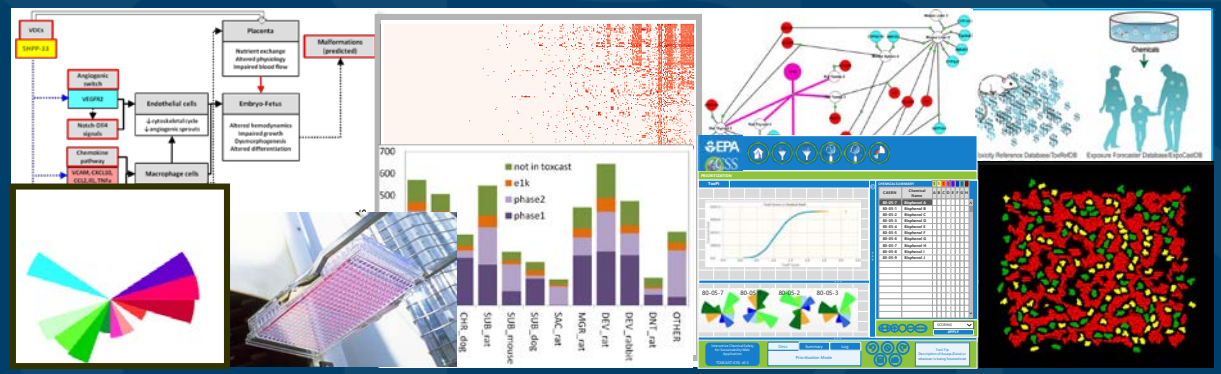
Building scientific confidence in metabolic similarity in read-across through the use of *in vitro*, *in silico* and analytical data



G Patlewicz[1], LE Lizarraga[2], EO Owens[2], J Lambert[2], SC Wesselkamper[2], QJ Zhao[2], B Hawkins[2], J Dean[2], A Williams[1], I Shah[1], KA Favela[3], A Yau[3], JA Bonzo[4], LR Moody[4], RS Thomas[1], JF Wambaugh[1]

[1]U.S. EPA, NCCT, RTP, NC, [2]U.S. EPA, NCEA, Cincinnati, OH, [3]Southwest Research Institute, San Antonio, TX, [4]Thermo Fisher Scientific, Madison, WI.

Outline

- Definitions of read-across
- Sources of uncertainty in read-across
- Pilot study
- Study workflow
- Progress to date
- Next steps

Definitions: Chemical grouping approaches

- Read-across describes one of the techniques for filling data gaps in either the analogue or category approaches i.e. not to be confused with the "analogue approach"

- "Analogue approach" refers to grouping based on a very limited number of chemicals (e.g. target substance) + source substance)

- "Category approach" is used when grouping is based on a more extensive range of analogues (e.g. 3 or more members)

- A chemical category is a group of chemicals whose physico-chemical and human heath and/or environmental toxicological and/or environmental fate properties are likely to be similar or follow a regular pattern as a result of structural similarity (or other similarity characteristics) e.g. metabolism similarity.

