Data Dictionary_Vliet et al. 2021

Tables & Figures

Figure 3. SeqAPASS level 1 analysis of transthyretin (TTR) conservation across taxonomic groups with available sequence information relative to the human protein.

Figure 4. SeqAPASS level 2 analysis of the transthyretin receptor (TTR) ligand binding domain (LBD) conservation across taxonomic groups with available sequence information relative to the human protein LBD.

Figure 5(A). SeqAPASS level 3 analysis of the conservation of amino acid residues important for TTR-chemical binding. Select species predicted to not be similarly susceptible relative to the human transthyretin TTR protein, demonstrating full, partial, and non-matching amino acids.

Figure 5(B). Level 3 summary table displaying the number of species with available sequence data across all taxonomic groups assessed in SeqAPASS.

Figure 6 (A). Mean effect concentrations across taxonomic groups with data available in the ECOTOXicology Knowledgebase for select chemicals known to bind to the human transthyretin (TTR) protein. Parentheses along x-axis indicate the number of query hits for which data was aggregated. Asterisks indicate pairs of significantly different effect concentrations between species groups within the context of a single chemical (Dunn test, $p<0.05$), where higher numbers of asterisks indicate stronger levels of significance (*: $p<0.05$; **: $p<0.01$; ***: $p<0.001$; ****: $p<1.0e-4$). Center lines within each box represent the median, with box edges demonstrating the interquartile range (IQR). Whiskers extend up to 1.5 times the IQR. Outliers falling outside of that range are shown as individual points.

Figure 6 (B). Overlap in the number of species included in each SeqAPASS analysis with the species for which ECOTOX data was available. Duplicates removed prior to analysis.

Figure 7. SeqAPASS level 1 analysis of mu-opioid receptor (MOR) conservation across taxonomic groups with available sequence information relative to the human protein.

Figure 8. SeqAPASS level 2 analysis of the mu-opioid receptor (µOR) ligand binding domain (LBD) conservation relative to the LBD in the human protein.

Table S1: Species with available ECOTOX data for the four chemicals of interest, aligned with SeqAPASS predictions of similar Susceptibility across level 1, 2 and 3. All SeqAPASS predictions relative to the human transthyretin sequence

Figure S1. Total number of species with available data across SeqAPASS and ECOTOX Evaluations

Figure S2. Level 3 summary table displaying conservation of critical amino acids important for µOR-ligand interactions across taxonomic groups for all species with available data. All species with available data are predicted to be similarly susceptible relative to the human receptor.

Supplemental Data

Tab 1   Transthyretin (hTTR) SeqAPASS Results- Level 1

- Column A- SeqAPASS Data Version
- Column B- Protein NCBI Accession Number, link to NCBI protein database
- Column C- Protein Count, the number of protein records per species in the NCBI protein database
- Column D- NCBI Species Taxonomic ID, link to NCBI Taxonomy Database
- Column E- Taxonomic Group to which species belongs
- Column F- Filtered Taxonomic Group, user-defined taxonomic group to which species belongs
- Column G- Species scientific name, link to NCBI Taxonomy Database
- Column H- Species Common name
- Column I- Protein name, link to NCBI protein database record
- Column J- BlastP Bitscore
- Column K- Ortholog candidate
- Column L- Ortholog Count
- Column M- Susceptibility Cutoff
- Column N- Species percent Similarity
- Column O- Susceptibility prediction
- Column P- Date and time analysis was completed
- Column Q- Eukaryote
- Column R- Link to species in Ecotoxicology Knowledgebase

Tab 2   Transthyretin (hTTR) SeqAPASS Results- Level 2

- Column A- SeqAPASS Data Version
- Column B- Protein NCBI Accession Number, link to NCBI protein database
- Column C- Protein Count, the number of protein records per species in the NCBI protein database
- Column D- NCBI Species Taxonomic ID, link to NCBI Taxonomy Database
- Column E- Taxonomic Group to which species belongs
- Column F- Filtered Taxonomic Group, user-defined taxonomic group to which species belongs
- Column G- Species scientific name, link to NCBI Taxonomy Database
- Column H- Species Common name
- Column I- Protein name, link to NCBI protein database record
- Column J- BlastP Bitscore
- Column K- Ortholog candidate
- Column L- Ortholog Count
- Column M- Susceptibility Cutoff
- Column N- Species percent Similarity
- Column O- Susceptibility prediction
- Column P- Date and time analysis was completed
- Column Q- Eukaryote
- Column R- Link to species in Ecotoxicology Knowledgebase

