



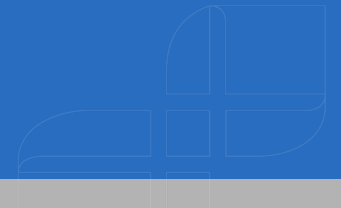
HPC for All of Us




©TGen

Copywrite 2021 TGEN

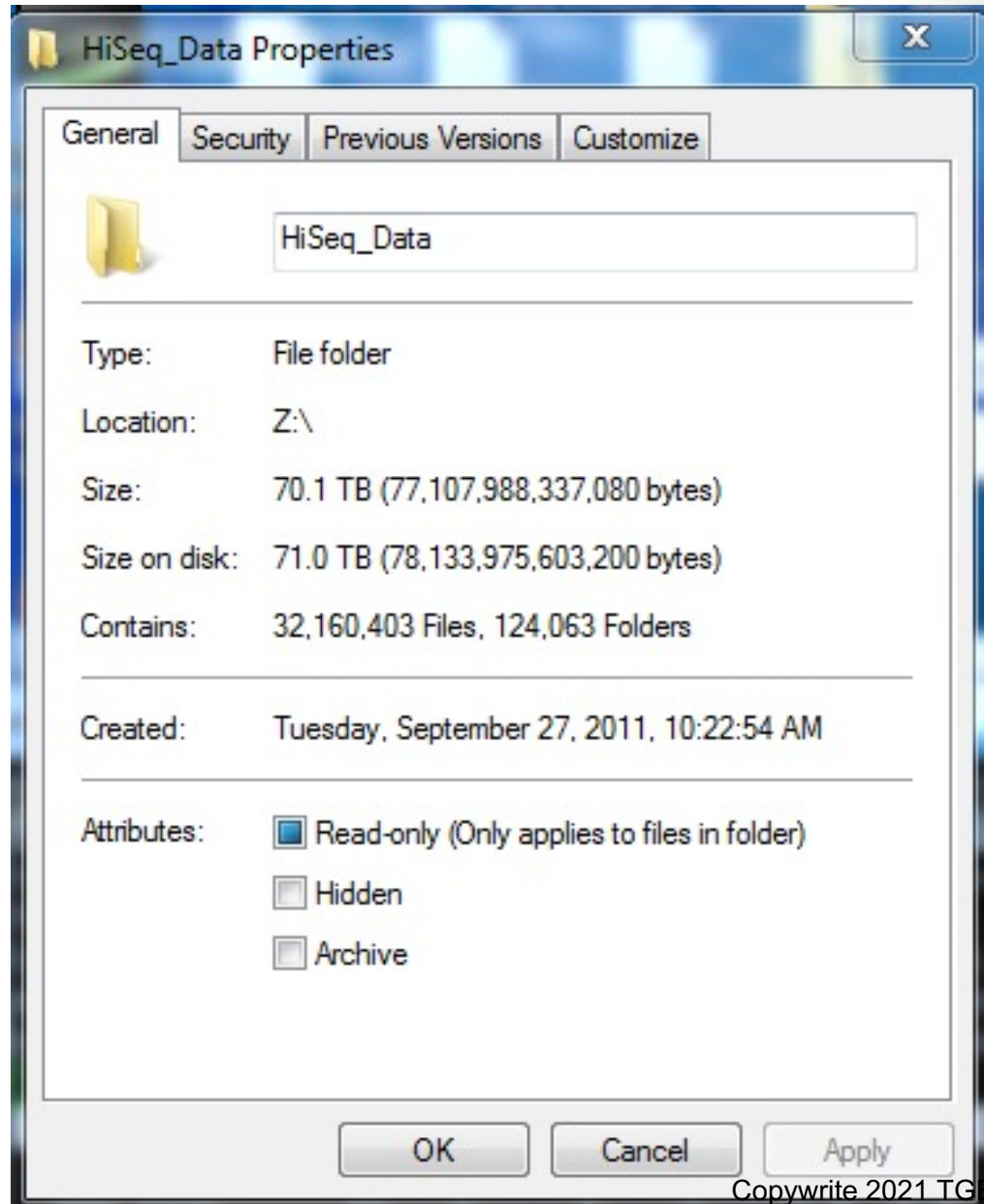
Translational Genomics Research Institute (TGen)



