



U01: Rice and Weir - From gene regions to whole chromosomes: scaling up association-finding for disease and omics outcomes in TOPMed

[Introduction slides from 9/1/17 Analysis Committee call](#) [1]

This application will bring unprecedented forms of analysis to TOPMed's already-rich data resources. The application uses recently-developed tools from numerical analysis, that quickly extract key features of the large data summaries used in high-throughput genetic studies. By using them, analyses can be performed on many more variants than is currently possible, so the methods remove a major "bottleneck" in the process of turning TOPMed's data into knowledge. The differences are dramatic; for example, instead of being able to analyze single genes at a time, the new methods scale up to (for example) use of whole chromosomes. By working together with TOPMed Working Groups, the funded work will employ these new and powerful tools where they can have greatest scientific impact.

The application has several specific aims. All of these will be implemented by our experienced and insightful team of investigators and staff, in collaboration with TOPMed Working Groups. Our group already has considerable collective experience in TOPMed, and so recognize the need for its constituent partners to produce high-quality, open, novel and collaborative science, rapidly.

First among the specific aims, we will provide the fast, large-scale methods described above, used to find associations with rare variants and TOPMed's disease outcomes. As an example, our new approach enables use of "topologically associated domains" (TADs) to aggregate variants for testing; TADs are much larger regions than those currently used for aggregation, but offer a highly compelling way to focus (and hence strengthen) the genetic association signals of interest in TOPMed. We will also collaborate with Working Groups to formulate other new and promising modes of association, and will work with TOPMed's computing infrastructure groups to ensure that these tools are broadly-available.

Second, we will further develop the available analysis tools to cover not just single disease outcomes, but high- dimensional "omics" outcomes. By using the same fast numerical methods, we will implement "canonical correlation"- based analyses that can scale up to the size of TOPMed's data. A first task here will be to summarize major patterns of association between genetic variants and the metabolomic and proteomic measures in TOPMed, but we will again be heavily involved with Working Groups, drawing on their expertise to prioritize the most promising analyses.

Third, we will also further develop the available analysis tools to assess problems of "stratification", that can confound and invalidate genetic association findings – particularly in TOPMed's cross-study work. By scaling up the extent of ancestry that can be identified and adjusted for in analyses, we will be able to make these forms of analysis more robust, and hence increase the proportion of replicable, well-understood scientific findings that TOPMed will produce.

NHLBI Program officer: Rebecca Beer

PI:	Ken Rice
Co-PI:	Bruce Weir
Award Type:	U01 NHLBI TOPMed Program: Integrative Omics Approaches for Analysis of TOPMed Data (RFA-HL-17-011)
Award number:	U01 HL137162-01
Start Year:	2017

Source URL (modified on 06/10/2019 - 4:28pm):<https://topmed.nhlbi.nih.gov/awards/2430>

Links

[1] <https://topmed.nhlbi.nih.gov/system/files/meetings/fastSKATanalysis0901.pdf>