progenetix oncogenomic resource

Qingyao Huang[12], Paula Carrio-Cordo[12], Bo Gao[12], Rahel Paloots[12], Ziying Yang[12], Michael Baudis[12]

[1] Swiss Institute of Bioinformatics, 8057 Zurich [2] University of Zurich, Winterthurerstrasse 190, 8057 Zurich

01 Introduction

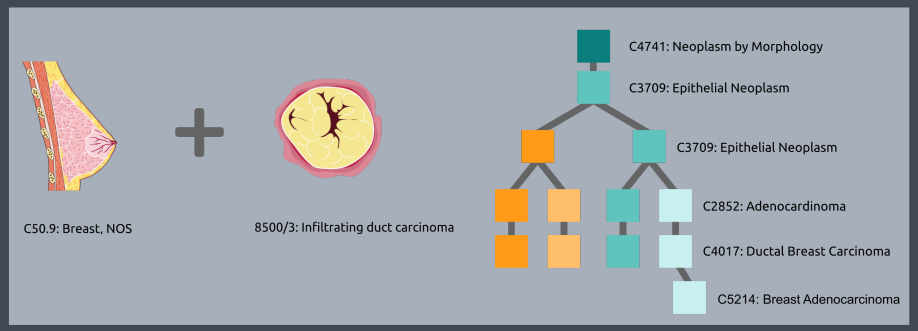
Progenetix resource

The Progenetix oncogenomics resource provides sample-specific cancer genome profiling data and biomedical annotations as well as provenance data for cancer studies. Especially through more than 100k genomic copy number number (CNV) profiles from over 500 cancer types, Progenetix is the most comprehensive reference resources for copy number aberration in cancer, empowers comparative analyses vastly exceeding individual studies and diagnostic concepts as well as supports development of data standards and exchange protocols through Global Alliance for Genomics and Health (GA4GH).
Database URL: progenetix.org

02 New meta-data

Domain-specific mapping

ICD-O system v.s. NCI thesaurus



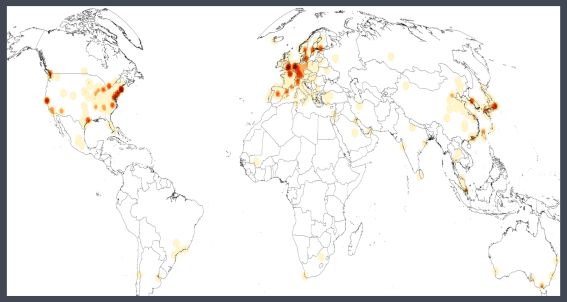
ICD topography

Clinical and diagnostic aspects

Uberon

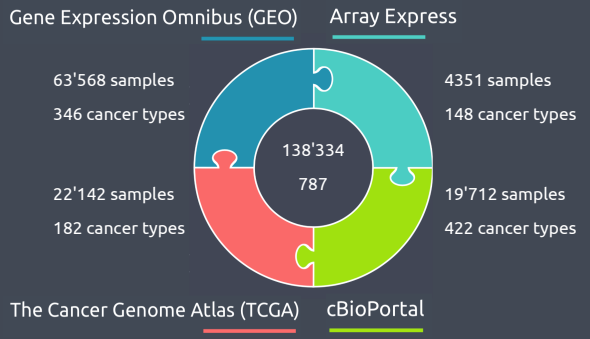
Cross-species anatomy
Developmental lineages
Cross-database reference
Spatial relations

Geographic provenance



03 New data sources

Sample expansion



04 Data standards

CURIE, GA4GH, Phenopackets schema

FAIR

Findability
Accessibility
Interoperability
Reusability

- Compact URI (CURIE) syntax, e.g. pgx:pgxbs-kftvgk8h
- GA4GH specification for variation data, i.e. individual -> biosample -> callset -> variant
- Phenopackets for phynotype data, i.e. id, phenotypes, EHR, provenance, biosample, biosample...

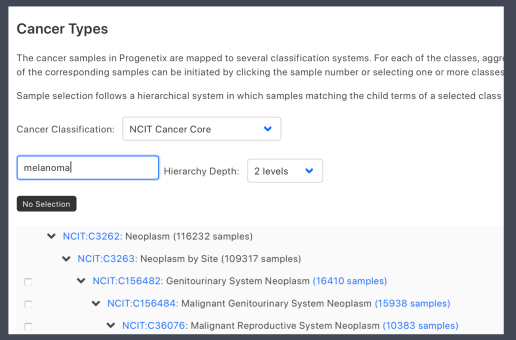
05 Beacon protocol

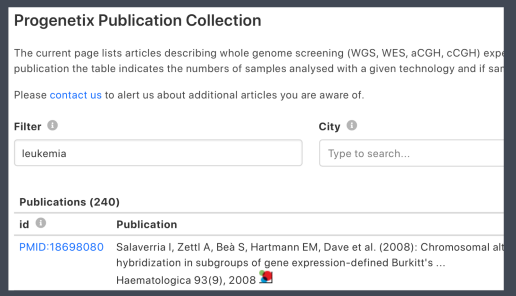
Features and prospects

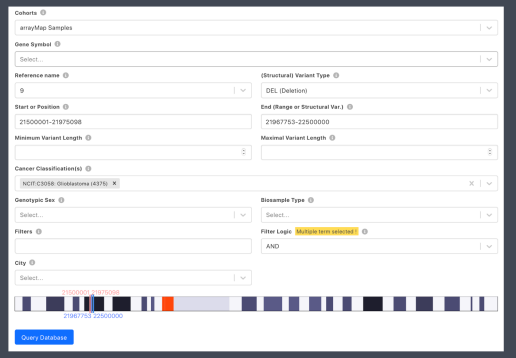
- A Beacon answers a query for a specific variant against an individual or aggregated collection
- Support single or range query
- Filter by disease, phenotype, ...
- Authentication (to come...)

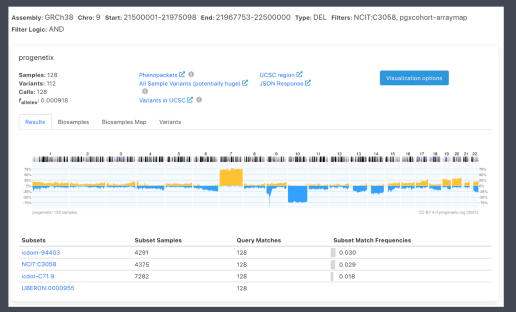
06 New web interface

Many features...









Just scan and visit ...