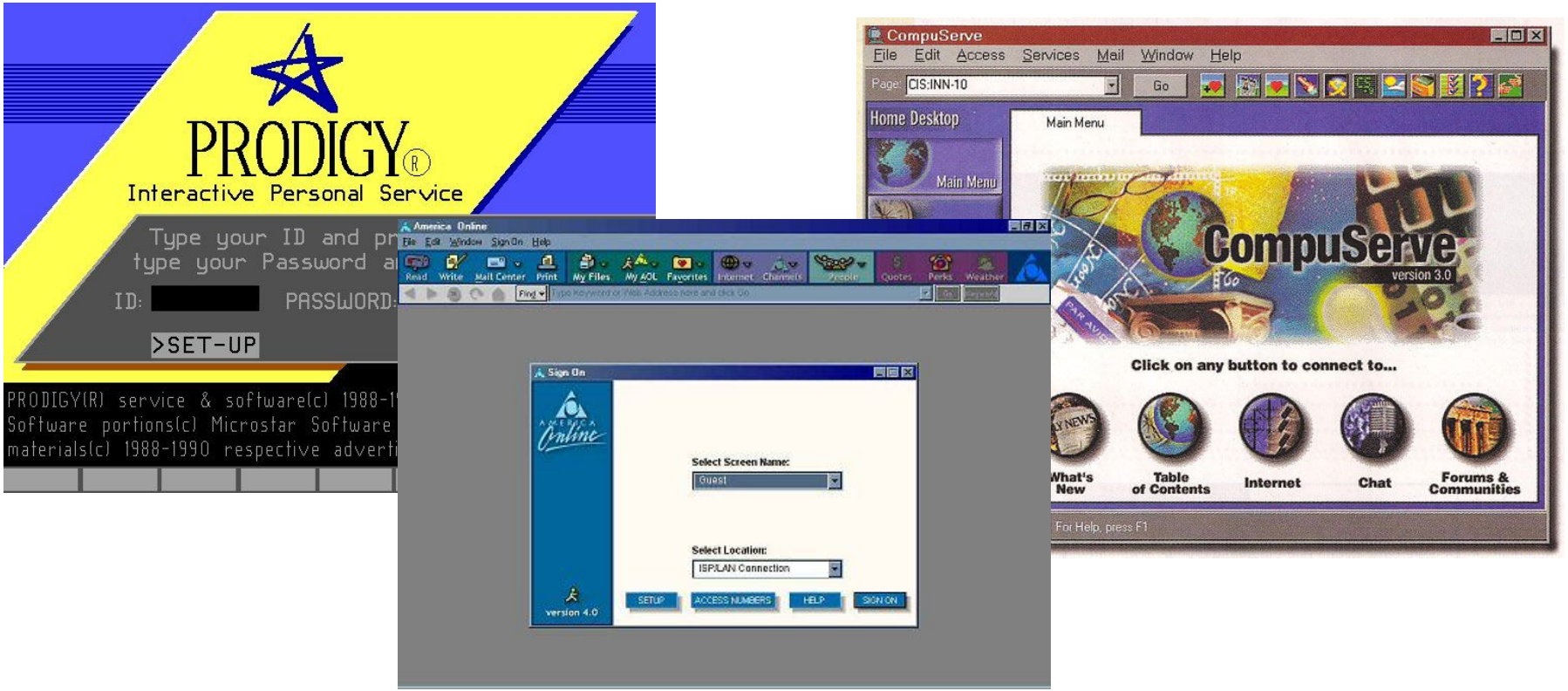
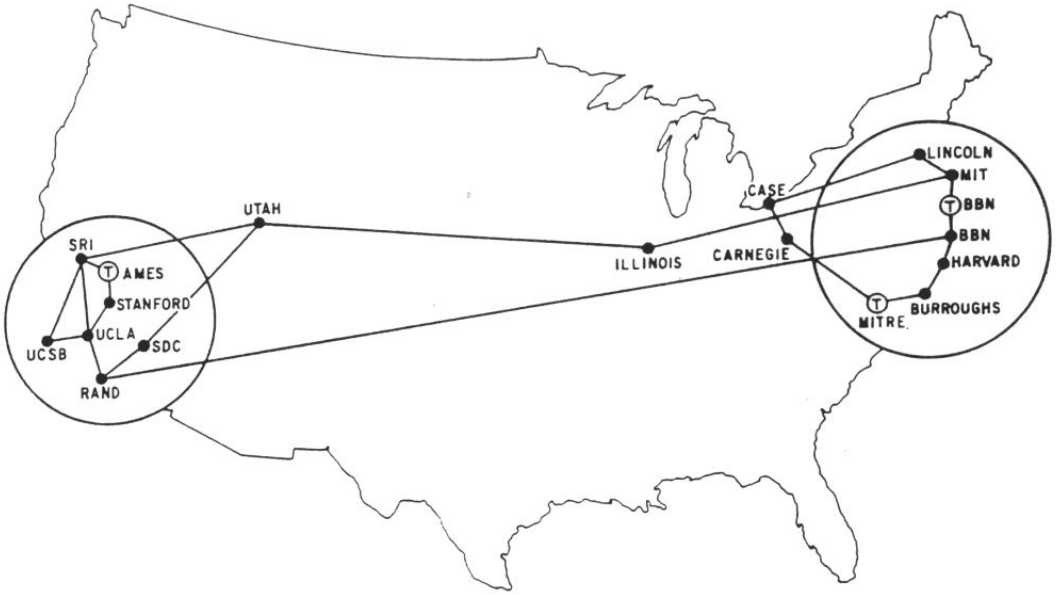
Global Alliance
for Genomics & Health

Collaborate. Innovate. Accelerate.