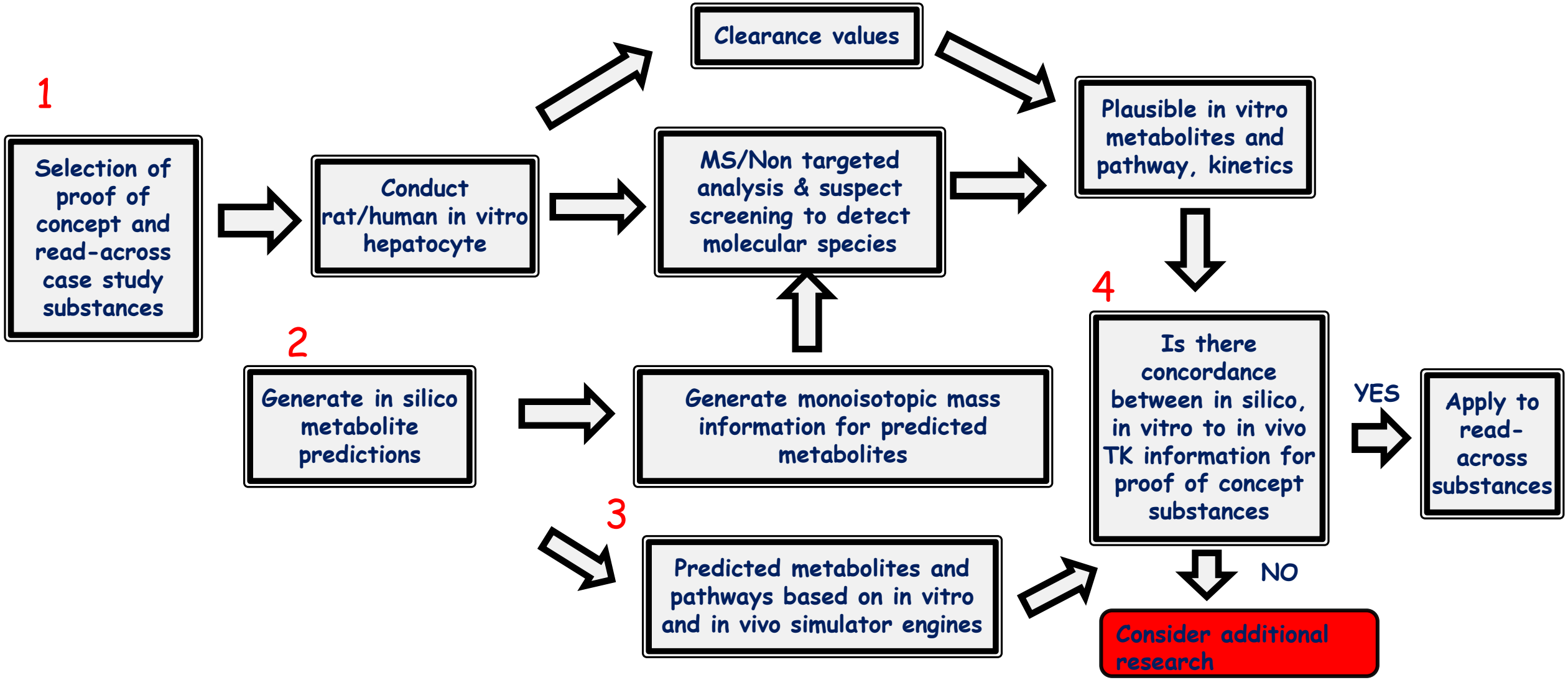
Sources of Uncertainty in Read-Across

- Analogue or category approach? (# analogues)
- Completeness of the data matrix – no. of data gaps
- Data quality for the underlying analogues for the target and source analogues
- Consistency of data across the data matrix – concordance of effects and potency across analogues
- Toxicokinetics – similarity in metabolism profile
- Overarching hypothesis/similarity rationale – how to identify similar analogues and justify their similarity for the endpoint of interest
- Address the dissimilarities and whether these are significant from a toxicological standpoint
- Presence vs. absence of toxicity

Pilot Study

- *In vivo* toxicokinetics information if available is often relied upon to substantiate biological similarity for the purposes of justifying a read-across
- This information is usually extracted from the literature
- In this pilot study we sought to investigate the feasibility and utility of using *in vitro* toxicokinetics data to substantiate biological similarity.

Pilot Study Workflow



Step 1: Target/Analogue identification

Chemicals	Category
Methyleugenol (CASRN 93-15-2), Estragole (CASRN 140-67-0)	**Proof-of-concept: Known similar metabolism; methyleugenol showed metabolic clearance in the Wetmore et., (2015) studies**
2-nitrotoluene (CASRN 88-72-2), 4-nitrotoluene (CASRN 99-99-0)	Proof-of-concept: Known different metabolism; 4-nitrotoluene showed metabolic clearance in Wetmore et al., (2015) studies
Target: 4-Methyl-2-Pentanol (CASRN 108-11-2) Analogues: 4-methyl-2-pentanone (CASRN 108-10-1), 2-propanol (CASRN 67-63-0), 2-propanone (CASRN 67-64-1)	Application to Read-across: Metabolism considerations form the basis for analogue identification and selection of most appropriate surrogate chemical
Target: 3,5-Dinitroaniline (CASRN 618-87-1) Analogues: 2-Nitroaniline (CASRN 88-74-4), 3-Nitroaniline (CASRN 99-09-2), 4-Nitroaniline (CASRN 100-01-6)	Application to Read-across: no information on target compound; exploring the utility of metabolism in informing analogue selection

