2015; Atanasov et al. 2021). Such a knowledge is however important when screening for new active ingredients in natural extracts or large databases, as well as for the development of nature-inspired pesticides.

The main goal of the present study is to bring new insights into the major differences present between synthetic and natural pesticides, based on a large number of variables and a representative dataset of both synthetic and natural pesticides, in order to address the limitations of previously published studies. Based on the physicochemical differences observed between drugs from synthetic and natural origins, and the constraints involved in the synthesis of drugs and pesticides, we aimed to test whether the observed pattern in drugs is also applicable to pesticides and to which extant natural pesticides are safer for the environment than synthetic ones. In more details, we hypothesized that (i) the physicochemical space occupied by synthetic pesticides is reduced compared to the space of natural ones, (ii) a clear physicochemical signature allows to differentiate natural and synthetic pesticides (iii) natural pesticides from different origins (i.e. from marine versus non-marine environment) differ in their physicochemical properties, (iv) the proportion of pesticides not respecting Lipinski's Rules-of-five and related rules is

higher in natural than in synthetic pesticides and (v) overall toxicity and bioaccumulation potential is reduced in natural pesticides.

For this purpose, we tested for differences in the physicochemical space occupied by synthetic (not found in nature) and natural (marine and non-marine) pesticides, including fungicides, insecticides and herbicides, as well as the possibility to assign them to these groups strictly based on their physicochemical signature. This latter being an important aspect when screening for new active ingredients. In this respect, we measured the mean and variance along the axes of synthetic and natural pesticides, as well as the overlap between these groups, in a reduced physicochemical space derived from a set of 44 physicochemical variables. We further tested for space equivalency between groups. In addition, bioaccumulation and toxicity potentials, systemicity and variables used in Lipinski's and further derived rules were also compared with linear models.

2. Materials and methods

A total of 195 and 147 synthetic and natural pesticides (fungicides, insecticides and herbicides from plants, fungi, algae, bacteriae, echinoderms, tunicates and sponges, i.e. 59 molecules from marine organisms and 88 from non-marine ones), respectively, was used in this study (Table S1, supplementary information). Natural pesticides included in the study were retrieved from scientific publications highlighting a clear toxicity of the compounds to fungi, plants or insects and synthetic ones included pesticides from major classes that have been or are still commercialized. Physicochemical data were obtained from the molecular formula of pesticides and the software DataWarrior (Sander et al. 2015). DataWarrior is an open-source programm, originally developped to visualize and analyze chemical library data. Here, we used the physicochemical properties prediction tool of the software to generates the numerical data for the various synthetic and natural pesticides used in this study. The predicition tool is based on the structure of the molecules, inputed in a SMILE notation. Our dataset is constituted of 44 variables, i.e. molecular weight (MW), number of carbon, hydrogen, nitrogen, sulfur, chlorine, oxygen, fluorine, phosphorus, bromine and silicium atoms, number of fragments (substructures), number of non-H atoms, number of non-C/H atoms, number of electronegative atoms, number of stereo

centers, number of rotatable bonds, number of rings closures, number of small rings, number of aromatic rings, number of aromatic atoms, number of sp3-atoms, number of symmetric atoms, number of amides, number of amines, number of alkyl-amines, number of aromatic amines, number of aromatic nitrogens, number of basic nitrogens, number of acidic oxygens, LogP, LogS, polar surface area (PSA), LogKow, boiling point, vapor pressure, Henry constant, LogKoa, H-acceptors, H-donors, total surface area, relative polar surface area, molecular flexibility (statistical torsion data based on rotatable bonds), and molecular complexity (sensitive to e.g. stereochemistry, heteroatoms, and symmetry). In addition, potential toxicity risks of compounds, i.e. mutagenicity, tumorigenicity, reproductive and irritating effects were also estimated with the software DataWarrior (Sander et al. 2015), a prediction process that relies on a precomputed set of structural fragments that give rise to toxicity alerts in case they are encountered in the structure of compounds.

To analyze the physicochemical signature of the different types of pesticides, we first reduced the dimensionality of the physicochemical space using a principal component analyses (PCA; function dudi.pca of the R package ade4 that provides tools for multivariate data analysis such as PCA; Dray and Dufour 2007) calibrated on the 44 original variables across all pesticides (natural and synthetic). We subsequently used the two first principal component axes that accounted respectively for 31.14% and 10.49% of the inertia of the PCA (eigen vectors are available in Table S2, supplementary information). The contribution of original variables to these axes can be seen in Fig. 1 and globally relate to the hydrophobicity and size of the molecules (see results section). Then we estimated the physicochemical "niche" of pesticides, i.e. the density and area occupied by the different types of pesticides in the reduced physicochemical space. For this we used the niche quantification functions provided by the R package ecospat (Di Cola et al. 2016). This method, which uses kernel density estimations (KDE; R package ks; Chacon and Duong 2018) over a gridded space is traditionally used to quantify the overlap between the environmental niche of pairs of species (Broennimann et al. 2012). Here it was used to estimate the density of occurrence of a focal pesticide over the physicochemical space. The associated test of niche equivalency (Broennimann et al. 2012) is used to assess whether two pesticides have a significantly different signature in the physicochemical space.

In addition, linear discriminant analyses (LDA) were performed with the software JMP 15 (SAS Institute Inc., Cary, NC, 1989-2019) and the 44 variables described above, in order to evaluate if, based on their physicochemical properties, it is possible to assign pesticides without ambiguity to defined groups, i.e. (i) synthetic and natural pesticides, (ii) fungicides, insecticides and herbicides (overall and within synthetic and natural pesticides), (iii) synthetic and natural within fungicides, insecticides and herbicides and (iv) origin of natural pesticides (algae, bacteriae, echinoderms, fungi, plants, sponges; marine or not).

We also tested for differences in the variables used in Lipinski's, Tice's and Hao's rules, that define the structural and physicochemical properties for the likeliness of a given compound to be used as a drug or pesticide (i.e., MW; LogKow; number of H-bond donors and receivers, number of rotatable and aromatic bonds) for synthetic and natural compounds (marine and non-marine), as well as for the compounds from the different quadrants of the PCA space (see Figure 1). The percentage of compounds not respecting the Lipinski's rule for the different variables (i.e. MW $\leq$ 500; LogKow $\leq$ 5; number of H-bond donors and receivers, respectively $\leq$ 5 and 10) was evaluated. We also tested for differences between synthetic and natural pesticides (marine or non-marine; explanatory variables) with ANOVA for variables used in Lipinski's, Tice's and Hao's rules (response variables) with the software JMP 15 (SAS Institute Inc., Cary, NC, 1989-2019). Significant differences between categories (synthetic, marine and non-marine) in ANOVAs were determined by Tukey's honestly significant difference (HSD) test. In addition, differences between categories (synthetic, marine and non-marine) in term of PSA, LogKoa and number of halogens (chlorine, fluorine and bromine), as well as phosphorus were tested the same way. Finally, differences in potential toxicity risks (mutagenicity, tumorigenicity, reproductive and irritating effects; classified as high, low or no potential risk) between natural and synthetic pesticides were evaluated with chi-squared tests.

3. Results

The first principal component relates mainly to the size of molecules (represented by various variables) as well as to a component of their bioaccumulation potential (LogKoa), whereas the second component relates to the hydrophobicity/hydrophilicity of molecules, their potential permeability to cell membranes, to another component of their bioaccumulation potential (LogKow) and the presence of halogens and aromatic bonds (Figure 1). Quadrant I represents small hydrophobic molecules with a high membrane permeability and rich in halogens and aromatic atoms, quadrant II small hydrophilic molecules with an overall low bioaccumulation potential (low LogKow and LogKoa), quadrant III large hydrophilic molecules with low membrane permeability and quadrant IV large hydrophobic molecules with low membrane permeability and rich in halogens and aromatic atoms. The centroid of synthetic pesticides (center of the occupied space determined by the first two PCA components; Table 1) is distributed within quadrant I, whereas the one of natural pesticides within quadrant III. Within natural pesticides, centroid position of molecules from marine versus non-marine organisms strongly differs and are respectively distributed within quadrants III and II.

Tables 1 and 2 summarize the results of the tests of niche equivalency and overlaps, as well as surface area occupied by the different types of pesticides. Natural and synthetic pesticides occupied respectively 87.4% and 37.7% (surface area) of the overall physicochemical space represented by the first two components (Table 1) and their space were significantly different (equivalency test, $P<0.05$; Table 2). The space of synthetic pesticides was included by 82% within the space of natural ones, whereas only 34% of the natural ones were included within the synthetic ones (Table 2).

Within natural pesticides, the space areas of marine and non-marine organisms were significantly different, but they occupied surface areas of similar sizes (marine organisms: 79%; non-marine: 67%) with an overlap of 70-83% (Tables 1 and 2).

Results of the LDA are summarized in Table 3, with a percentage of correct assignment varying from 100% (natural versus synthetic origin of insecticides and herbicides: entropy $r^2 = 1.0$) to 76% (type of pesticides in overall pesticides: entropy $r^2 = 0.48$).

We estimated that among the four physicochemical variables involved in the Lipinski's rule, between 0 to 10% of tested synthetic pesticides did not follow the rule (i.e. MW ≤ 500; LogKow ≤ 5; number of H-bond donors and receivers, respectively ≤ 5 and 10; Lipinski et al. 2001). Concerning natural pesticides, between 0 and 16.7% for non-marine molecules and between 0 and 58.3% for marine ones. ANOVAs revealed overall significant differences between compounds of different origin (natural from marine origin, natural not from marine origin and synthetic) for MW ($P < 0.0001$; $F_{2,341} = 77.9$), logKow ($P < 0.0001$; $F_{2,341} = 11.6$), number of H-bond donors ($P < 0.0001$; $F_{2,354} = 52.5$), number of H-bond receivers ($P < 0.001$; $F_{2,341} = 44.9$), number of rotatable bonds ($P < 0.001$; $F_{2,354} = 26.8$) and aromatic bonds ($P < 0.001$; $F_{2,341} = 56.7$), number of chlorine ($P < 0.0001$; $F_{2,341} = 29.0$) and fluorine ($P < 0.0001$; $F_{2,341} = 11.6$). No significant difference was observed for bromine ($P = 0.2$; $F_{2,341} = 1.6$) and result was only marginally significant for phosphorus ($P = 0.03$; $F_{2,341} = 3.5$). Results of Tukey's HSD tests are summarized in Table S3, supplementary information.

In term of potential toxicity risks, significant differences were observed between natural and synthetic pesticides, with risks being lower in natural pesticides (tumorigenicity : $df = 2$, $\chi^2 = 23.75$, $P < 0.0001$; reproductive effect : $df = 2$, $\chi^2 = 29.5$, $P < 0.0001$; mutagenicity : $df = 2$, $\chi^2 = 21.6$, $P < 0.0001$; irritating effect : $df = 2$, $\chi^2 = 17.2$, $P = 0.0002$). Within natural pesticides, potential risks were significantly lower in pesticides from marine than non-marine organisms for tumorigenicity ($df = 2$, $\chi^2 = 13.2$, $P = 0.0014$), reproductive effect ($df = 2$, $\chi^2 = 24.6$, $P < 0.0001$), mutagenicity ($df = 2$, $\chi^2 = 17.2$, $P = 0.0002$), but not for irritating effect ($df = 2$, $\chi^2 = 0.7$, $P = 0.70$; see Figure S2, supplementary information for more details).

4. Discussion

Our study allowed to pinpoint major physicochemical differences between natural and synthetic pesticides, a difference that was also observed within natural ones, depending of their origin (from marine versus non-marine organisms). Such a pattern is comparable to the one observed in drugs (Muigg et al. 2013; Stratton et al. 2015; Chen et al. 2020) and is a direct consequence of design and synthesis constraints of synthetic compounds, albeit they both are used for different purposes.