1993


PRODIGY®
Interactive Personal Service
Type your ID and pr
type your Password a
ID:
PASSWORD:
>SET-UP
PRODIGY(R) service & software(c) 1988-1
Software portions(c) Microstar Software
materials(c) 1988-1990 respective adverti
America Online
File Edit Window Sign On Help
Read Write Mail Center Print My Files My AOL Favorites Internet Channels People Quotes Perks Weather
Sign On
Select Screen Name:
Guest
Select Location:
ISP/LAN Connection
SETUP
ACCESS NUMBERS
HELP
SIGN ON
version 4.0
CompuServe
File Edit Access Services Mail Window Help
Page: CIS:INN-10
Go
Home Desktop
Main Menu
Main Menu
CompuServe
version 3.0
Click on any button to connect to...
What's New
Table of Contents
Internet
Chat
Forums & Communities
For Help, press F1

Internet & TCP/IP

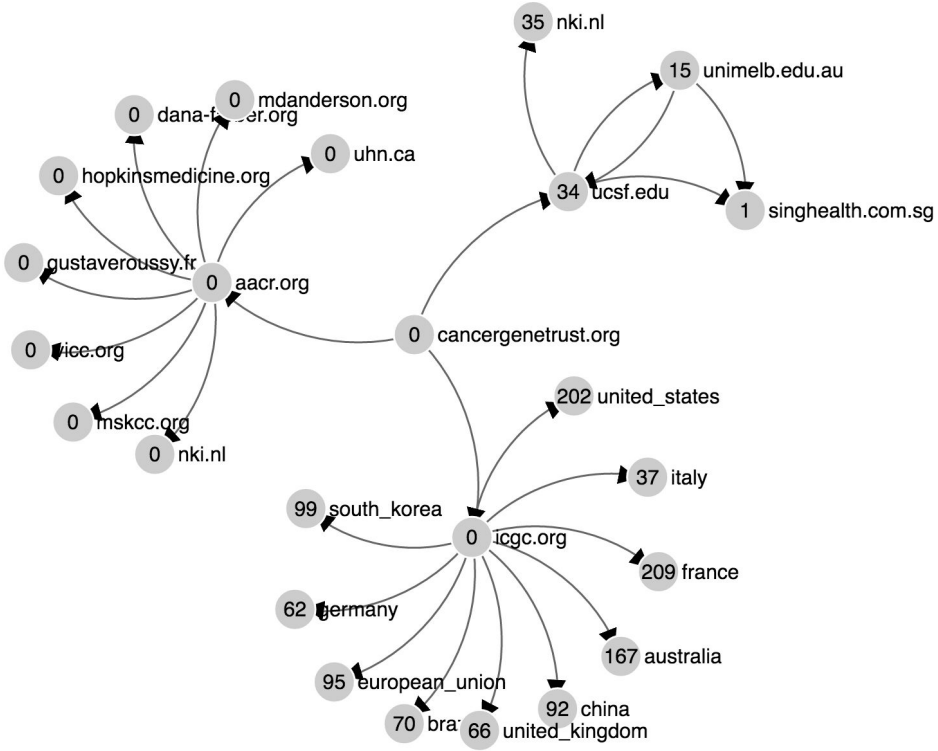


Not yet dreamed of

ebay™ YouTube

GitHub

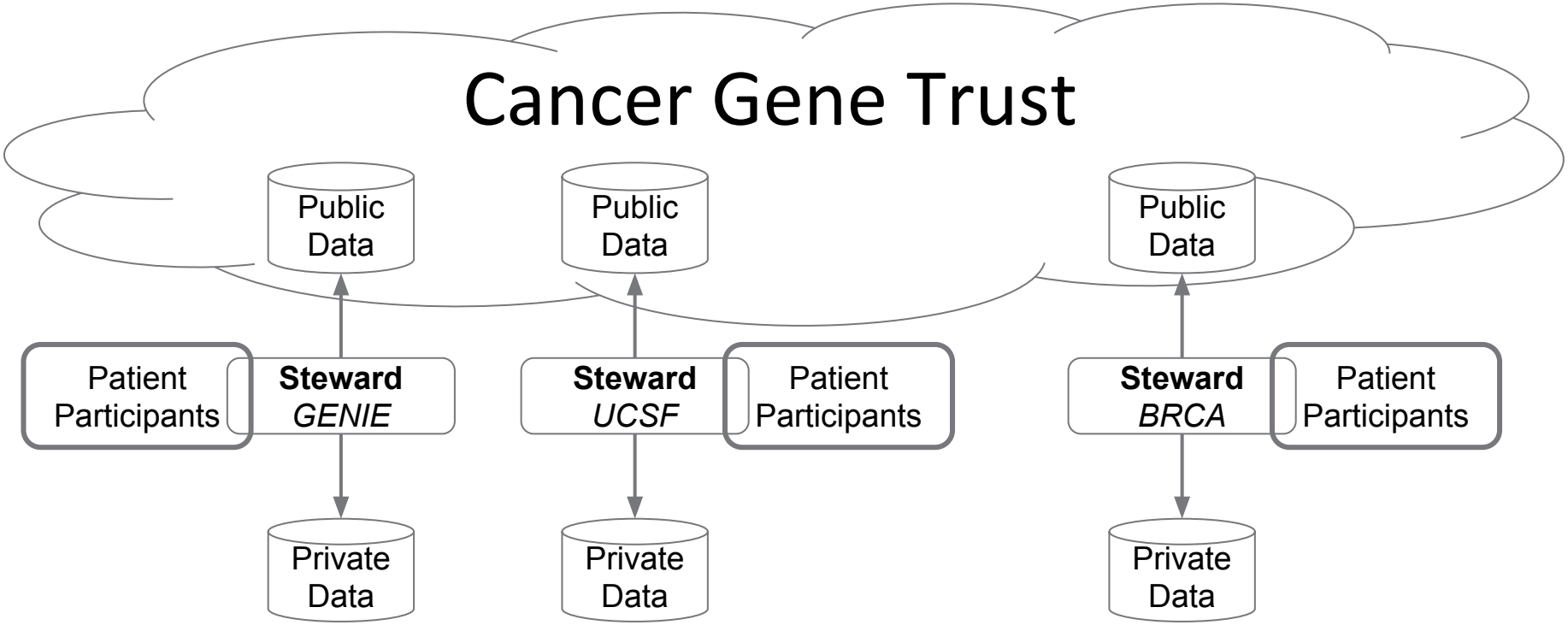
Starting a Genomic Internet



Simple, real-time, global network for sharing somatic cancer data and associated clinical information