Workflow steps

- Perform *in vitro* rat and human hepatocyte study to determine intrinsic clearance
- Apply analytical spectroscopy (MS) for the detection of molecular species and non-targeted analysis for metabolite identification
- Use third party expert systems for the prediction of potential metabolites and their pathways to facilitate MS analysis
- Evaluate concordance of *in vitro* metabolism data relative to existing *in vivo* data
- Evaluate concordance of *in silico* metabolism to both *in vitro* metabolism and *in vivo* metabolism data for the proof of concept substances
- Use the predicted and experimental metabolism data to determine which source analogue(s) are valid for read-across

Step 2: *Generate in silico* metabolism predictions

- There are a handful of metabolism prediction tools.
- Examples include: MetaPrint 2D, Meteor Nexus, TIMES and the simulators contained within the OECD Toolbox.
- Some are freely available, others are commercial.
- In this pilot study we selected Meteor Nexus, TIMES and the OECD Toolbox.

TIMES = TIssue MEtabolism System

- Commercial hybrid expert system
- Unique platform to facilitate toxicity predictions whilst accounting for metabolism
- Endpoints that have been modelled include skin sensitisation, Ames mutagenicity, *in vitro* chromosomal aberration, *in vivo* micronucleus
- The metabolism and autoxidation simulators can be used *in vacuo* though they have been "trained" to reproduce the metabolic maps and their associated metabolites for these endpoints
- Of specific interest are the *in vitro* and *in vivo* rat liver metabolism simulators

TIMES – *in vitro*

- Rat liver S9 v11.15 (technically rodent rather than rat)
- Metabolism training set contains experimentally observed (documented) *in vitro* metabolic pathways for 261 parent chemicals of wide structural diversity, and 978 observed metabolites compiled into a searchable electronic database.
- Published data on the metabolism of these chemicals in rodent liver microsomes and S9 fraction were mainly collected from the literature
- Current *in vitro* rat liver metabolic simulator (transformation table) represents electronically designed set of 517 structurally generalised, hierarchically arranged biotransformation reactions.

TIMES – *in vitro*

- The following types of molecular transformations are included in the *in vitro* simulator:
  - 25 - 30 abiotic (non-enzymatic) and, also, a few enzyme-controlled reactions believed to occur at a very high rate as compared to the duration of the tests.
  - 450 – 470 enzymatic phase I (mostly CYP450-catalysed) transformations such as aliphatic C-oxidation, aromatic C-hydroxylation, oxidative N- and O-dealkylation etc.
  - 15 – 20 enzymatic phase II transformations, such as glucuronidation, sulfation, glutathione conjugation, N-acetylation, etc.

TIMES – *in vivo*

- Rat v07.11
- Metabolism training set contains experimentally observed (documented) *in vivo* metabolic pathways for 647 structurally different parent chemicals, and 4382 observed metabolites.
- The current *in vivo* rat metabolic simulator (transformation table) represents electronically designed set of 622 structurally generalised, hierarchically arranged biotransformation reactions.
- The following types of molecular transformations are included:
  - 26 abiotic (non-enzymatic) reactions.
  - 479 enzymatic phase I transformations
  - 104 enzymatic phase II transformations

OECD Toolbox

- Contains *in vitro* and *in vivo* rat liver simulators which are taken from the commercial TIMES system
- Provides a list of metabolites but not their hierarchy or the associated metabolic tree

Meteor Nexus

- Developed by Lhasa Ltd
- Several options - default is a site of metabolism scoring with molecular mass variance

