



HAL
open science

Knowledge Graphs And Bayesian Networks Analyses To Elucidate Transformation Products Filiations During Wastewater Treatment: Application To Sulfamethoxazole (SMX)

Antonin Azaïs, Ali Badara Minta, J.M. Choubert, Rachid Ouaret, Claire Albasi

► To cite this version:

Antonin Azaïs, Ali Badara Minta, J.M. Choubert, Rachid Ouaret, Claire Albasi. Knowledge Graphs And Bayesian Networks Analyses To Elucidate Transformation Products Filiations During Wastewater Treatment: Application To Sulfamethoxazole (SMX). 12th Micropol and Ecohazard Conference, Jun 2022, Saint-Jacques-de-Compostelle, Spain. . hal-04906549

HAL Id: hal-04906549

<https://hal.science/hal-04906549v1>

Submitted on 22 Jan 2025

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.

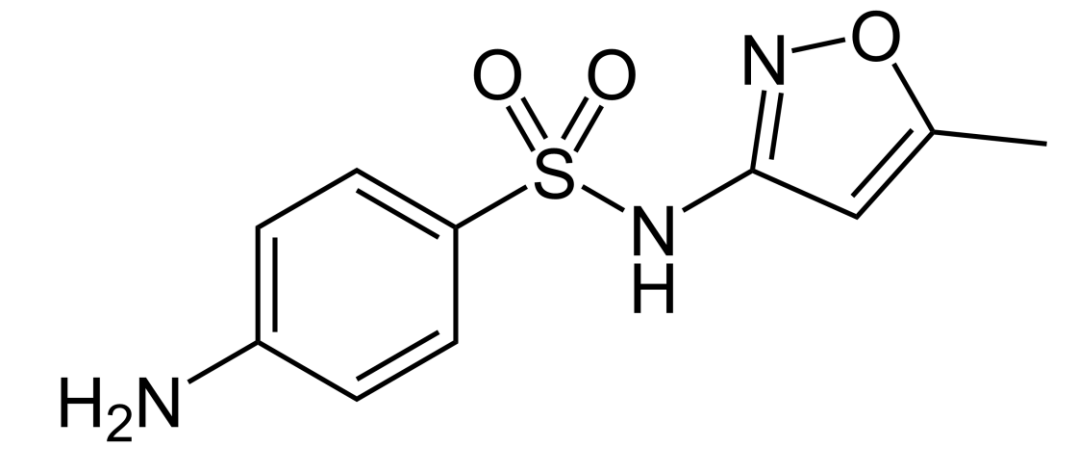
UR REVERSAAL

Antonin AZAÏS *
Ali Badara MINTA
Jean-Marc CHOUBERT

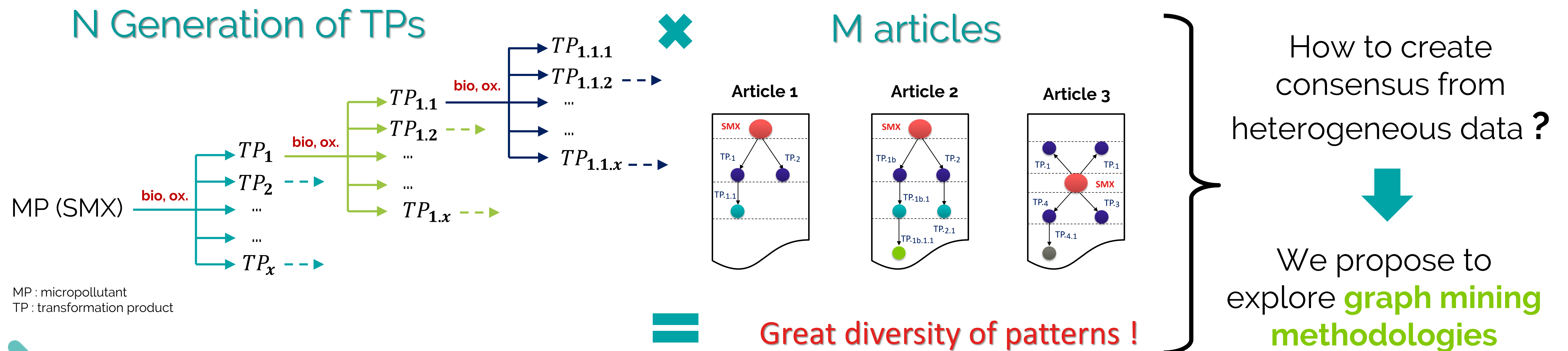
LGC TOULOUSE

Rachid OUARET
Claire ALBASI

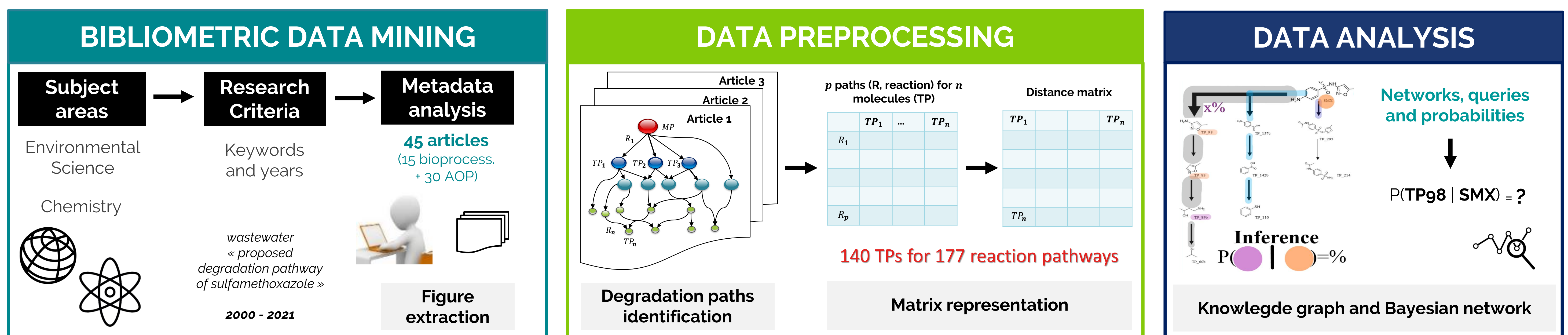
Knowledge Graphs And Bayesian Networks Analyses To Elucidate Transformation Products Filiations During Wastewater Treatment: Application To Sulfamethoxazole (SMX)