Users

Researchers, Clinicians, Pharma, Data Analysts, Citizen Scientists



Submissions

Public/CGT

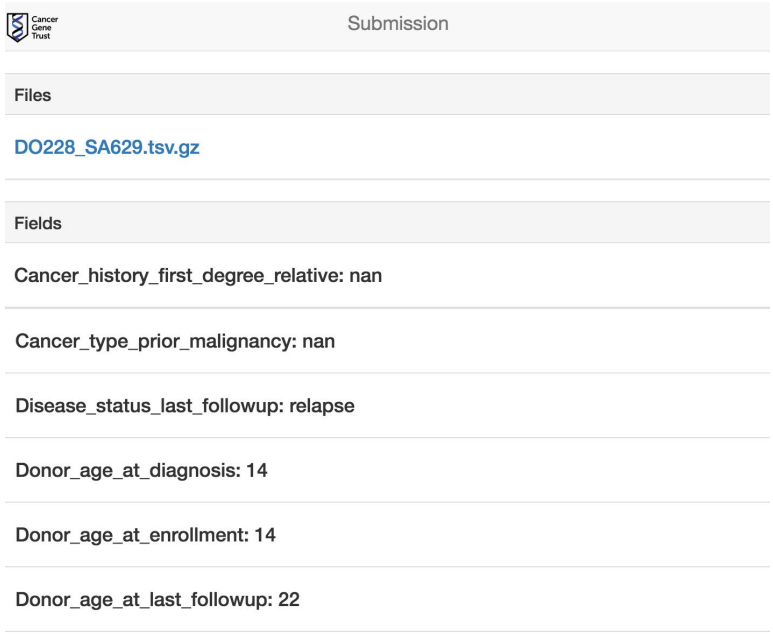
- Somatic Tumor Variants
- Gene Expression Levels
- Non-identifying Clinical Data
- Steward's Identity
- Public Random Participant ID

Private/Steward

- Germline Genome
- Clinical History/EHR
- Signed Consent
- Participant's Instructions
- Internal Patient Participant ID

Submissions

View on web app



Under the hood


```
{
  "files": [
    {"multihash": "QmVeX3fBSa6EELtFSjii8hYCQwV2M2rxaw6zDPWFhaZSxu",
     "name": "DO228_SA629.tsv.gz"}
  ],
  "fields": {
    "cancer_history_first_degree_relative": "nan",
    "cancer_type_prior_malignancy": "nan",
    "disease_status_last_followup": "relapse",
    "donor_age_at_diagnosis": "14",
    "donor_age_at_enrollment": "14",
    "donor_age_at_last_followup": "22",

    ...

    "donor_relapse_type": "progression (liquid tumours)",
    "donor_sex": "male",
    "donor_survival_time": "3060"
  }
}
```

