Integrative Chemical-Biological Read-Across Approach

for Chemical Hazard Classification

Yen Low[†,‡], Alexander Sedykh[†], Denis Fourches[†], Alexander Golbraikh[†],

Maurice Whelan[‖], Ivan Rusyn[‡*], and Alexander Tropsha[†*]

[†]Laboratory for Molecular Modeling, Division of Chemical Biology and Medicinal Chemistry, and [‡]Department of Environmental Sciences and Engineering, University of North Carolina, Chapel Hill, NC 27599, USA;

[‖]Systems Toxicology Unit, Institute for Health and Consumer Protection, Joint Research Centre, European Commission, via E. Fermi 2749, 21027 Ispra, Italy

Supplemental File 2 (Figure S1)

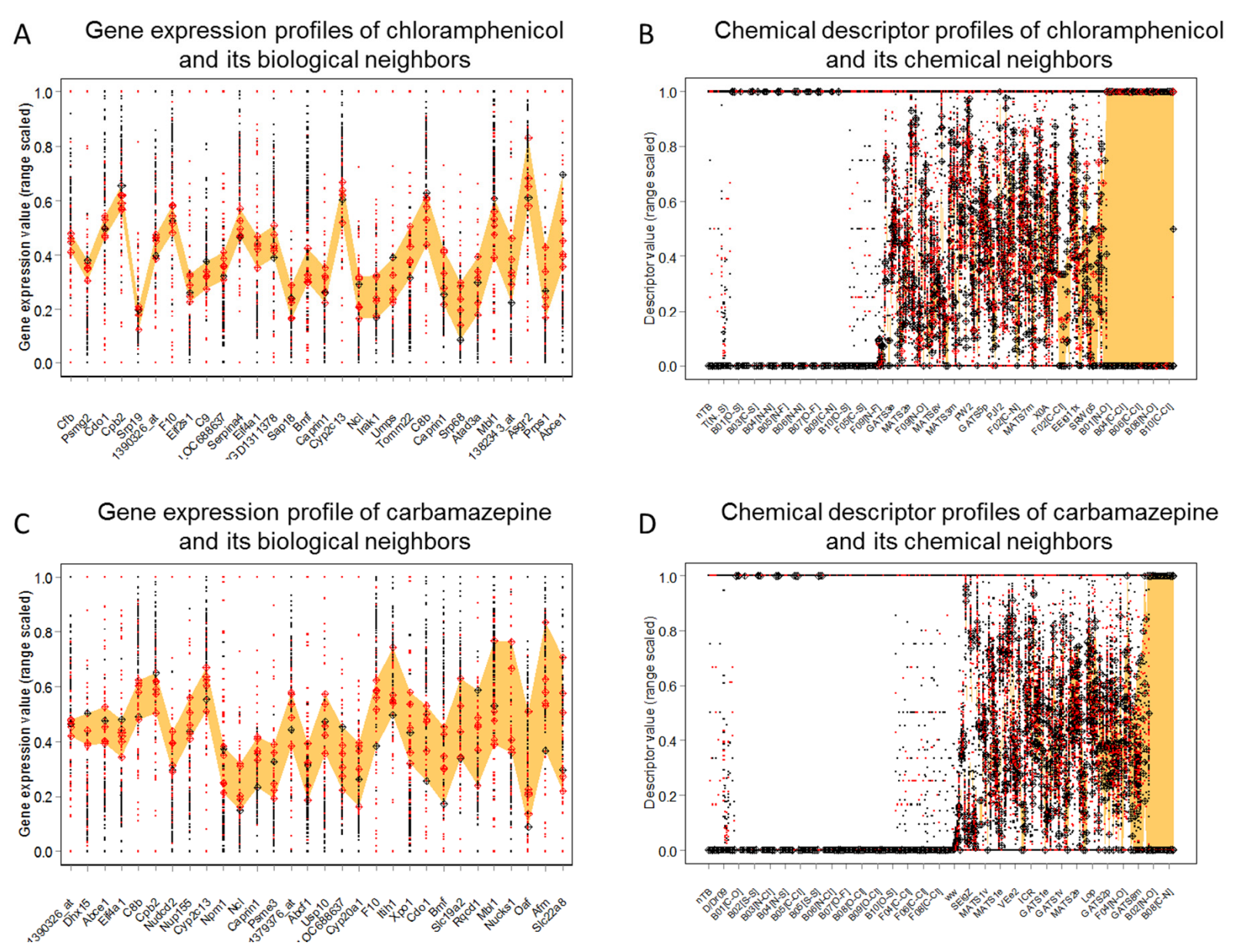

Figure S1. Descriptor profiles of chloramphenicol and (A) its biological neighbors across 30 genes, (B), its chemical neighbors across 304 chemical descriptors, and carbamazepine and (C) its biological neighbors across 30 genes, and (D) its chemical neighbors across 304 chemical descriptors. Each diamond marks a descriptor value of the target compound or its neighbors used for RA-*k*NN (red for toxic, black for nontoxic). For each descriptor, these values within the neighborhood form a range (orange ribbon). Descriptors are ranked in increasing range (i.e. orange bands widen along the x-axis). Target compounds with more similar neighbors are expected to exhibit narrower orange bands (e.g. panel A shows the narrow band formed by chloramphenicol and its highly similar biological neighbors). Smaller dots mark descriptor values of all other compounds and show their distribution along each descriptor dimension. Likewise, they are colored according to toxicity (red for toxic, black for nontoxic). All 304 chemical descriptors are shown although only every other 10th descriptor is labeled on the horizontal axis.