In more details, the chemical space of commercialized synthetic pesticides and natural ones strongly differs (Figure 1 and Table 1) and their overall assignment to one of these two categories or within fungicides, herbicides and insecticides could be determined with 93-100% certainty (Table 3). Importantly, the reduced physicochemical space based on 44 variables occupied by synthetic pesticides was 2.6 fold smaller than the one of natural pesticides, with 82% of its space being included with the space of natural pesticides. Among natural pesticides a significant difference was also observed between those originating from marine and non-marine organisms (assigned at 91.2% certainty). However, physicochemical breadths were comparable. In addition, the assignment of compounds to a type of pesticides (fungicides, insecticides and herbicides) could be determined at 95% for natural compounds, revealing important physicochemical differences between them. Interestingly, it dropped to less than 80% in synthetic ones (Table 3), a difference that is reflected by the smaller physicochemical space of synthetic pesticides and to some extent to their overall lower specificity (Marrone 2019).

Overall, synthetic pesticides are mainly small hydrophobic compounds (mean MW: 303 Da; mean LogKow: 3.3) rich in halogens (chlorine, fluorine and bromine) and aromatic atoms with a capacity to go through membranes (mean PSA: 66 Å), whereas non-marine natural pesticides exhibit a similar molecular weight (mean: 305 Da), but are more hydrophilic (LogKow: 1.7) with a slightly reduced capability to go through membranes (mean PSA: 91Å) and a lower bioaccumulation potential compared to synthetic and marine compounds (in term of LogKow and LogKoa; Figure 1; supplementary online materials 3). In addition, halogens are absent (except in one compound; Strobilurin B from the fungi *Strobilurus tenacellus*), and the number of aromatic atoms is reduced. Results that contrast with natural pesticides from marine origin which are larger compounds (mean MW: 588Da) either hydrophobic or hydrophilic with some being halogenated (strictly chlorine and bromine) but are showing a reduced number of aromatic bonds.

Studies focusing on differences between natural and synthetic molecules (without considering their function or mode of action), revealed a similar difference in term of molecular weight, marine ones being larger than the other groups of pesticides (Muigg et al. 2013). Hence, such an overall difference

is not only valid for pesticides, but also for any kind of compounds. While the small size of synthetic molecules is driven by the way they are designed and the ease of synthetizing such small molecules compared to larger ones, the size difference between marine and non-marine natural compounds is not fully understood. From a general point of view, this pattern could be driven by (i) differences in the marine and non-marine environments exerting different selective pressures on organisms, (ii) differences in the availability of particular elements and/or (iii) taxa specific, i.e. some types of organisms being strictly present in marine or non-marine environments (e.g. Fuge 1988; Jha and Zi-rong 2004; Palumbi et al. 2019; Macedo et al. 2021).

Similarly, the number of rotatable bonds present in a molecule (correlated to its flexibility; e.g. Wicker and Cooper 2016; also considered as an important limiting factor in Tice's and Hao rules for the design of pesticides) is higher in marine pesticides (mean: 11.7) than in non-marine (mean: 6.5) and synthetic ones (mean: 4.3; this study), as well as in any kind of compounds from these three origins (Muigg et al. 2013). This result suggests that marine compounds are more flexible than non-marine and synthetic compounds, independently of their activities. Linear regressions between MW and the number of rotatable bonds revealed that the number of bonds increase with MW in marine ($P=0.001$) and synthetic pesticides ($P<0.0001$), but not in non-marine ones ($P=0.88$). Hence such a difference between synthetic and marine pesticides is mainly a direct consequence of a larger MW observed in these latter. However, it is not the case in non-marine compounds, meaning that the observed difference is associated to structural differences (i.e., lower flexibility) and not to differences in MW. The observed number of aromatic bonds (upper limit fixed at ≤ 17 by Hao's rule; Hao et al. 2011) was very low in natural pesticides (mean for marine and non-marine respectively: 2.0 and 3.1) compared to synthetic ones (7.9). It is interesting to note that the number of aromatic bonds increased significantly with the MW in synthetic pesticides (linear regression; $P<0.0001$), such an increase being absent in natural pesticides. A difference that should be attributed to the starting compounds used for producing synthetic compounds, particularly pesticides, that are rich in aromatic bonds (mainly aromatic nitrogens) compared to natural compounds ($P>0.05$) (Ju and Parales et al. 2010; Debarati 2020). The same is true for halogens that are frequently used in the design of synthetic pesticides (Jeschke 2017), but are less

common in natural compounds, except in some marine organisms (Gribble 2015). A large proportion of halogens (chlorine, bromine and Iodine) of the Earth being present in the oceans (Fuge 1988), a higher availability of this compounds for marine organisms likely explains the observed difference.

Lipinski's rule, that define the structural and physicochemical properties for drug-likeliness of given compounds, has been widely used for the development of pesticides and further adapted specifically to pesticides with e.g., Tice's and Hao's rules (Lipinski et al. 2001; Tice 2001; Hao et al. 2011). Avram et al. (2014; from GVKBio agrochemical patents collection) revealed that 83.65% of pesticides passed Lipinski's rule with no violation. Based on our dataset, we found a comparable result for synthetic pesticides (78%). However, values for natural pesticides were lower with 68.9% for non-marine and 26.7% for marine compounds, with the highest violation rate being for MW (58.3%) in this latter, marine pesticides being overall larger.

Lipinsiki's rule has been primarily developed for drugs and then adapted for the design of synthetic pesticides (Lipinski et al. 2001; Tice 2001; Hao et al. 2011). For that reason, Lipinsiki's rule is strongly biased, as efficient large natural compounds were not considered for establishing it. In this respect, more and more Lipinski violators, i.e., large natural molecules, are being described in drugs (Clardy and Walsh 2004; Ganesan 2008; Benet et al. 2016). A result that can be linked to long evolutionary processes shaping natural compounds and optimizing molecular flexibility with molecular weight, hydrophobicity and H-bonds (Ganesan 2008), synthetic compounds lacking such a fine-tuning. The same might also apply to natural pesticides.

Bioaccumulation of compounds strongly differs between aquatic and terrestrial food webs, whereas a combination of low LogKow and LogKoa is characteristic of compounds with low bioaccumulation potential in both food webs, higher values can have different impacts on the organisms. In aquatic ecosystems (using fish as model system), a maximum bioaccumulation is reached at a value of 7 for LogKow and compounds with values exceeding ~8.5 or 9 do not seem to bioaccumulate (Gobas et al. 2003). In terrestrial food web (from data on mammals), compounds with LogKow >8.5 can still

bioaccumulate and for values of log Koa >5, a bioaccumulation potential cannot be excluded if log Kow is >2 and the rate of chemical transformation is low (Gobas et al. 2003).

In our dataset, logKoa was significantly much higher in marine compounds (mean: 22.2) compared to non-marine and synthetic ones (respectively 12 and 11), whereas values of LogKow tend to be comparable with synthetic pesticides showing the highest mean value (3.3) and non-marine the lowest (1.7). These results suggest that non-marine pesticides have the lower probabilities of bioaccumulating compared to both marine and synthetic pesticides. It is important to note that the higher bioaccumulation potential of marine compounds, to some extant linked to their large size (significant linear regression between MW and LogKoa: $P<0.0001$), is also counterbalanced by the fact that bioaccumulation is limited in large compounds. Indeed, bioaccumulation of large compounds is reduced in fish through gills for compounds smaller than 700Da and only compounds <550Da showed a very high potential (Sakuratani et al. 2008; Arnot et al. 2010), as to human skin penetration (and related irritation) only for those smaller than 500Da (Bos and Meinardi 2001), providing advantages to large size pesticides compared to small ones in term of indirect effect on non-target organisms.

Beside the potential detrimental effect of pesticide bioaccumulation (see above), the evaluation of toxicity risks to non-targeted organisms, based on their molecular substructures is a key point. In this present study, we revealed differences between synthetic and natural pesticides with an overall reduced risks in these latter. Toxicity risks can be triggered by various substructures identified in compounds, such as e.g. halogens and more particularly polyhalogenic (mainly Chlorine) aromatic rings that are common in synthetic pesticides (Patel et al. 2020), but absent in natural pesticides from this study. However, few natural polybrominated compounds have been documented from marine (cyano-) bacteriae (Duell et al. 2020).

The large systemicity of natural pesticides in plants (logKow < 4) offers advantages compared to contact ones which remain on the surface of the plants due to their hydrophobicity and hence have limited modes of action (Russell 2005), systemicity being also the norm in synthetic pesticides. Both marine

and non-marine natural pesticides also fall within this category. The larger size of marine pesticides should not be an obstacle to systemicity. Indeed, some proteins present in phloem and xylem saps are as large as 90kDA (Balachandran et al. 1997; Neumann et al. 2010; Schröder and Collins, 2010) and cell wall permeability is only limited for molecules larger than 60kDa (Lodish et al. 2012). However, cell to cell exchanges in plants is limited to smaller molecules (<874Da; Goodwin 1983), which are still fully in the range of large marine pesticides (mean 588Da, lower and upper 95% mean: 522-654Da).

In conclusion, our study highlights that the physicochemical space of natural pesticides is much larger that the space of commercialized synthetic pesticides, which is mainly included within the one of natural pesticides. Hence, the design of synthetic pesticides has not yet realized the full potential of the possible physicochemical space of pesticides (with e.g., no halogens and a reduced number of aromatic bonds) which is constrained by specific physicochemical rules. It is also important to note that the cost and the complexity to synthetize and design efficient large molecules mimicking natural ones, or failures to pass certification processes, might be responsible for this observed difference. Such a limitation should trigger the development of efficient natural pesticides (from natural extracts or synthetized based on their molecular structure) avoiding as much as possible detrimental effects on non-target organisms and lacking toxic components found in many currently used synthetic pesticides, in order to be more in line with a sustainable future.

Acknowledgments

We thank S. Pellaud for helpful discussions and the European Innovation Council - European Union (EIC-EU) for funding.

References

Annand, P.N. 1944. Tests conducted by the Bureau of Entomology and Plant Quarantine to appraise the usefulness of DDT as an insecticide. J Econ Entomol. 37, 125–126.

Akamatsu, M. 2011. Importance of physicochemical properties for the design of new pesticides. J Agri Food Chem. 59, 2909–2917.

Arnot, J.A., Arnot, M.I., Mackay, D., Couillard, Y., MacDonald, D., Bonnell, M., Doyle, P. 2010. Molecular size cutoff criteria for screening bioaccumulation potential: Fact or fiction? Int. Env. Ass. Manag. 6, 210–224.

Atanasov, A.G., Zotchev, S.B., Dirsch, V.M., the International Natural Product Sciences Taskforce, Supuran, C.T. 2021. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov. 20, 200–216.

Avram, S., Funar-Timofei, S., Borota, A., Chennamaneni, S.R., Manchala, A.K., Muresan, S. 2014. Quantitative estimation of pesticide-likeness for agrochemical discovery. J Cheminform. 6, 42.

Balachandran, S., Xiang, Y., Schobert, C., Thompson, G.A., Lucas, W.J. 1997. Phloem sap proteins from *Cucurbita maxima* and *Ricinus communis* have the capacity to traffic cell to cell through plasmodesmata. Proc Natl Acad Sci USA. 94,14150–14155.

Bahlai, C.A., Xue, Y., McCreary, C.M., Schaafsma, A.W., Hallett, R.H. 2010. Choosing organic pesticides over synthetic pesticides may not effectively mitigate environmental risk in soybeans. PLoS ONE. e11250.

Benet, L.Z., Hosey, C.M., Ursu, O., Oprea, T.I. 2016. BDDCS, the Rule of 5 and Drugability. Adv Drug Deliv Rev. 101, 89–98.

