

Exploring an ecotoxicity database with the OECD (Q)SAR Toolbox and DRAGON descriptors in order to prioritise testing on algae, daphnids, and fish

Cléo Tebby, Enrico Mombelli, Pascal Pandard, Alexandre R.R. Pery

▶ **To cite this version:**

Cléo Tebby, Enrico Mombelli, Pascal Pandard, Alexandre R.R. Pery. Exploring an ecotoxicity database with the OECD (Q)SAR Toolbox and DRAGON descriptors in order to prioritise testing on algae, daphnids, and fish. Science of the Total Environment, 2011, 409 (18), pp.3334-3343. 10.1016/j.scitotenv.2011.05.029 . ineris-00961761

HAL Id: ineris-00961761

https://ineris.hal.science/ineris-00961761

Submitted on 20 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exploring an ecotoxicity database with the OECD (Q)SAR Toolbox and DRAGON descriptors in order to prioritise testing on algae, daphnids, and fish.

Cleo Tebby[a]*, Enrico Mombelli[a], Pascal Pandard[b], and Alexandre R.R. Péry[a]

[a] Unité Modèles pour l'Ecotoxicologie et la Toxicologie (METO), INERIS, Parc Technologique Alata, BP2, 60550 Verneuil-en-Halatte, France
[b] Unité Expertise et Essais en Ecotoxicologie (EXES), INERIS, Parc Technologique Alata, BP2, 60550 Verneuil-en-Halatte, France
* Corresponding author. Tel.: +33 3 44 61 82 22; fax: +33 3 44 55 68 00
e-mail address: cleo.tebby@ineris.fr

Abstract:
The European regulation on chemicals (REACh) places emphasis on reduction of systematic toxicity testing, thus fostering development of alternative methods. Consequently, we analysed acute toxicity data gathered by the Japanese Ministry of Environment for three species belonging to three different trophic levels (i.e., *Pseudokirchneriella subcapitata* 72-hour EC50, *Daphnia magna* 48-hour EC50 and *Oryzias latipes* 96-hour LC50). This paper investigates the relationships between the chemical structure and both the toxicity of the chemicals and the cross-species differences in sensitivity. The physicochemical properties of the chemicals were represented by the categories they belonged to in several widely-used categorisation schemes implemented by the freely available OECD (Q)SAR Toolbox and by quantitative molecular descriptors using DRAGON software. The outputs of these software products were analysed and compared in terms of quality of prediction and biological interpretation. Amongst the categorisations implemented by the OECD Toolbox, those focussing on bioaccumulation or biotransformation appeared to be the most interesting in terms of environmental prediction on a whole set of chemicals, in particular as the predicted biotransformation half-life is strongly dependent on hydrophobicity. In predicting toxicity towards each species, simple linear regression on logP performed better than PLS regression of toxicity on a very large set of molecular descriptors. However, the predictions based on the interspecies correlations performed better than the QSAR predictions. The results in terms of cross-species comparisons encourage the use of test strategies focussing on reducing the number of tests on fish.

Keywords:
Cross-species comparison; Acute aquatic toxicity; Sensitivity; Quantitative Structure-Activity Relationships (QSAR); OECD (Q)SAR Toolbox

1 Introduction

According to the European Community Regulation No 1272/2008 (CLP) on classification, labelling and packaging of substances and mixtures, substances are classified according to their acute aquatic toxicity and environmental fate. Acute aquatic toxicity data are usually determined using an algal

species 72- hour EC50, a crustacean species 48-hour EC50, and a fish 96-hour LC50. Most of the time, these tests are carried out according to the Organisation for Economic Cooperation and Development (OECD) test guidelines (i.e. TG201 (OECD, 2006), TG202 (OECD, 2004), and TG203 (OECD, 1992)). Countries outside Europe have also adopted similar regulations: in Japan for instance, the Ministry of the Environment is responsible for evaluating toxicity of chemicals on ecosystems, and recommends the use of tests involving aquatic organisms such as *Oryzias latipes* (fishes) or *Daphnia magna* (daphnia), in addition to algae data (Furuhama et al., 2010). The choice of these three trophic levels (primary producers, primary and secondary consumers) is considered to be relevant in order to protect aquatic ecosystems (Wei et al., 2006). Furthermore, algae, daphnids, and fish are often used for cross-species comparisons of sensitivity (de Roode et al., 2006; Dom et al., 2010; Henegar et al., 2011; Lessigiarska et al., 2004; Tremolada et al., 2004; Weyers et al., 2000).

The European regulation on chemicals (REACh) places emphasis on the reduction of systematic toxicity testing, thus fostering the development of alternative methods, such as testing strategies or statistical methods based on existing data. In this context, Quantitative Structure-Activity Relationships (QSAR) methods relate the physicochemical properties of chemicals with their toxicity on the basis that similar compounds have similar biological activities or properties (Tropsha, 2010). Many QSAR models predict toxicity with specific descriptors such as the n-octanol/water partition coefficient (logP) alone (Hsieh et al., 2006), associated with a chemical class as in EcoSAR software, or, depending on the mechanism of toxic action (Duchowicz and Ocsachoque, 2009), associated with other parameters such as the energy of the lowest unoccupied molecular orbital (Chen et al., 2007; Huang et al., 2007), or the dissociation constant, Hammett Co constant, index of valence molecular connectivity, perimeter of the efficient cross-section of molecule, and melting point (Shigeoka et al., 1988). Other QSAR models rely on generating large sets of descriptors and using statistical methods to reduce dimensionality and identify underlying structural factors influencing toxicity. QSAR models have been made available in various software packages (de Roode et al., 2006).

Structure-Activity Relationships (SAR) can take the physicochemical properties into account by considering the functional or chemical class of molecules. Besides considering the overall level of toxicity of a chemical towards each species, studying species' relative sensitivity is particularly relevant in the regulatory context of environmental risk assessment and when considering the avoidance of tests on vertebrates (Hoekzema et al., 2006; Hutchinson et al., 2003; Jeram et al., 2005). Tremolada et al. (2004) developed Quantitative Inter-specific Chemical Activity Relationships (QICAR) for pesticides on algae, daphnids, and fish. The toxicity was related to the functional and chemical class of the compounds and the physicochemical properties also contributed to predicting the toxicity towards one species based on the toxicity towards a different species.

In this study, we analysed an ecotoxicity database produced by the Japanese Ministry of Environment and available online (Japanese Ministry of Environment, 2010).

Additional data on physicochemical properties of these chemicals were collected in two different ways. The chemicals were classified according to several widely-used classifications implemented by the OECD (Q)SAR Toolbox. Quantitative physicochemical properties were obtained by calculating a set of molecular descriptors using DRAGON software and the results were compared with those obtained using the OECD Toolbox. Since the latter is freely available for download and a free version of the DRAGON descriptors is available online (Tetko et al., 2005; VCCLAB. Virtual Computational Chemistry Laboratory, 2005) our methodology can be easily adopted and reproduced by all the stakeholders involved in regulations on chemicals.

The aims of this paper are two-fold. Firstly, it studies structure-activity relationships between the acute toxicity of the chemicals for three aquatic species and the physicochemical properties. Secondly, it relates the cross-species differences in sensitivity with the physicochemical properties.

2 Material and methods

2.1 Data Set

The data in the ecotoxicity database were gathered by the Japanese Ministry of Environment and available online (Japanese Ministry of Environment, 2010). The tests were conducted according to the OECD test guidelines performed under Good Laboratory Practices (GLP). We selected the toxicity data of organic chemicals that had been collected in tests in or after 2002. Data from tests using vehicles to dissolve the test substance were discarded.

