# DRIVING DISCOVERY IN IMMUNO-ONCOLOGY THROUGH DATA SHARING

*Presented By*

Shawn M. Sweeney, PhD

Director, AACR Project GENIE Coordinating Center

- WHAT IS GENIE
  - GOAL: link clinical genotypes to clinical outcomes to improve clinical decision making, and drive clinical and translational research.
  - The GENIE registry was built by aggregating clinical-grade sequencing data from 8 international sites.
  - Virtual cohorts are then built to answer clinical questions and the data abstracted from the EHR through a federated model.
  - Driven by openness, transparency, and inclusion.

- PARTNERING MODELS
  - Philanthropy
  - Sponsored research of single studies
  - Broader collaborative projects (disease registries, etc.)

American Association for Cancer Research

FINDING CURES TOGETHER℠

American Association for Cancer Research

FINDING CURES TOGETHER℠

DFCI
GRCC
JHU
MDA
MSK
NKI
UHN
VICC

**A**

Clinical Sequencing

regular data uploads

Synapse
CONTRIBUTE to the CURE

- Data mapped to common ontology and harmonized
- Limited PHI removed
- Data governance, provenance, and versioning in a secure, HIPAA-compliant environment.

cBioPortal for Cancer Genomics

Institution-only access
6 months

Consortium-only access
6 months

www.aacr.org/genie/data

**B** clinical queries are posed based on registry content

clinical data required to answer the question are manually abstracted

genomic and clinical data linked

Consortium/sponsor-only access
6 months to time of publication

American Association
for Cancer Research

**FINDING CURES TOGETHER℠**

**GENOMICS**

✓ **Somatic Tumor DNA**

**PHENOMICS**

**Tumor type**
**Histology**
**Demographics**
**Vital status**

**Sponsored Research**

**PHENOMICS**

**Tumor type**
**Histology**
**Demographics**
**Vital status**
**Detailed Clinicopathology**
**Prior Tx**
**Outcomes**

**47,500 Tumors**
**8 Cancer Centers**

**Data made publicly available 12 months after date of sequencing**

**Specific Cohorts**
**Variable # of Centers**

**Data made public at time of publication**

# Plus Associated Biospecimens

- BAM files
  - Develop & test new analytic models/pipelines
- Extracted nucleic acid libraries
  - Perform new analyses (WES)
  - TCRseq
- Tissue blocks/cores
  - RNA
  - Additional IHC/other staining protocols
  - Additional tissue processing
- Stained slides
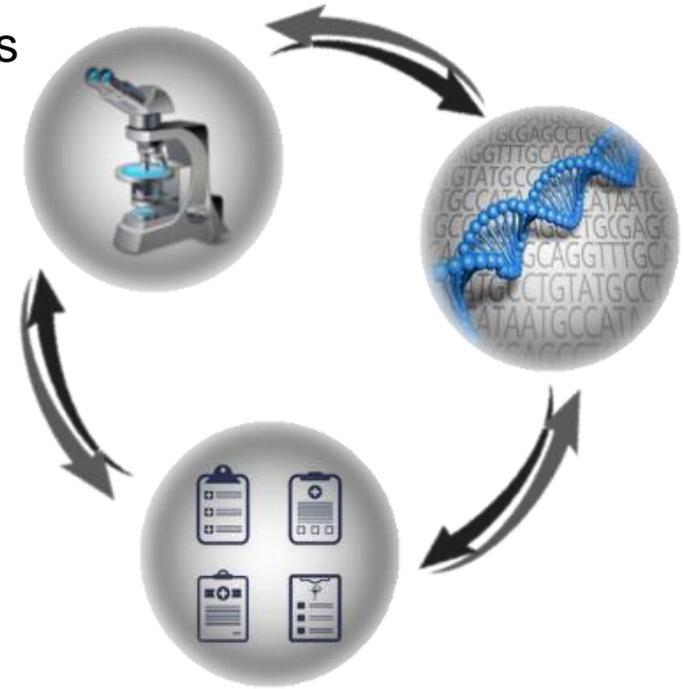
- Microsatellite Instability (MSI)
  - MSI Sensor; could apply other algorithms

- DNA Mismatch Repair Deficiency (dMMR)
  - **MLH1 (1.5%), PMS2 (1.9%), MSH2 (2.2%), MSH6 (2.7%)**, MLH3, MSH3, PMS1, Exo1, and POLE
  - Mutated in ~6% of patients in the GENIE cohort

- Tumor Mutation Burden (TMB)
  - Based on panels sequencing ≥ 750 kb
  - Reported as score per patient and the TMB range for the cancer type

- Currently have active projects correlating calculated results with SOC testing (PCR and IHC) as well as outcomes to immune checkpoint blockade.