- Non-profit medical research Institute, Phoenix, Arizona
 - Founded 2002 by Jeffrey Trent Ph.D., F.A.C.M.G.
TGen President & Research Director
 - Four Arizona locations
 - Precision medicine pioneer
 - >300 employees
 - 38 Investigators across 8 Divisions
 - Collaborates with 447 academic, medical and industry partners worldwide in 28 countries and US territories
- **Joined City of Hope November 2016**

	Oncology & Phase I Trials
	Rare Genetic Disease
	Neurologic Disorders
	Pathogens & Microbiome
	Quantitative Medicine
	Drug Discovery
	Big Data and AI

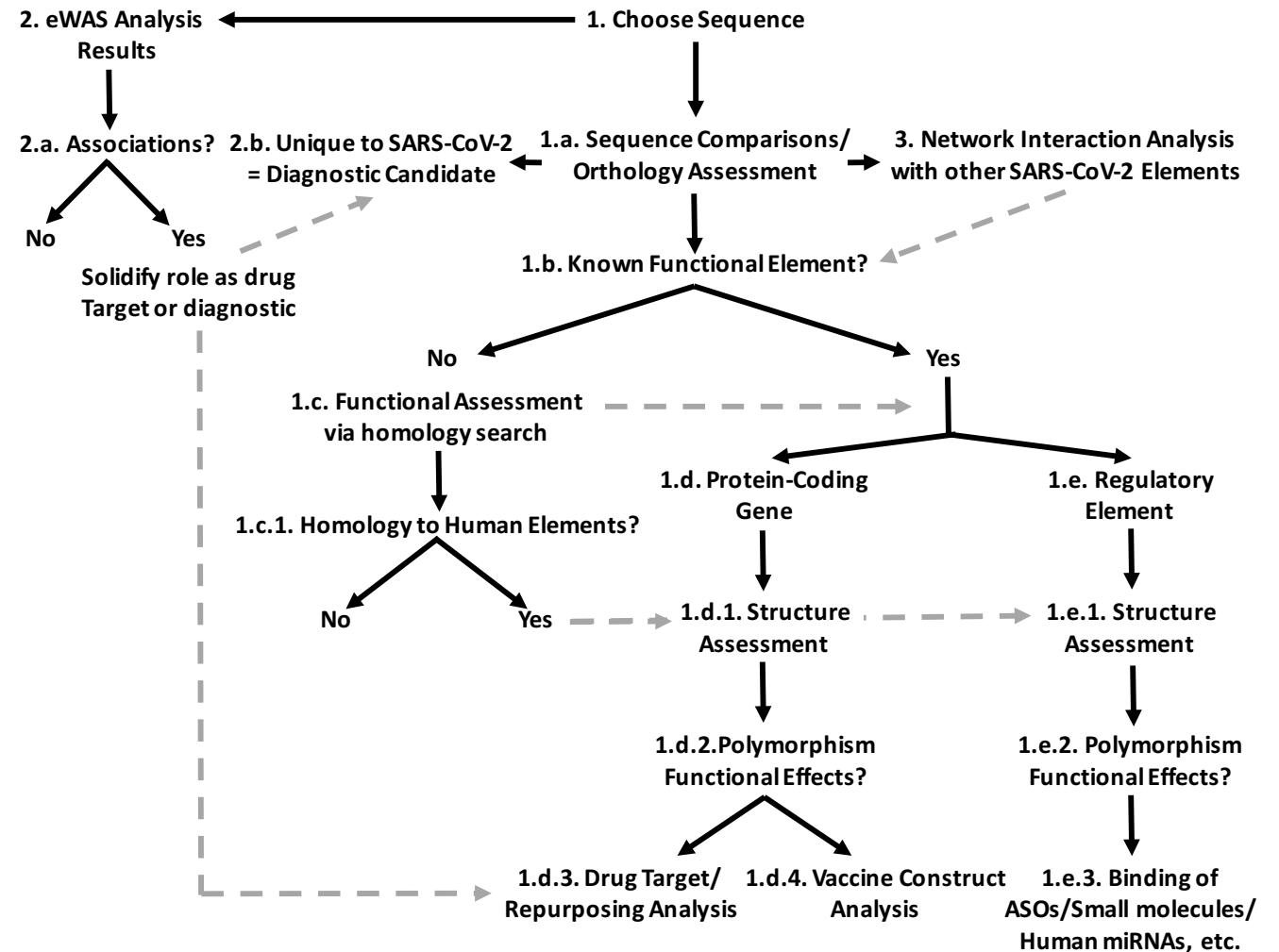
Memories



Comprehensive Interaction Annotation (CIA) Analysis of the SARS-CoV-2 Genome: Basic Workflow