We studied the acute toxicity expressed as the EC50 or LC50 for three aquatic species from three different trophic levels, one alga, one daphnid and one fish. The species selected were *Pseudokirchneriella subcapitata*, previously known as *Selenastrum capricornutum* (72-hour EC50, calculated from average specific growth rate), *Daphnia magna* (48-hour EC50), and *Ozyrias latipes* (96-hour LC50) also known as the Japanese Medaka. *P. subcapitata* and *D. magna* are two widely used and relatively sensitive species; the sensitivity of *O. latipes* is comparable to that of other fish species (Wei et al., 2006). The toxicity information for all three species was available for 85 chemicals in the database. The Chemical Abstract Service (CAS) registration numbers and names of these chemicals are reported in Table 1 of the supplementary material. The final dataset is chemically heterogeneous, containing aliphatic and aromatic organic chemicals, and excluding mixtures. At least eight of the chemicals are pesticides (CAS numbers 117-80-6, 132-27-4, 556-61-6, 56-23-5, 88-85-7, 93-15-2, 127-90-2, 504-24-5. There were no highly hydrophobic chemicals, as the MlogP (Moriguchi's octanol-water partition coefficient obtained by multiple regression analysis on 13 structural parameters (Moriguchi et al., 1992)) ranged from -1.4 to 4.5 and the AlogP (Ghose-Crippen octanol-water partition coefficient based on hydrophobic atomic constants measuring the lipophilic contribution of atoms in the molecule (Viswanadhan et al., 1989)) ranged from -0.9 to 5.8.

The two-dimensional structure of the chemicals was coded in SMILES notation (Weiniger, 1988; Weiniger et al., 1989) in the database. This was used as an input in the freely available OECD Q(SAR) Toolbox version 2.0 (OECD, 2010) and used for profiling the chemicals according to several predefined classification schemes implemented by the Toolbox. The chemicals were thus profiled according to their organic functional groups, the EcoSAR Classification, the mutagenicity/Carcinogenicity alerts by Benigni/Bossa, the US EPA Categorization, the Lipinski Rule, the OASIS Acute Toxicity Mode Of Action, the OASIS protein binding, the superfragment profiling, the OASIS DNA Binding, the Estrogen Receptor binding, the bioaccumulation metabolism alerts (BioWIN MITI fragments, developed to predict biodegradation by microorganisms (Tunkel et al., 2000), with some additional classes, US EPA), and the primary biotransformation half-lives in fish (US EPA), developed partly to improve design of bioaccumulation and toxicity testing (Arnot et al., 2009). These categorisations were used in order to analyse the relationships between the presence of particular structural fragments and the toxicity of chemicals. Attention is drawn to the fact that, even inside a given categorisation scheme, a chemical can belong to several classes. In view of further analysis, the data extracted from the Toolbox was thus recoded into a table where each column was a binary variable indicating whether a given chemical belonged or not to a given class (one column for each chemical class).

The three-dimensional models of the molecules were built with hydrogens and their stereochemistry coincides with the chirality information displayed by the website ChemIDplus lite (http://chem.sis.nlm.nih.gov/chemidplus/chemidlite.jsp). Molecules with less than 6 rotatable bonds were submitted to systematic conformational analysis otherwise they were submitted to a conformational search by Boltzman jump (50000 steps, T=300 K, RMSD=60 degrees). At the end of this search, a conjugate gradients minimization was performed with the MM2 force field and the MOPAC/AM1 semiempirical quantum mechanical calculations were used to generate atom partial charges and further optimize molecular conformations. All the described procedures were carried out in vacuum thanks to the software VegaZZ (Pedretti et al., 2004). A set of 797 two- and three-

dimensional molecular descriptors were calculated using the software DRAGON version 5.3 licensed by Talete srl, Milan (Italy).

2.2 Statistical analysis

The analyses aimed to answer three main questions:

- In the database, are there overall cross-species differences in sensitivity?

- How are classes defined according to chemical structure related to the toxicity of chemicals and to the cross-species differences?

- When considering a large set of DRAGON descriptors, how are these related to the categorisations implemented by the OECD (Q)SAR Toolbox, to toxicity, and to cross-species differences in sensitivity?

2.2.1 Data set verification

Basic properties of the EC50 and LC50 distributions were investigated to ensure the relevance of our statistical analysis. Distribution of log(EC50) or log(LC50) of each species was checked for large gaps (exceeding 15% of the total range) between two consecutive values (Tropsha, 2010). 15% is a relatively high threshold but accounts for the small size of the dataset.

The EC50 or LC50 values were compared with the estimated solubility (logS) calculated with software ALOGPS version 2.1 (Tetko et al., 2001; VCCLAB, 2005).

2.2.2 Cross-species comparison

The interspecies toxicity correlations were measured by the square of Pearson's correlation coefficient between the log(EC50) or log(LC50) of all pairs of species.

The most and the least sensitive species towards each chemical were identified. χ^2 tests on the counts of most or least sensitive species were used to determine whether overall sensitivity was significantly

different between species. The proportion of chemicals which were more toxic towards *P. subcapitata* or *D. magna* than towards *O. latipes* was also computed.

2.2.3 Relationships between chemical class and toxicity

Either Student's t-test or the Mann-Whitney test on the log(EC50) or log(LC50) was used to determine whether the toxicity of a chemical class was significantly different from that of the rest of the chemicals. Student's t-test was used when the log(EC50) or log(LC50) was distributed normally in each population; the Mann-Whitney test was used if this were not the case. Normality in subpopulations was assessed by the Shapiro-Wilk test.

The chemical class was cross-tabulated against the most sensitive species and against the least sensitive species, thus producing one contingency table for each chemical class. Fisher's exact test was used to determine whether sensitivity to a given chemical class was significantly different between trophic levels, taking into account the possible overall cross-species differences in sensitivity. This test was repeated for each chemical class. The expected counts under the null hypothesis of no association between the chemical classes and the species' relative sensitivity were compared with the observed counts in order to identify which species (either one or two) were the most sensitive.

2.2.4 Relationships between molecular descriptors and toxicity

The molecular descriptors were first of all explored by Principal Component Analysis (PCA) in order to identify the main structural differences between the chemicals in the dataset that are highlighted by the set of descriptors, and to relate these with the categorisations implemented by the OECD (Q)SAR Toolbox (with Mann-Whitney's test), with the toxicity (test of nullity of Pearson's correlation coefficient), and the species' relative sensitivity (Fisher's exact test).

Either χ^2 tests or Fisher's exact tests were used to identify the chemical classes that were characteristic of specific subgroups revealed by the PCA. Fisher's exact test was used when the χ^2 approximations could have been incorrect.

The database was randomly split into a training set and a test set (for external validation) with a three to one size ratio (63 chemicals in the training set and 22 chemicals in the test set, Table 1 of the supplementary material), using random sampling without replacement. The distribution of log(EC50) or log(LC50) in both sets was then compared and a Mann-Whitney test was performed for each endpoint to check that the values in the test set were not significantly different from the ones in the training set. Partial Least Squares (PLS) regression was performed on the training set on the whole set of molecular descriptors. Prediction quality was evaluated with the cross-validated R^2 with 7 segments and with the external validation Q^2_{ext} calculated using the average log(EC50) or log(LC50) on the test set (Schüürmann et al., 2008). Linear regression of toxicity on two different calculations of the logP (MlogP and AlogP) and their squares was performed. The quality was evaluated by calculating the coefficient of determination (R^2), both on the training and on the test set (Q^2_{ext}).

All calculations and graphs were made with R software version 2.12.0 (R Development Core Team, 2010), using contributed packages FactoMineR and pls.