A Submission Can be Cited in a Scientific Paper

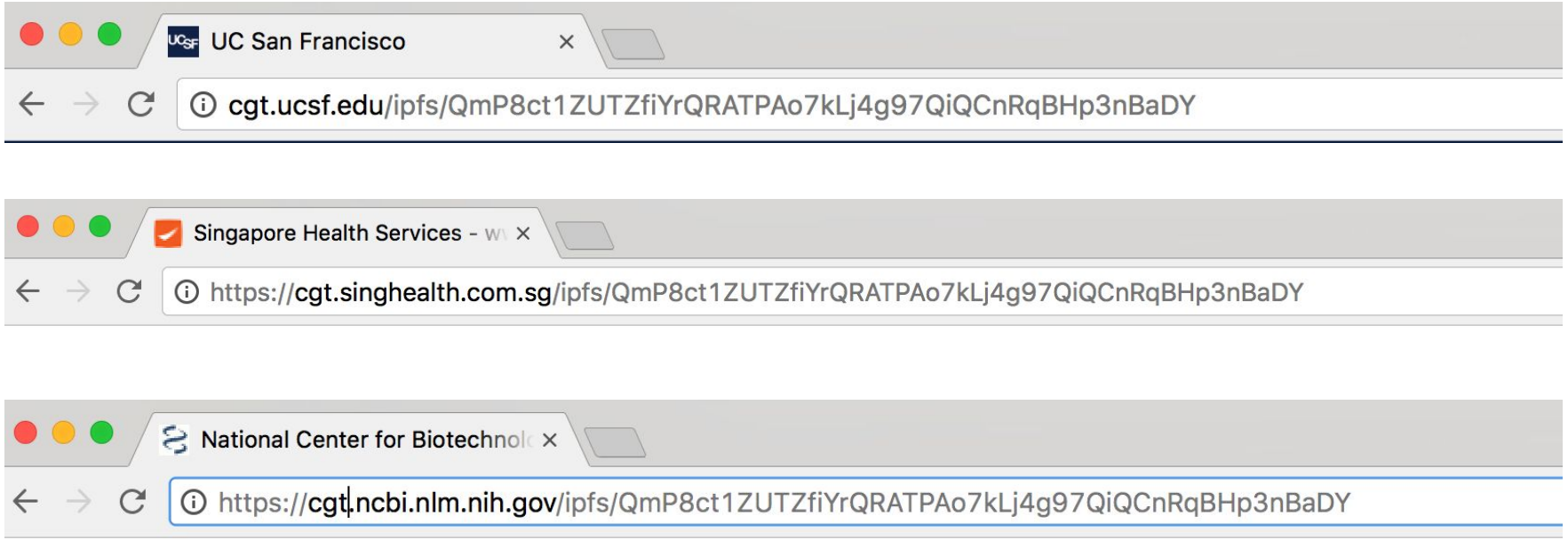
Materials and Methods

For the bulk cancer samples, we downloaded RNA-Seq data for 4 samples from the Cancer Gene Trust steward https://cgt.singhealth.com.sg/ipfs/

```
QmP8ct1ZUTZfiYrQRATPAo7kLj4g97QiQCnRqBHp3nBaDY
QmPGrYpAoAJqbetzq2TxefCTcCfBoeDbSVcvaRM5PLxfh1
QmPMbmhvN4B4XuVgnzyE8KpvRr1pDp13gqWCbqhcxdgCxi
QmPMx67cuUgTgWTUiAAeFmaAJKb43me8QuAF6mj7mygfU6
```

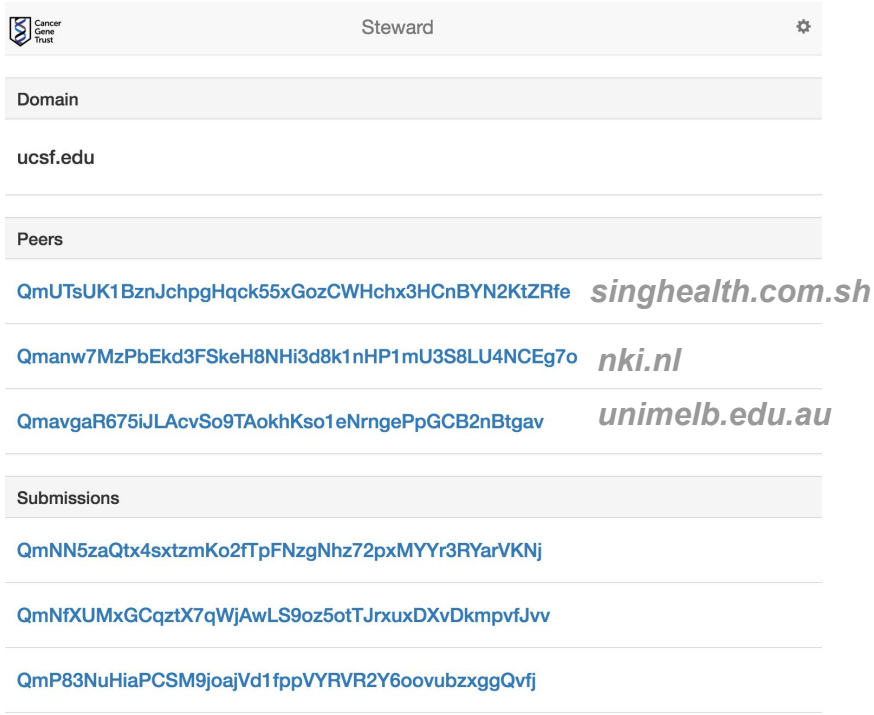
These samples were then processed using the rna pipeline on quay.io with tag 1b46bfcb73e9….

Mirroring


UC San Francisco
cgt.ucsf.edu/ipfs/QmP8ct1ZUTZfiYrQRATPAo7kLj4g97QiQCnRqBHp3nBaDY
Singapore Health Services - w
https://cgt.singhealth.com.sg/ipfs/QmP8ct1ZUTZfiYrQRATPAo7kLj4g97QiQCnRqBHp3nBaDY
National Center for Biotechnol
https://cgt.ncbi.nlm.nih.gov/ipfs/QmP8ct1ZUTZfiYrQRATPAo7kLj4g97QiQCnRqBHp3nBaDY

Stewards

View on web app



Under the hood


```
{
  "domain": "ucsf.edu",
  "peers": [
    "QmUTsUK1BznJchpgHqck55xGozCWHchx3HCnBYN2KtZRfe",
    "Qmanw7MzPbEkd3FSkeH8NHi3d8k1nHP1mU3S8LU4NCEg7o",
    "QmavgaR675iJLAcvSo9TAokhKso1eNrngePpGCB2nBtgav"
  ],
  "submissions": [
    "QmNN5zaQtx4sxtzmKo2fTpFNzgNhz72pxMYYr3RYarVKNj",
    "QmNN5zaQtx4sxtzmKo2fTpFNzgNhz72pxMYYr3RYarVKNj",
    "QmNfXUMxGCqztX7qWjAwLS9oz5otTJrxuxDXvDkmpvfJvv",
    "QmP83NuHiaPCSM9joajVd1fppVYRVR2Y6oovubzxggQvfj",
    ...
    "QmPmcWesfsbBVV9AqWd9XhKhd5S7tvrzeSN2xh4rJGC4WP",
    "QmfNntRgEPyF6Ly9bWDxczZngy5y8t4rip8g48sfksSVBx"
  ]
}
```

GA4GH: What? Why? How?