Bennett, R., Walsgrove, R.M. 1994. Secondary metabolites in plants defence mechanisms. New Phytol. 127, 617–633.

Bos, J.D., Meinardi, M.H.M. 2001. The 500 Dalton rule for the skin penetration of chemical compounds and drugs. Exp Dermat. 9, 165–169.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H. 2011. Measuring ecological niche overlap from occurrence and spatial environmental data. Global Ecol Biogeogr. 21, 481–497.

Chacon, J.E., Duong, T. 2018. Multivariate Kernel Smoothing and its Applications. Chapman and Hall/CRC, Boca Raton. pp248.

Chen, S., Cashman, J.R. 2013. Organophosphate Exposure. Adv Molecular Toxicol. Pp. 207–233.

Chen, Y., de Lomana, M.G., Friedrich, N.-O., Kirchmair, J. 2018. Characterization of the chemical space of known and readily obtainable natural products. J Chem Inf Model. 8, 1518–1532.

Chen, Y., de Lomana M.G., Friedrich N.-O., Kirchmair, J. 2018. Characterization of the Chemical Space of Known and Readily Obtainable Natural Products. J Chem Inf Model. 58, 1518–1532.

Chen, Y., Kirchmair, J. 2020. Cheminformatics in Natural Product-based Drug Discovery. Mol Inf. 32, 2000171.

Cheung, R.C.F., Wong, J.H., Pan, W.L., Chan, Y.S., Yin, C.M., Dan, X.L., Wang, H.X., Fang, E.F., Lam, S., Ngai, P.H.K., Xia, L.X., Liu, F., Ye, X.Y., Zhang, G.Q., Liu, Q.H., Sha, O., Lin, P., Ki, C., Bekhit, A.A., Bekhit, A.E., Wan, D.C.C., Ye, X.J., Xia, J., Ng, T.B. 2014. Antifungal and antiviral products of marine organisms. Appl Microbiol Biotechnol. 98, 3475–94.

Clardy, J., Walsh, C. 2004. Lessons from naturals molecules. Nature. 432, 829–837.

Copping, L.G., Duke, S.O. 2007. Natural products that have been used commercially as crop protection agents. Pest Manag Sci. 63, 524–554.

Cordier, C., Morton, D., Murrison, S., Nelson, A., O'Leary-Steele, C. 2008. Natural products as an inspiration in the diversity-oriented synthesis of bioactive compound libraries, Nat Prod Rep. 25, 719–737.

Cortes-Hernandez, P., Nunez, R.V., Dominguez-Ramirez, L. 2020. Docking and Molecular Dynamics Predictions of Pesticide Binding to the Calyx of Bovine β-Lactoglobulin. Int J Mol Sci. 21, 1988.

Debarati, P. 2020. Metabolism of Nitroaromatic Compounds by Microbes and Study of Chemotaxis Toward These Compounds. Microbial Technology for Health and Environment. pp 303–325.

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F.T., D'Amen, M., Randin, C., Engler, R., Pottier, J., Pio, D., Dubuis, A., Pellissier, L., Mateo, R.G., Hordijk, W., Salamin, N., Guisan, A. 2016. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. Ecography. 40, 774–787.

Doak B.C., Over, B., Giordanetto, F., Kihlberg J. 2014. Oral druggable space beyond the rule of 5: insights from drugs and clinical candidates. Chem. Biol. 21:1115–1142.

Dray, S., Dufour, A.B. 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. Journal of Statistical software 22. 10.18637/jss.v022.i04

Dubey, S., Guignard, F., Pellaud, S., Pedrazzetti, M., van der Schuren, A., Gaume, A., Schnee, S., Gindro, K., Dubey, O. 2020. Isothiocyanate derivatives of glucosinolates as efficient natural fungicides. PhytoFrontiers. 1, 40–50.

Duell, E.R., Milzarek, T.M., Omari, M.E., Linares-Otoya, L.J., Schäberle, T.F., König, G.M., Gulder, T.A.M. 2020. Identification, Cloning, Expression and Functional Interrogation of the Biosynthetic Pathway of the Polychlorinated Triphenyls Ambigol A–C from *Fischerella ambigua* 108b. Org Chem Frontiers. 7, 3193-3201.

Duke, S.O., Powles, S.B. 2008. Glyphosate: a once-in-a-century herbicide. Pest Manag Sci. 64, 319–325.

Duran-Lara, E.F., Valderrama, A., Marican, A. 2020. Natural organic compounds for application in organic farming. Agriculture. 10, 41.

Erb, M., Kliebenstein, D.J. 2020. Plant Secondary Metabolites as Defenses, Regulators, and Primary Metabolites: The Blurred Functional Trichotomy. Plant Physiol. 184, 39–52.

Ertl, P., Schuffenhauer, A. 2008. Cheminformatics analysis of natural products: lessons from nature inspiring the design of new drugs. Prog Drug Res. 66, 219–235.

Fuge, R. 1988. Sources of halogens in the environment, influences on human and animal health. Environ Geochem Health. 10, 51–61.

Ganesan, A. 2008. The impact of natural products upon modern drug discovery. Curr Opin Chem Biol. 12, 306–317.

Goodwin, P.B. 1983. Molecular size limit for movement in the symplast of the *Elodea* leaf. Planta. 157, 124-130.

Gribble, G. 2015. A recent survey of naturally occurring organohalogen compounds. Environ Chem. 12, 396-405.

Hao, G., Dong, Q., Yang, G. 2011. A comparative study on the constitutive properties of marketed pesticides. Mol Inf. 30, 614–622.

Harvey, A.L., Edrada-Ebel, R., Quinn, R. J. 2015. The re-emergence of natural products for drug discovery in the genomics era. Nat. Rev. Drug Discov. 14, 111–129.

Jeschke, P. 2017. Latest generation of halogen-containing pesticides. Pest Manag Sci. 73, 1053–1066.

Jha, R.K., Zi-rong, X. 2004. Biomedical Compounds from Marine organisms**.** Mar Drugs. 2: 123–146.

Ju, K.S., Parales, R.E. 2010. Nitroaromatic compounds, from synthesis to biodegradation. Microbiol Mol Biol Rev. 74, 250–272.

Künzler, M. 2018. How fungi defend themselves against microbial competitors and animal predators. PLoS Pathog. 14, e1007184.

Lengai, G.M.W., Muthomi, J.W., Mbega, E.R. 2020. Phytochemical activity and role of botanical pesticides in pest management for sustainable agricultural crop production. Sci Afr. 7, e00239.

Lindell, S.D., Pattenden, L.C., Shannon, J. 2009. Combinatorial chemistry in the agrosciences. Bioorg Med Chem. 17, 4035-46

Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J. 2001. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Delivery Rev. 46, 3–26.

Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J. 2012. Molecular cell biology. W. H. Freeman. pp.973.

Macedo, M.W.F.S., da Cunha, N.B., Carneiro, J.A., da Costa, R.A., de Alencar, S.A., Cardoso, M.H., Franco, O.L., Dias, S.C. 2021. Marine Organisms as a Rich Source of Biologically Active Peptides. Mar. Sci. 8.

Magnus, W.W. 2002. Structure-based design of agrochemical, Nat. Prod. Rep. 19, 278–291.

Marrone, P.G. 2019. Pesticidal natural products – status and future potential. Pest Manag Sci. 75, 2325–2340.

Muigg, P., Rosen, J., Bohlin, L., Backlund, A. 2013. In silico comparison of marine, terrestrial and synthetic compounds using ChemGPS-NP for navigating chemical space. Phytochem Rev. 12, 449–457.

Nawaz, M., Mabudu, J.I., Hua, H. 2016. Current status and advancement of biopesticides: Microbial and botanical pesticides. J Entomol and Zool Stud. 4, 241–246.

Ndakidemi, B., Mtei, K., Ndakidemi, P. 2016. Impacts of synthetic and botanical pesticides on beneficial insects. Agr Sci. 7, 364–372.

Neeraj, G.S., Kumar, A., Ram, S., Kumar, V. 2017. Evaluation of nematicidal activity of ethanolic extracts of medicinal plants to *Meloidogyne incognita* (kofoid and white) chitwood under lab conditions. Int J Pure Appl Biosci. 1, 827–831.

Neumann, P.M., Weissman, R., Stefano, G., Mancuso, S. 2010. Accumulation of xylem transported protein at pit membranes and associated reductions in hydraulic conductance. J Exp Bot. 61, 1711–1717.

Ndolo, D., Njuguna, E., Adetunji, C.O., Harbor, C., Rowe, A., Breeyen, A.D., Sangeetha, J., Singh, G., Szewczyk, B., Anjorin, T.S., Thangadurai, D., Hospet, R. 2019. Research and Development of Biopesticides: Challenges and Prospects. Outlooks Pest Manag. 30, 267–276.

Nofiani, R., de Mattos-Shipley, K., Lebe, K.E., Han, L.-C., Iqbal, Z., Bailey, A.M., Willis, C.L., Simpson, T.J., Cox, R.J. 2018. Strobilurin biosynthesis in Basidiomycete fungi. Nat Commun 9, 3940.

Nören-Müller, A., Reis-Corrêa, I., Prinz, H., Rosenbaum, C., Saxena, K., Schwalbe, H.J., Vestweber, D., Cagna, G., Schunk, S., Schwarz, O., Schiewe, H., Waldmann, H. 2006. Discovery of protein phosphatase inhibitor classes by biology-oriented synthesis, Proc Natl Acad Sci USA. 103, 10606–10611.

Oguh, C.E., Okpaka, C.O., Ubani, C.S., Okekeaji, U., Joseph, P.S., Amadi, E.U. 2019. Natural pesticides (biopesticides) and uses in pest management- A critical review. Asian J Biotech Genet Eng. 2, 1–18.

Oršolić, D., Pehar, V., Šmuc, T. Stepanic V. 2021. Comprehensive machine learning based study of the chemical space of herbicides. Sci Rep. 11, 11479.

Palumbi, S.R., Evans, T.G., Pespeni, M.H., Somero, G.N. 2019. Present and Future Adaptation of Marine Species Assemblages: DNA-Based Insights into Climate Change from Studies of Physiology, Genomics, and Evolution. Oceanography. 32, 82 – 93.

Patel, A.B., Shaikh, S., Jain, K.R., Desai, C., Madamwar, D. 2020. Polycyclic Aromatic Hydrocarbons: Sources, Toxicity, and Remediation Approaches. Front. Microbiol. 11, 562813.

Poongavanam, V., Doak, B.C., Kihlberg, J. 2018. Opportunities and guidelines for discovery of orally absorbed drugs in beyond rule of 5 space. Curr Opin Chem Biol. 44, 23–29.

Rahman, S., Hossain, M. 2003. Pesticide demand in hybrid seed production technology. J Food Agric Environ. 1, 174-179.

Rattan, R.S. 2010. Mechanism of action of insecticidal secondary metabolites of plant origin. Crop Prot. 29, 913e920.

Rijswijk, K., Bewsell, D., O'Callaghan, M., Turner, J. 2018. The Next Generation of Biopesticides: Institutional barriers and enablers to co-innovation in a science and commercialisation programme. Rural Ext Innov Systems J. 14, 52-61.

Reganold, J.P., Glover, J.D., Andrews, P.K., Hinman, H.R. 2001. Sustainability of three apple production systems. Nature. 410, 926–930.

Rodrigues, E.T, Lopes, I., Pardal, M.A. 2013. Occurrence, fate and effects of azoxystrobin in aquatic ecosystems: A review. Env. Int. 53, 18-28.

Russell, P.E. 2005. A century of fungicide evolution. J Agric Sci. 143, 11–25.

Sakuratani, Y., Noguchi, Y., Kobayashi, K., Yamada, J., Nishihara, T. 2008. Molecular size as a limiting characteristic for bioconcentration in fish. J Environ Biol. 29, 89–92.