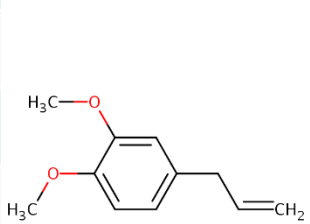
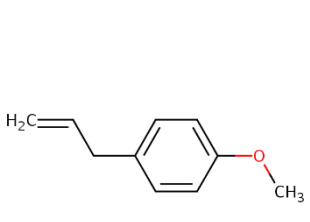
Site of Metabolism Scoring (with Molecular Mass Variance) - default method

Sites of metabolism are identified for the query structure. A Site of Metabolism (SoM) is a sub-set of the atoms of a structure where a biotransformation takes place. For every example compound in the model, the SoM is retrieved. The examples' Sites of Metabolism are sorted in order of their similarity to the calculated SoM of the query structure using the Tanimoto Coefficient. The most similar SoMs are used as the supporting examples that determine the score.

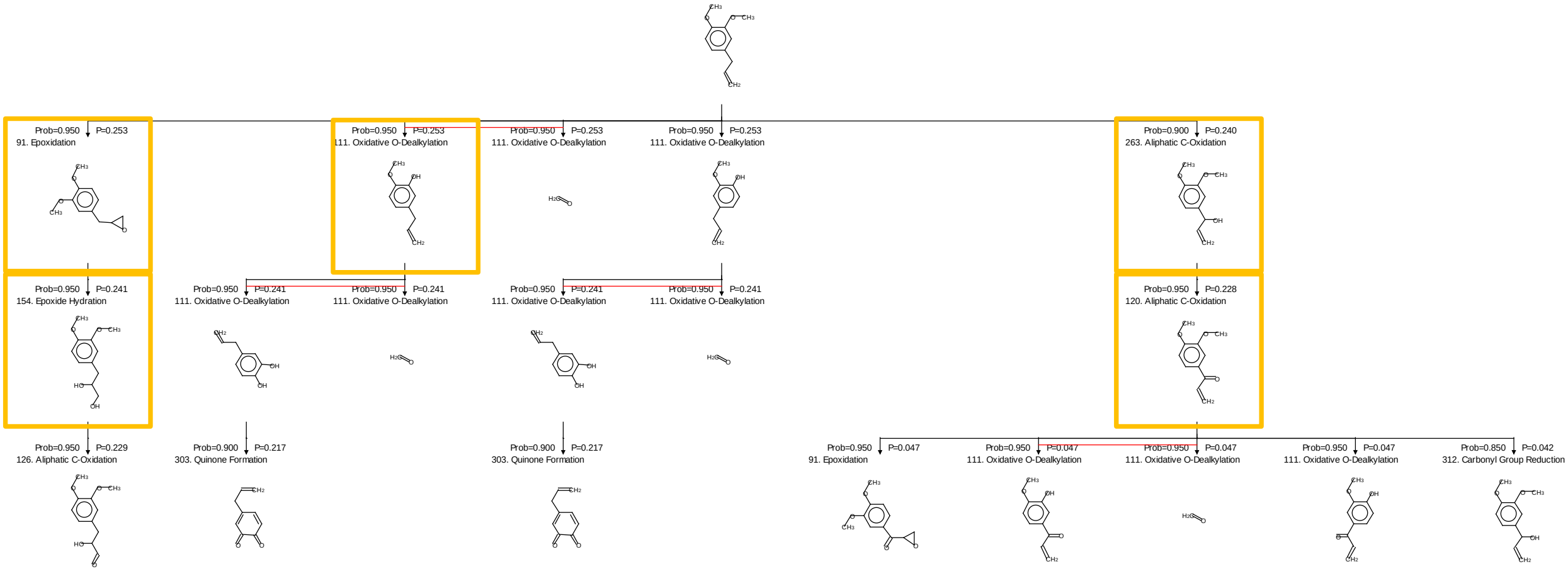
Processing Options for the Site of Metabolism Scoring (with Molecular Mass Variance) method are:

