

# Implementing the GA4GH Beacon API for Genomic Discovery and Sharing in an Open Data Paradigm



Michael Baudis, Hangjia Zhao & baudisgroup @ UZH  
... with members of the Beacon Developer Community



Department of Molecular Life Sciences and Swiss Institute of Bioinformatics, University of Zurich, Switzerland

The Beacon protocol provides a framework for metadata-rich queries but also data delivery response options through its data model and flexible *handovers*

- The Beacon protocol was first developed for federated discovery of genomic variation data
- Recent versions (v2.0 +) support variant, biomedical and metadata query parameters as well as **data delivery** through the Beacon model and *handover* solutions
- Progenetix, the open reference resource for oncogenomic variation data, implements complete data access through the Beacon data model, handover endpoints and custom extensions for currently **240'000 CNV profiles** from **>1'100 cancer codes** as well as SNVs (e.g. TCGA) and associated **biomedical data**
- A complete solution for Beacon implementations is provided through the **bycon** project, based on a MongoDB database solution with a Python software stack and an optional JavaScript/React front end
- Access to the Beacon API for data analysis projects is facilitated through the **pgxRpi** library

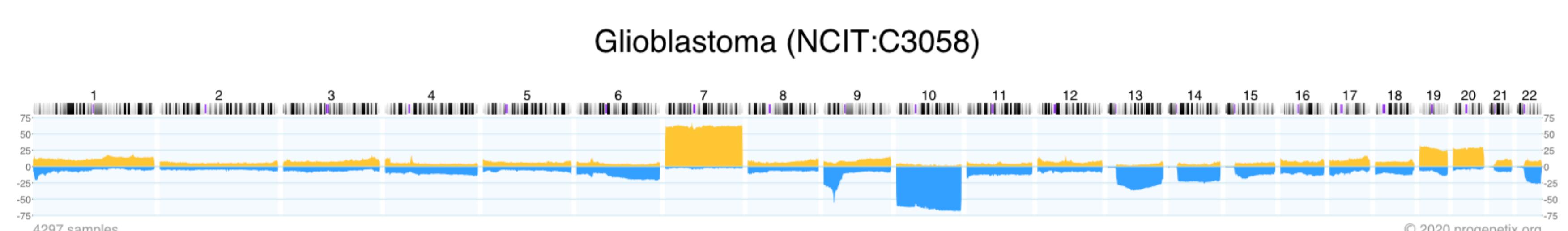
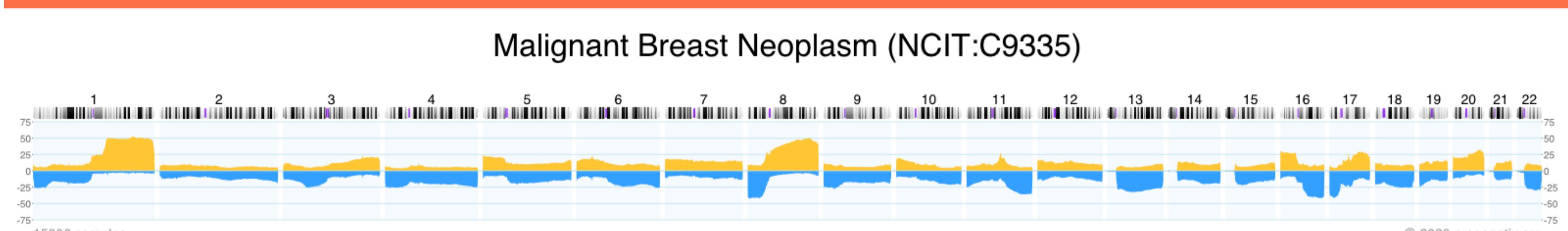
Progenetix is a reference CNV resource built on top of the **bycon** code stack

**bycon** - Beacon full stack open source project

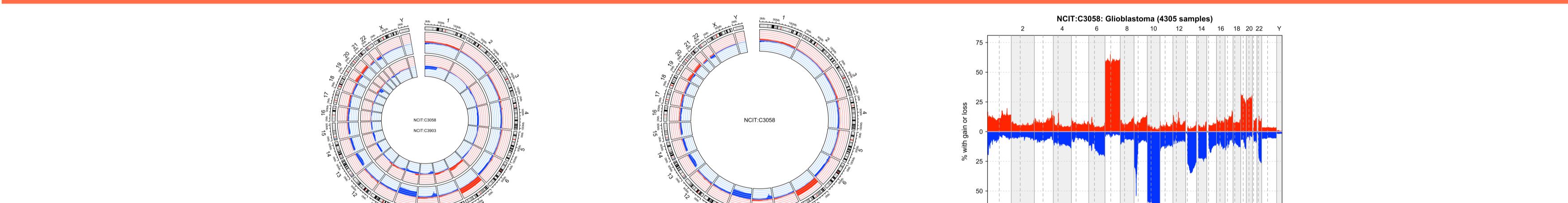
CNV Frequencies provided over Beacon *handover* objects