Sander, T., Freyss, J., von Korff, M., Rufener, C. 2015. DataWarrior: An open-source program for chemistry aware data visualization and analysis. J Chem Inf Model. 55, 460–473.

Schröder, P., Collins, C. 2010. Organic xenobiotics and plants. Springer Science & Business Media, Berlin, Germany.

Seifert, J. 2005. Encyclopedia of toxicology (second edition). Elsevier Inc. p. 196–200.

Suteu, D., Rusu, L., Zaharia, C., Badeanu, M., Daraban, G.M. 2020. Challenge of utilization vegetal extracts as natural plant protection products. Appl. Sci. 10, 8913.

Shang, J., Hu, B., Wang, J., Zhu, F., Kang, Y., Li, D., Sun, H., Kong, D.X., Hou, T. 2018. A cheminformatic insight into the differences between terrestrial and marine originated natural products. J Chem Inf Model. 58, 1182–1193.

Smith, C.J., Perfetti, T.A. 2020. A comparison of the persistence, toxicity, and exposure to high-volume natural plant-derived and synthetic pesticides. Tox Res App. 4, 2397847320940561.

Speck-Planche, A., Kleandrova, V.V., Rojas-Vargas, J.A. 2011. QSAR model toward the rational design of new agrochemical fungicides with a defined resistance risk using substructural descriptors. Mol Divers. 15, 901–909.

Stratton, C.F., Newman, D.J., Tan, D.S. 2015. Cheminformatic comparison of approved drugs from natural product versus synthetic origins. Bioorg Med Chem Lett. 25, 4802–4807.

Tice, C.M. 2001. Selecting the right compounds for screening: does Lipinski's Rule of 5 for pharmaceuticals apply to agrochemicals? Pest Manag Sci. 57, 3–16.

Wicker, J.G.P., Cooper, R.I. 2016. Beyond rotatable bond counts: Capturing 3D conformational flexibility in a single descriptor. J Chem Inf Model. 56, 2347–2352.

Yamamuro, M., Komuro, T., Kamiya, H., Kato, T, Hasegawa, H., Kameda, Y. 2019. Neonicotinoids disrupt aquatic food webs and decrease fishery yields. Science. 366, 620–623.

Zaid, H., Raiyn, J., Nasser, A., Saad, B., Rayan A. 2010. Physicochemical properties of natural based products versus synthetic chemicals. Open Nutraceuticals J. 3, 194-202.

Table 1 Scores and breadths of the principal components 1 and 2 of different categories (synthetic and natural pesticides; synthetic and natural insecticides, fungicides and herbicides, marine and non-marine organisms) with associated quadrant and surface area (%) compared to the entire physicochemical space of analysed pesticides.

Categories	Score PC1 (breadth)	Score PC2 (breadth)	Quadrant	Surface Area (%)
Synthetic pesticides	1.6 (1.5)	0.7 (4.3)	I	37.7
Natural pesticides	-1.0 (14.6)	-0.3 (2.6)	III	87.4
Synthetic insecticides	1.2 (1.0)	0.8 (6.0)	I	33.4
Synthetic fungicides	1.2 (1.3)	1.1 (3.1)	I	32.6
Synthetic herbicides	1.6 (2.0)	-0.3 (0.86)	II	17.0
Natural insecticides	2.2 (13.3)	-0.7 (1.9)	II	28.9
Natural fungicides	-1.1 (25.9)	-0.8 (3.1)	III	87.8
Natural herbicides	-0.9 (5.0)	-0.6	III	36.1
Marine organisms	-3.8 (14.7)	-0.1 (1.5)	III	78.8
Non-marine organisms	1.6 (7.5)	-1.3 (2.9)	II	67.1

Table 2 Results of equivalency test (*: P<0.05; NS: P<0.05; lower matrix) and % overlap (upper matrix) between natural (N.) and synthetic (S.) insecticides, fungicides and herbicides (a->b/b->a); Results of equivalency test (*: P<0.05; NS: P<0.05) and % overlap for natural versus synthetic pesticides and marine versus not marine pesticides.

	S. insecticides (b)	**S. fungicides (b)**	**S. herbicides (b)**	**N. insecticides (b)**	**N. fungicides (b)**	**N. herbicides (b)**
S. insecticides (a)		76/84	38/75	48/55	76/29	48/44
S. fungicides (a)	*		45/86	55/62	82/30	53/48
S. herbicides (a)	*	*		85/50	100/19	91/43
N. insecticides (a)	*	*	*		100/29	56/45
N. fungicides (a)	*	*	*	*		41/100
N. herbicides (a)	*	*	*	*	NS	

	P-value - Niche equivalence	**% overlap (a->b/b->a)**
N. pesticides (a) vs S. pesticides (b)	*	35/82
Marine (a) versus not marine (b)	*	70/83

Table 3 Results of the discriminant analyses (number of categories compared and molecules [count], overall % of misclassified molecules and entropy R^2).

Within Pesticides	**Categories**	**Count**	**% miscl.**	**Entropy R^2**
Natural versus synthetic	2	341	6.5	0.74
Type of pesticides (fungicide, insecticide, herbicide; natural or synthetic)	6	341	24.6	0.57
Natural versus synthetic (within fungicides)	2	229	7.0	0.78
Natural versus synthetic (within insecticides)	2	61	0	0.99
Natural versus synthetic (within herbicides)	2	51	0	1
Within synthetic pesticides				
Type of pesticides (fungicide, insecticide, herbicide)	3	205	23.6	0.48
Within natural pesticides				
Type of pesticides (fungicide, insecticide, herbicide)	3	147	5.4	0.71
Origin (Algae, Bacteriae, Echinoderms, Fungi, Plants, Sponges, Tunicates)	7	147	19.0	0.64
Origin (Marine, Not marine)	2	147	8.8	0.65

Figure legend

Figure 1 Principal component analyses (PCA) on 44 variables representing the physicochemical space of synthetic and natural pesticides, including the origin of these latter (marine versus non marine). The red dots represent the center of the occupied space (centroid) for synthetic and natural pesticides and those from marine and non-marine organisms (A). Their contributions to the first two axes displayed are indicated by vectors distributed in four quadrants (I-IV) (B).

Supplementary online materials

Figure S1, supplementary information

Illustrated examples of natural (Spongistatin I from sponges, polyoxin B from fungi, and eugenol and cineole from plants) and synthetic pesticides (fipronil, chlordane, glyphosate, fenitrothion, fluazinam, triazoxide).

Figure S2, supplementary information

Potential toxicity risks of natural (marine and non-marine) and synthetic fungicides, i.e. mutagenicity, tumorigenicity, reproductive and irritating effects.

Table S1, supplementary information

List of pesticides (fungicides, insecticides, herbicides) used in this study.

Table S2, supplementary information

Eigen vectors for the first two components (PC1 & 2) of the principal component analysis.

Table S3, supplementary information

Mean values with 95% confidence interval for molecular weight (MW in Da), logKow, number of H-donors and receivers, polar surface area (PSA; Å) and number of rotatable and aromatic bonds for Quadrant I to IV, synthetic and natural pesticides (from marine and non-marine origins). Ranges and limits for these variables according to Lipinski's rule for drugs, tice's rule for herbicides and insecticides and Hao's rule for pesticides; Different letters represent significant differences between pairs as determined by Tukey's HSD tests for synthetic, marine and non-marine pesticides.

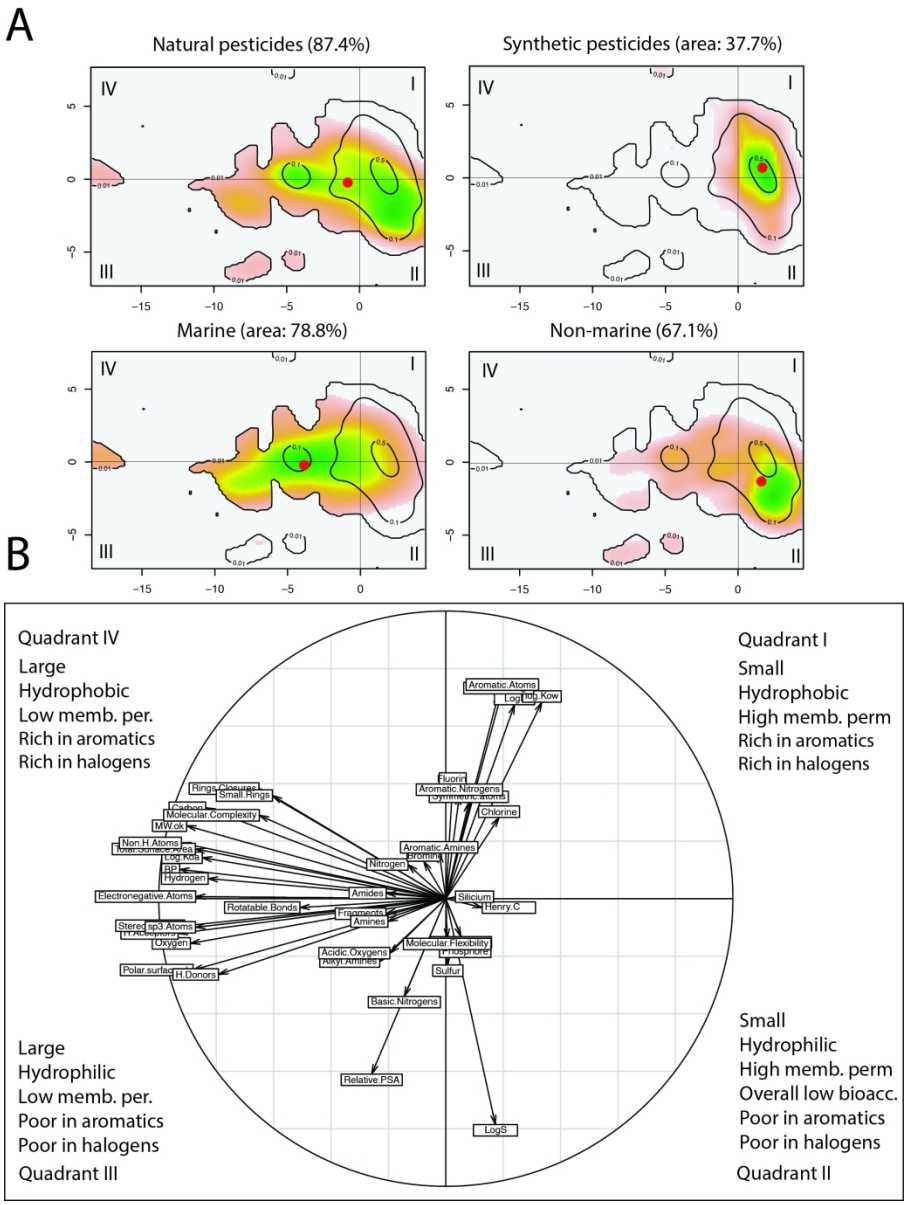
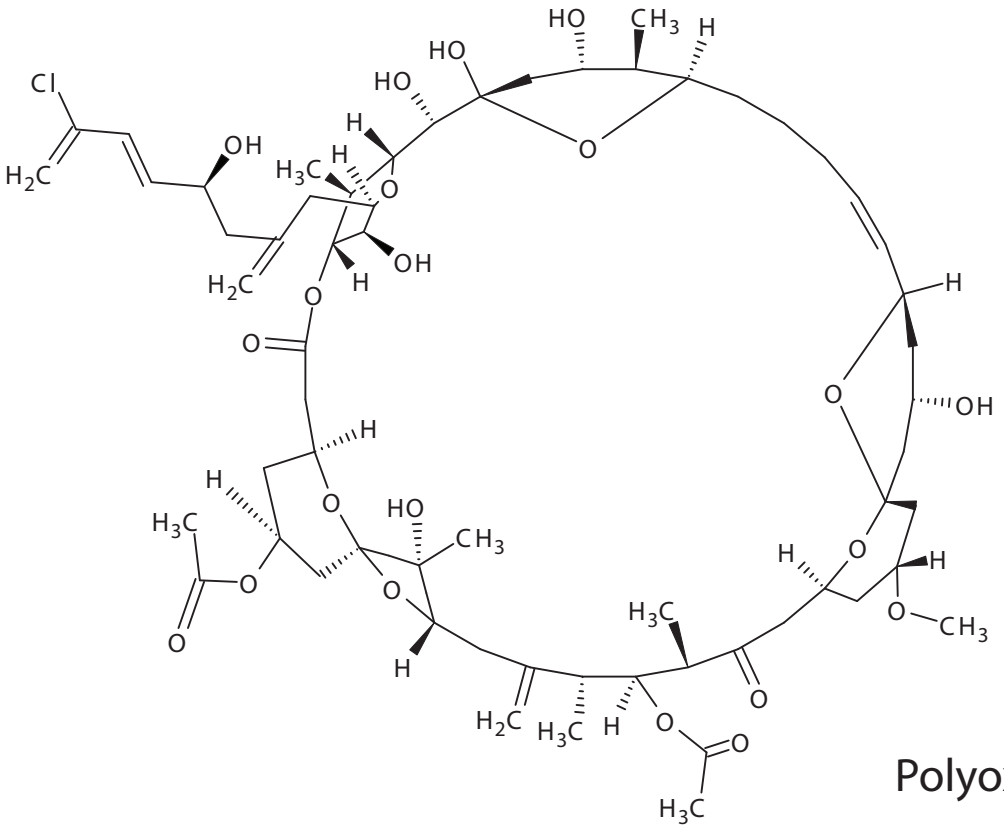