CONTEXT AND IDEA



METHODOLOGY

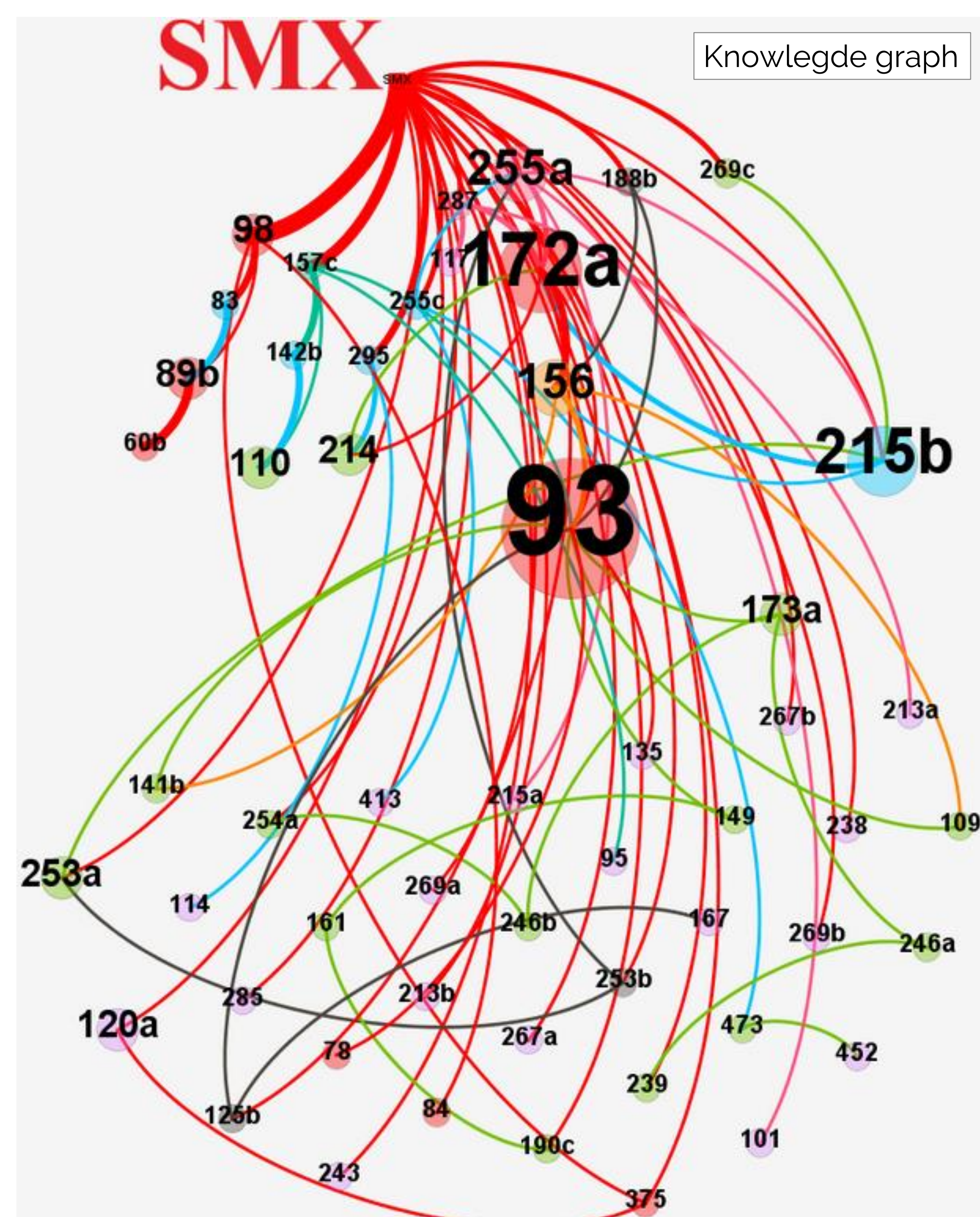


RESULTS AND PERSPECTIVES

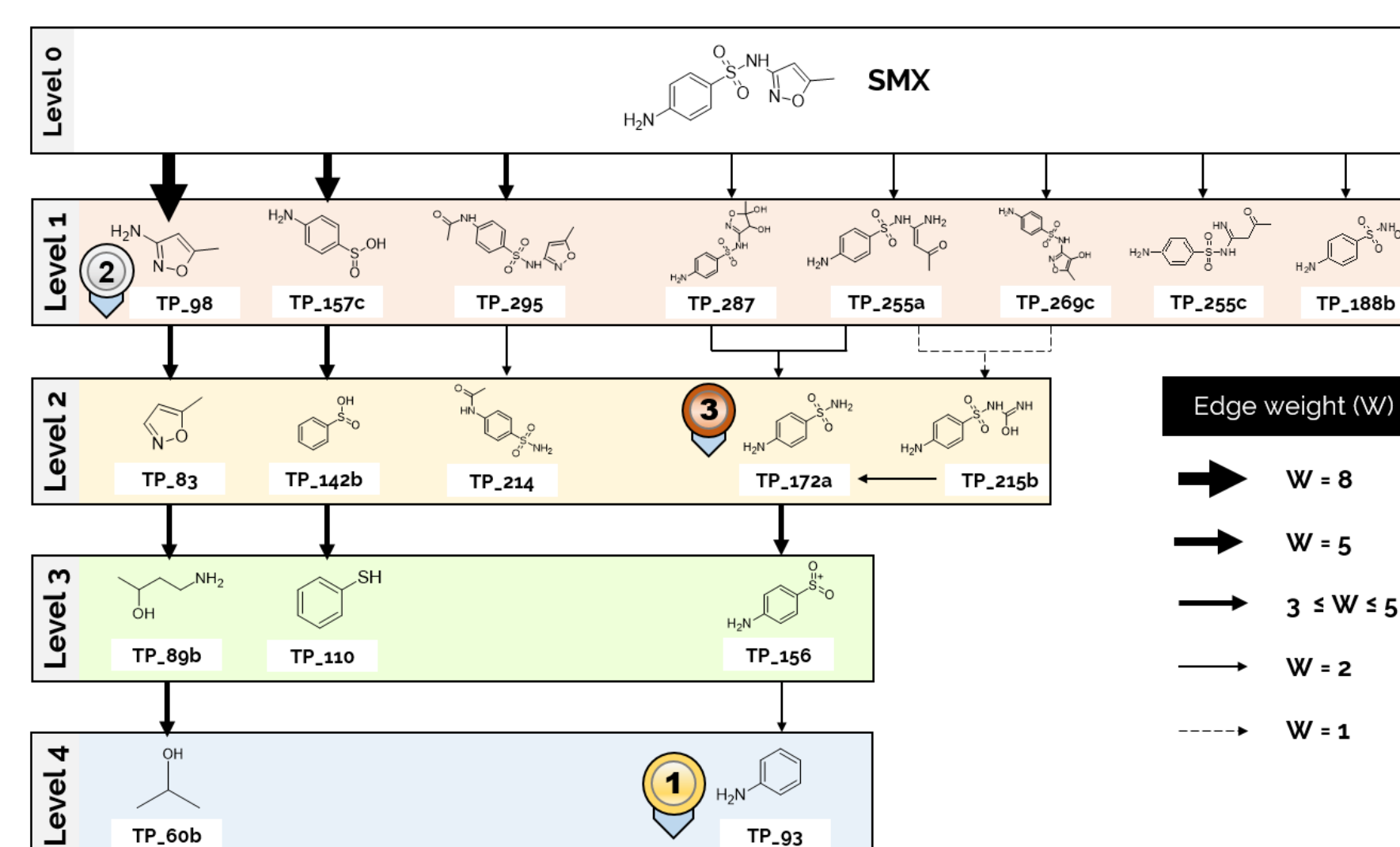
Degradation network of the SMX obtained for biological processes - based on bibliometric analysis - was plotted with Gephi software.

The size of each link (or edge) reflects the recurrence of the pathways considered, giving an indication of the **most reported reaction pathways (SMX > TP₉₈)**.