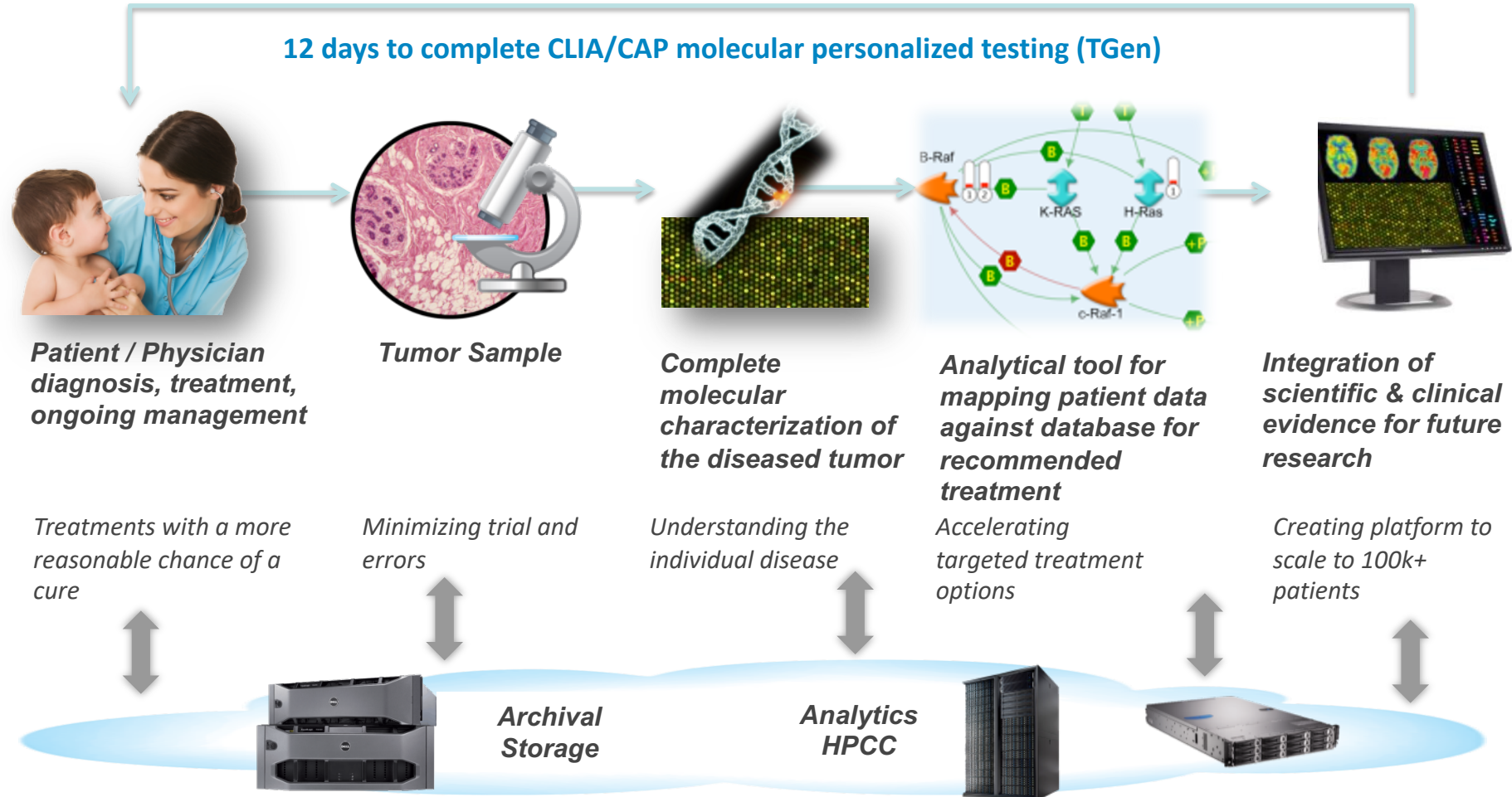
Integrated Computational Analysis

1. Massive data: 10,000 SARS-CoV-2 genomes + 1-2 billion other comparator species genomes
 - 1.a. Orthology/homology assessments
 - 1.b. Known or unknown functional element?
 - 1.c. If not known, evaluate similarity to other elements
 - 1.d. Protein coding gene characterization
 - 1.d.1. Protein structure determination
 - 1.d.2. Protein coding variant annotation
 - 1.d.3. *In silico* drug target studies
 - 1.d.4. Vaccine construct consideration
 - 1.e. Regulatory element characterization
 - 1.e.1. Regulatory element (RNA) structure determination
 - 1.e.2. Regulatory variant annotation
 - 1.e.3. *In silico* RNA targeting studies
 - 1.e. Evaluate polymorphism effects
2. eWAS and phylogenetic signal analysis