3 Results

3.1 Data set verification

72-hour EC50 for *P. subcapitata* ranged from 0.0064 mg/L to 82 mg/L. 48-hour EC50 for *D. magna* ranged from 0.016 mg/L to 87 mg/L. 96-hour LC50 for *O. latipes* ranged from 0.031 mg/L to 90 mg/L. All the chemicals could thus be ranked in one of the classes “very toxic”, “toxic”, or “harmful” defined for the classification and labelling of chemicals according to the European Community Regulation No 1272/2008 (CLP). Subsequent calculations were performed on the logarithm of EC50 or LC50.

For all three species, there were no gaps in log(EC50) or log(LC50) greater than 15% of the total log(EC50) and log(LC50) ranged.

The log(EC50) and log(LC50) expressed in mol/L were compared with the estimated logS for 81 of the chemicals (the logS could not be estimated for the four remaining chemicals). The calculated EC50 or LC50 values were greater than the estimated solubility for two chemicals for *P. subcapitata* (0.43 log units in the worst case) and for two chemicals for *O.latipes* (0.12 log units in the worst case). According to these comparisons, we found no reason to discard any of the substances due to solubility aspects.

3.2 General cross-species comparisons

The interspecies toxicity correlation was highest between *D. magna* and *O. latipes* (R^2=68.3%). The correlation was smaller between the toxicity towards *P. subcapitata* and *D. magna* (R^2=53.8%) and between the toxicity towards *P. subcapitata* and *O. latipes* (R^2=41.3%) (Figure 1).

Insert Figure 1

For each chemical, the EC50 or LC50 were compared between the three species in order to identify the most and least sensitive species. There was one tie in the most sensitive species and two ties in the least sensitive species; in those cases the calculations were done for all possible combinations (two combinations for the test on the most sensitive species and four combinations for the test on the least sensitive species).

All three species were about equally often the most sensitive species (χ^2 =2.85 or 2.14, p-value=0.24 or 0.34, depending on the assignment in case of ties). *D. magna* was rarely the least sensitive species (to 14 or 15 chemicals) (χ^2 =10.54 to 13.01, p-value=0.0015 to 0.0051, depending on the assignment in case of ties). Toxicity towards *P. subcapitata* or *D. magna* was higher than towards *O. latipes* for 75.3% of the chemicals.

3.3 Chemical class and toxicity

The OECD (Q)SAR Toolbox identified a total of 243 classes. Only classes that comprised at least 10% of the whole dataset were selected: as there were 85 chemicals in the dataset, classes with 9 to 76 chemicals were selected for further analysis. Some categorisations did not yield classes with sufficient chemicals: the Lipinski rule and the superfragment profiling were not utilised in this study. In the dataset, 80 out the 85 chemicals were bioavailable according to the Lipinski rule and 77 had no superfragment. Some of the classes were identical in several categorisations, such as phenols or anilines (with some differences in the US EPA classification).

3.3.1 Degree of toxicity according to the chemical class

Few chemical classes show a difference in toxicity towards any of the species considered independently. The ratios of EC50 or LC50 in each class on EC50 or LC50 out of the class are reported for each species in the columns entitled “Toxicity of chemical class” in Table 1. Ratios greater than 1 indicate that the class is less toxic towards a given species than the other chemicals in the data set. The ratios also allow cross-species comparisons to a certain extent, as one class of chemicals can be more toxic than the other classes towards one particular species.

Insert Table 1

Only the categorisations regarding bioaccumulation and biotransformation yielded classes of chemicals that were both represented by a sufficient number of chemicals to be tested and were significantly more or less toxic than the other chemicals in the dataset (i.e. presented a significant difference in log(EC50) or log(LC50)). The chemicals with methyl substituents on an aromatic ring (identified in the bioaccumulation metabolism alerts) were less toxic towards all species than other chemicals (EC50 or LC50 2.7 to 6.0 times greater). Chemicals with an unsubstituted phenyl group were less toxic towards *D. magna* and *O. latipes*. The chemicals with several fused 6C aromatic rings (bioaccumulation metabolism alert) were more toxic towards *P subcapitata* (EC50 11 times smaller)

and *D. magna* (EC50 7.1 times smaller) than other chemicals. Seven out of nine of these chemicals were naphthalene derivatives. The biotransformation half-lives were undetermined for two chemicals, and were moderate (1 to 10 days), fast (0.1 to 1 day) or very fast (under 0.1 day) for all other chemicals. Chemicals with moderate biotransformation half-lives (i.e. longer half-lives than the other chemicals) were more toxic towards all species (EC50 or LC50 3.4 to 4.7 times smaller) and chemicals with very short half-lives were less toxic towards *D. magna* and *O. latipes* (EC50 or LC50 2.8 times greater). Closer inspection of the 19 chemicals with very short biotransformation half-lives revealed that 8 of them belonged to the aniline class (in the EcoSAR classification), and none of them were neutral organics (in the EcoSAR classification) or basesurface narcotics (OASIS Acute Toxicity Mode Of Action).

A more specific toxicity of chemicals with methyl groups towards *O. latipes* was also shown. Also, chemicals with an alkyl substituent on an aromatic ring were less toxic than other chemicals towards *P. subcapitata*.

3.3.2 Cross-species differences in sensitivity

Each chemical class was cross-tabulated separately against the most sensitive species and against the least sensitive species, thus producing one contingency table for each chemical class. Fisher's exact test was applied to each contingency table to determine whether sensitivity to a given chemical class was significantly different between the species. To account for ties, the test was carried out on all combinations for each chemical class. The number of tests with a p-value smaller than 0.05 are reported in Table1 in two separate columns for the test on the most sensitive species and on the least sensitive species. When $p<0.05$, the observed counts were compared with the counts expected under the null hypothesis of no associations between the structural categories and ecotoxicological sensitivity in order to determine which species was (or were) significantly the most sensitive. The significantly most sensitive species to each chemical class are reported in Table 1 in the third to last and in the last column.

When considering the most sensitive species, there was a difference in sensitivity between the three species for 7 out of the 40 chemical classes investigated. *O. latipes* was the most sensitive species towards chemicals with an impaired OH or NH_2 group (Estrogen Receptor binding) or a methyl metabolism alert (BioWIN MITI fragment) (most sensitive species to 41% or 45% of the chemicals with methyl alerts and most sensitive to only 15% of the chemicals without methyl alerts). *D. magna* was the most sensitive species towards arenes (most sensitive species to 45% of the arenes and most sensitive to only 12.5% or 17% of the non-arenes). In the organic functional group classification, arenes are chemicals with one to three six-membered aromatic rings bearing diverse substituents. *D. magna* was also the least sensitive towards chemicals with a methylene functional group (most sensitive species to only 14% of the chemicals with methylene and most sensitive to 44% or 46% of the chemicals without methylene). *P. subcapitata* was less sensitive than the other species towards chemicals with an alkyl substituent on an aromatic ring (BioWIN MITI fragment) (most sensitive species to only 11% of the chemicals with an alkyl substituent on an aromatic ring and most sensitive to 45% of the other chemicals).

When considering the least sensitive species, there was a difference in sensitivity between the three species for phenols, which were identified as a class by four of the categorisations investigated. *P. subcapitata* was shown to be less sensitive towards phenols. Because of ties in EC50 or LC50, it was unclear whether *P. subcapitata* was less sensitive towards chemicals with a methyl functional group and whether *O. latipes* was less sensitive to anilines (in the US EPA classification) and more sensitive towards chemicals with a carbon with four single bonds and no hydrogen (BioWIN MITI fragment).

The chemical classes for which the toxicity towards *P. subcapitata* or *D. magna* was notably higher than towards *O. latipes* in over 85% of cases were identified. This was observed only for a small number of chemical classes: chemicals with aryl halide organic functional groups (aromatic chloride BioWIN MITI fragments for example), and chemicals with several fused 6-carbon aromatic rings (BioWIN MITI fragments).

