

---

# Modèles de méta-analyse pour la détection de la pléiotropie au niveau des gènes et des variants génétiques tenant compte de la structure de groupe des données génétiques

Pierre-Emmanuel Sugier<sup>\*1,2,3</sup>, Yazdan Asgari<sup>4</sup>, Thérèse Truong<sup>4</sup>, and Benoit Liquet<sup>5,6</sup>

<sup>1</sup>Laboratoire de Mathématiques et de leurs Applications de Pau – Université de Pau et des Pays de l’Adour, UMR CNRS 5142, E2S-UPPA – France

<sup>2</sup>Equipe “Exposome et Hérité” – Paris-Saclay University, UVSQ, Gustave Roussy, Inserm, CESP – France

<sup>3</sup>Mathématiques Appliquées Paris 5 – Université Paris Cité, CNRS, MAP5, F-75006, Paris – France

<sup>4</sup>Equipe “Exposome et Hérité” – Paris-Saclay University, UVSQ, Gustave Roussy, Inserm, CESP – France

<sup>5</sup>Laboratoire de Mathématiques et de leurs Applications de Pau – Université de Pau et des Pays de l’Adour, UMR CNRS 5142, E2S-UPPA – France

<sup>6</sup>School of Mathematical and Physical Sciences, Macquarie University, Sydney – Australie

## Résumé

Les études d’association pangénomique (GWAS) qui se concentrent sur les associations entre un phénotype et plusieurs millions de marqueurs génétiques (SNPs) testés de manière indépendante ont permis d’identifier des centaines de milliers de SNPs associés à de nombreux phénotypes complexes. L’une des découvertes majeures de l’ère GWAS est que la pléiotropie – qui est le fait qu’un même gène affecte plusieurs traits distincts – est un phénomène très répandu dans les maladies complexes chez l’Homme. L’étude des facteurs de risque génétiques partagés entre plusieurs maladies peut améliorer notre compréhension de ces maladies en identifiant de nouveaux gènes et *pathways* (voies biologiques) impliqués. Un nombre toujours croissant de statistiques résumées issues de résultats de GWAS sont mis à disposition de la communauté scientifique. Ainsi, combiner les résultats des GWAS sur différents phénotypes pourrait permettre d’identifier de nouvelles associations de phénotypes pléiotropes (1). Les méthodes de sélection existantes examinent les associations pléiotropes une par une à l’échelle soit du variant génétique soit du gène, mais ne permettent pas de considérer simultanément l’ensemble de l’information génétique. Pour pallier cette limitation, nous proposons une nouvelle approche appelée MPSG de méta-analyse multivariée pénalisée adaptée pour la sélection de la pléiotropie, en tenant compte des informations de structure de groupe imbriquées dans les données, afin de sélectionner les variants et gènes (ou *pathways*) pertinents simultanément sur l’ensemble du génome. Pour ce faire, nous avons implémenté un algorithme des directions alternées (ADMM). Nous avons comparé les performances de la méthode avec d’autres approches de méta-analyse de référence en utilisant différents types de statistiques résumées en entrée (2-4). Notre méthode a été appliquée à l’identification de

---

\*Intervenant

gènes pléiotropes potentiels entre les cancers du sein et de la thyroïde à partir de *pathways* candidats.

## References

Hackinger S, Zeggini E. Statistical methods to detect pleiotropy in human complex traits. *Open biology* 2017;7(11):170125. Baghfalaki, T., Sugier, P. E., Truong, T., Pettitt, A. N., Mengersen, K., & Liqueur, B. (2021). Bayesian meta-analysis models for cross cancer genomic investigation of pleiotropic effects using group structure. *Statistics in Medicine*, 40(6), 1498-1518. Bhattacharjee, P. Rajaraman, K. B. Jacobs, W. A. Wheeler, B. S. Melin, P. Hartge, M. Yeager, C. C. Chung, S. J. Chanock, and N. Chatterjee. A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. *The American Journal of Human Genetics*, 90(5):821–835, 2012. Ray and N. Chatterjee. A powerful method for pleiotropic analysis under composite null hypothesis identifies novel shared loci between type 2 diabetes and prostate cancer. *PLoS genetics*, 16(12):e1009218, 2020