- **Molecular Mass Similarity Threshold** - A threshold by which the molecular weight of the query compound is used to decide which examples are considered. The molecular weight of the example compounds, as a percentage of the query compound's molecular weight, must be equal to or above this value. The default is 70.0.
- **Number of Nearest Neighbours** - Specifies the number of examples shown in the results. As examples also determine the score for a metabolite, this setting can also strongly affect what metabolites are actually predicted. You can set this value to between 6 and 12 examples, but the default has been set to 8 following investigations. We advise you to leave this at 8 if possible.
- **Keep Nearest Neighbours with Equal Similarities** - When ticked, prevents Meteor from discarding supporting examples where there is a tie in similarity rank at the cut-off point, even if this means that the number of examples kept exceeds the Number of Nearest Neighbours specified (see ▼ below for an example - using the Top N filter).
- **Scoring Filters** - Records how predictive a structure is when evaluated against the biotransformations in Meteor, and orders them according to their occurrence ratios. The options are:
  - `Relative` - Only displays metabolites with scores at or above the percentage value set in the **Score Threshold** field (between 0% and 100% with a default of 70). For example if your highest score is 449 and the threshold is set to 50% then anything with a score of 224.5 will be displayed.

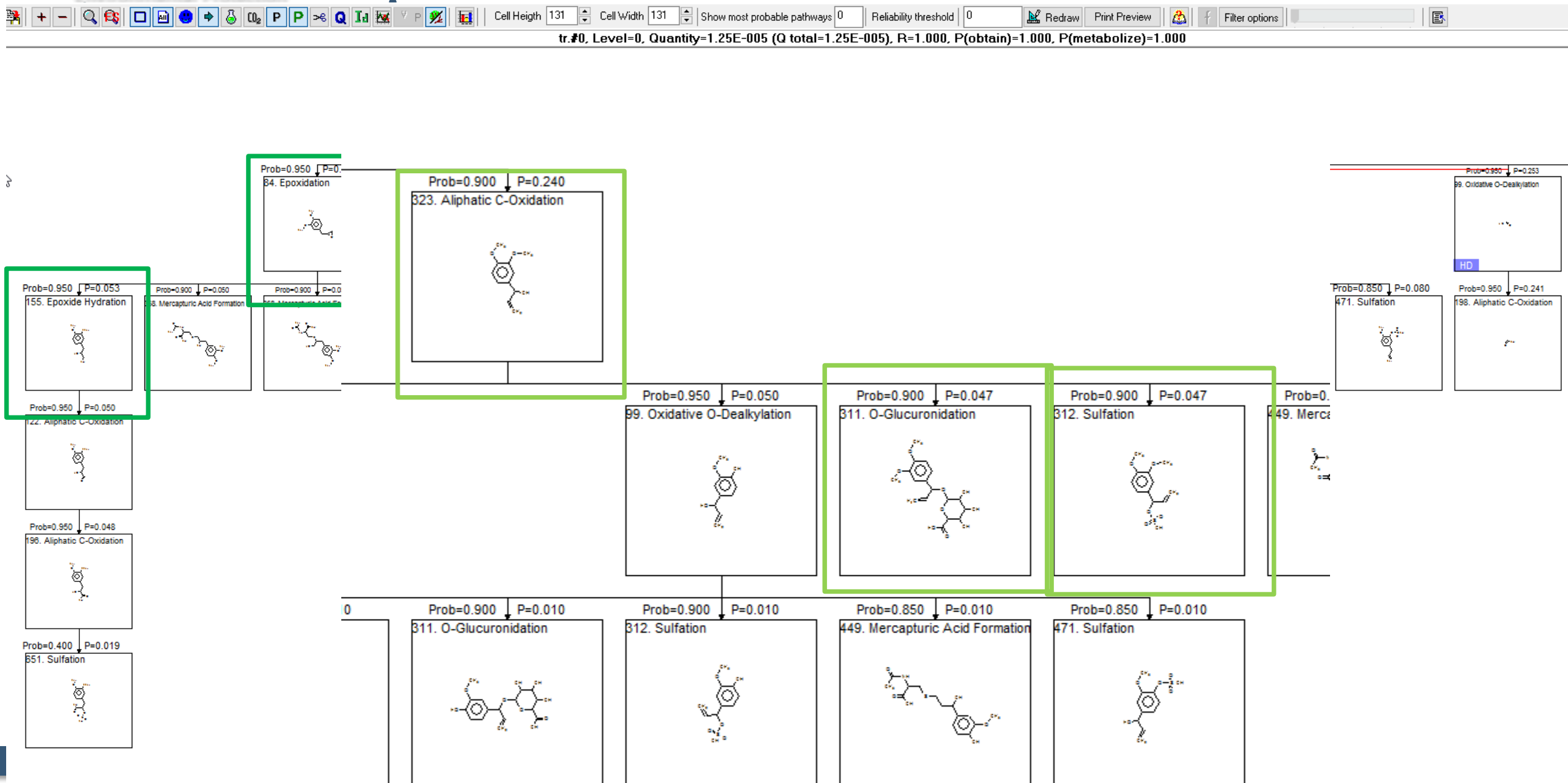
Step 3: Derive *in silico* metabolism pathways