3.4 Molecular descriptors and toxicity

3.4.1 Exploratory analysis

Amongst the set of 797 molecular descriptors, the ten that correlated most with the log(EC50) or log(LC50) of each species are shown in Table 2. The signification of the descriptors is reported in Table 2 of the supplementary material. The descriptor that correlated best with the toxicity was the squared logP (AlogP² with *P. subcapitata*: r=-0.43; MlogP² with *D. magna*: r=-0.57; MlogP² with *O. latipes*: r=-0.53). AlogP was amongst the top ten for *P. subcapitata* and both MlogP and AlogP were amongst the top ten for *D. magna* and *O. latipes*. As expected, toxicity of chemicals increased with hydrophobicity.

Insert Table 2

Principal Component Analysis (PCA) was applied to the 797 standardised molecular descriptors in order to identify the main structural differences between the chemicals in the dataset that are highlighted by the set of descriptors, without focussing on toxicity. Pearson's correlation coefficient was calculated between the four first principal components and the two different calculations of the logP and their squared values that were included in the molecular descriptors. The correlations between MlogP, MlogP², AlogP, and AlogP² and the first and fourth principal components were all significantly different from zero ($p<0.05$) and relatively high with the fourth component (Pearson's r ranged from 0.34 to 0.52 with the first component and ranged from 0.50 to 0.56 for the fourth component). This shows that hydrophobicity is one of the main characteristics reflected by the DRAGON molecular descriptors and that differentiates the chemicals in this database.

The coordinates of the chemicals of each class on the four first principal components were compared with the coordinates of chemicals that were not in the class (Table 2 of the supplementary material). Most of the classes of chemicals were represented. The two first principal components discriminated chemicals with methyl functional groups (Figure 2). The chemicals in this class were more

hydrophobic (Student's t test, $p<0.05$ for MlogP and AlogP). This could be related to the fact that methyl groups are small hydrophobic units. The second principal component strongly discriminated chemicals with arene functional groups (aromatic hydrocarbons) from non-cyclic structures (Figure 2).

Insert Figure 2

The differences in biotransformation half-life were represented on the fourth component, therefore the relationship between logP and biotransformation half-life was investigated. There was a significant relationship between logP and biotransformation half-life (one-factor analysis of variance without the data with an undefined half-life; p-value= 4.705e-11): the chemicals with a very short half-life were less hydrophobic (Figure 3).

Insert Figure 3

Some chemical classes that were not represented on the first five principal components were actually pointed out earlier in this paper as being important regarding overall toxicity (chemicals with a methyl substituent on an aromatic ring) or relative sensitivity of species (chemicals with an impaired OH or NH_2 group). A number were not represented on the first three. Also there were no significant correlations between log(EC50) or log(LC50) and the two first PCA components (test of nullity of Pearson's correlation coefficient, $p<0.05$). The 2D and 3D DRAGON molecular descriptors used were thus representative of several structural and chemical properties that are related to toxicity, but not of all the structural fragments pointed out earlier in this paper.

The most and the least sensitive species to each chemical are shown on the scatter plot in Figure 4. In case of ties, the most/least sensitive species to a chemical was randomly assigned. The patterns tended to show that relative toxicity towards the various species were at least related to the two first principal components. Approximately one quarter of the scatter plot contained 30 chemicals that were mostly more toxic towards *D. magna* or *O. latipes* and less toxic towards *P. subcapitata* or *O. latipes*. This area (coordinate on 1st component smaller than 1 and coordinate on 2nd component smaller than -1.85)

was examined by a series of Fisher's exact tests on the chemical class (results in Table 4 of the supplementary material). The chemicals in this area were all arenes (organic functional group classification) but displaying a wide variety of substituents; compared to the rest of the dataset, a large proportion had an aryl halide functional group with chlorine or were classified as anilines in the US EPA classification. None of them had a methylene functional group. The chemicals in this area were on average as hydrophobic as the other chemicals (Mann-Whitney test on MlogP, p=0.051), however the range and variance of logP were smaller (MlogP ranged from 1.2 to 3.8; Levene's Test for Homogeneity of Variance on MlogP: p=0.0040). The chemicals in this area had noticeably lower values for some DRAGON descriptors belonging to the groups of Edge Adjacency indices Eigenvalues, of lowest eigenvalues of the Burden matrix, 3rd component size or accessibility directional WHIM index as well as a lower spherosity (SPH descriptor values were characteristic of aromatic chemicals).

Insert Figure 4

The first principal components are thus related to cross-species differences in sensitivity as well as to some chemical classes which in turn are related to the degree of toxicity of the chemicals. Therefore the molecular descriptors might be predictors of the acute toxicity of the chemicals. The relationships between DRAGON descriptors on one side and toxicity and relative sensitivity on the other side were further investigated by PLS regression during QSAR modelling.

3.4.2 QSAR models of toxicity

Table 1 of the supplementary material shows which chemicals were in the training set and in the test set. The log(EC50) or log(LC50) in the training set and test set were not significantly different for any of the species (Mann-Whitney test, p>0.05), although the lowest value for *P. subcapitata* and for *D. magna* was in the test set. A visual inspection of the position of the test set chemicals on the first two axes of the PCA with all chemicals did not reveal any differences in distribution of the test set and training set.

PLS1 regression was performed for each of the three endpoints on the training set with 7-segment cross-validation on the 782 descriptors that were non-constant both on the training and test set. The PLS models had negative cross-validated coefficients of determination (Q^2_{CV}) and very low predictive validity (results not shown).

Regarding external validation quality of prediction, PLS modelling did not perform as well as a simple linear regression of log(EC50) or log(LC50) on either calculations of the logP or their squares (R^2, R^2 adjusted for the number of variables, and Q^2 are reported in Table 3), where the Q^2_{ext} ranged from 0.12 (MlogP, *P. subcapitata*) to 0.61 (AlogP², *D. magna*) although the R^2 values obtained on the training set were smaller. These predictions were also of poor predictive validity. Multiple linear regression on both the logP and logP² (Patrick, 2001) did not perform better regarding adjusted R^2 or Q^2_{ext} (results not shown).

Insert Table 3

Compared to linear regression on AlogP², the predictions obtained considering pairwise interspecies correlations (Table 3) were better for all three species. For *P. subcapitata*, Q^2_{ext} with linear regression on AlogP² was equal to 0.47, whereas on $\log(EC50)_{D.magna}$, Q^2_{ext} was equal 0.55. For *D. magna*, the increase in Q^2_{ext} was greater: on AlogP², Q^2_{ext} was equal to 0.61 whereas on $\log(LC50)_{O.\ latipes}$, Q^2_{ext} was equal to 0.80. For *O. latipes*, on AlogP², Q^2_{ext} was equal to 0.57; on $\log(EC50)_{D.\ magna}$, Q^2_{ext} reached 0.77. Predicting toxicity towards one species with toxicity from both other species (instead of only one) only improved the predictions for *D. magna* when considering the adjusted R^2 (R^2_{adj}=0.64 on $\log(LC50)_{O.\ latipes}$; R^2_{adj}=0.72 on both $\log(EC50)_{P\ subcapitata}$ and $\log(LC50)_{O.\ latipes}$). Adding the AlogP² measure of hydrophobicity to the later model did not further improve the predictions when considering the adjusted R^2 (Table 3).

PLS discriminant analysis of the most sensitive species according to the physico-chemical descriptors performed poorly (results not shown).

4 Discussion

The dataset analysed was chemically heterogeneous, containing aliphatic and aromatic organic chemicals, but there were no highly hydrophobic chemicals (with AlogP>6) and no persistent chemicals. Its representativeness is quite limited, considering the physicochemical properties of chemicals notified in Europe.

