

GEM3 Graduate Student Helps Bring Efficiency to Genome to Phenome Research



John M. A. Wojahn

John M. A. Wojahn (aka Michael), a PhD student in Ecology, Evolution, and Behavior with Global Change Emphasis at Boise State University, recently led efforts to develop a new tool (together with Dr. Stephanie Galla, Dr. Anthony Melton and Dr. Sven Buerki) that allows users to formulate hypotheses about genes to phenotypes (G2P) interactions in their preferred study system through an automated, efficient, and unbiased literature review approach.

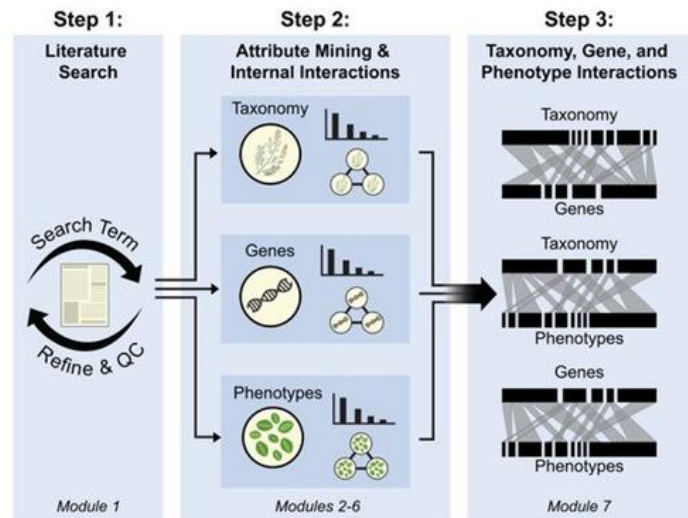
The G2PMineR project goal was to produce a tool (the G2PMineR R-package) to address a gap in the conceptual framework linking G2P for non-model organisms, as most non-model organisms do not yet have genomic resources readily available. Researchers often have to perform literature reviews to understand G2P linkages by developing a list of likely gene candidates, hinging upon other studies already conducted in closely related systems. This can be a tedious process and also has the potential to introduce bias into the research.

The G2PMineR tool is also free and available on the web. This means it is accessible to anyone regardless of ability to pay, whether they are rural or urban, or if they live in the developed or developing world. It is also an open-source design, meaning it can be modified and shared due to its public accessibility. In addition;

G2PMineR



- G2PMineR allows for users to review thousands of abstracts efficiently (in a day) without human bias and extracts information about possible G2P interactions from them so that the user can formulate hypotheses
- G2PMineR can be used by people at any point in their scientific career—from high-school students to professors
- G2PMineR is flexible: users are not restricted to a single kingdom—it can be used on Plantae, Animalia, and Fungi, and the user can customize functions to look for any set of words. In addition, the package can process text from any source, not just abstracts.



These attributes of the G2PMineR allow for the literature review process to be greatly quickened, speeding up research on important topics such as climate change.

Wojahn, who began his GEM3 involvement as a mentor in the Summer Authentic Research Experience (SARE) program in Summer 2020 working alongside Dr. Sven Buerki, helped introduce bioinformatics to the students. There, the G2PMineR project began to take shape and efforts continued through the development of a GEM3 Vertically Integrated Project (VIP) course in the Fall of 2020, where Wojahn took on the role of mentee and then VIP mentor in Spring 2021, providing bioinformatics support to new cohort of VIP students. Through his mentorship, a cornerstone of the GEM3 program, Wojahn provided mentorship to many students including Carlos Dumaguit (SARE undergraduate) who was inspired by the experience to join the Buerki lab as a graduate student (and being involved in conducting sagebrush genotype-by-environment experiments).

Wojahn is currently serving dual roles. He is a research assistant in the Buerki lab where he is working on developing a neural network AI to infer the phenotypes of sagebrush seedlings, and also working on extending G2PMineR's analytical capacity for a future 2.0 release. He is also continuing his work within his doctoral program, studying the evolutionary history and genomics of the genus *Pandanus*, which is a genus of tropical monocotyledonous trees in the Old World tropics.

Wojahn's future plans include studying the evolutionary history and genomics of plants in service of designing better conservation strategies, as well as developing free and open-source software for biodiversity research. Ultimately, after graduating from the EEB doctoral program and finishing a postdoc, Wojahn hopes to be a principal investigator at either a botanical garden, a museum, a university, or a non-profit organization.

Wojahn and team recently published a manuscript entitled "[G2PMineR: A Genome to Phenome Literature Review Approach](#)" in *Genes*. The GEM3 team is also using the G2PMineR to study drought tolerance in plants. To learn more about G2PMineR visit: https://buerkilabteam.github.io/G2PMineR_Web/



The project described was supported by NSF award number OIA-1757324 from the NSF Idaho EPSCoR Program and by the National Science Foundation.