Tab 3   Transthyretin (hTTR) SeqAPASS Results- Level 3

- Column A- SeqAPASS Data Version
- Column B- Protein NCBI Accession Number, link to NCBI protein database
- Column C- Protein Count, the number of protein records per species in the NCBI protein database
- Column D- NCBI Species Taxonomic ID, link to NCBI Taxonomy Database
- Column E- Taxonomic Group to which species belongs
- Column F- Species scientific name, link to NCBI Taxonomy Database
- Column G- Species Common name
- Column H- Protein name, link to NCBI protein database record
- Column I- Date and time analysis was completed
- Column J- Susceptibility prediction
- Column K- First critical amino acid, position
- Column L- First critical amino acid, one letter abbreviation
- Column M- First critical amino acid, total match status
- Column N- Second critical amino acid, position
- Column O- Second critical amino acid, one letter abbreviation
- Column P- Second critical amino acid, total match status
- Column Q- Third critical amino acid, position
- Column R- Third critical amino acid, one letter abbreviation
- Column S- Third critical amino acid, total match status
- Column T- Fourth critical amino acid, position
- Column U- Fourth critical amino acid, one letter abbreviation
- Column V- Fourth critical amino acid, total match status
- Column W- Fifth critical amino acid, position
- Column X- Fifth critical amino acid, one letter abbreviation
- Column Y- Fifth critical amino acid, total match status
- Column Z- Sixth critical amino acid, position
- Column AA- Sixth critical amino acid, one letter abbreviation
- Column AB- Sixth critical amino acid, total match status

Tab 4 Ecotoxicology Knowledgebase (EcoTox) data for Diethylstilbestrol (DES)

- Column A- CAS Number
- Column B- Chemical Name
- Column C- Chemical Grade, if applicable
- Column D- Chemical Analysis, if applicable
- Column E- Chemical Purity Mean, if applicable
- Column F- Chemical Purity Min, if applicable
- Column G- Chemical Purity Max, if applicable
- Column H- Species Scientific Name
- Column I- Species Common Name
- Column J- Species Group
- Column K- Organism Life stage
- Column L- Organism Age Mean, if applicable
- Column M- Organism Age Min, if applicable
- Column N- Organism Age Max, if applicable
- Column O- Age Units
- Column P- Exposure Type
- Column Q- Media Type
- Column R- Test Location
- Column S- Conc Mean
- Column T- Number of Doses
- Column U- Conc 1 Type
- Column V- Conc 1 Min, if applicable
- Column W- Conc 1 Max, if applicable
- Column X- Conc Units
- Column Y- Effect Type
- Column Z- Effect Measurement
- Column AA- Endpoint
- Column AB- Response Site
- Column AC- Observed Duration Mean
- Column AD- Observed Duration Min, if applicable
- Column AE- Observed Duration Max, if applicable
- Column AF- Observed Duration Units
- Column AG- Author
- Column AH- Reference Number, ECOTOX ID
- Column AI- Reference Title
- Column AJ- Reference Journal
- Column AK- Reference Publication Year

Tab 5   Ecotoxicology Knowledgebase (EcoTox) data for Perfluorooctane sulfonic acid (PFOS)

- Column A- CAS Number
- Column B- Chemical Name
- Column C- Chemical Grade, if applicable
- Column D- Chemical Analysis, if applicable
- Column E- Chemical Purity Mean, if applicable
- Column F- Chemical Purity Min, if applicable
- Column G- Chemical Purity Max, if applicable
- Column H- Species Scientific Name
- Column I- Species Common Name
- Column J- Species Group
- Column K- Organism Life stage
- Column L- Organism Age Mean, if applicable
- Column M- Organism Age Min, if applicable
- Column N- Organism Age Max, if applicable
- Column O- Age Units
- Column P- Exposure Type
- Column Q- Media Type
- Column R- Test Location
- Column S- Conc Mean
- Column T- Number of Doses
- Column U- Conc 1 Type
- Column V- Conc 1 Min, if applicable
- Column W- Conc 1 Max, if applicable
- Column X- Conc Units
- Column Y- Effect Type
- Column Z- Effect Measurement
- Column AA- Endpoint
- Column AB- Response Site
- Column AC- Observed Duration Mean
- Column AD- Observed Duration Min, if applicable
- Column AE- Observed Duration Max, if applicable
- Column AF- Observed Duration Units
- Column AG- Author
- Column AH- Reference Number, ECOTOX ID
- Column AI- Reference Title
- Column AJ- Reference Journal
- Column AK- Reference Publication Year