The interspecies correlations between fish and *Daphnia* toxicity values were higher than the correlation with algae, thus confirming previous research on other ecotoxicity databases (Henegar et al., 2011; Lessigiarska et al., 2004; Tremolada et al., 2004; Weyers et al., 2000; Zhang et al., 2010). Zhang et al. (2010) suggest the difference is due to bio-uptake and mode of action. The magnitude of the correlations are similar to those calculated by Tremolada et al. (2004) for pesticides between *Daphnia* and trout (R^2=59%), and for unspecific pesticides between algae and *Daphnia* (R^2=52%) and algae and trout (R^2=49%).

In an analysis of 164 chemicals from a database provided by the French Ministry of Ecology and Sustainable Development, Henegar et al. (2011) showed that fish were less sensitive than algae and *Daphnia*. Weyers et al. (2000) showed that alga was the most sensitive trophic level in a database of 694 chemicals, extracted from the New Chemicals Database which comprises information about new chemical substances manufactured or imported in the European Union since 1981. In a study on 1,439 chemicals, Jeram et al. (2005) showed that alga was the most sensitive and fish was the least sensitive. However, their database included many herbicides, which are likely to target algae. The database analysed in this study was smaller and contained only eight pesticides, of which one herbicide, one insecticide, and several non-specific pesticides. Part of our results suggests that daphnids might be the most sensitive in general, but the small size of the dataset leads to limit the significance of this conclusion. Moreover, the other published studies do not support this conclusion.

OECD, based on the several studies (Hoekzema et al., 2006; Hutchinson et al., 2003; Jeram et al., 2005), proposed a threshold approach to reduce the number of fish for the determination of acute

aquatic toxicity in a regulatory context. The authors report that algae or daphnia were more sensitive than fish for 80% to 88% of the chemicals in their databases, compared to 75% in our dataset. The testing strategy based on that developed by Hutchinson et al. (2003) would however still have enabled reducing the number of fish used, by designing the fish tests on the basis of the results obtained with algae or daphnia and proceeding sequentially (limit test in a first step and full test if lethal effects are observed at the lowest EC50 determined in the daphnia and algae tests).

When using the OECD (Q)SAR Toolbox to identify structural fragments related to toxicity, only the categorisations focussing on bioaccumulation or biotransformation yielded classes of chemicals that were both represented by a sufficient number of chemicals to be tested and were significantly more or less toxic than the other chemicals in the dataset. In particular, chemicals with very short biotransformation half-lives in fish were less toxic to all species. In this data set, the chemicals with a very short biotransformation half-life were less hydrophobic, as the QSAR model which was used to predict the biotransformation half-life (Arnot et al., 2009) was partly based on the logP. Linear regression of log(EC50) or log(LC50) on logP confirmed previous research showing that toxicity of chemicals increased with hydrophobicity (Duchowicz and Ocsachoque, 2009; Hsieh et al., 2006; Shigeoka et al., 1988). It is unclear whether the difference in toxicity observed in the dataset is due to biotransformation half-life or to hydrophobicity.

Linear regression on a measure of logP performed better than PLS regression on a large set of DRAGON molecular descriptors. According to Golbraikh and Tropsha (2002), two of the necessary conditions for a model to be considered as predictive are that $R^2>0.6$ and $Q^2>0.5$. None of the QSAR models developed in this paper by PLS regression or by linear regression on a measure of logP satisfy these criteria. The results from the PLS regression do not encourage the selection of any particular molecular descriptors apart from the logP. Interspecies correlations between toxicity tests were better than between toxicity and logP, as shown by Weyers et al. (2000) in a study on part of the New Chemicals Database. The external validation Q^2 values were often greater than the R^2 values and

therefore the distribution of the chemical classes in the original data set in the training set and test set was checked. In the test set, there were a slightly higher proportion of aromatic chlorides and chemicals with moderate biotransformation half-lives, but this did not result in any significant difference in toxicity.

The conclusions in terms of relative sensitivity of species are not straightforward, as the results obtained with the test on the most sensitive species are not confirmed by the test on the least sensitive species. On the other hand, tests generally focus on the most sensitive species rather than on the least sensitive, meaning that the data concerning the most sensitive species are more reliable. Basing their research on the Verhaar classification, Vaal et al. (1997b) concluded that chemicals with an unspecific mode of action were the least toxic and demonstrated smaller differences in interspecies sensitivity. They also concluded (1997a) that, except for anilines, compounds with the highest overall toxicity also have the largest variation in toxicity for different species. In our dataset however, the most toxic chemical classes (i.e. chemicals with the longest biotransformation half-lives) did not display any cross-species differences in toxicity.

Our analysis with the OECD Toolbox outputs was repeated on the same dataset with data collected on 141 additional chemicals before 2002, which might be less reliable as in some cases a dispersant was used, thus increasing almost three-fold the number of chemicals. The results obtained on the larger database confirmed the differences in sensitivity towards arenes (*D. magna* being more sensitive than the other species), methylene functional groups (*D. magna* being less sensitive), and phenols (*P. subcapitata* being less sensitive). Some of the classes identified by the Toolbox, such as anilines (aromatic amines), are subject to differences between categorisation schemes (US EPA versus organic functional groups or EcoSAR). On the other hand, arenes appear to be easily identifiable and not subject to differences between categorisation schemes. It has to be noted that arenes are a rather heterogeneous class of chemicals and further research focussing on this class would be required to better define the structural characteristics that could make daphnids more sensitive.

Anilines (aromatic amines) are pinpointed by a number of the classifications implemented by the OECD (Q)SAR Toolbox as an important class of chemicals regarding toxicity and have been shown to be particularly toxic towards daphnids due to a specific mode of action (Chen et al., 2007; Dom et al., 2010; Neuwoehner et al., 2010; Vaal et al., 1997a). However, in this dataset, chemicals belonging to the aniline class were not significantly more toxic towards *D. magna*.

Overall, the results regarding classes of chemicals representing structural alerts were obtained by performing a large number of tests with a high significance threshold (5%) in an exploratory analysis. The number of significant relationships is relatively small, and due to the fact that several classifications yield some identical or very similar classes, it is difficult comparing the proportion of significant results to the chance level. However, the results of the analysis of the larger database confirmed the results in terms of relative sensitivity of species and also generalised to all three species the effect of biotransformation half-lives observed with the 85-chemical dataset.

5 Conclusions

Although this dataset is relatively small, it is nevertheless valuable in particular as 85 chemicals were tested with the same methodology, in a relatively short span of time, and on three aquatic species belonging to three different trophic levels. Also the main results obtained on this small dataset were confirmed by the analysis of a larger dataset. The effort put into sharing the data by the Japanese Ministry of Environment is thus very valuable and should encourage the publication of this type data in general as it could increase the relevance of comparative studies. Datasets relative to more homogeneous chemicals would also be welcome.

Exploratory analyses both on presence of structural fragments and on quantitative molecular descriptors illustrated relationships between toxicity and cross-species differences in sensitivity on the one hand and physico-chemical characteristics, especially hydrophobicity, on the other hand. In predicting toxicity towards each species, simple linear regression on logP performed better than PLS

regression of toxicity on a very large set of molecular descriptors, thus confirming the importance of hydrophobicity in predicting ecotoxicity. The predictions of toxicity based on the interspecies correlations performed better than the QSAR predictions. In the scope of reducing the number of in vivo tests, the cross-species differences in sensitivity obtained with the outputs of the OECD (Q)SAR Toolbox suggest that aromatic chemicals should be tested first of all on *D. magna*, as part of a testing strategy. The concentrations tested on fish could then be adjusted according to the toxicity towards *D. magna*, in order to avoid unnecessary tests and thus reduce the number of vertebrates used in ecotoxicity testing.

