

	<b>Natacha Koenig (2018-2022)</b>
	Approches bioinformatiques pour l'exploration sans a priori des voies moléculaires chez les espèces non modèles : le cas du métabolisme lipidique chez <i>Gammarus fossarum</i>
	Encadrants : Olivier Geffard & Davide Degli Esposti , ECOTOX, RiverLy, INRAE
	Ecole Doctorale: E2M2

## Résumé

L'utilisation des espèces modèles en science environnementale se confronte à plusieurs verrous scientifiques, tels que leur distribution géographique et leur représentativité de la diversité des espèces dans milieux aquatiques. En effet, les connaissances liées aux organismes modèles ne permettent pas d'extrapoler et prédire la ou les réponse(s) des organismes présents dans les milieux. Néanmoins, les avancées des nouvelles technologies de séquençage et de spectrométrie de masse (MS) permettent d'étendre l'acquisition des données moléculaires à large échelle aux organismes de pertinence environnementale, comme l'amphipode *Gammarus fossarum*. L'objectif général de cette thèse consiste au développement d'approches bioinformatiques sans a priori et multiomique (transcriptomique et protéomique) pour exploiter les nombreuses données omiques disponibles chez les espèces non modèles et contourner les limitations de l'annotation fonctionnelle à partir des bases de données d'espèces éloignées. Pour répondre à cet objectif, deux stratégies ont été mises en œuvre. Premièrement, une analyse de réseaux de coexpression a été réalisée à l'aide de la méthode WGCNA sur deux jeux de données protéomiques, dans le but d'identifier les acteurs moléculaires (*i.e.* protéines) régulant la physiologie et/ou la réponse au contaminant. Le premier jeu de données a combiné les profils protéomiques des gonades mâles et femelles à différents stades de maturation et les embryons au cours de leur développement. Le second jeu de données provenait d'une étude précédente, et était composé de données protéomiques issues de testicules de gammarex exposés à trois contaminants : le pyriproxyfène, le cadmium et le méthoxyfénoside. Ces deux études ont permis d'identifier des modules de protéines coexprimées, spécifiquement associés aux différents stades de développement et de reproduction d'une part, et à l'exposition aux différents contaminants d'autre part. Les analyses d'enrichissement des modules de protéines ont révélé et appuyé de nouvelles hypothèses pour comprendre les processus biologiques impliqués dans la physiologie du gammarex, ainsi que les mécanismes d'actions (MoA) des contaminants. De plus, cette stratégie nous a permis d'identifier des protéines clés impliquées dans des voies physiologiques et associées aux contaminants, que l'analyse différentielle utilisée n'avait pas permis de révéler. Dans un second temps, nous appliquons une stratégie multiomique pour caractériser les voies métaboliques impliquées dans le ML chez *G. fossarum*. En effet, nous avons adapté un outil d'annotation génomique (CycADS) qui permet de reconstruire des voies métaboliques, aux données transcriptomiques du mâle et de la femelle, en réduisant la redondance des isoformes des transcrits. Par la suite, nous avons intégré les données de protéogénomique de différents organes pour valider les enzymes identifiées par MS avec les données d'annotation fonctionnelle et analyser l'organotropisme du ML. La reconstruction des voies métaboliques a permis d'identifier plus de 70 voies impliquées dans le ML. Les données de MS ont permis de valider la détection d'une centaine d'enzymes catalysant les réactions des voies identifiées dans les différents organes, ainsi que leurs profils d'expression qui sont caractéristiques de certains organes. Dans l'ensemble, ce travail de thèse constitue une base solide pour l'utilisation des données omiques d'organismes non modèles dans une stratégie d'exploration sans a priori ou multiomique. Les résultats ont permis de mettre en évidence des protéines potentiellement impliquées dans les processus biologiques liés aux stades de développement et de reproduction de *G. fossarum*, ainsi que dans la toxicité testiculaire. Ces résultats soulignent également l'intérêt de l'application d'approches omiques au niveau des organes des espèces sentinelles pour identifier et évaluer les possibles différences de MoA de contaminants en fonction de l'organe cible.

## Financement

Bourse doctorale : INRAE,

Fonctionnement : ANR & ANSES

## Publications et communications

- **Natacha Koenig**, Patrice Baa–Puyoulet, Amélie Lafont, Isis Lorenzo–Colina, Vincent Navratil, et al.. Proteogenomic reconstruction of organ–specific metabolic networks in an environmental sentinel species, the amphipod *Gammarus fossarum*. *Comparative Biochemistry and Physiology – Part D: Genomics and Proteomics*, 2024, 52, pp.101323. (10.1016/j.cbd.2024.101323). (hal–04704294)
- **Natacha Koenig**. Approches bioinformatiques pour l’exploration sans a priori des voies moléculaires chez les espèces non modèles : le cas du métabolisme lipidique chez *Gammarus fossarum*. *Bio–informatique [q–bio.QM]*. Université Claude Bernard – Lyon I, 2022. Français. (NNT : 2022LYO10006). (tel–04199684)
- **Natacha Koenig**, Patrice Baa–Puyoulet, Maxime Leprêtre, Jean–Charles Gaillard, J. Armengaud, et al.. Intégration de données multi–omiques pour l’identification d’enzymes clés du métabolisme lipidique chez l’espèce sentinelle *Gammarus fossarum*. *GDR Ecotoxicologie*, Nov 2021, Rennes, France. (hal–03810311)
- **Natacha Koenig**, Christine Almunia, Aurore Bonnal–Conduzorgues, Jean Armengaud, Arnaud Chaumot, et al.. Co–expression network analysis identifies novel molecular pathways associated with cadmium and pyriproxyfen testicular toxicity in *Gammarus fossarum*. *Aquatic Toxicology*, 2021, 235, (10.1016/j.aquatox.2021.105816). (hal–03487859)
- Davide Degli–Esposti, **Natacha Koenig**, Yannick Cogne, Duarte Gouveia, Arnaud Chaumot, et al.. From proteogenomics to systems biology in the freshwater amphipod *G. fossarum*. 18th ICA – International Colloquium on Amphipoda, Thierry Rigaud & Rémi Wattier – University of Burgundy, Aug 2019, Dijon, France. (hal–04476766)
- Davide Degli Esposti, Christine Almunia, Marc–Antoine Guery, **Natacha Koenig**, J. Armengaud, et al.. Co–expression network analysis identifies gonad– and embryo–associated protein modules in the sentinel species *Gammarus fossarum*. *Scientific Reports*, 2019, 9 (1), pp.1–10. (10.1038/s41598–019–44203–5). (hal–02622270)