	Methyl eugenol	Estragole
Role	Target	Source
CAS	[93-15-2]	[140-67-0]
DTXSID	DTXSID5025607	DTXSID0020575
Structure		
Average Mass	178.23	148.21
LogKow (predicted)	2.61	3.02
Boiling pt (predicted)	252.16 deg C	219.74 deg C
Melting pt (predicted)	30.59 deg C	-7.93 deg C
Reactivity	Potential to be activated to a Michael acceptor	Potential to be activated to a Michael acceptor

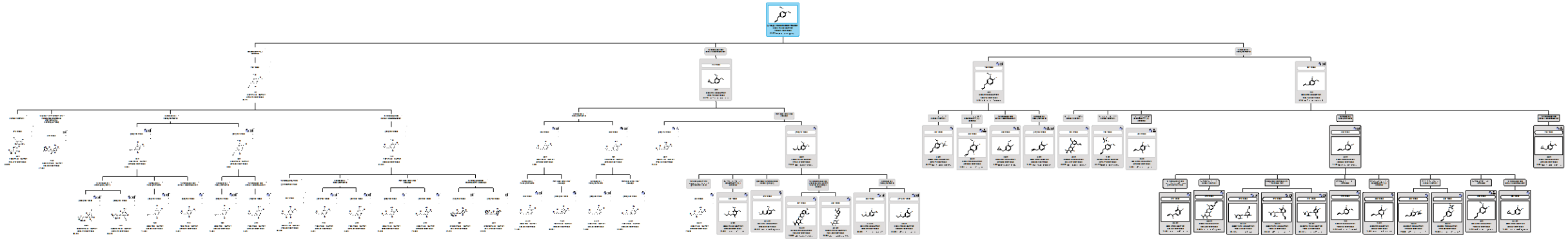
Step 3: *In vitro* TIMES metabolic map of methyl eugenol