- Full stack, modular Beacon implementation used for Progenetix, [cancercelllines.org](#) ...
- Implements the Beacon v2.n default model with extensions, e.g. supporting the GA4GH **Phenopackets** format as well as various aggregated exports over **handover** objects
- Forward looking extension of Beacon queries, e.g. with multi-variant queries and **GeoJSON** geographic proximity matches

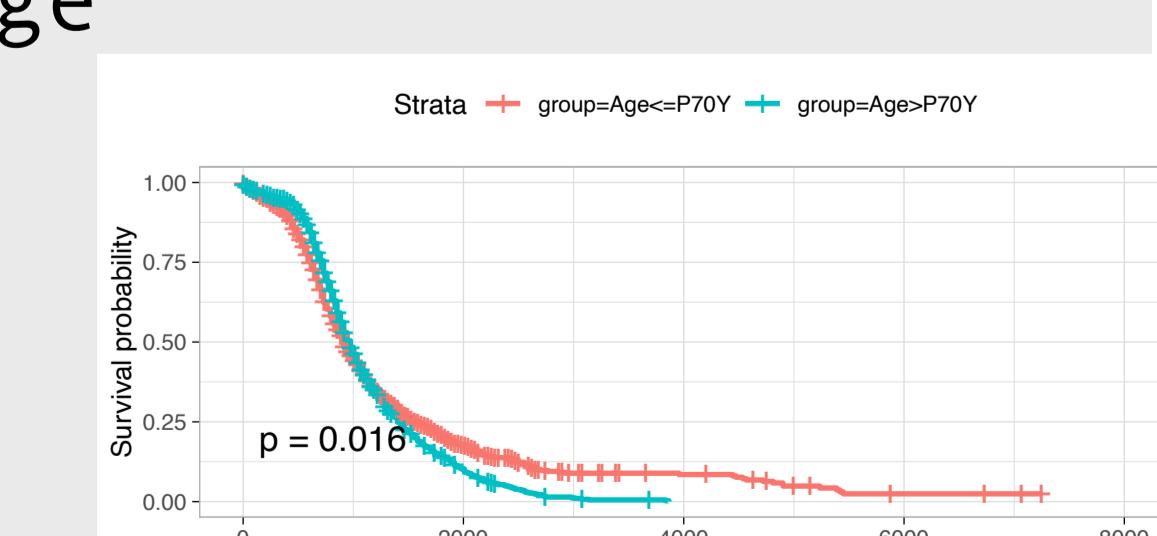


**pgxRpi** - user friendly R/Bioconductor Beacon API package



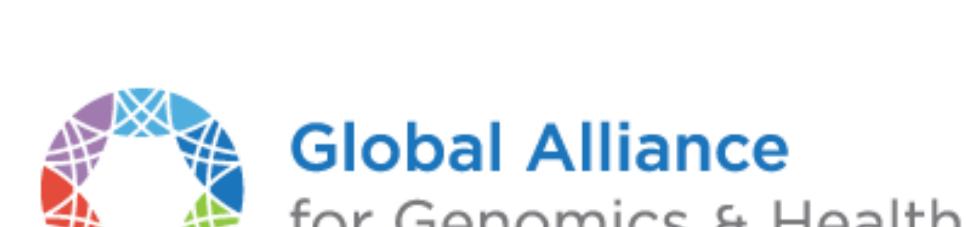
```
# Retrieve metadata of individuals with lung adenocarcinoma
luad_individuals <- pgxLoader(
  type = "individuals",
  filters = "NCIT:C3512",
  domain = c("progenetix.org", "cancercelllines.org"))

# Visualize survival data stratified by age
pgxMetaplot(
  luad_individuals[["progenetix.org"]],
  group_id = "age_iso",
  condition = "P70Y",
  pval = TRUE)
```



Links and Info

[progenetix.org](http://progenetix.org) | [docs.genomebeacons.org](http://docs.genomebeacons.org) | [github.com/progenetix/pgxRpi](https://github.com/progenetix/pgxRpi) | [bycon.progenetix.org](http://bycon.progenetix.org)



University of  
Zurich<sup>UZH</sup>