Though a number of statistically significant associations were revealed in this paper, the interpretation in terms of biological mode of action is not straightforward due to a lack of knowledge. Though the OECD Toolbox does implement a categorisation based on the mode of action, the classes obtained did not differ in toxicity. Amongst the categorisations implemented by the OECD Toolbox, those focussing on bioaccumulation or biotransformation appeared to be the most interesting in terms of environmental prediction on a whole set of chemicals, partly as they yielded classes of chemicals that were both represented by a sufficient number of chemicals to be tested. However, some categorisations may pinpoint particularly relevant structural fragments but may not be sufficiently represented in this dataset to allow drawing any conclusions. The use of the categorisation schemes implemented in the OECD (Q)SAR Toolbox and of DRAGON molecular descriptors showed that the molecular descriptors are related to some well-defined chemicals classes, such as arenes or chemicals with a methyl functional group, which are relevant considering toxicity or relative sensitivity of species.

Further research is required to investigate the relationship between hydrophobicity, biotransformation half-life and toxicity towards aquatic species. Further developments in QSAR methods could include nonlinear models to improve predictions or the integration of uncertainty on the experimental values

(EC50or LC50) in order to estimate the uncertainty on the predictions. The integration of molecular reactivity represented by quantum descriptors in the QSAR models might improve the predictions.

6 Acknowledgements

This study was supported by the French Ministry in charge of Ecology and Sustainable Development, within the framework of Programme 181 and Programme 190 and by the National Research Agency (ANR) within the project AMORE (Contract number : no. 2009 CESA 15 01).

7 References

Arnot JA, Meylan W, Tunkel J, Howard PH, Mackay D, Bonnell M, et al. A quantitative structure–activity relationship for predicting metabolic biotransformation rates for organic chemicals in fish. Environmental Toxicology and Chemistry 2009; 28: 1168–1177.

Chen C-Y, Ko C-W, Lee P-I. Toxicity of substituted anilines to *Pseudokirchneriella subcapitata* and quantitative structure–activity relationship analysis for polar narcotics. Environmental Toxicology and Chemistry 2007; 26: 1158–1164.

de Roode D, Hoekzema C, de Vries-Buitenweg S, van de Waart B, van der Hoeven J. QSARs in ecotoxicological risk assessment. Regulatory Toxicology and Pharmacology 2006; 45: 24-35.

Dom N, Knapen D, Benoot D, Nobels I, Blust R. Aquatic multi-species acute toxicity of (chlorinated) anilines: Experimental versus predicted data. Chemosphere 2010; 81: 177–186.

Duchowicz PR, Ocsachoque MA. Quantitative Structure-Toxicity Models for Heterogeneous Aliphatic Compounds. QSAR & Combinatorial Science 2009; 28: 281 – 295.

Furuhama A, Toidaa T, Nishikawaa N, Aokia Y, Yoshiokab Y, Shiraishi H. Development of an ecotoxicity QSAR model for the KAshinhou Tool for Ecotoxicity (KATE) system, March 2009 version. SAR and QSAR in Environmental Research 2010; 21: 403-413.

Golbraikh A, Tropsha A. Beware of q^2! Journal of Molecular Graphics and Modelling 2002; 20: 269–276.

Henegar A, Mombelli E, Pandard P, Péry ARR. What can be learnt from an ecotoxicity database in the framework of the REACh regulation? Science of the Total Environment 2011; 409: 489-494.

Hoekzema CC, Murk AJ, van de Waart BJ, van der Hoeven JCM, de Roode DF. Alternative approaches can greatly reduce the number of fish used for acute toxicity testing. Environmental Toxicology and Chemistry 2006; 25: 1322-1325.

Hsieh S-H, Hsu C-H, Tsai D-Y, Chen C-Y. Quantitative structure–activity relationships for toxicity of nonpolar narcotic chemicals to *Pseudokirchneriella subcapitata*. Environmental Toxicology and Chemistry 2006; 2006: 2920–2926.

Huang C-P, Wang Y-J, Chen C-Y. Toxicity and quantitative structure–activity relationships of nitriles based on Pseudokirchneriella subcapitata. Ecotoxicology and Environmental Safety 2007; 67: 439–446.

Hutchinson TH, Barrett S, Buzby M, Constable D, Hartmann A, Hayes E, et al. A strategy to reduce the numbers of fish used in acute ecotoxicity testing of pharmaceuticals. Environmental Toxicology and Chemistry 2003; 22: 3031-3036.

Japanese Ministry of Environment. Japan ecotoxicity tests data, March 2010, 2010.

Jeram S, Riego Sintes JM, Halder M, Baraibar Fentanes J, Sokull-Klüttgen B, Hutchinson TH. A strategy to reduce the use of fish in acute ecotoxicity testing of new chemical substances notiWed in the European Union. Regulatory Toxicology and Pharmacology 2005; 42: 218-224.

Lessigiarska I, Worth AP, Sokull-Klüttgen B, Jeram S, Dearden JC, Netzeva TI, et al. QSAR investigation of a large data set for fish, algae and Daphnia toxicity. SAR and QSAR in Environmental Research 2004; 15: 413-431.

Moriguchi I, Hirono S, Liu Q, Nakagome I, Matsushita Y. Simple method of calculating Octanol/Water Partition Coefficient. Chemical & Pharmaceutical Bulletin 1992; 40: 127-130.

Neuwoehner J, Zilberman T, Fenner K, Escher BI. QSAR-analysis and mixture toxicity as diagnostic tools: Influence of degradation on the toxicity and mode of action of diuron in algae and daphnids. Aquatic Toxicology 2010; 97: 58–67.

OECD. Organisation for Economic Cooperation and Development. Guidelines for Testing of Chemicals 203, Fish Acute Toxicity Test, 1992.

OECD. Organisation for Economic Cooperation and Development. OECD Guidelines for Testing of Chemicals, Guideline 202. Daphnia sp., Acute Immobilisation Test, 2004.

OECD. Organisation for Economic Cooperation and Development. OECD Guidelines for Testing of Chemicals, Guideline 201. Freshwater Alga and Cyanobacteria, Growth Inhibition Test, 2006.

OECD. Organisation for Economic Cooperation and Development. QSAR Toolbox User Manual and Installation Notes, 2010.

Patrick GL. Quantitative Structure-Activity Relationships. In: Oxford University Press, editor. An Introduction to Medicinal Chemistry, 2001, pp. 622.

Pedretti A, Villa L, Vistoli G. VEGA – An open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. Journal of Computer-Aided Molecular Design 2004; 18: 167-173.

R Development Core Team. R: A language and environment for statistical computing, 2010.

Schüürmann G, Ebert R-U, Chen J, Wang B, Kühne R. External Validation and Prediction Employing the Predictive Squared Correlation Coefficient - Test Set Activity Mean vs Training Set Activity Mean. Journal of Chemical Information and Computer Sciences 2008; 48: 2140–2145.

Shigeoka T, Sato Y, Takeda Y, Yoshida K, Yamauchi F. Acute toxicity of chlorophenols to green algae, *Selenastrum capricornutum* and *Chlorella vulgaris*, and quantitative structure-activity relationships. Environmental Toxicology and Chemistry 1988; 7: 847-854.

Tetko I, Tanchuk V, Kasheva T, Villa A. Estimation of aqueous solubility of chemical compounds using E-state indices. Journal of Chemical Information and Computer Sciences 2001; 41: 1488-1493.

Tetko IV, Gasteiger J, Todeschini R, Mauri A, Livingstone D, Ertl P, et al. Virtual computational chemistry laboratory – design and description. Journal of Computer-Aided Molecular Design 2005; 19: 453–463.

Tremolada P, Finizio A, Villa S, Gaggi C, Vighi M. Quantitative inter-specific chemical activity relationships of pesticides in the aquatic environment. Aquatic Toxicology 2004; 67: 87-103.

Tropsha A. Best Practices for QSAR Model Development, Validation, and Exploitation. Molecular Informatics 2010; 29: 476 – 488.