Tab 6   Ecotoxicology Knowledgebase (EcoTox) data for Perfluorohexanoic acid (PFHxA)

- Column A- CAS Number
- Column B- Chemical Name
- Column C- Chemical Grade, if applicable
- Column D- Chemical Analysis, if applicable
- Column E- Chemical Purity Mean, if applicable
- Column F- Chemical Purity Min, if applicable
- Column G- Chemical Purity Max, if applicable
- Column H- Species Scientific Name
- Column I- Species Common Name
- Column J- Species Group
- Column K- Organism Life stage
- Column L- Organism Age Mean, if applicable
- Column M- Organism Age Min, if applicable
- Column N- Organism Age Max, if applicable
- Column O- Age Units
- Column P- Exposure Type
- Column Q- Media Type
- Column R- Test Location
- Column S- Conc Mean
- Column T- Number of Doses
- Column U- Conc 1 Type
- Column V- Conc 1 Min, if applicable
- Column W- Conc 1 Max, if applicable
- Column X- Conc Units
- Column Y- Effect Type
- Column Z- Effect Measurement
- Column AA- Endpoint
- Column AB- Response Site
- Column AC- Observed Duration Mean
- Column AD- Observed Duration Min, if applicable
- Column AE- Observed Duration Max, if applicable
- Column AF- Observed Duration Units
- Column AG- Author
- Column AH- Reference Number, ECOTOX ID
- Column AI- Reference Title
- Column AJ- Reference Journal
- Column AK- Reference Publication Year

Tab 7   Ecotoxicology Knowledgebase (EcoTox) data for Tetrabromobisphenol A (TBBPA)

- Column A- CAS Number
- Column B- Chemical Name
- Column C- Chemical Grade, if applicable
- Column D- Chemical Analysis, if applicable
- Column E- Chemical Purity Mean, if applicable
- Column F- Chemical Purity Min, if applicable
- Column G- Chemical Purity Max, if applicable
- Column H- Species Scientific Name
- Column I- Species Common Name
- Column J- Species Group
- Column K- Organism Life stage
- Column L- Organism Age Mean, if applicable
- Column M- Organism Age Min, if applicable
- Column N- Organism Age Max, if applicable
- Column O- Age Units
- Column P- Exposure Type
- Column Q- Media Type
- Column R- Test Location
- Column S- Conc Mean
- Column T- Number of Doses
- Column U- Conc 1 Type
- Column V- Conc 1 Min, if applicable
- Column W- Conc 1 Max, if applicable
- Column X- Conc Units
- Column Y- Effect Type
- Column Z- Effect Measurement
- Column AA- Endpoint
- Column AB- Response Site
- Column AC- Observed Duration Mean
- Column AD- Observed Duration Min, if applicable
- Column AE- Observed Duration Max, if applicable
- Column AF- Observed Duration Units
- Column AG- Author
- Column AH- Reference Number, ECOTOX ID
- Column AI- Reference Title
- Column AJ- Reference Journal
- Column AK- Reference Publication Year

Tab 8   EcoTox Group Mean Calculations

- Column A- Chemical
- Column B- Species Group
- Group Mean

Tab 9   SeqAPASS-EcoTox Data Comparisons

- Column A- Taxonomic Group
- Column B- Species Common Name, from ECOTOX Knowledgebase
- Column C- Data for DES present in ECOTOX Knowledgebase?
- Column D- Data for PFHxA present in ECOTOX Knowledgebase?
- Column E- Data for PFOS present in ECOTOX Knowledgebase?
- Column F- Data for TBBPA present in ECOTOX Knowledgebase?
- Column G- SeqAPASS Level 1 prediction
- Column H- SeqAPASS Level 2 prediction
- Column I- SeqAPASS Level 3 prediction