Step 3: *In vivo* TIMES metabolic map of methyl



Step 3: Meteor metabolic map of methyl eugenol



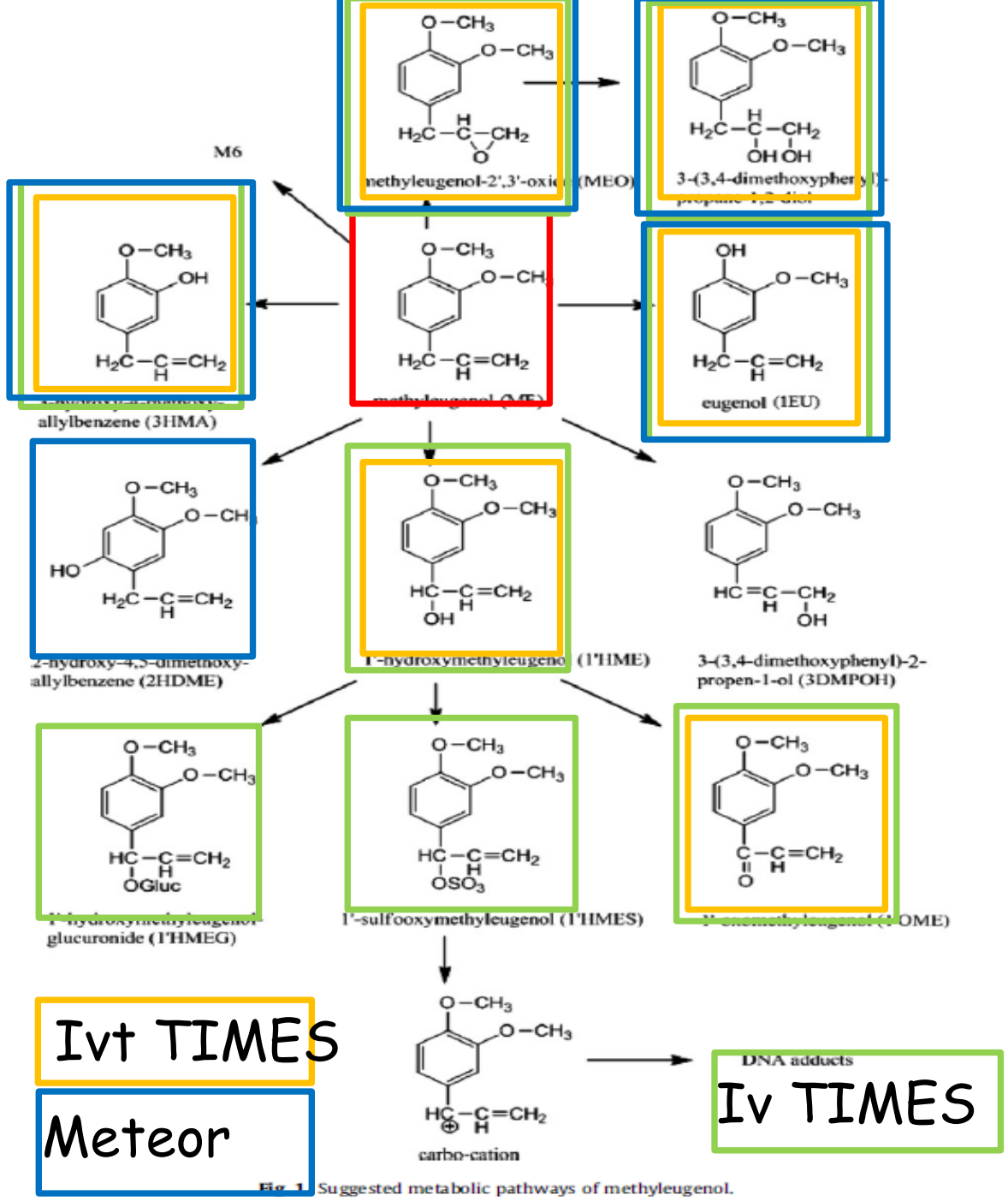


Fig. 1 Suggested metabolic pathways of methyleugenol.

Step 4: Concordance between *in silico* and *in vivo* data for target Methyl Eugenol

Step 4: Concordance between *in silico* and *in vivo* data for source analogue Estragole