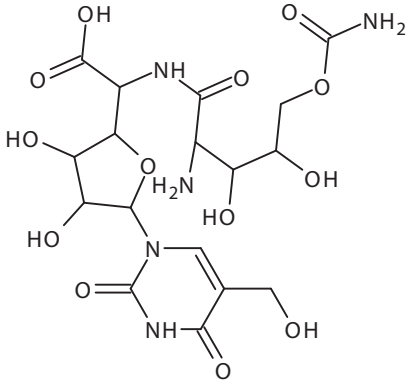
Figure 1 Principal component analyses (PCA) on 44 variables representing the physicochemical space of synthetic and natural pesticides, including the origin of these latter (marine versus non marine). The red dots represent the center of the occupied space (centroid) for synthetic and natural pesticides and those from marine and non-marine organisms (A). Their contributions to the first two axes displayed are indicated by vectors distributed in four quadrants (I-IV) (B).

197x263mm (300 x 300 DPI)

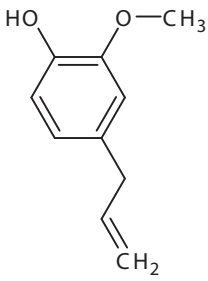
Spongistatin I



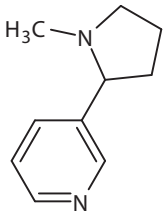
Polyoxin B



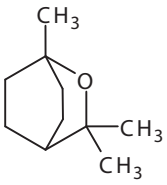
Eugenol



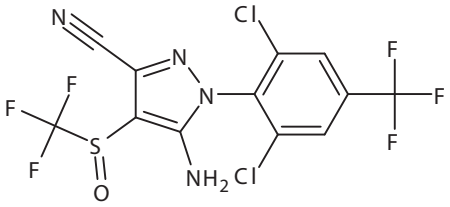
Nicotine



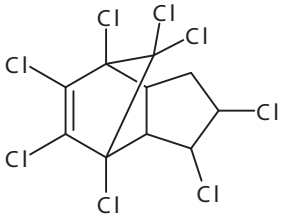
Cineole



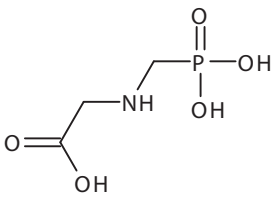
Fipronil



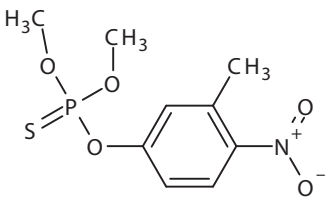
Chlordane



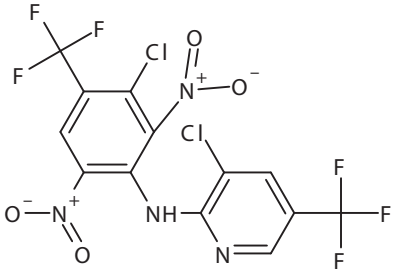
Glyphosate



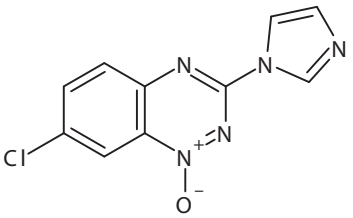
Fenitrothion

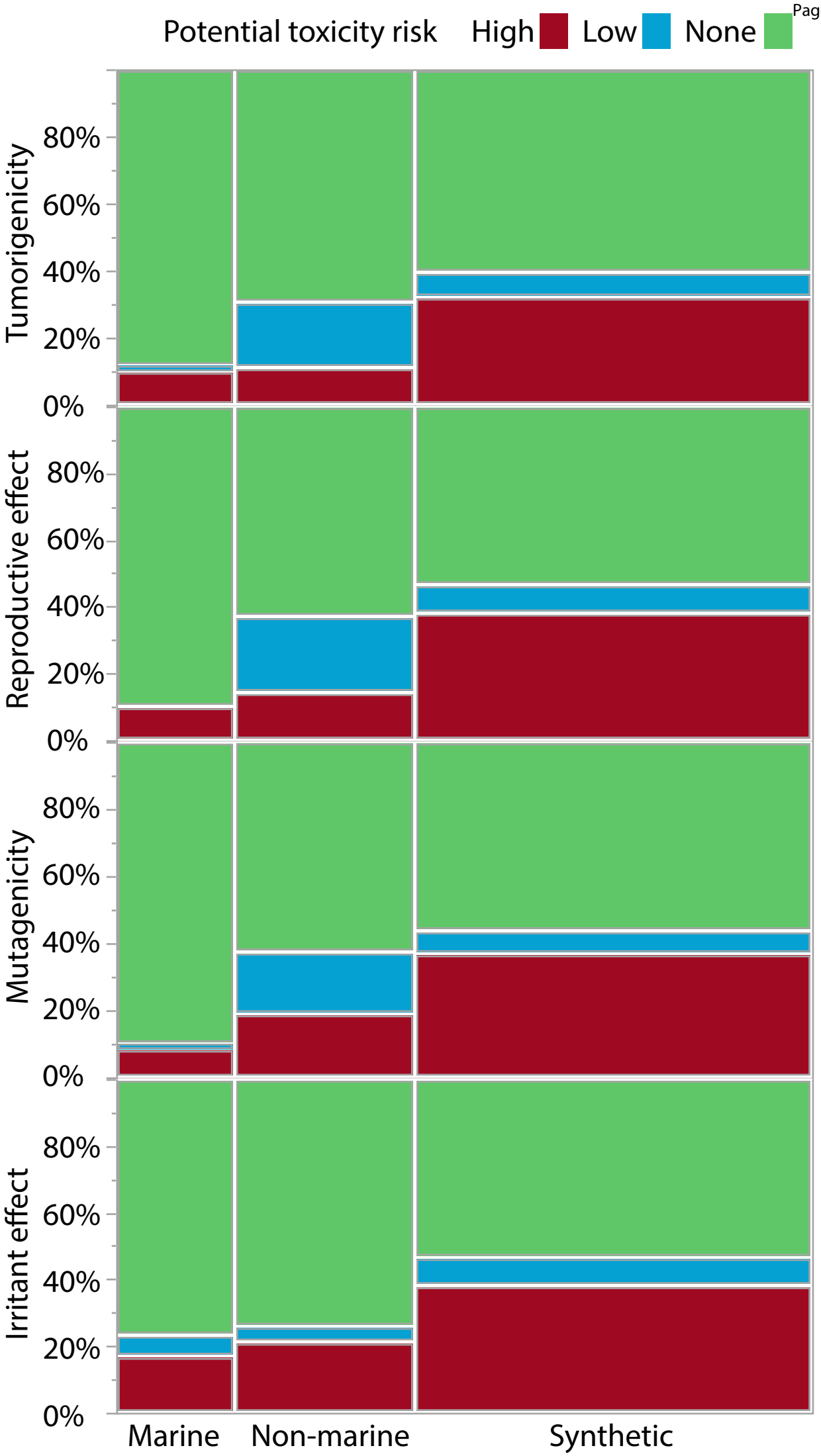


Fluazinam



Triazoxide




Potential toxicity risk
High
Low
None
Tumorigenicity
Reproductive effect
Mutagenicity
Irritant effect
80%
60%
40%
20%
0%
Marine
Non-marine
Synthetic

