

***AMARETTO for network biology and medicine:
linking diseases, drivers, targets and drugs
via graph-based fusion of multi-omics, clinical, imaging and perturbation data**

We present *AMARETTO as a software toolbox for network biology and medicine, towards developing a data-driven platform for diagnostic, prognostic and therapeutic decision-making in cancer. *AMARETTO links diseases, drivers, targets and drugs via network graph-based fusion of multi-omics, clinical, imaging, and driver and drug perturbation data across model systems and patient studies of cancer.

The *AMARETTO platform features a modular approach to incorporating prior biological knowledge based on multimodal and multiscale network-structured modeling: (1) The **AMARETTO** algorithm learns networks of regulatory circuits - circuits of drivers and target genes - from functional genomics or multi-omics data and associates these circuits to clinical, molecular and imaging-derived phenotypes within each biological system (e.g., model systems or patients); (2) The **Community-AMARETTO** algorithm learns subnetworks of regulatory circuits that are shared or distinct across networks derived from multiple biological systems (e.g., model systems and patients, cohorts and individuals, diseases and etiologies, *in vitro* and *in vivo* systems); (3) The **Perturbation-AMARETTO** algorithm maps genetic and chemical perturbations in model systems onto patient-derived networks for driver and drug discovery, respectively, and prioritizes lead drivers, targets and drugs for follow-up with experimental validation; and (4) The **Imaging-AMARETTO** algorithm maps radiographic and histopathology imaging data onto the patient-derived multi-omics networks for non-invasive and histopathology imaging diagnostics.

We demonstrate the utility of *AMARETTO via Jupyter Notebook workflows for several Use Cases integrating multi-omics, clinical, imaging, and driver and drug perturbation data across model systems and patient studies of cancer: (1) A study of hepatitis C and B virus-induced hepatocellular carcinoma (LIHC) with driver and drug discovery for chemoprevention across pan-etiologicals of hepatocellular carcinoma, experimentally validated in rat models; (2) A study of glioblastoma multiforme (GBM) and low-grade glioma (LGG) with driver discovery for diagnostic and prognostic molecular subclasses associated with imaging-derived features for non-invasive imaging diagnostics; and (3) A pan-cancer study of squamous cell carcinoma (SCC) across five SCC cancer sites, in particular, lung (LUSC), head and neck (HNSC), esophageal (ESCA), cervical (CESC) and bladder (BLCA).

Software availability: *AMARETTO is available as user-friendly tools via GitHub, Bioconductor and R Jupyter Notebook to enable further algorithm and software development and via GenePattern, GenomeSpace and GenePattern Notebook to reach a broad audience of biomedical researchers.

References, Tools and Resources: <http://portals.broadinstitute.org/pochetlab/amaretto.html>

Funding: NIH NCI ITCR R21 CA209940 (Pochet), NIH NIAID R03 AI131066 (Pochet), NIH NCI ITCR U01 CA214846 collaborative suppl. (Carey/Pochet)