Tunkel J, Howard PH, Boethling RS, Stiteler W, Loonen H. Predicting ready biodegradability in the Japanese ministry of international trade and industry test. Environmental Toxicology and Chemistry 2000; 19: 2478–2485.

Vaal M, van der Wall JT, Hoekstra J. Pattern analysis of the variation in the sensitivity of aquatic species to toxicants. Chemosphere 1997a; 35: 1291-1309.

Vaal M, van der Wall JT, Hoekstra J, Hermens J. Variation in the sensitivity of aquatic species in relation to the classification of environmental pollutants. Chemosphere 1997b; 35: 1311-1327.

VCCLAB. Virtual Computational Chemistry Laboratory. 2005.

Viswanadhan VN, Ghose AK, Revankar GR, Robins RK. Atomic physicochemical parameters for 3 dimensional structure directed quantitative structure - activity relationships .4. Additional parameters for hydrophobic and dispersive interactions and their application for an automated superposition of certain naturally-occurring nucleoside antibiotics. Journal of Chemical Information and Computer Sciences 1989; 29: 163-172.

Wei D, Kisuno A, Kameya T, Urano K. A new method for evaluating biological safety of environmental water with algae, daphnia and fish toxicity ranks. Science of the Total Environment 2006; 371: 383–390.

Weiniger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. Journal of Chemical Information and Computer Sciences 1988; 28: 31-36.

Weiniger D, Weiniger A, Weiniger JL. SMILES. 2. Algorithm for Generation of Unique SMILES Notation. Journal of Chemical Information and Computer Sciences 1989; 29: 97-101.

Weyers A, Sokull-Klüttgen B, Baraibar-Fentanes J, Vollmer G. Acute toxicity data: a comprehensive comparison of results of fish, daphnia, and algae tests with new substances notified in the European Union. Environmental Toxicology and Chemistry 2000; 19: 1931–1933.

Zhang XJ, Qin HW, Su LM, Qin WC, Zou MY, Sheng LX, et al. Interspecies correlations of toxicity to eight aquatic organisms: Theoretical considerations. Science of the Total Environment 2010; 408: 4549-4555.

Table 1: Ratios of mean EC50 or LC50 of chemical class over mean EC50 or LC50 of other chemicals; number of cases where the Fisher's exact test p-value was smaller than 0.05; most sensitive species.

			Toxicity of chemical class: ratio EC50 or LC50			Species' relative sensitivity			
						Fisher's exact test on most sensitive species		Fisher's exact test on least sensitive species	
Classification	Class	Number of chemicals	EC50 P. subcapitata	EC50 D. magna	LC50 O. latipes	Number of cases were the p-value is <0.05 (out of 2 cases)	Most sensitive species	Number of cases were the p-value is <0.05 (out of 4 cases)	Most sensitive species
Organic functional group	Anilines	14	1.74	1.61	2.37				
	Arene	61	1.33	0.724	1.03	2/2	D. m		
	Aryl halide	12	1.57	1.28	1.44				
	Ether	11	1.78	2.42	2				
	Methyl	39	1.06	0.616	0.511			2/4	D. m O. l
	Methylene	22	0.81	0.895	0.612	2/2	P. s O. l		
	Phenol	28	1.39	1.03	0.772			3/4	D. m O. l
EcoSAR Classification	Anilines (Aromatic Amines)	14	1.74	1.61	2.37				
	Neutral Organics	21	0.808	0.951	1.14				
	Phenols	25	1.36	0.834	0.589			4/4	D. m O. l
Mutagenicity/ Carcinogenicity alerts by Benigni/Bossa	No alerts for carcinogenic activity	53	0.842	1.08	1.05				
	Primary aromatic amine,hydroxyl amine and its derived esters	12	1.47	1.42	2.02				
	Structural alert for genotoxic carcinogenicity	29	0.95	0.759	0.737				
US EPA Classification	N/A)	26	0.796	0.945	0.824				
	Anilines (Acute toxicity)	11	1.52	1.03	2.44			2/4	P. s D. m
	Phenols (Acute toxicity)	25	1.36	1.01	0.691			4/4	D. m O. l
Acute Toxicity MOA by OASIS	Basesurface narcotics	21	0.615	0.659	0.871				
	Phenols and Anilines	34	1.82	1.01	1.33				
	Reactive unspecified	25	0.649	1.1	0.622				
Protein binding	No Binding	71	1.34	1.57	2.99				
DNA binding	No Binding	60	1.28	1.11	0.952				
Estrogen Receptor Binding	Non-cyclic structure	15	0.418	0.986	0.881	1/2	P. s O. l		
	Weak binder, OH	10	1.86	1.5	1.14				
	With impaired OH or NH2 group	11	2.43	2	0.89	2/2	O. l		
	Without OH or NH2 group	27	0.751	0.634	0.754				
Bioaccumulation – metabolism alerts	-CH2- [linear]	17	0.821	0.869	0.57	2/2	P. s O. l		
	Alkyl substituent on aromatic ring	19	3.32	1.59	1.31	2/2	D. m O. l		

	Aromatic-CH3	12	6.02	2.75	2.88				
	Aromatic-H	65	1.66	0.929	1.05				
	Aromatic alcohol [-OH]	28	1.39	1.03	0.772			3/4	D. m O. l
	Aromatic amine [-NH2 or -NH-]	17	1.16	1.21	1.73				
	Aromatic chloride [-CL]	11	1.8	1.34	1.5				
	Benzene	51	2.26	1.36	1.37				
	Carbon with 4 single bonds & no hydrogens	12	0.881	0.645	0.528			2/4	O. l
	Methyl [-CH3]	29	0.6	0.465	0.31	2/2	O. l		
	Number of fused 6-carbon aromatic rings	9	0.0922	0.14	0.293				
	Unsubstituted phenyl group (C6H5-)	11	0.352	0.273	0.324				
Bioaccumulation – metabolism half-lives	Fast	46	1.88	1.72	1.42				
	Moderate	18	0.29	0.211	0.264				
	Very fast	19	1.84	2.84	2.8				

Highlighted cells relate to p-values under 0.05 meaning that the chemical class is either more or less toxic than the rest of the chemicals (Student tests or Mann-Whitney tests depending on normality of EC50 or LC50 distributions). Tests on most sensitive species were carried out twice, allowing for the two possible cases due to one tie. Tests on least sensitive species were carried out four times, allowing for the four possible cases due to two ties.

Table 2: Top ten molecular descriptors with highest Pearson correlation coefficient (r) with log(EC50) or log(LC50) values

	log(EC50) *P. subcapitata*		log(EC50) *D. magna*		log(LC50) *O. latipes*	
Rank	Molecular descriptor	r	Molecular descriptor	r	Molecular descriptor	r
1	ALOGP²	-0.43	MLOGP²	-0.57	MLOGP²	-0.53
2	MLOGP²	-0.41	MLOGP	-0.55	ALOGP²	-0.52
3	Mor27u	-0.4	BLTF96	0.55	ALOGP	-0.49
4	D/Dr10	-0.39	ALOGP²	-0.54	MLOGP	-0.47
5	L/Bw	-0.38	ALOGP	-0.5	BLTF96	0.47
6	nR10	-0.35	TPSA(NO)	0.46	EEig14d	0.39
7	piPC09	-0.35	nHAcc	0.45	Hy	0.37
8	IC1	0.35	nO	0.43	MAXDN	0.36
9	ALOGP	-0.35	SEigv	-0.41	TPSA(NO)	0.36
10	EEig14d	0.34	TPSA(Tot)	0.41	Ms	0.35

Signification of the abbreviations are given in Table 2 of the supplementary material.