Compounds	Molecular formula	Type of pesticides	Synthetic or natural	Marine or not marine	Organism of origin
2.4-Dichlorophenoxyacetic acid	C8H6O3Cl2	Herbicide	Synthetic	/	/
2.4.5-Trichlorophenoxyacetic acid	C8H5O3Cl3	Herbicide	Synthetic	/	/
3-decen-2-one	C10H18O	Herbicide	Synthetic	/	/
Acéphate	C4H10NO3PS	Insecticide	Synthetic	/	/
Acetamiprid	C10H11N4Cl	Insecticide	Synthetic	/	/
Acibenzolar-S-methyl	C8H6N2OS2	Fungicide	Synthetic	/	/
Aldicarb	C7H14N2O2S	Insecticide	Synthetic	/	/
Aldimorph	C18H37NO	Fungicide	Synthetic	/	/
Ametoctradin	C15H25N5	Fungicide	Synthetic	/	/
Ametryn	C9H17N5S	Herbicide	Synthetic	/	/
Aminocarb	C11H16N2O2	Insecticide	Synthetic	/	/
Aminopyralid	C6H4N2O2Cl2	Herbicide	Synthetic	/	/
Amisulbrom	C13H13N5O4BrFS2	Fungicide	Synthetic	/	/
Amitraz	C19H23N3	Insecticide	Synthetic	/	/
Amitrole	C2H4N4	Herbicide	Synthetic	/	/
Anilazine	C9H5N4Cl3	Fungicide	Synthetic	/	/
Atrazine	C8H14N5Cl	Herbicide	Synthetic	/	/
Azinphos-methyl	C10H12N3O3PS2	Insecticide	Synthetic	/	/
Azoxystrobin	C22H17N3O5	Fungicide	Synthetic	/	/
Benalaxyl	C20H23NO3	Fungicide	Synthetic	/	/
Bendiocarb	C11H13NO4	Insecticide	Synthetic	/	/
Benomyl	C14H18N4O3	Fungicide	Synthetic	/	/
Beta-cypermethrin	C22H19NO3Cl2	Insecticide	Synthetic	/	/
Bifenthrine	C23H22O2ClF3	Insecticide	Synthetic	/	/
Binapacryl	C15H18N2O6	Fungicide	Synthetic	/	/
Bitertanol	C20H23N3O2	Fungicide	Synthetic	/	/
Bromophos	C8H8O3BrCl2PS	Insecticide	Synthetic	/	/
Bupirimate	C13H24N4O3S	Fungicide	Synthetic	/	/
Butralin	C14H21N3O4	Herbicide	Synthetic	/	/
Captafol	C10H9NO2Cl4S	Fungicide	Synthetic	/	/
Captan	C9H8NO2Cl3S	Fungicide	Synthetic	/	/
Carbaryl	C12H11NO2	Insecticide	Synthetic	/	/
Carbendazime	C9H9N3O2	Fungicide	Synthetic	/	/
Carbofuran	C12H15NO3	Insecticide	Synthetic	/	/
Carbosulfan	C20H32N2O3S	Insecticide	Synthetic	/	/
Carpropamid	C15H18NOCl3	Fungicide	Synthetic	/	/
Chinomethionat	C10H6N2OS2	Fungicide	Synthetic	/	/
Chlorbenzilate	C16H14O3Cl2	Insecticide	Synthetic	/	/
Chlordane	C10H6Cl8	Insecticide	Synthetic	/	/
Chlordimeform	C10H13N2Cl	Insecticide	Synthetic	/	/
Chlorfenapyr	C15H11N2OBrClF3	Insecticide	Synthetic	/	/
Chloroneb	C8H8O2Cl2	Fungicide	Synthetic	/	/
Chlorothanonil	C8N2Cl4	Fungicide	Synthetic	/	/
Chlorpyrifos	C9H11NO3Cl3PS	Insecticide	Synthetic	/	/
Chlozolinate	C13H11NO5Cl2	Fungicide	Synthetic	/	/
Clopyralid	C6H3NO2Cl2	Herbicide	Synthetic	/	/
Clothianidin	C6H8N5O2ClS	Insecticide	Synthetic	/	/
Cyazofamid	C13H13N4O2ClS	Fungicide	Synthetic	/	/
Cyflufenamid	C20H17N2O2F5	Fungicide	Synthetic	/	/
Cyfluthrin	C22H18NO3Cl2F	Insecticide	Synthetic	/	/
Cyhalothrin	C23H19NO3ClF3	Insecticide	Synthetic	/	/
Cymoxanil	C7H10N4O3	Fungicide	Synthetic	/	/
Cyprodinil	C14H15N3	Fungicide	Synthetic	/	/
Deltamethrine	C22H19NO3Br2	Insecticide	Synthetic	/	/
Diazinon	C12H21N2O3PS	Insecticide	Synthetic	/	/
Dicamba	C8H6O3Cl2	Herbicide	Synthetic	/	/
Dichlobenil	C7H3NCl2	Herbicide	Synthetic	/	/
Dichlofluanid	C9H11N2O2Cl2FS2	Fungicide	Synthetic	/	/
Dichlorodiphenyldichloroethane	C14H10Cl4	Insecticide	Synthetic	/	/
Diclocymet	C15H18N2OCl2	Fungicide	Synthetic	/	/
Diclomezine	C11H8N2OCl2	Fungicide	Synthetic	/	/
Dicloran	C6H4N2O2Cl2	Fungicide	Synthetic	/	/
Dicofol	C14H9OCl5	Insecticide	Synthetic	/	/
Diethofencarb	C14H21NO4	Fungicide	Synthetic	/	/
Diflubenzuron	C14H9N2O2ClF2	Insecticide	Synthetic	/	/
Diflumetorim	C15H16N3OClF2	Fungicide	Synthetic	/	/
Dimethirimol	C11H19N3O	Fungicide	Synthetic	/	/
Dimethoate	C5H12NO3PS2	Insecticide	Synthetic	/	/
Dimethomorph	C21H22NO4Cl	Fungicide	Synthetic	/	/
Dinotefuran	C7H14N4O3	Insecticide	Synthetic	/	/
Dithianon	C14H4N2O2S2	Fungicide	Synthetic	/	/
Edifenphos	C14H15O2PS2	Fungicide	Synthetic	/	/
Endosulfan	C9H6O3Cl6S	Insecticide	Synthetic	/	/
Etaconazole	C14H15N3O2Cl2	Fungicide	Synthetic	/	/
Ethaboxam	C14H16N4OS2	Fungicide	Synthetic	/	/
Ethalfluralin	C13H14N3O4F3	Herbicide	Synthetic	/	/
Ethirimol	C11H19N3O	Fungicide	Synthetic	/	/
Fenazaquin	C20H22N2O	Fungicide	Synthetic	/	/
Fenhexamid	C14H17NO2Cl2	Fungicide	Synthetic	/	/
Fenitrothion	C9H12NO5PS	Insecticide	Synthetic	/	/
Fenoxanil	C15H18N2O2Cl2	Fungicide	Synthetic	/	/
Fenpiclonil	C11H6N2Cl2	Fungicide	Synthetic	/	/
Fenpicoxamid	C31H38N2O11	Fungicide	Synthetic	/	/
Fenpropidin	C19H31N	Fungicide	Synthetic	/	/
Fenpropimorph	C20H33NO	Fungicide	Synthetic	/	/
Fenpyrazamide	C17H21N3O2S	Fungicide	Synthetic	/	/
Ferimzone	C15H18N4	Fungicide	Synthetic	/	/
Fipronil	C12H4N4OCl2F6S	Insecticide	Synthetic	/	/
Flazasulfuron	C13H12N5O5F3S	Herbicide	Synthetic	/	/
Fluazifop	C15H12NO4F3	Herbicide	Synthetic	/	/
Fluazinam	C13H4N4O4Cl2F6	Fungicide	Synthetic	/	/

Fluconazole	C13H12N6OF2	Fungicide	Synthetic	/	/
Flucytosine	C4H4N3OF	Fungicide	Synthetic	/	/
Fludioxonil	C12H6N2O2F2	Fungicide	Synthetic	/	/
Fluopicolide	C14H8N2OCl3F3	Fungicide	Synthetic	/	/
Fluoroimide	C10H4NO2Cl2F	Fungicide	Synthetic	/	/
Fluroxypyr	C7H5N2O3Cl2F	Herbicide	Synthetic	/	/
Flusulfamide	C13H7N2O4Cl2F3S	Fungicide	Synthetic	/	/
Flutianil	C19H14N2OF4S2	Fungicide	Synthetic	/	/
Fluxapyroxad	C18H12N3OF5	Fungicide	Synthetic	/	/
Folpet	C9H4NO2Cl3S	Fungicide	Synthetic	/	/
Folprocarb	C16H21N2O3F3	Fungicide	Synthetic	/	/
Formetanate	C11H15N3O2	Insecticide	Synthetic	/	/
Formothion	C6H12NO4PS2	Insecticide	Synthetic	/	/
Fthalide	C8H2O2Cl4	Fungicide	Synthetic	/	/
Fuberidazole	C11H8N2O	Fungicide	Synthetic	/	/
Furalaxyl	C17H19NO4	Fungicide	Synthetic	/	/
Glyphosate	C3H8NO5P	Herbicide	Synthetic	/	/
Hymexazol	C4H5NO2	Fungicide	Synthetic	/	/
Imazamox	C15H19N3O4	Herbicide	Synthetic	/	/
Imazapic	C14H17N3O3	Herbicide	Synthetic	/	/
Imazapyr	C13H15N3O3	Herbicide	Synthetic	/	/
Imidacloprid	C9H10N5O2Cl	Insecticide	Synthetic	/	/
Iminoctadine	C18H41N7	Fungicide	Synthetic	/	/
Iodocarb	C8H12NO2I	Fungicide	Synthetic	/	/
Iprodione	C13H13N3O3Cl2	Fungicide	Synthetic	/	/
Iprovalicarb	C18H28N2O3	Fungicide	Synthetic	/	/
Isavuconazole	C22H17N5OF2S	Fungicide	Synthetic	/	/
Isofenphos	C15H24NO4PS	Insecticide	Synthetic	/	/
Isofetamid	C20H25NO3S	Fungicide	Synthetic	/	/
Isoprothiolane	C12H18O4S2	Fungicide	Synthetic	/	/
Isoproturon	C12H18N2O	Herbicide	Synthetic	/	/
Isotianil	C11H5N3OCl2S	Fungicide	Synthetic	/	/
Itraconazole	C35H38N8O4Cl2	Fungicide	Synthetic	/	/
Lindane	C6H6Cl6	Insecticide	Synthetic	/	/
Linuron	C9H10N2O2Cl2	Herbicide	Synthetic	/	/
Mandestrobin	C19H23NO3	Fungicide	Synthetic	/	/
MCPA (2-methyl-4-chlorophenoxyacetic acid	C9H9O3Cl	Herbicide	Synthetic	/	/
Mepanipyrim	C14H13N3	Fungicide	Synthetic	/	/
Meptyl dinocap	C18H24N2O6	Fungicide	Synthetic	/	/
Methasulfocarb	C9H11NO4S2	Fungicide	Synthetic	/	/
Metolachlor	C15H22NO2Cl	Herbicide	Synthetic	/	/
Metrafenone	C19H21O5Br	Fungicide	Synthetic	/	/
Metsulfuron-methyl	C14H15N5O6S	Herbicide	Synthetic	/	/
Naftiline	C21H21N	Fungicide	Synthetic	/	/
Nitenpyram	C11H15N4O2Cl	Insecticide	Synthetic	/	/
Octhilinone	C11H19NOS	Fungicide	Synthetic	/	/
Ofurace	C14H16NO3Cl	Fungicide	Synthetic	/	/
Orysastrobin	C18H25N5O5	Fungicide	Synthetic	/	/
Oxadiargyl	C15H14N2O3Cl2	Herbicide	Synthetic	/	/
Oxadixyl	C14H18N2O4	Fungicide	Synthetic	/	/
Oxathiapiprolin	C24H22N5O2F5S	Fungicide	Synthetic	/	/
Oxolinic acid	C13H11NO5	Fungicide	Synthetic	/	/
Paraquat	C12H14N2	Herbicide	Synthetic	/	/
Pencycuron	C19H21N2OCl	Fungicide	Synthetic	/	/
Pendimethalin	C13H19N3O4	Herbicide	Synthetic	/	/
Phenamacril	C12H12N2O2	Fungicide	Synthetic	/	/
Picarbutrazox	C20H23N7O3	Fungicide	Synthetic	/	/
Picloram	C6H3N2O2Cl3	Herbicide	Synthetic	/	/
Picoxystrobin	C18H16NO4F3	Fungicide	Synthetic	/	/
Piperonyl butoxyde	C19H30O5	Insecticide	Synthetic	/	/
Posaconazole	C37H42N8O4F2	Fungicide	Synthetic	/	/
Probenazole	C10H9NO3S	Fungicide	Synthetic	/	/
Procymidone	C13H11NO2Cl2	Fungicide	Synthetic	/	/
Propamocarb	C9H20N2O2	Fungicide	Synthetic	/	/
Propanil	C9H9NOCl2	Herbicide	Synthetic	/	/
Propargite	C19H26O4S	Insecticide	Synthetic	/	/
Propisochlor	C15H22NO2Cl	Herbicide	Synthetic	/	/
Proquinazid	C14H17N2O2I	Fungicide	Synthetic	/	/
Prothiocarb	C8H18N2OS	Fungicide	Synthetic	/	/
Pyrazophos	C14H20N3O5PS	Fungicide	Synthetic	/	/
Pyributicarb	C18H22N2O2S	Fungicide	Synthetic	/	/
Pyrifenox	C14H12N2OCl2	Fungicide	Synthetic	/	/
Pyrimethanil	C12H13N3	Fungicide	Synthetic	/	/
Pyrimorph	C22H25N2O2Cl	Fungicide	Synthetic	/	/
Pyriofenone	C18H20NO5Cl	Fungicide	Synthetic	/	/
Pyroquilon	C11H11NO	Fungicide	Synthetic	/	/
Quinoxyfen	C15H8NOCl2F	Fungicide	Synthetic	/	/
Resmethrine	C22H26O3	Insecticide	Synthetic	/	/
Silthiofam	C13H21NOSSi	Fungicide	Synthetic	/	/
Simazine	C7H12N5Cl	Herbicide	Synthetic	/	/
Tebufloquin	C17H20NO2F	Fungicide	Synthetic	/	/
Teclofthalam	C14H5NO3Cl6	Fungicide	Synthetic	/	/
Teflubenzuron	C14H6N2O2Cl2F4	Insecticide	Synthetic	/	/
Tepraloxydim	C17H24NO4Cl	Herbicide	Synthetic	/	/
Terbinafine	C21H25N	Fungicide	Synthetic	/	/
Tetradifon	C12H6O2Cl4S	Insecticide	Synthetic	/	/
Tetrasul	C12H6Cl4S	Insecticide	Synthetic	/	/
Thiacloprid	C10H9N4ClS	Insecticide	Synthetic	/	/
Thiamethoxam	C8H10N5O3ClS	Insecticide	Synthetic	/	/
Thifluzamide	C13H6N2O2Br2F6S	Fungicide	Synthetic	/	/
Thiocyclam	C5H11NS3	Insecticide	Synthetic	/	/
Thiodicarb	C10H18N4O4S3	Insecticide	Synthetic	/	/
Thiophanate	C14H18N4O4S2	Fungicide	Synthetic	/	/

