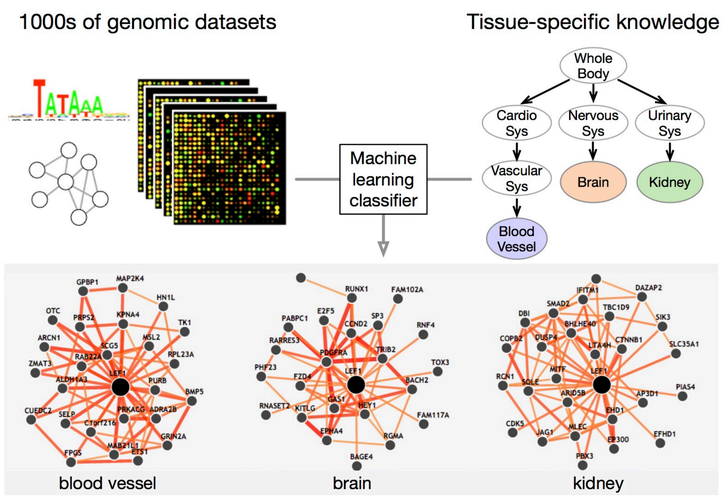
Bringing genomic data into focus for studying complex diseases in specific biological contexts

Arjun Krishnan[*,1], Ran Zhang[*,2], Victoria Yao[3], Chandra Theesfeld[1], Aaron Wong[4], Alicja Tadych[1], Natalia Volfovsky[4], Alan Packer[4], Alex Lash[#,4], Olga Troyanskaya[#,1,3,4]

[1]Lewis-Sigler Inst. for Integrative Genomics, [2]Dept. of Molecular Biology, [3]Dept. of Computer Science, Princeton University; [4]Simons Center for Data Analysis, Simons Foundation

Overview

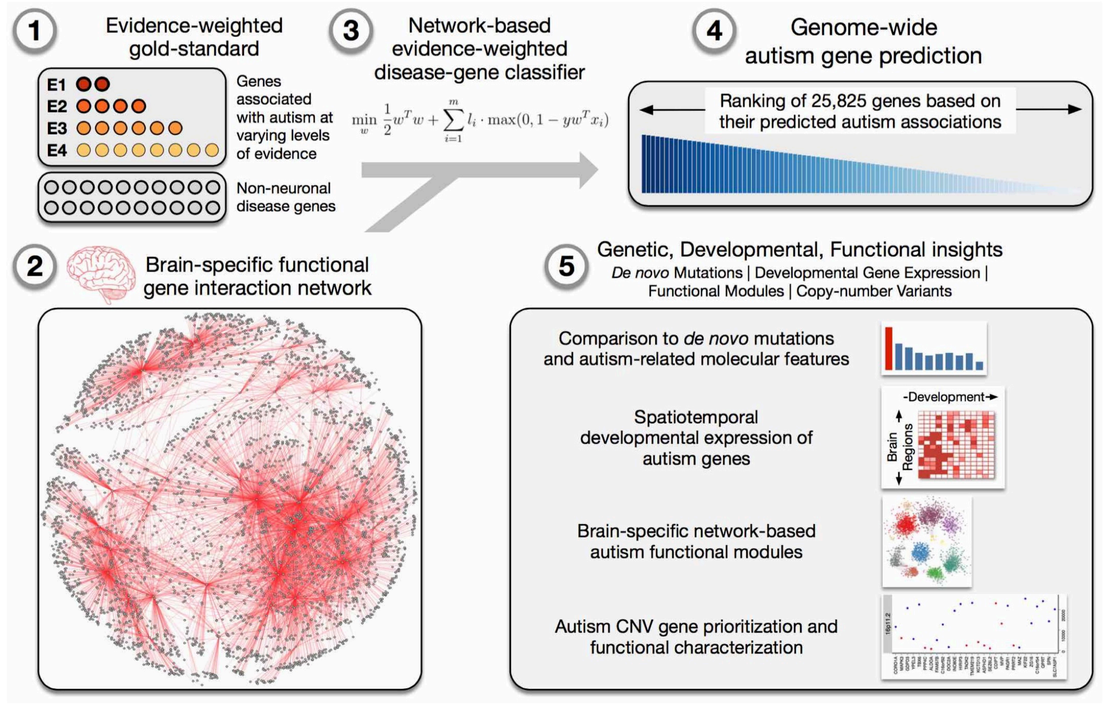
Tissue-specific networks are powerful tools for studying complex diseases like autism spectrum disorder (ASD).



We present the first **genome-wide prediction of autism-associated genes** in the context of the human brain-specific gene network.

The genome-wide predictions, associated brain developmental signatures, functional modules, and prioritized copy-number-variants (CNVs) are all available at **asd.princeton.edu**.

Studying autism genetics in the context of the human brain



Our ASD-gene predictions are based on a machine learning approach that (1) uses a gold-standard of known disease genes – those linked to autism with varying levels of evidence (E1-4) as positives, and other genes linked to non-neuronal diseases as negatives – in the context of a (2) human brain-specific functional interaction network to (3) build an evidence-weighted network-based classifier capturing autism-specific gene interaction patterns, and (4) predict the level of autism-association of all the genes in the genome. A number of subsequent analyses (5) demonstrate the accuracy and utility of our genome-wide complement of autism-associated genes.

arjunk@princeton.edu @compbiologist