3. Network analysis of SARS-CoV-2 elements and human host interacting elements



Applying Precision Treatments for Cancer

- Infrastructure & processes support data generated at every point of the personalized diagnosis and treatment workflow
- Every application optimized to work with the cloud and accelerate analytics and treatment recommendations



COH GEMINI CLUSTER

COMPUTE

28 R640 nodes - 1008 Xeon Gold 6154 cores
18 C4140 (GPU) nodes – 576 Xeon Gold 6142 cores
6 R640 nodes (Mgmt) nodes
46 nodes - 1584 Intel x86 cores

GPU

18 C4140 (4NVIDIA V100)

Virtual Environment

4 Dell VxRail HCI servers
224 Intel x86 cores
1.5 TB total RAM

NETWORKING

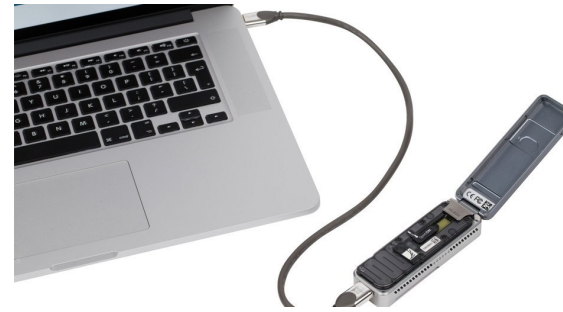
10Gb Ethernet to clients
100Gb Ethernet Core switching



STORAGE

288 TB F800 Flash DellEMC Isilon
960 TB H500 DellEMC Isilon
600 TB Weka NVMe Flash

ITS COMING....