Thiram	C6H12N2S4	Fungicide	Synthetic	/	/
Tolfenpyrad	C21H22N3O2Cl	Fungicide	Synthetic	/	/
Tolyfluanid	C10H13N2O2Cl2FS2	Fungicide	Synthetic	/	/
Triasulfuron	C14H16N5O5ClS	Herbicide	Synthetic	/	/
Triazoxide	C10H6N5OCl	Fungicide	Synthetic	/	/
Triclopyr	C7H4NO3Cl3	Herbicide	Synthetic	/	/
Tricyclazole	C9H7N3S	Fungicide	Synthetic	/	/
Trifluralin	C13H16N3O4F3	Herbicide	Synthetic	/	/
Triforine	C10H14N4O2Cl6	Fungicide	Synthetic	/	/
Voriconazole	C16H14N5OF3	Fungicide	Synthetic	/	/
Zoxamide	C14H16NO2Cl3	Fungicide	Synthetic	/	/
Amphidinin G	C22H38O4	Fungicide	Natural	Marine	Algae
Amphidinol 3	C70H118O23	Fungicide	Natural	Marine	Algae
Capisterone A	C32H50O8S	Fungicide	Natural	Marine	Algae
Capisterone B	C30H48O7S	Fungicide	Natural	Marine	Algae
Laminarin	C18H32O16	Fungicide	Natural	Marine	Algae
Bahamaolide A	C39H64O11	Fungicide	Natural	Marine	Bacteriae
Basiliskamide A	C23H31NO4	Fungicide	Natural	Marine	Bacteriae
Basiliskamide B	C23H31NO4	Fungicide	Natural	Marine	Bacteriae
Dihydromaltophilin	C29H40N2O6	Fungicide	Natural	Marine	Bacteriae
Forazoline A	C43H69N4O10ClS2	Fungicide	Natural	Marine	Bacteriae
Lyngbyabellin B	C28H40N4O7Cl2S2	Fungicide	Natural	Marine	Bacteriae
Majusculoic acid	C15H23O2Br	Fungicide	Natural	Marine	Bacteriae
Rocheicoside A	C30H42N6O9	Fungicide	Natural	Marine	Bacteriae
Tanikolide	C17H32O3	Fungicide	Natural	Marine	Bacteriae
Natamycin	C33H47NO13	Fungicide	Natural	Not marine	Bacteriae
Amphotericin B	C47H73NO17	Fungicide	Natural	Not marine	Bacteriae
Blasticidin -S	C17H26N8O5	Fungicide	Natural	Not marine	Bacteriae
Kasugamycin	C14H25N3O9	Fungicide	Natural	Not marine	Bacteriae
Oxytetracycline	C22H24N2O9	Fungicide	Natural	Not marine	Bacteriae
Phosphinotricine	C11H22N3O6P	Herbicide	Natural	Not marine	Bacteriae
Polyoxin B	C17H25N5O13	Fungicide	Natural	Not marine	Bacteriae
Streptomycin	C21H39N7O12	Fungicide	Natural	Not marine	Bacteriae
Thaxtomine A	C22H22N4O6	Herbicide	Natural	Not marine	Bacteriae
validamycin	C20H35NO13	Fungicide	Natural	Not marine	Bacteriae
7-deoxy-sedoheptulose	C7H14O6	Herbicide	Natural	Marine	Bacteriae
Holothurin B	C41H63O17S.Na	Fungicide	Natural	Marine	Echinoderms
Holotoxin F	C59H96O26	Fungicide	Natural	Marine	Echinoderms
Holotoxin G	C58H94O25	Fungicide	Natural	Marine	Echinoderms
Neothyonidioside	C53H82O24S	Fungicide	Natural	Marine	Echinoderms
Didymellamide A	C24H29NO7	Fungicide	Natural	Marine	Fungi
Ilicicolin H	C27H31NO4	Fungicide	Natural	Marine	Fungi
P. meleagrinum macrolide (PF1163A)	C27H43NO6	Fungicide	Natural	Marine	Fungi
P. meleagrinum macrolide (PF1163B)	C27H43NO5	Fungicide	Natural	Marine	Fungi
Xestodecalactone B	C14H16O6	Fungicide	Natural	Marine	Fungi
(3R.4R)-4-hydroxymellein	C10H10O4	Fungicide	Natural	Not marine	Fungi
3-hydroxyfumiquinazoline A	C24H23N5O5	Fungicide	Natural	Not marine	Fungi
Acide aspterrique	C15H22O4	Herbicide	Natural	Not marine	Fungi
Afritoxinone A	C8H10O4	Fungicide	Natural	Not marine	Fungi
Compound 21	C7H6O4	Fungicide	Natural	Not marine	Fungi
Compound 36	C7H6O4	Fungicide	Natural	Not marine	Fungi
Compound 72	C35H56O9	Fungicide	Natural	Not marine	Fungi
Compound 74	C34H58O9	Fungicide	Natural	Not marine	Fungi
Compound 75	C34H60O9	Fungicide	Natural	Not marine	Fungi
Cytochalasin A	C29H35NO5	Fungicide	Natural	Not marine	Fungi
Fumiquinazonlie A	C24H23N5O4	Fungicide	Natural	Not marine	Fungi
Fumiquinazonlie F	C21H18N4O2	Fungicide	Natural	Not marine	Fungi
Fumitremorgin B	C27H33N3O5	Fungicide	Natural	Not marine	Fungi
Fusapyrone	C34H54O9	Fungicide	Natural	Not marine	Fungi
Macrocidine A	C20H23NO5	Herbicide	Natural	Not marine	Fungi
Oxysporone	C7H8O4	Fungicide	Natural	Not marine	Fungi
Papulacandin B	C47H64O17	Fungicide	Natural	Not marine	Fungi
R-(-)-Mellein	C10H10O3	Fungicide	Natural	Not marine	Fungi
Safingol	C18H39NO2	Fungicide	Natural	Not marine	Fungi
Sphaeropsidin A	C20H26O5	Fungicide	Natural	Not marine	Fungi
Sphaeropsidin B	C20H28O5	Fungicide	Natural	Not marine	Fungi
Sphaeropsidin C	C20H28O4	Fungicide	Natural	Not marine	Fungi
Sphaeropsidone	C7H8O4	Fungicide	Natural	Not marine	Fungi
Strobilurin A	C16H18O3	Fungicide	Natural	Not marine	Fungi
Strobilurin B	C17H19O4Cl	Fungicide	Natural	Not marine	Fungi
Strobilurin C	C21H26O4	Fungicide	Natural	Not marine	Fungi
Strobilurin E	C26H32O7	Fungicide	Natural	Not marine	Fungi
Strobilurin F	C21H26O5	Fungicide	Natural	Not marine	Fungi
(E)-1-isothiocyanato-8-(methylsulfinyl)oct-2-ene	C10H17NOS2	Fungicide	Natural	Not marine	Plants
1-(ethylsulfinyl)-8-isothiocyanatooctane	C11H21NOS2	Fungicide	Natural	Not marine	Plants
1-(ethylsulfonyl)-8-isothiocyanatooctane	C11H21NO2S2	Fungicide	Natural	Not marine	Plants
1-(isothiocyanatomethyl)-3-(4-(methylsulfinyl)butyl)benzene	C13H17NOS2	Fungicide	Natural	Not marine	Plants
1-(isothiocyanatomethyl)-3-(4-(methylsulfonyl)butyl)benzene	C13H17NO2S2	Fungicide	Natural	Not marine	Plants
1-Isothiocyanato-3-(methylsulfinyl)-propane	C5H9NOS2	Fungicide	Natural	Not marine	Plants
1-Isothiocyanato-3-(methylsulfonyl)-propane	C5H9NO2S2	Fungicide	Natural	Not marine	Plants
1-Isothiocyanato-6-(methylsulfinyl)-hexane	C8H15NOS2	Fungicide	Natural	Not marine	Plants
1-Isothiocyanato-7-(methylsulfinyl)-heptane	C9H17NOS2	Fungicide	Natural	Not marine	Plants
1-Isothiocyanato-8-(methylsulfinyl)-octane	C10H19NOS2	Fungicide	Natural	Not marine	Plants
1-Isothiocyanato-8-(methylsulfonyl)-octane	C10H19NO2S2	Fungicide	Natural	Not marine	Plants
1-Isothiocyanato-9-(methylsulfinyl)-nonane	C11H21NOS2	Fungicide	Natural	Not marine	Plants
1-Isothiocyanato-9-(methylsulfonyl)-nonane	C11H21NO2S2	Fungicide	Natural	Not marine	Plants
1.8-cineole	C10H18O	Herbicide	Natural	Not marine	Plants
2-Phenethyl Propionate	C11H14O2	Herbicide	Natural	Not marine	Plants
2-phenylethyl-isothiocyanate	C9H9NS	Insecticide	Natural	Not marine	Plants
2.4-dihydroxy-7-methoxy-1.4-benzoxazin-3-one	C9H9NO5	Herbicide	Natural	Not marine	Plants
2(S)-hydroxy-3-butenyl-nitrile	C4H5NO	Insecticide	Natural	Not marine	Plants
3-methoxybenzyl-isothiocyanate	C9H9NOS	Insecticide	Natural	Not marine	Plants
3.4-epithiobutyl-nitrile	C4H5NS	Insecticide	Natural	Not marine	Plants