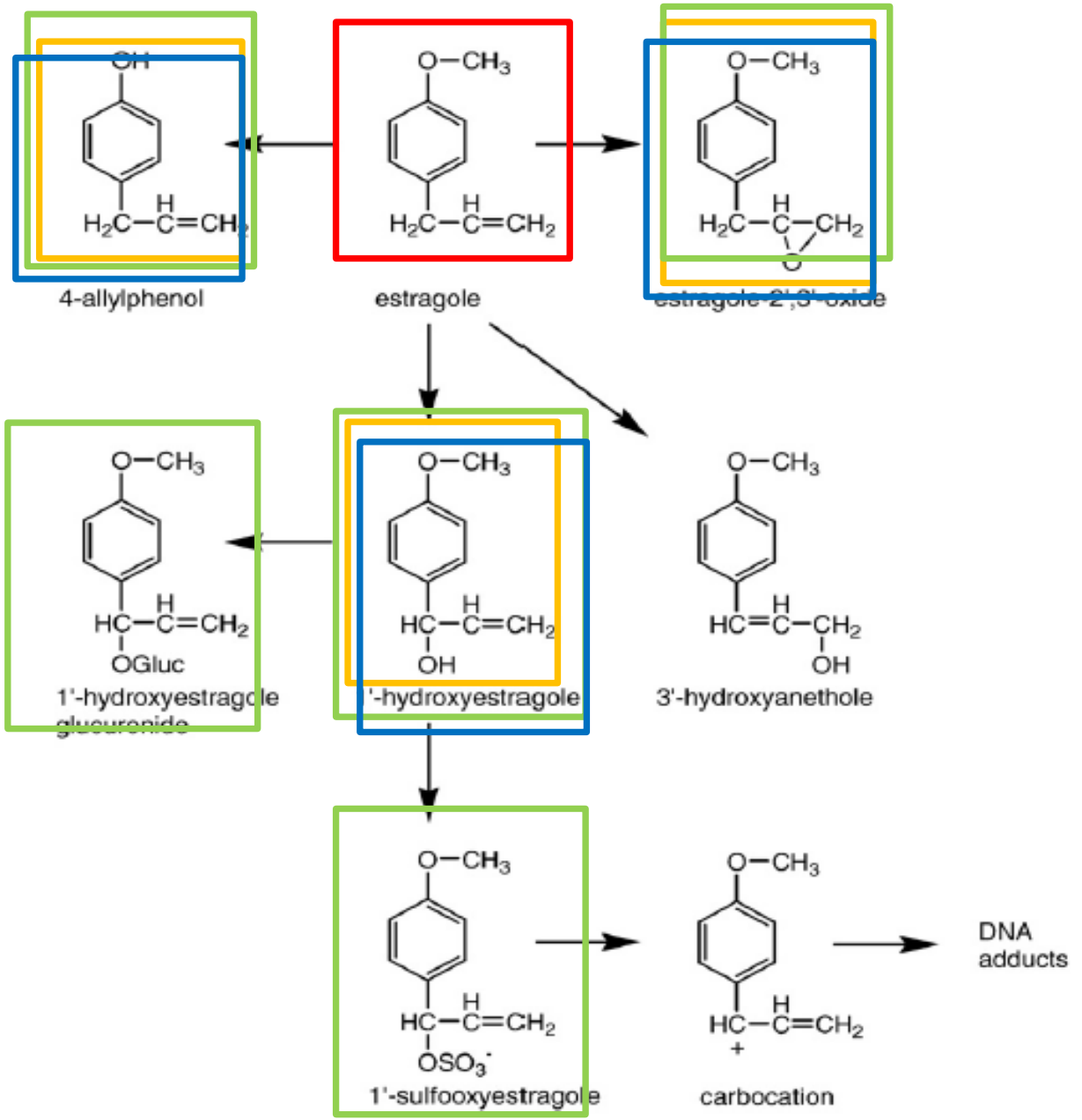


Fig. 1. Metabolism of estragole (adapted from Smith et al., 2002).

Step 4: Concordance between *in silico* and *in vivo* data

- The OECD toolbox metabolites presented a subset of those identified by TIMES (results not shown).
- *In vitro* TIMES predictions presented an incomplete picture of the metabolism pathway of target and analogue substances.
- *In vivo* TIMES metabolism simulator was able to replicate the majority of the experimental in vivo metabolites for the proof of concept substances. Data only shown for methyl eugenol/estragole pair but similar findings found for the nitrotoluenes pair.

Step 4: Concordance between *in silico* and *in vivo* data

- These pathways are relevant to the genotoxicity of methyleugenol/estragole which highlights the significance and relevance of these findings for understanding the role of metabolism plays in the expected toxicity of these chemicals.

Next steps

- Compare concordance with *in vitro* experimental data to evaluate whether a combination of *in silico* & *in vitro* data best represents the *in vivo* metabolism profile of a given target/analogue.

United States Environmental Protection Agency

- Questions?
- Contact: Patlewicz.grace@epa.gov