

The *AMARETTO Framework: Multimodal and Multiscale Network Inference for Driver and Drug Discovery in Cancer

Mohsen Nabian¹, Jayendra Shinde², Celine Everaert¹, Shaimaa Bakr², Ted Liefeld³, Thorin Tabor³, Charles Blatti⁴, Thomas Baumert⁵, Michael Reich³, Mikel Hernaez⁴, Jill Mesirov³, Vincent Carey¹, Olivier Gevaert², Nathalie Pochet^{1*}

¹Brigham and Women's Hospital, Harvard Medical School, Broad Institute of MIT and Harvard, ²Stanford University, ³University of California, San Diego, ⁴University of Illinois, Urbana-Champaign, ⁵INSERM, University of Strasbourg *npochet@broadinstitute.org

We developed the ***AMARETTO framework** that learns networks of regulatory circuits - circuits of target genes and their shared drivers - across biological systems with applications ranging from disease subtyping to driver and drug discovery in cancer. The ***AMARETTO framework** features: **(1) Modular approach to incorporating prior biological knowledge.** Core algorithms yield multimodal and multiscale reformulations to network inference: *AMARETTO* infers regulatory circuits within each biological system from functional genomics or multi-omics data; *Community-AMARETTO* identifies subnetworks of regulatory circuits shared and distinct across biological systems (model systems and patient studies); *LINKER* provides gene-level network refinements; each accompanied by tools for optimizing generalized prediction performances. Future algorithmic reformulations include: *Consensus-AMARETTO*; *Perturbation-AMARETTO*; *SingleCell-AMARETTO*; *Imaging-AMARETTO*. **(2) Downstream utility of circuits for interpreting experimental and clinical outcomes.** Downstream analytics with visualization to interpret regulatory circuits and subnetworks: functional and clinical characterization for clinical, molecular and imaging-derived phenotypes; prioritization and validation of drivers using genetic perturbations; discovering drug treatments that modulate drivers and targets using chemical perturbations. **(3) Multimodal and multiscale inference of regulatory circuits in cancer.** Driver prediction for: pan-cancer multi-omics subnetworks across 12 cancer (sub)types validated in cell lines; hepatitis C and B virus-induced hepatocellular carcinoma across subnetworks derived from >6 systems validated in cell lines, and prediction of chemopreventive treatments modulating disease-associated subnetworks using perturbations in cell lines, experimentally validated in rat models; multi-omics subnetworks associated with non-invasive imaging-derived features representing prognostic molecular subclasses of glioblastoma multiforme. **Software availability:** *AMARETTO is available via GitHub, Bioconductor, Jupyter Notebook (<http://bit.ly/2v0Elbr>), GenePattern, GenomeSpace, GenePattern Notebook (<http://bit.ly/2IqJ2ZC>).