HIGH-DIMENSIONAL PROFILING OF TRANSCRIPTION FACTOR ACTIVITY DIFFERENTIATES TOXCAST CHEMICAL GROUPS

Wilson, A[1,2]; Martin, M[1]; Frame, A[1]; Kothiya, P[1]; Judson, R[1]; Richard, A[1]; Medvedev, A[3]; Makarov, S[3]; Houck, K[1]; Reif, D[1]

[1]National Center for Computational Toxicology, US EPA, RTP, NC
[2]Dept. of Statistics, NC State University, Raleigh, NC
[3]Attagene Inc., RTP, NC

The ToxCast™ project at the U.S. EPA uses a diverse battery of high throughput screening assays and informatics models to rapidly characterize the activity of chemicals. A central goal of the project is to provide empirical evidence to aid in the prioritization of chemicals for additional toxicity testing. Chemicals can be differentiated and prioritized based on the observed activity profiles as reported by ToxCast assays. Here, we provide a statistical framework to quantify and visualize activity patterns across assays and chemicals. We demonstrate this method by comparing broadly-defined chemical use groups amongst the 1060 unique compounds in the combined ToxCast Phase I and II sets. The groups were obtained by querying usage information across the ACToR database. The large chemical set represents a diverse landscape including groups such as pharmaceuticals, pesticides, food additives and fragrances. We used a set of 81 ToxCast assays that measure chemical induced transcription factor activity in a human hepatic cell line. The multiplexed assay enables high-content, functional assessment of transcription factor activities, which are a core component of cellular gene regulatory networks. Using our statistical profiling framework, we highlight a subset of these assays that best differentiate chemical use groups and report similarities amongst transcription factor activity profiles of chemicals used for radically different purposes. We then drill down into these profiles using a Bayesian approach to infer the degree of relatedness between all chemicals within and across groups—shedding light on informative subgroups having similar transcription factor activity patterns. Our methods provide a flexible framework for understanding high-dimensional toxicological data and are extensible to diverse definitions of “group”: e.g., alternate chemical usages, known *in vivo* toxicity endpoints, or structural classes.
This abstract does not necessarily reflect U.S. EPA policy.

Categories: Bioinformatics, Computational Toxicology, Risk Assessment