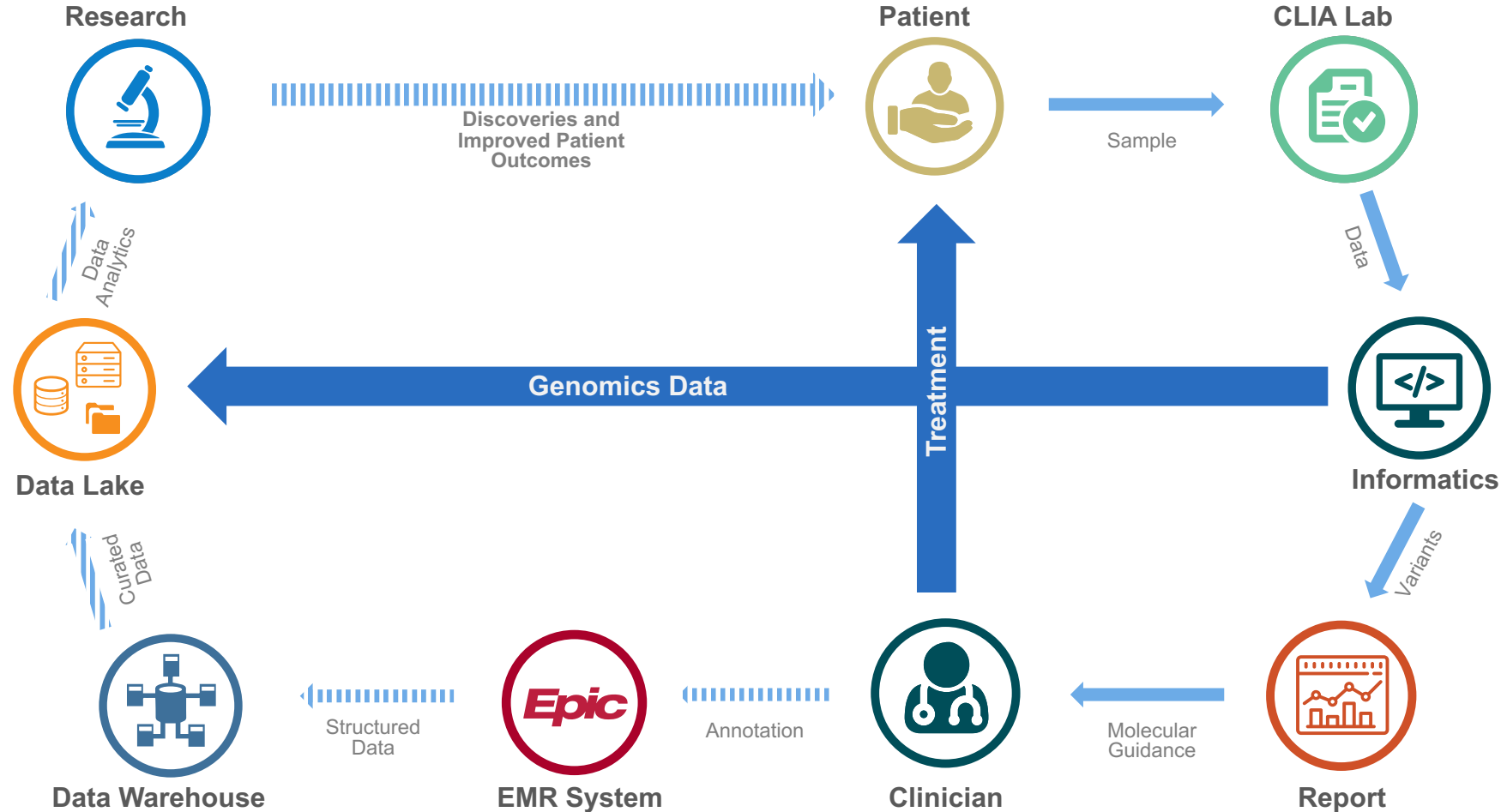


“as much as 2–40 exabytes of storage capacity will be needed by 2025 ”

“Variant calling on 2 billion genomes per year, with 100,000 CPUs in parallel, would require methods that process 2 genomes per CPU-hour”

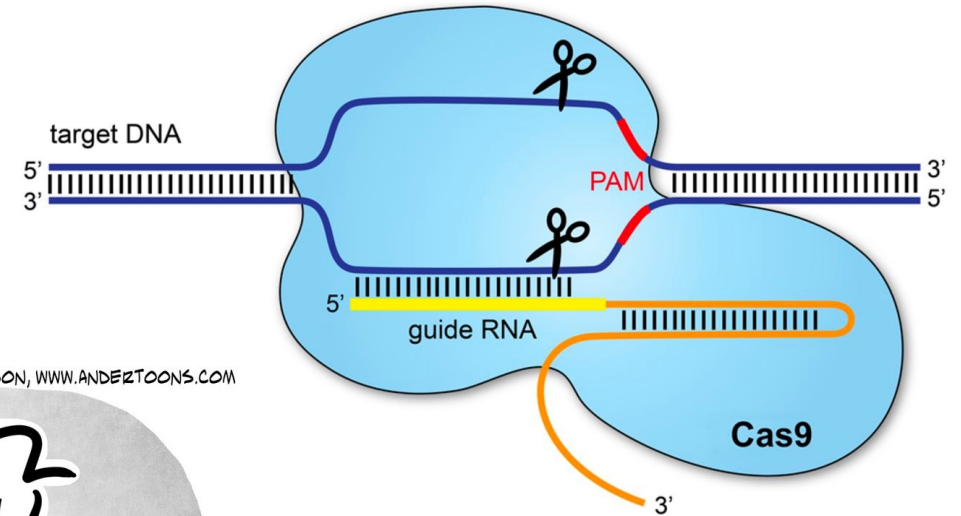
<https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002195#pbio.1002195.s004>