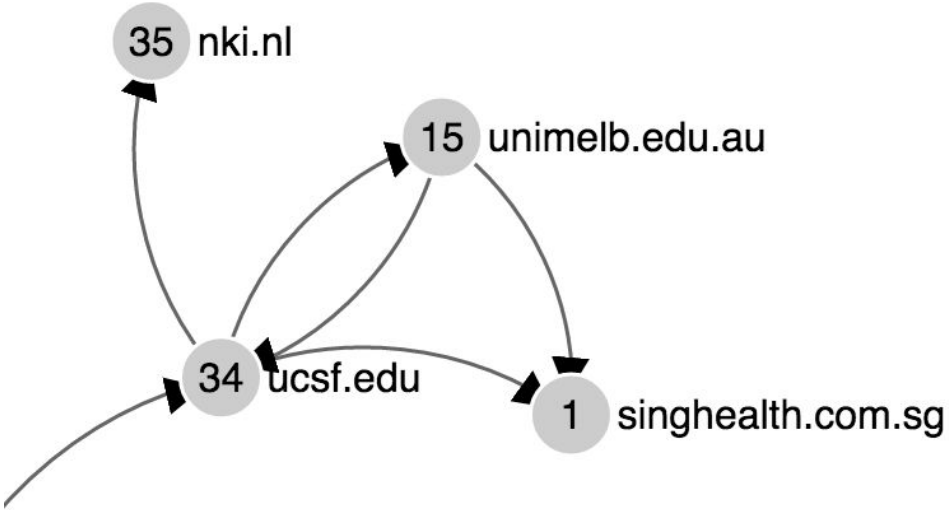
Over the next decade, healthcare will begin to change the way it collects molecular measurements from patients. 'Genomics entering the clinic' means that it will soon become a routine matter to gather DNA, RNA, protein and metabolite data from patients

The goal of the GA4GH is to enable a future in which secondary use of healthcare-generated genomics data is routine and practical

- Ewan Birney

http://genomeinformatician.blogspot.com/2016/10/ga4gh-what-why-how.html

Clinical Stewards



Steward: UCSF

Eric Collisson, MD

INTRODUCTION AND OBJECTIVES

1- Abstract

This project aims to share information about biomarkers () with research groups for the purposes of correlating it with response to treatment and outcome. The vision is that by sharing and combining all subject biomarker data, researchers will be rewarded with a richer data source than if each participant maintained their separate marker data in isolation.

As the data is anonymized a
their privacy, this project sh
expedited review ("minimal
http://irb.ucsf.edu/levels-reviev
approval is not required for

2- Background

DNA sequencing is
lung cancer(lynch), melanor

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO
CONSENT TO DONATE DATA FOR RESEARCH

This is a request that you donate information about your health and clinical care to medical research. The researcher, Eric Collisson, M
Family Comprehensive Cancer Center will
you may ask the study doctor.

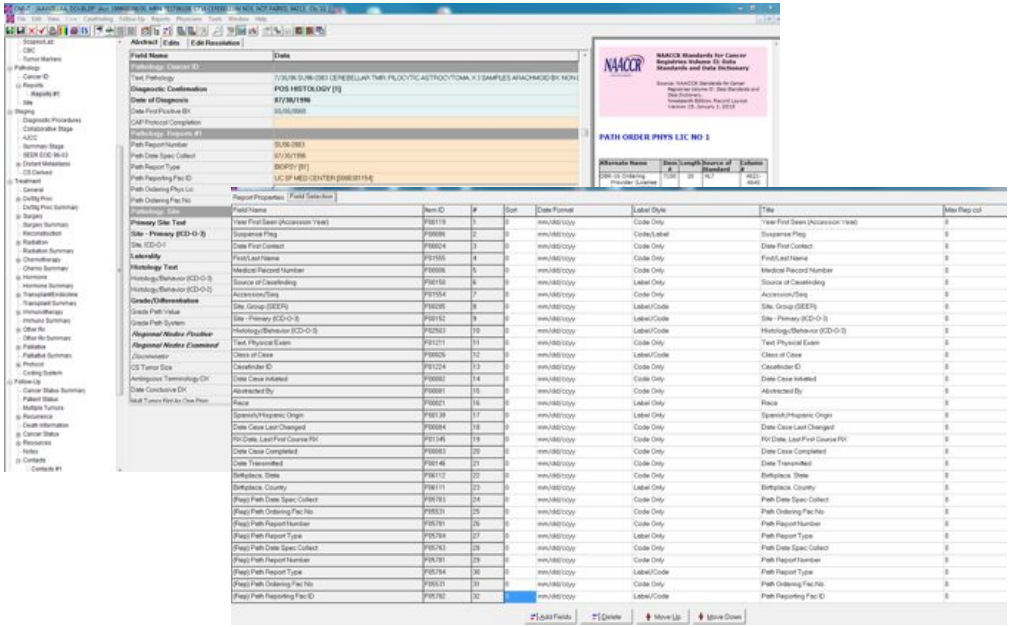
Medical research includes only people who
decision about participating. You may discu
with your healthcare team. If you have any

You are being asked to participate in this st
and your cancer has been genetically analy