Tab 10   μ-Opiod Receptor (hMOR) SeqAPASS Results- Level 1

- Column A- SeqAPASS Data Version
- Column B- Protein NCBI Accession Number, link to NCBI protein database
- Column C- Protein Count, the number of protein records per species in the NCBI protein database
- Column D- NCBI Species Taxonomic ID, link to NCBI Taxonomy Database
- Column E- Taxonomic Group to which species belongs
- Column F- Filtered Taxonomic Group, user-defined taxonomic group to which species belongs
- Column G- Species scientific name, link to NCBI Taxonomy Database
- Column H- Species Common name
- Column I- Protein name, link to NCBI protein database record
- Column J- BlastP Bitscore
- Column K- Ortholog candidate
- Column L- Ortholog Count
- Column M- Susceptibility Cutoff
- Column N- Species percent similarity
- Column O- Susceptibility prediction
- Column P- Date and time analysis was completed
- Column Q- Eukaryote
- Column R- Link to species in Ecotoxicology Knowledgebase

Tab 11 μ-Opiod Receptor (hMOR) SeqAPASS Results- Level 2

- Column A- SeqAPASS Data Version
- Column B- Protein NCBI Accession Number, link to NCBI protein database
- Column C- Protein Count, the number of protein records per species in the NCBI protein database
- Column D- NCBI Species Taxonomic ID, link to NCBI Taxonomy Database
- Column E- Taxonomic Group to which species belongs
- Column F- Filtered Taxonomic Group, user-defined taxonomic group to which species belongs
- Column G- Species scientific name, link to NCBI Taxonomy Database
- Column H- Species Common name
- Column I- Protein name, link to NCBI protein database record
- Column J- BlastP Bitscore
- Column K- Ortholog candidate
- Column L- Ortholog Count
- Column M- Susceptibility Cutoff
- Column N- Species percent similarity
- Column O- Susceptibility prediction
- Column P- Date and time analysis was completed
- Column Q- Eukaryote
- Column R- Link to species in Ecotoxicology Knowledgebase

Tab 12 μ-Opiod Receptor (hMOR) SeqAPASS Results- Level 3

- Column A- SeqAPASS Data Version
- Column B- Protein NCBI Accession Number, link to NCBI protein database
- Column C- Protein Count, the number of protein records per species in the NCBI protein database
- Column D- NCBI Species Taxonomic ID, link to NCBI Taxonomy Database
- Column E- Taxonomic Group to which species belongs
- Column F- Species scientific name, link to NCBI Taxonomy Database
- Column G- Species Common name
- Column H- Protein name, link to NCBI protein database record
- Column I- Date and time analysis was completed
- Column J- Susceptibility prediction
- Column K- First critical amino acid, position
- Column L- First critical amino acid, one letter abbreviation
- Column M- First critical amino acid, total match status
- Column N- Second critical amino acid, position
- Column O- Second critical amino acid, one letter abbreviation
- Column P- Second critical amino acid, total match status
- Column Q- Third critical amino acid, position
- Column R- Third critical amino acid, one letter abbreviation
- Column S- Third critical amino acid, total match status
- Column T- Fourth critical amino acid, position
- Column U- Fourth critical amino acid, one letter abbreviation
- Column V- Fourth critical amino acid, total match status
- Column W- Fifth critical amino acid, position
- Column X- Fifth critical amino acid, one letter abbreviation
- Column Y- Fifth critical amino acid, total match status
- Column Z- Sixth critical amino acid, position
- Column AA- Sixth critical amino acid, one letter abbreviation
- Column AB- Sixth critical amino acid, total match status
- Column AC- Seventh critical amino acid, position
- Column AD- Seventh critical amino acid, one letter abbreviation
- Column AE- Seventh critical amino acid, total match status
- Column AF- Eighth critical amino acid, position
- Column AG- Eighth critical amino acid, one letter abbreviation
- Column AH- Eighth critical amino acid, total match status
- Column AI- Ninth critical amino acid, position
- Column AJ- Ninth critical amino acid, one letter abbreviation
- Column AK- Ninth critical amino acid, total match status