Constant Feedback



Compliance, Ethics and Regulations

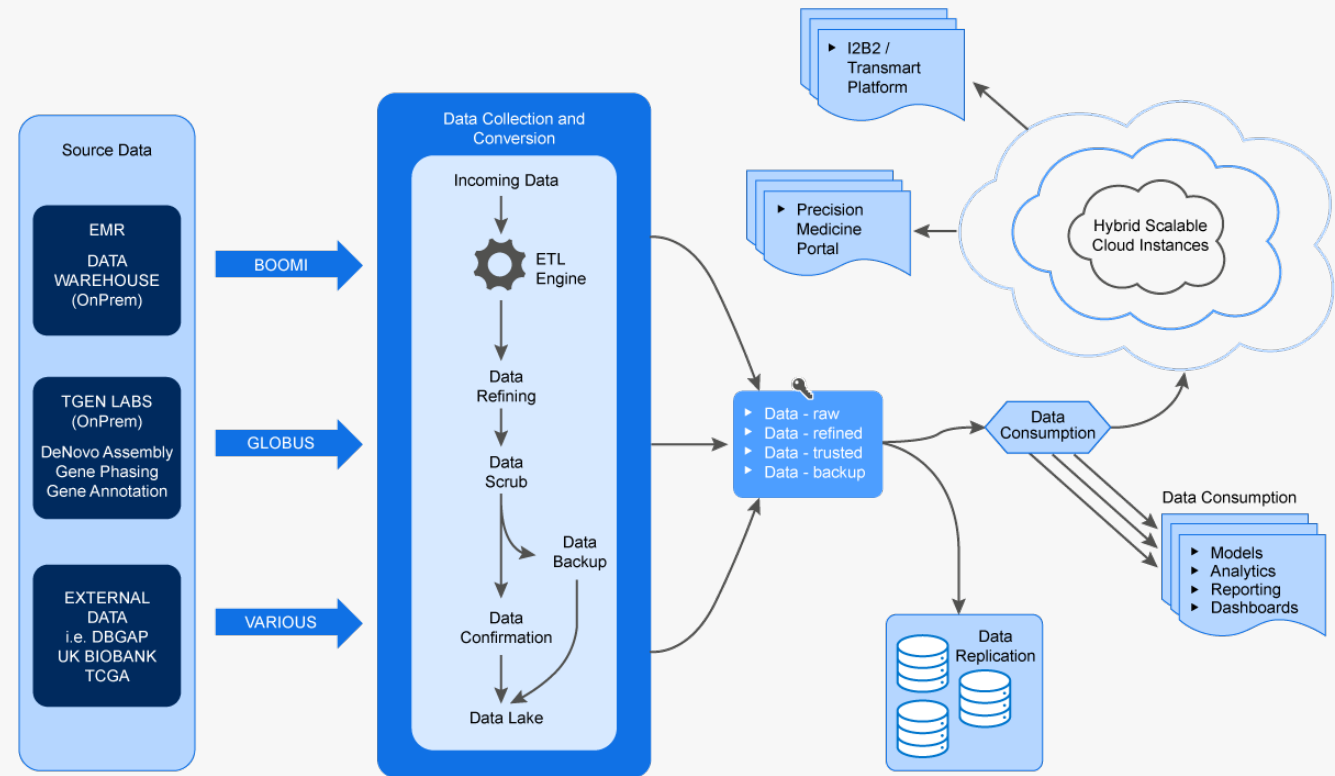
- Gene Editing
- CAR-T
- Rare disorders
- The pace of change
- Government(s)
- Privacy concerns



"I like the privacy, but it does make it hard to see."

Multi-Cloud Ecosystem

- Expand TGen's current hyperconverged and High-Performance Computing (HPC)
 - Integrate cloud-based technologies: **NGS pipeline ported to "cloud native" form, running on AWS, Azure and Google**
- Process vast quantities of data: **Working with PetaGene for data compression and encryption to facilitate storage and transfer of large data sets**
 - Exploiting data that resides in any cloud-based infrastructure (including public clouds and hybrid or private clouds)
- Support, e.g., Kubernetes and container-based computing:
 - **Tools for Deploying hybrid Slurm and Kubernetes clusters for container-based computing**
- Leverage different computational architectures including x86, ARM, GPU, and FPGA based processors: **NGS pipeline ported to run on GPU**
- Storage is critical, different tiers for different data / processes: **Cloud native software defined storage is desirable, i.e. Weka on AWS**
- Tech partners; e.g., Intel, NVIDIA, etc.: **Awarded Intel COVID-19 grant for computational resources to build out SARS-CO-V2 data processing portal at PhoenixNAP (PNAP)**



THANK YOU

TGen HPC Team

Weka

Forward thinking
Scientists and
Doctors