4-acetyltropolone	C8H6O3	Fungicide	Natural	Not marine	Plants
Acide pélargonique	C9H18O2	Herbicide	Natural	Not marine	Plants
Aconitine	C34H47NO11	Insecticide	Natural	Not marine	Plants
Allyl thiocyanate	C4H5NS	Insecticide	Natural	Not marine	Plants
Azadirachtine	C35H44O16	Insecticide	Natural	Not marine	Plants
Benzoquinone	C6H4O2	Herbicide	Natural	Not marine	Plants
Benzyl isothiocyanate	C8H7NS	Insecticide	Natural	Not marine	Plants
Benzyl thiocyanante	C8H7NS	Insecticide	Natural	Not marine	Plants
Beta-dolabrin	C10H10O2	Fungicide	Natural	Not marine	Plants
Citral	C10H16O	Herbicide	Natural	Not marine	Plants
Citric Acid	C6H8O7	Herbicide	Natural	Not marine	Plants
Dehydro-a-lapachone	C15H12O3	Fungicide	Natural	Not marine	Plants
Resveratrol	C14H12O3	Fungicide	Natural	Not marine	Plants
Eugenol	C10H12O2	Herbicide	Natural	Not marine	Plants
Gamma-thujaplicin	C10H12O2	Fungicide	Natural	Not marine	Plants
Geraniol	C10H18O	Fungicide	Natural	Not marine	Plants
Hinokitiol	C10H12O2	Fungicide	Natural	Not marine	Plants
Hydroquinone	C6H6O2	Herbicide	Natural	Not marine	Plants
Indoleacetonitrile	C10H8N2	Fungicide	Natural	Not marine	Plants
Jamaicin	C22H18O6	Insecticide	Natural	Not marine	Plants
Juglone	C10H6O3	Herbicide	Natural	Not marine	Plants
Lauric acid	C12H24O2	Fungicide	Natural	Not marine	Plants
Licetin	C21H18O7	Insecticide	Natural	Not marine	Plants
Nicotine	C10H14N2	Insecticide	Natural	Not marine	Plants
Pipernonaline	C21H27NO3	Fungicide	Natural	Not marine	Plants
Quassine	C22H28O6	Insecticide	Natural	Not marine	Plants
Radulanine A	C19H20O2	Herbicide	Natural	Not marine	Plants
Rotenone	C23H22O6	Insecticide	Natural	Not marine	Plants
Ryanodine	C25H35NO9	Insecticide	Natural	Not marine	Plants
Thymol	C10H14O	Fungicide	Natural	Not marine	Plants
Batzelladine L	C39H68N6O2	Fungicide	Natural	Marine	Sponges
Bengamide A	C31H56N2O8	Fungicide	Natural	Marine	Sponges
Bengamide B	C32H58N2O8	Fungicide	Natural	Marine	Sponges
Bengamide E	C17H30N2O6	Fungicide	Natural	Marine	Sponges
Bengamide F	C18H32N2O7	Fungicide	Natural	Marine	Sponges
Bengamide L	C32H58N2O8	Fungicide	Natural	Marine	Sponges
Callipeltin J	C31H58N8O11	Fungicide	Natural	Marine	Sponges
Corticatic acid A	C31H44O3	Fungicide	Natural	Marine	Sponges
Corticatic acid E	C31H46O3	Fungicide	Natural	Marine	Sponges
Crambescin A2 392	C20H36N6O2	Fungicide	Natural	Marine	Sponges
Crambescin A2 406	C21H38N6O2	Fungicide	Natural	Marine	Sponges
Crambescin A2 420	C22H40N6O2	Fungicide	Natural	Marine	Sponges
Hippolachnin A	C19H30O3	Fungicide	Natural	Marine	Sponges
Hippolide j	C25H36O3	Fungicide	Natural	Marine	Sponges
Leucascandrolide A	C38H56N2O10	Fungicide	Natural	Marine	Sponges
Massadine	C22H24N10O5Br4	Fungicide	Natural	Marine	Sponges
Oceanapiside	C34H68N2O9	Fungicide	Natural	Marine	Sponges
Roridin A	C29H40O9	Fungicide	Natural	Marine	Sponges
Roridin D	C29H38O9	Fungicide	Natural	Marine	Sponges
Spongistatin I	C63H95O21Cl	Fungicide	Natural	Marine	Sponges
Theopederin F	C27H47NO10	Fungicide	Natural	Marine	Sponges
Topsentiasterol sulfate D (H salt)	C30H46O14S3	Fungicide	Natural	Marine	Sponges
Topsentiasterol sulfate D (sodium salt)	C30H43O14Na3S3	Fungicide	Natural	Marine	Sponges
Topsentiasterol sulfate E (H salt)	C31H52O13S3	Fungicide	Natural	Marine	Sponges
Topsentiasterol sulfate E (Sodium salt)	C31H49O13Na3S3	Fungicide	Natural	Marine	Sponges
Woodylide A	C18H34O4	Fungicide	Natural	Marine	Sponges
Acanthosterol J (H salt)	C30H48O8S	Fungicide	Natural	Marine	Sponges
Acanthosterol J (sodium salt)	C30H47O8NaS	Fungicide	Natural	Marine	Sponges
Agelasidine A	C18H34N3O2ClS	Fungicide	Natural	Marine	Sponges
Agelasine F	C26H40N5Cl	Fungicide	Natural	Marine	Sponges
Ageline B	C31H43N6O2	Fungicide	Natural	Marine	Sponges
Aurantoside K	C33H43N2O15Cl	Fungicide	Natural	Marine	Sponges
Avarol	C21H30O2	Fungicide	Natural	Marine	Sponges
Dolastatin 10	C42H68N6O6S	Fungicide	Natural	Marine	Tunicates
(2S.3R)-2-aminododecan-3-ol	C12H27NO	Fungicide	Natural	Marine	Tunicates

Eigen vectors for the first two components (PC1 & PC2) of the principal component analysis

Variables	PC1	PC2
Log Kow	0.33446636	0.6830587
Log P	0.23937831	0.67433492
Aromatic Atoms	0.19690057	0.72454628
Chlorine	0.1854977	0.28272934
Aromatic Rings	0.18382579	0.71214152
Log S	0.17272928	-0.7855087
Henry C	0.12683809	-0.0310142
Symmetric.atoms	0.08000801	0.33424919
Phosphorus	0.06731637	-0.1647166
VP1	0.0524389	-0.1319129
Aromatic Nitrogens	0.04713873	0.35925527
Silicium	0.03369719	0.00670116
Fluorine	0.02377687	0.39765527
Sulfur	0.0070592	-0.2302982
Molecular Flexibility	0.00476745	-0.137733
Aromatic Amines	-0.0231852	0.158676
Bromine	-0.07698	0.13010148
Nitrogen	-0.1317613	0.11872575
Basic Nitrogens	-0.1445401	-0.3395737
Acidic Oxygens	-0.1949558	-0.1893562
Amines	-0.2019795	-0.0791169
Amides	-0.2079069	0.01905675
Fragments	-0.2086661	-0.0523105
Alkyl Amines	-0.2305793	-0.2187028
Relative PSA	-0.258151	-0.6096639
Rotatable Bonds	-0.5082707	-0.0307589
Small Rings	-0.604036	0.35858239
Molecular Complexity	-0.6498752	0.28983721
Rings Closures	-0.6510462	0.38253168
H Donors	-0.792077	-0.2634333
Hydrogen	-0.82804	0.06983381
Carbon	-0.8369091	0.3160647
Log Koa	-0.8486332	0.14261532
Electronegative Atoms	-0.8717418	0.00716589
Non C H Atoms	-0.8717807	0.00724998
Sp3 Atoms	-0.8722177	-0.0995463
Total Surface Area	-0.8724984	0.17287071
Polar Surface	-0.8804495	-0.2478448
Oxygen	-0.889839	-0.1554328
MW	-0.9030034	0.25380394
Stereo Centers	-0.9096884	-0.0940876
Non H Atoms	-0.9213871	0.19384145
BP	-0.9272972	0.10163543
H Acceptors	-0.9327043	-0.1209428

Mean values with 95% confidence interval for molecular weight (MW in Da), logKow, number of H-donors and receivers, polar surface area (PSA; Å) and number of rotatable and aromatic bonds for Quadrant I to IV, synthetic and natural pesticides (from marine and non-marine origins). Ranges and limits for these variables according to Lipinski's rule for drugs, tice's rule for herbicides and insecticides and Hao's rule for pesticides; Different letters represent significant differences between pairs as determined by Tukey's HSD tests for synthetic, marine and non-marine pesticides.

	MW (Da)	LogKow	H-donors	H-receiver	Rot. Bonds	Arom. bonds	PSA (Å)	LogKoa	Chlorine	Fluorine	Bromine	Phosphorus	Reference
Quadrant I	317 (305-329)	4.2 (4.0-4.5)	0.5 (0.4-0.6)	3.6 (3.3-3.8)	3.9 (3.4-4.3)	9.1 (8.4-9.8)	52 (48-57)	11.2 (10.7-11.7)	1.52 (0.87-1.43)	0.57 (0.34-0.79)	0.03 (0-0.07)	0.02 (0-0.05)	This study
Quadrant II	227 (215-238)	2.2 (1.9-2.5)	0.9 (0.7-1.0)	3.8 (3.5-4.1)	4.3 (3.7-4.8)	3.2 (2.6-3.7)	67 (62-72)	8.9 (8.4-9.3)	0.40 (0.24-0.56)	0.02 (0-0.04)	0.01 (0-0.02)	0.07 (0.02-0.12)	This study
Quadrant III	512 (445-580)	-0.11 (-1.1-0.9)	4.9 (3.9-5.9)	11.9 (10.5-13.3)	11.1 (7.1-15.1)	1.2 (0.5-1.9)	177 (152-203)	20.5 (17.9-23.0)	0.06 (0-0.13)	0 (0)	0 (0)	0.02 (0-0.06)	This study
Quadrant IV	571 (510-632)	3.6 (2.9-4.4)	3.3 (2.2-4.3)	9.5 (8.2-10.8)	10.8 (8.5-13.0)	7.5 (5.6-9.4)	141 (119-163)	20.9 (18.9-22.9)	0.19 (0.04-0.33)	0.46 (0.08-0.84)	0.13 (0-0.30)	0.01 (0-0.05)	This study
Synthetic pesticides	303 (290-315) B	3.3 (3.1-3.6) A	0.7 (0.6-0.8) C	4.4 (4.1-4.7) C	4.3 (3.8-4.6) C	7.9 (7.2-8.5) A	66 (61-71) C	11.0 (10.6-11.5) B	0.97 (0.09-1.42) A	0.48 (0.31-0.65) A	0.034 (0-0.07) A	0.063 (0.03-0.1) A	This study
Marine pesticides	588 (522-654) A	2.6 (1.7-3.5) AB	4.3 (3.2-5.3) A	10.1 (8.5-11.6) A	11.7 (9.7-13.8) A	2 (1.1-2.9) B	161 (136-186) A	22.2 (20.0-24.5) A	0.13 (0.02-0.23) B	0 (0) B	0.08 (0-0.22) A	0 (0) A	This study
Non-marine pesticides	305 (268-343) B	1.7 (1.1-2.4) B	2.1 (1.5-2.7) B	5.8 (4.7-6.8) B	6.5 (4.2-8.8) B	3.1 (2.2-4.0) B	91 (76-105) B	12.0 (10.5-13.4) B	0.01 (0-0.03) B	0 (0) B	0 (0) A	0.01 (0-0.03) A	This study
Natural pesticides	418 (378-459)	2.1 (1.6-2.6)	3.0 (2.4-3.5)	7.5 (6.6-8.4)	8.6 (7.0-10.2)	2.64 (2.0-3.2)	119 (104-133)	16.1 (14.6-17.6)	0.06 (0.02-0.1)	0 (0)	0.03 (0-0.09)	0.01 (0-0.02)	This study
Lipinski's rule for drugs	<500	≤5	≤5	≤10									Lipinski et al. (2001)
Tice's rule for herbicides	150-500	≤3.5	≤3	2-12	<12								Tice (2001); Avram et al. (2014)
Tice's rule for insecticides	150-500	0-5	≤2	1-8	<12								Tice (2001); Avram et al. (2014)
Hao's rule for pesticides	≤435	≤6	≤2	≤6	≤9	≤17							Hao et al. (2011); Avram et al. (2014)

	% of compounds not respecting Lipinski's rule				
	MW	**LogKow**	**H-acceptor**	**H-donor**	**Overall violation**
Synthetic pesticides	2.9	10.0	2.0	0.0	22
Marine pesticides	58.3	28.3	38.3	21.7	73.3
Non-marine pesticides	12.2	12.2	16.7	14.4	31.3

References

Avram et al. (2014) Quantitative estimation of pesticide-likeness for agrochemical discovery. J Cheminform. 2014; 6: 42.

Hao et al. (2011) A comparative study on the constitutive properties of marketed pesticides. Mol. Inf. 30, 614–622.

Lipinski et al. (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Advanced Drug Delivery Reviews. 46, 3–26.

Tice (2001) Selecting the right compounds for screening: does Lipinski's Rule of 5 for pharmaceuticals apply to agrochemicals? Pest management Science 57: 3–16.