GA4GH Beacon v2 - Evolving Reference Standard for Genomic Data Exchange

Gary Saunders, Jordi Rambla de Argila, Anthony Brookes, Juha Törnroos and Michael Baudis

For the ELIXIR Beacon project, GA4GH Discovery work stream and the international network of Beacon API developers

The Beacon Protocol - From the first approved GA4GH standard towards supporting an "Internet of Genomics"



Beacon v2 - Areas of Change and **Future Developments**

The Beacon driver project was one of the earliest initiatives of the Global Alliance for Genomics and Health with the Beacon v1.0 API as first approved GA4GH standard.

Version 2 of the protocol is slated to provide fundamental changes, towards a Internet of Genomics foundational standard:

- requests beyond genomic variants ("filters")
- payload responses, secured through open AAI
- aligning w/ GA4GH standards (Phenopackets, VRS, DUO...) through SchemaBlocks {S}[B]
- Working with international partners on deployment of advanced implementations

		/irol Bc	1	K"laCaixa"	<i>elixir</i> ° ine	* iE	States and States		Beacon* About Progenetix Help				
A Home	- Sum	mary	acon						Search Samples				
• SNID	Filters 🗸	8964 C>T						Search	CNV Request Allele Request Range Q	uery A	All Fields		
SIN	Variant Anr	notation							CNV Example				
Region	VARIANT TYPE	GENE REGION	EFFE	ст	FUNCTIONAL CLASS	LOCUS NAME	LOCUS ID	AMINOACID CHANGE	This example shows the guary for CNV data	tion voria	ante quertanning the CDKN2A genele anding region with		
Feature	SNP CODING NON_SYNONYMOUS_CODING MISSENSE ORF1ab GU280_gp01, S137L, at least a arrayMap							at least a single base, but limited to "focal" arrayMap collection and can be modified e.	at least a single base, but limited to "focal" hits (here i.e. <= ~2Mbp in size). The query is against the arrayMap collection and can be modified e.g. through changing the position parameters or data source.				
AA	Variant Fre	quency per [Dataset										
minoacid	E	NA-ILLUMINA		ENA-ONT	ENA-ONT GISAID			CONSENSUS	This query type is for copy number queries ("variantCNVrequest"), e.g. using fuzzy ranges for start and end positions to capture a set of similar variants.				
Motifs		0.0029		0.0006	G	.0002	Θ	.0010					
? Info	Metadata	# of results: 24)						Cytoband(s)				
~	FREQUENCY	PER RUN ≑	HOST AGE 🖨	HOST SEX 🖨		GEO ORIGIN	÷	DATASET 🗢	Dataset				
ipeline	0.0	0.0707		-	-			ENA-Illumina	Ilumina arraymap X				
A	Θ.Θ	010	-			-		ENA-Illumina	Genome Assembly				
/arnings	0.0010		 26 female		- Oceania:Australia:Victoria			ENA-Illumina	GRCh38 / hg38		· ·		
							:Victoria	ENA-Illumina	Reference name		(Structural) Variant Type		
	0.0076		49	female 00		Oceania:Australia:Victoria		ENA-Illumina	9		DEL (Deletion)		
	0.0142		50	50 male		Oceania:Australia:Victoria		ENA-Illumina	Start or Position		End (Range or Structural Var.)		
	0.0141 0.0246 0.0046		83	83 female		Oceania:Australia:Victoria		ENA-Illumina	21000001-21975098		21967753-23000000		
			60 female		Oceania:Australia:Victoria		:Victoria	ENA-Illumina	Cancer Classification(s)				
			20	20 female		Oceania:Australia:Victoria		ENA-Illumina	NCIT:C3058: Glioblastoma (2119) ×				

Examples of Beacon v2 test implementations: The CRG Viral Beacon provides extensive information about COVID19 variants (left). The Beacon+ server allows query on cancer genomes using disease code *filters* (right).











Have you seen this variant? It came up in my patient and we don't know if this is a common SNP or worth following up.

A Beacon network federates genome variant queries across databases that support the **Beacon API**



Here: The variant has been found in **few** resources, and those are from **disease** specific **collections**.

9:1800000,21975098 21967753.26000000:DE ncit:C3058 DUO:000004 HP:0003621

Have you seen deletions in this region on chromosome 9 in Glioblastomas from a juvenile patient, in a dataset with unrestricted access?



	C '	n e e w e r k	
RCh38 오	17 : 7577121 G > A	Search	
xample variant	query	Advanced Search	
progenet x	baudisgroup at UZH and SIB Progenetix Cancer Genomics Beacon+ Beacon+ provides a forward looking implementation of the Beacon API, with focus on structural variants and metadata based on the cancer and reference genome profiling data represented in the Progenetix oncogenomic data resource (https://progenetix.org). Visit Us · Beacon API · Contact Us	NBSSNational Bioinformatics Infrastructure Sweden SweFreq BeaconBeacon API Web Server based on the GA4GH Beacon APIVisit Us · Beacon API · Contact Us	
csc	CSC - IT Center for Science Development Beacon Beacon API Web Server based on the GA4GH Beacon API Visit Us • Beacon API • Contact Us	LCSB at University of Luxembourg ELIXIR.LU Beacon ELIXIR.LU Beacon Visit Us · Beacon API · Contact Us	
DisgeNET	Research Programme on Biomedical Informatics DisGeNET Beacon Variant-Disease associations collected from curated resources and the literature Visit Us • Beacon API • Contact Us	Eventuation Eventuation Eqa beacon This Beacon is based on the GA4GH Beacon v1.1.0 + Visit Us · Beacon API · Contact Us	
102 - UNIT HEREITY OF AREL SEA	University of Tartu Institute of Genomics, Estonia Beacon at the University of Tartu, Estonia Beacon API Web Server based on the GA4GH Beacon API Visit Us • Beacon API • Contact Us	CSC - IT Center for Science Production Beacon Beacon API Web Server based on the GA4GH Beacon API Visit Us · Beacon API · Contact Us	
ne Bea dividu ggrega ich as eacon	acon protocol draws it al resources with fede ation. Current impleme developed by ELIXIR query types, and supp	s power from networking rated queries and response ntation of Beacon networks enable the latest versions of port open authentication and	

Service Registry API VALID {--} build passing license Apache 2 GA4GH service-info specification VALID [1] build passing license Apache 2

Networking Beacon resources is supported through GA4GH standards such as the Service Registry and Service Info specifications.









typed queries, to facilitate the detection of diverse or incompletely defined genome variants. Scenarios can include the scanning for any variant in a genomic range (A); the retrieval of copy number variants duplicating a gene's CDR (B) or deletion CNVs affecting a gene in any extent, but limiting this to "focal" events (C).