	UCSF Internal Cancer Registrar Field	Anonymized Public Field
4	Birthplace	Birthplace (County)
5	Casefinding Source	Casefinding Source
6	Cause of Death	Cause of Death
7	Chemotherapy at This Hospital	Chemotherapy at This Hospital
8	Chemotherapy Summary	Chemotherapy Summary
9	Class of Case	Class of Case
10	Coding Procedure	Coding Procedure
77	Payment Source (Secondary)	
78	Payment Source Text	
79	Pediatric Stage	Pediatric Stage
80	Pediatric Stage Coder	Pediatric Stage Coder
81	Pediatric Stage System	Pediatric Stage System
82	Phone Number (Patient)	
83	Physician (Following)	
84	Physician (Managing)	
89	Physician (Referring)	
90	Physician (Surgeon)	
91	Place of Death	Place of Death (County)
92	Place of Diagnosis	Place of Diagnosis (County)

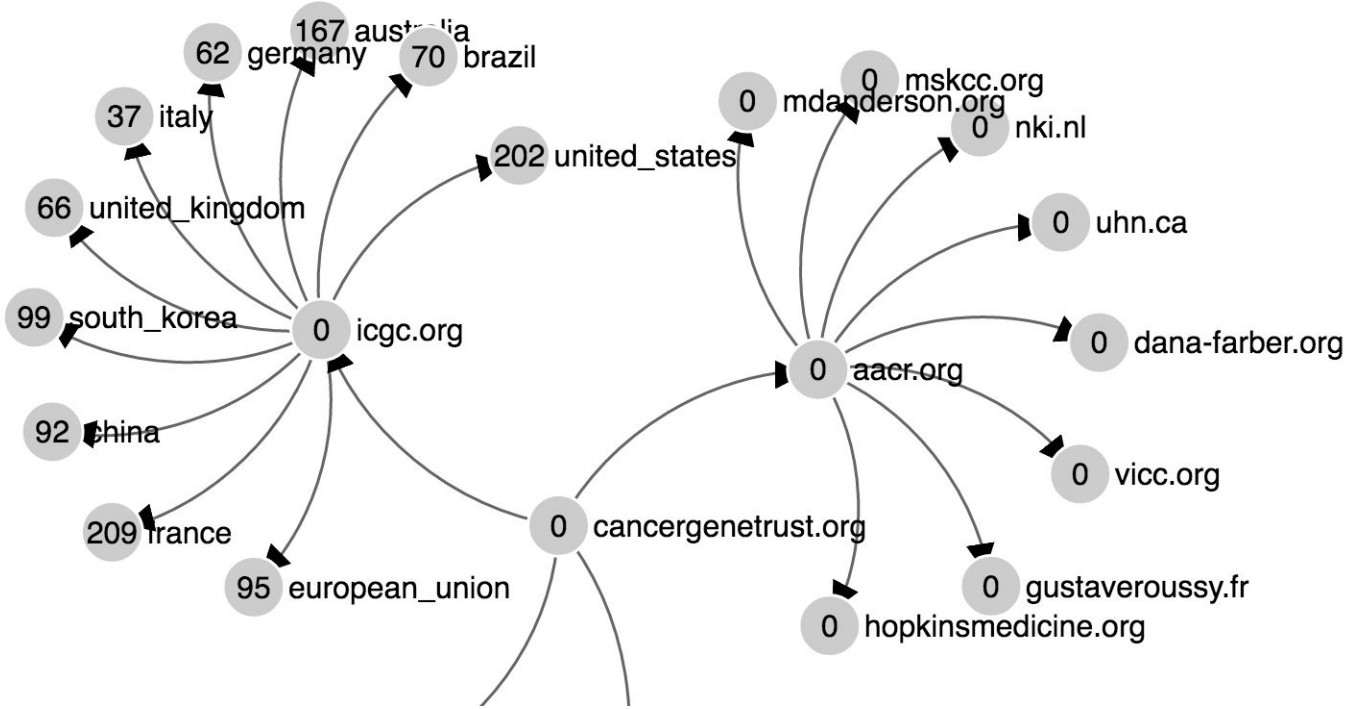
Cancer Registrar

Ann Griffin PhD, CTR