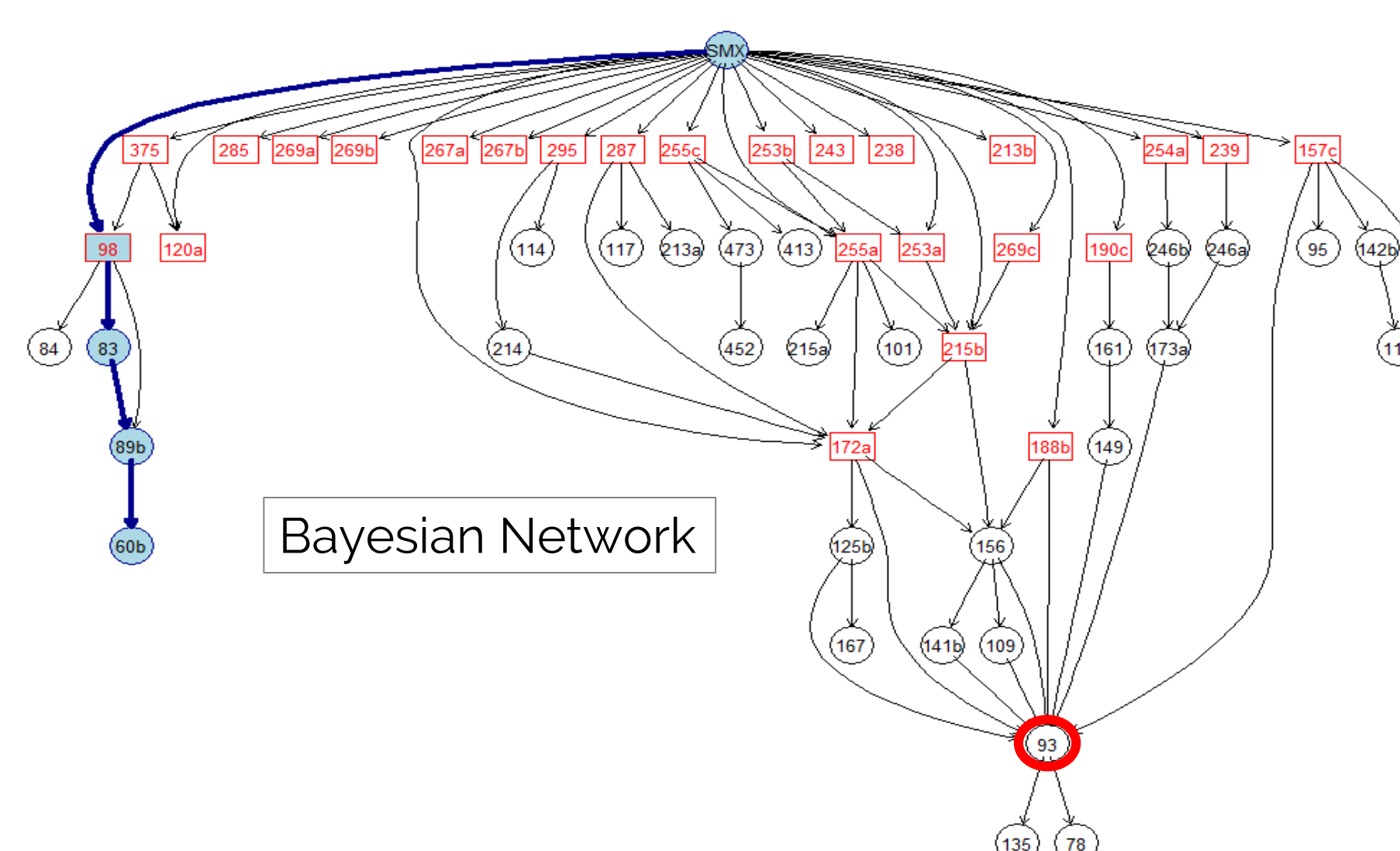
The size of the nodes represents the probability of producing a molecule from several reaction pathways (TP₉₃).



Simplified degradation scheme obtained from the knowledge graph. The three more reported TPs (TP₉₃ > TP₉₈ > TP_{172a}) are represented with medals. Edge weights are reported.



Bayesian networks (BN) were used for inference queries and to draw degradation tree after with structure initialization. Highlighted path (blue) is the most recurrent degradation way.



Queries (react. pathway)	Probability
P (TP ₉₈ SMX)	19 %
P (TP ₈₃ SMX, TP ₉₈)	37 %
P (TP _{89b} SMX, TP ₉₈)	50 %
P (TP _{60b} SMX, TP ₉₈ , TP ₈₃ , TP _{89b})	100 %

CONCLUSION

No methodology for processing and consolidating data from degradation pathways are considered in the literature.

Inspired by social networks, we proposed to explore graph methodologies and attempts to respond to this lack, to create **consensus** on the basis of studies where data are widely dispersed.

Graphical models shed new light on TP filiations, allowing, to visualize the degradation network, determine reactions probabilities and preferred reaction pathways.

To date, the part of the database enrichment is done manually by updating the articles and identifying new molecules and reaction pathways. **The most ambitious perspective is the automation of the database enrichment phase.**