Table 3: Coefficient of determination (R^2, first row), adjusted coefficient of determination (2nd row) on training set, and externally validated coefficient of determination (Q^2_{ext}, 3rd row) for linear regression models of log(EC50) or log(LC50) for each species on logP (two different calculations by Dragon and their squares), on each other species, on both other species, and on both other species and $AlogP^2$.

	Log(EC50) *P. subcapitata*	Log(EC50) *D. magna*	Log(LC50) *O. latipes*	Log(EC50) or log(LC50) 2 species	Log(EC50)or log(LC50) 2 species + $AlogP^2$	MlogP	$MlogP^2$	AlogP	$AlogP^2$
		0.525	0.382	0.529	0.529	0.084	0.166	0.037	0.068
Log(EC50) *P. subcapitata*		0.518	0.371	0.513	0.505	0.069	0.153	0.022	0.052
		0.555	0.477	0.560	0.560	0.117	0.117	0.258	0.474
	0.525		0.644	0.729	0.731	0.269	0.299	0.129	0.119
Log(EC50) *D. magna*	0.518		0.638	0.720	0.717	0.257	0.288	0.115	0.104
	0.540		0.761	0.799	0.816	0.318	0.363	0.465	0.615
	0.382	0.644		0.647	0.655	0.138	0.185	0.111	0.132
Log(LC50) *O. latipes*	0.371	0.638		0.635	0.638	0.124	0.172	0.097	0.118
	0.474	0.775		0.777	0.784	0.352	0.459	0.464	0.575

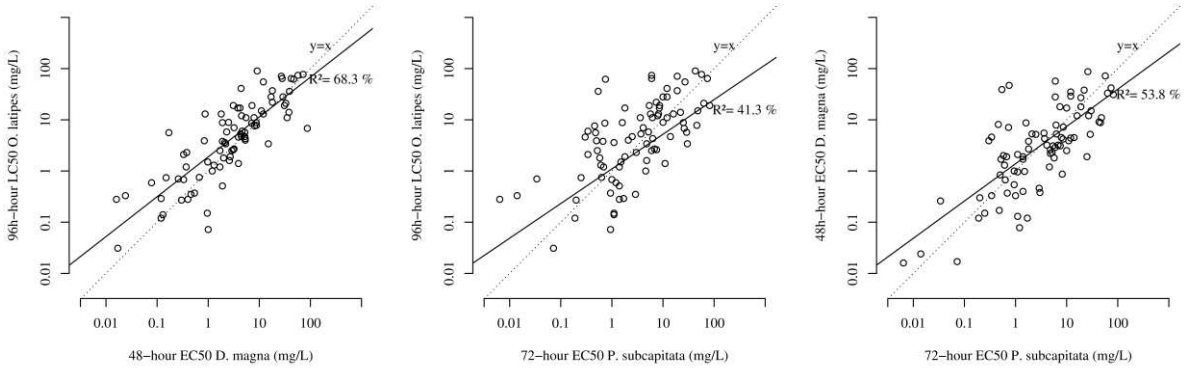


Figure 1: Pairwise interspecies correlations between experimental EC50 or LC50 values for *O. latipes*, *D. magna*, and *P. subcapitata*. The solid line is the regression line.

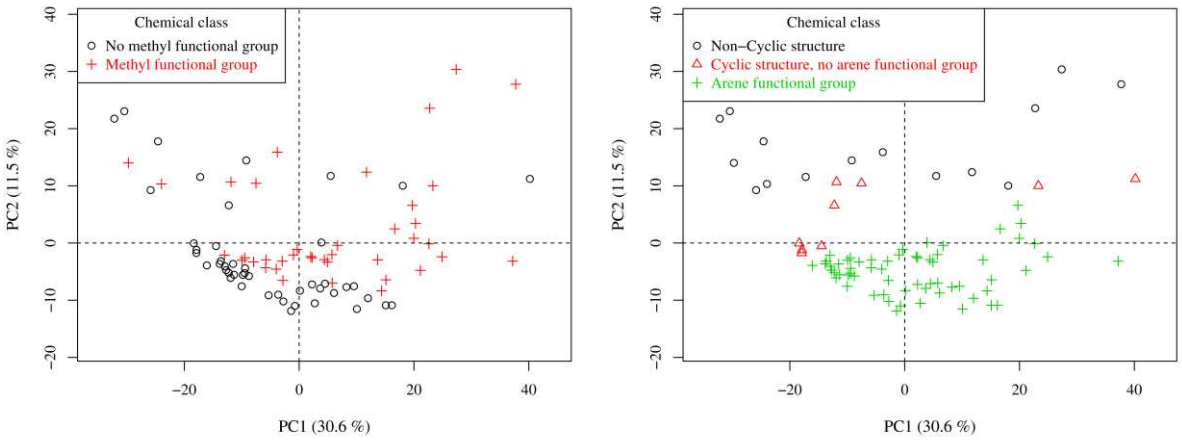


Figure 2: Scatter plots of the two first principal components, highlighting chemicals with a methyl organic functional group (left) and aromatic vs. non-cyclic chemicals (right)

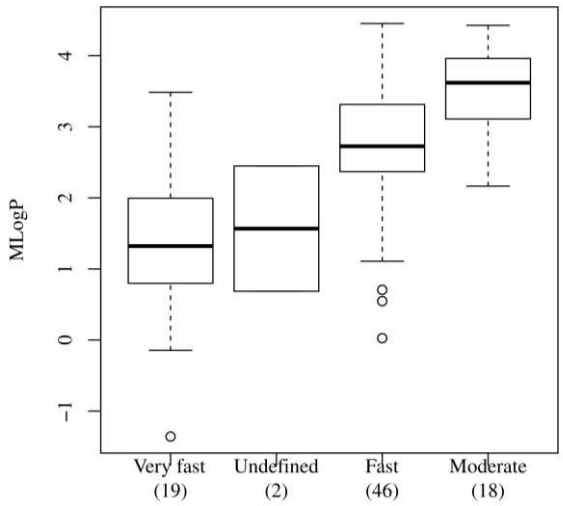


Figure 3: Boxplot of MLogP according to the biotransformation half-life in fish. Number of chemicals in brackets.

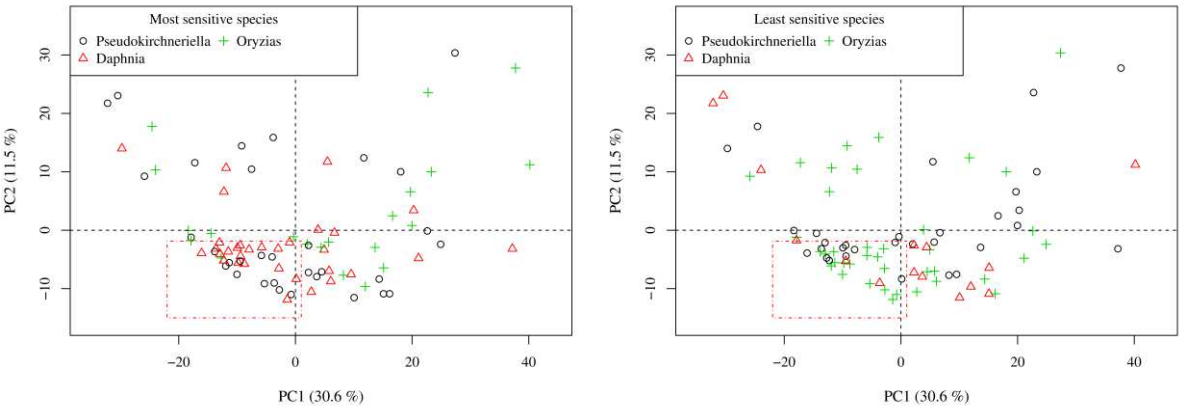


Figure 4: Scatter plots of the two first principal components, showing most sensitive species to each chemical (left), least sensitive species to each chemical (right) and specific area investigated.