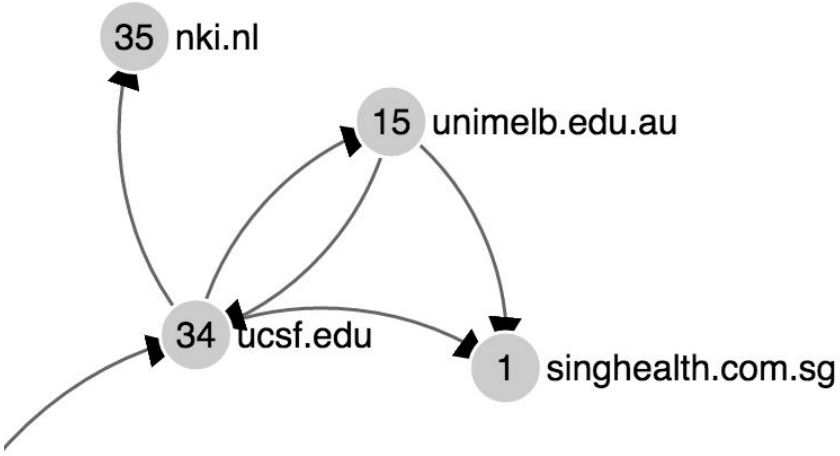


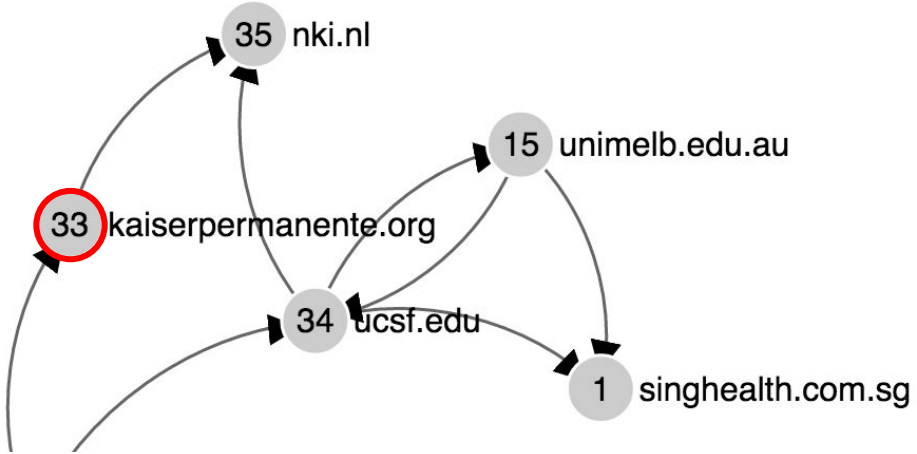
FOUNDATION MEDICINE®

Reference Stewards

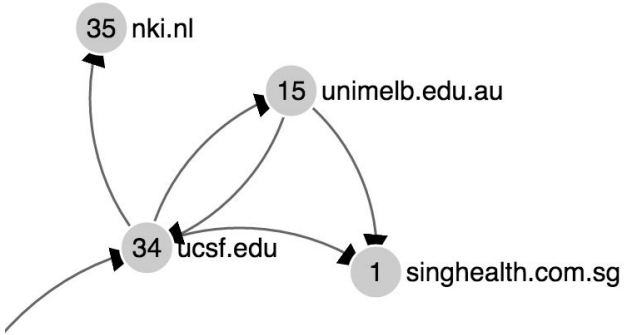


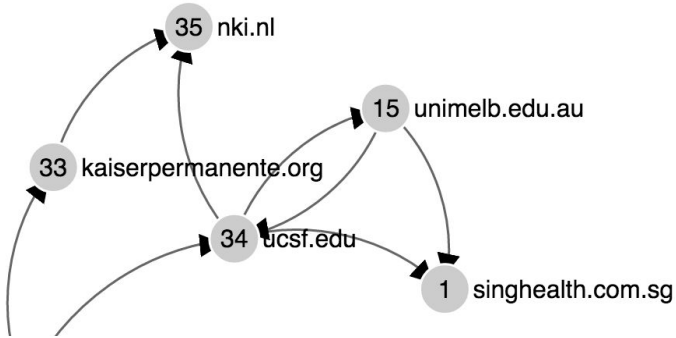
Peers


35 nki.nl
15 unimelb.edu.au
34 ucsf.edu
1 singhealth.com.sg


35 nki.nl
15 unimelb.edu.au
33 kaiserpermanente.org
34 ucsf.edu
1 singhealth.com.sg

Realtime & Organic





Cancer Gene Trust API

RESTful API for the Cancer Gene Trust Daemon (cgtd)

Note that most operations involve validating public keys, signing, and publishing the updated signature to other stewards. As a result some operations can take several seconds if the cached resolution from pki to multihash has expired.

Arrays are sorted by value and dictionaries by key so that hashes are independent of the order the data is added.

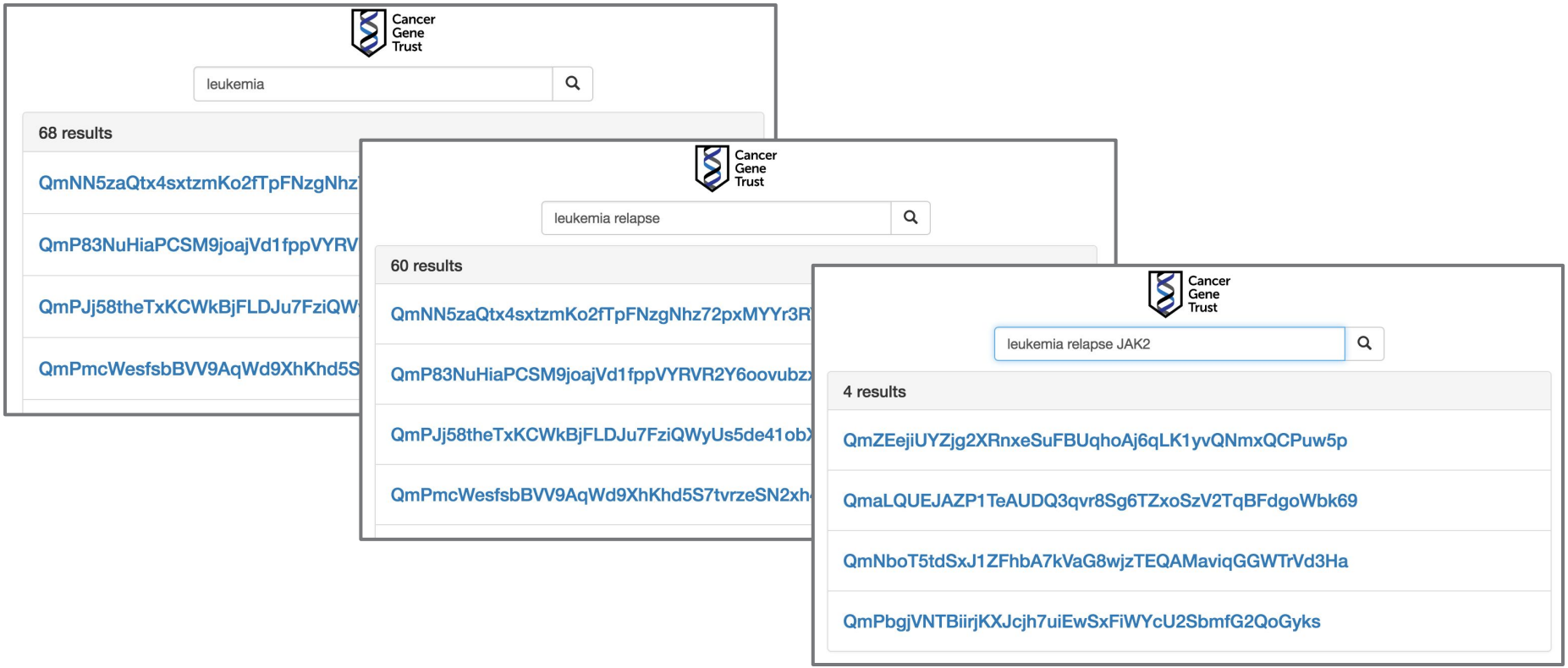
default : Default namespace

Show/Hide | List Operations | Expand Operations

GET	/v0/	Get steward's index
GET	/v0/address	Get steward's address
DELETE	/v0/peers/{address}	Delete address from peer list
POST	/v0/peers/{address}	Add address to this steward's peer list
GET	/v0/stewards/{address}	Get steward index
GET	/v0/submissions	Get list of all submissions
POST	/v0/submissions	Make a submission
PUT	/v0/submissions	Add a list of existing submissions
DELETE	/v0/submissions/{multihash}	Delete submission
GET	/v0/submissions/{multihash}	Get submission

[BASE URL: / , API VERSION: v0]

Application: Search



Application: OncoWatch

To: Charles Sawyers <******@mskcc.org>

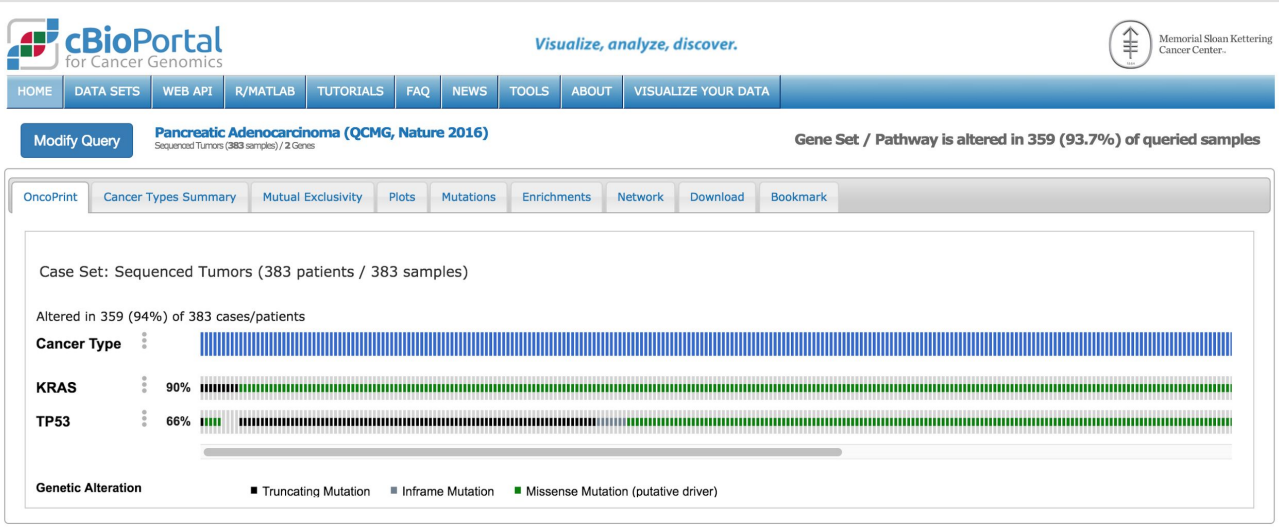
From: OncoWatch

Subject: CGT Submission Match

A new Submission to nki.nl matched your oncowatch for:

metastatic and **erlotinib** and **TP53**

Application: Virtual portal



- Can’t make sense of Variants of Uncertain Significance from small cohorts
- Interactively explore remote data

Application: Somatic Variant Frequencies

Variant	Chrom	Position	Consequence	Filter	Annotation	Flags	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
12:25368370 C / T	12	25368370	c.*4+1G>A	PASS	splice donor		1	121324	0	0.000008242
12:25368395 T / A	12	25368395	p.Lys184Ter	PASS	stop gained	LC LoF	1	121336	0	0.000008242
12:25368405 A / T	12	25368405	p.Cys180Ter	PASS	stop gained		7	121334	0	0.00005769
12:25368442 A / T	12	25368442	p.Leu168Ter	PASS	stop gained		2	121340	0	0.00001648
12:25368452 G / A	12	25368452	p.Gln165Ter	PASS	stop gained		3	121348	0	0.00002472
12:25368455 G / A	12	25368455	p.Arg164Ter	PASS	stop gained		3	121336	0	0.00002472

- Live summary of somatic variants, mutation frequencies

Acknowledgements

International Leadership
David Haussler
Charles Sawyers
Bin Tean Teh
Clare Turnbull
Emile Voest

Collaborators
Eric Collisson
Ann Griffin
Obi Griffith
Malachi Griffith
Sean Grimmond
Nuria Lopez-Bigas
David Tamborero

Technical Development
Rob Currie
Ted Goldstein
Max Haeussler
Oliver Hofmann
James Hughes
Lim Weng Khong
Bisheng Liu
Steve Rosen
Jelle ten Hoeve
Roman Valls

GA4GH Support
Peter Goodhand
Mark Lawler
Rachel Liao
John Mattison
Lillian Siu
Julia Wilson
Peter Yu
CWG & DWG Members

Global Alliance
for Genomics & Health