**13, 375 patients
351 MSI-H**

| Cancer Type | Number | Average | MSI-High | Fraction |
|---|---|---|---|---|
| Endometrial Cancer | 525 | 4.63 | 90 | 17.1% |
| Colorectal Cancer | 1108 | 3.74 | 105 | 9.5% |
| Soft Tissue Sarcoma | 593 | 1.44 | 30 | 5.1% |
| Esophagogastric Cancer | 326 | 2.11 | 16 | 4.9% |
| Bladder Cancer | 369 | 1.42 | 14 | 3.8% |
| Prostate Cancer | 996 | 0.97 | 19 | 1.9% |
| Cancer of Unknown Primary | 583 | 1.24 | 11 | 1.9% |
| Germ Cell Tumor | 284 | 1.38 | 5 | 1.8% |
| Gastrointestinal Stromal Tumor | 172 | 1.12 | 3 | 1.7% |
| Mesothelioma | 128 | 0.66 | 2 | 1.6% |
| Thyroid Cancer | 215 | 0.85 | 3 | 1.4% |
| Hepatobiliary Cancer | 379 | 0.73 | 5 | 1.3% |
| Non-Small Cell Lung Cancer | 2137 | 0.75 | 27 | 1.3% |
| Ovarian Cancer | 412 | 1.48 | 5 | 1.2% |
| Glioma | 627 | 0.70 | 6 | 1.0% |
| Melanoma | 648 | 0.72 | 5 | 0.8% |
| Pancreatic Cancer | 840 | 0.55 | 4 | 0.5% |
| Breast Cancer | 2404 | 0.81 | 11 | 0.5% |
| Head and Neck Cancer | 206 | 0.46 | 0 | 0.0% |
| Renal Cell Carcinoma | 292 | 0.40 | 0 | 0.0% |
| Skin Cancer, Non-Melanoma | 131 | 0.37 | 0 | 0.0% |
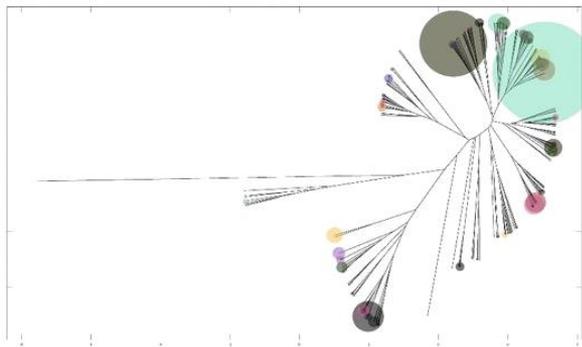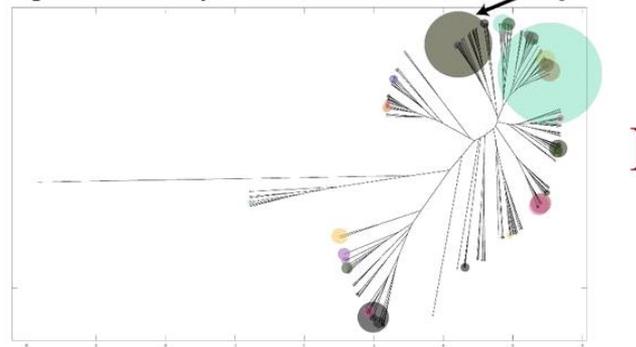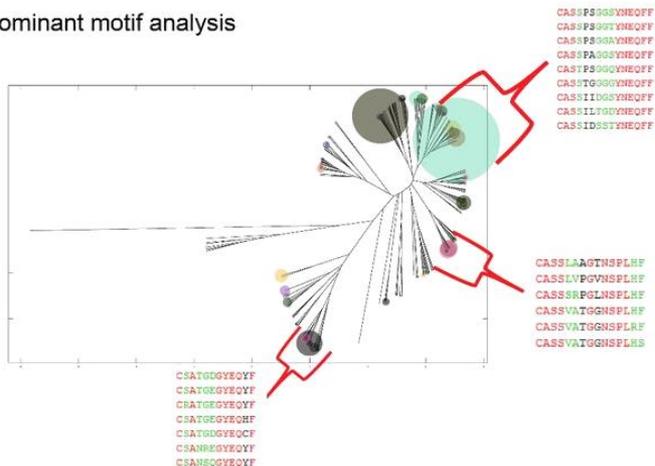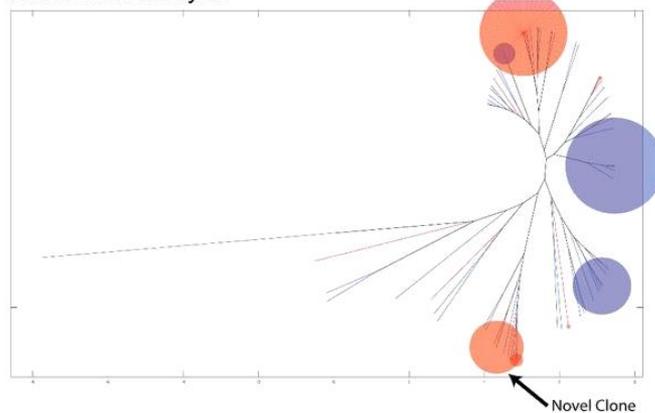
Generates a clinically predictive signature as compared to TCRseq alone

# Extracting New Information From Existing Slides
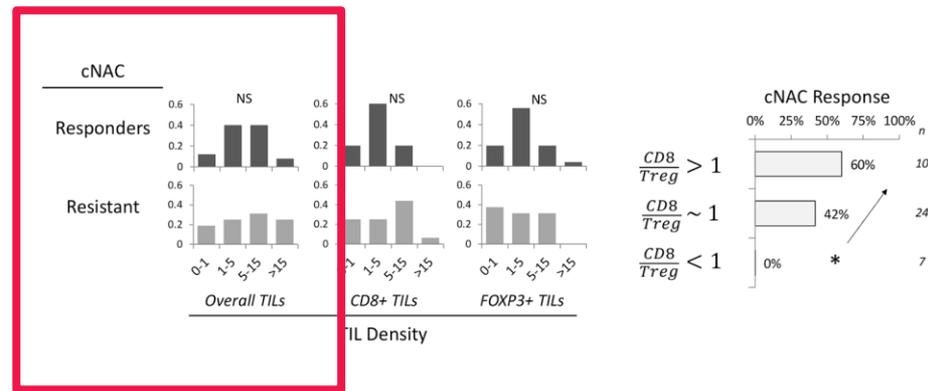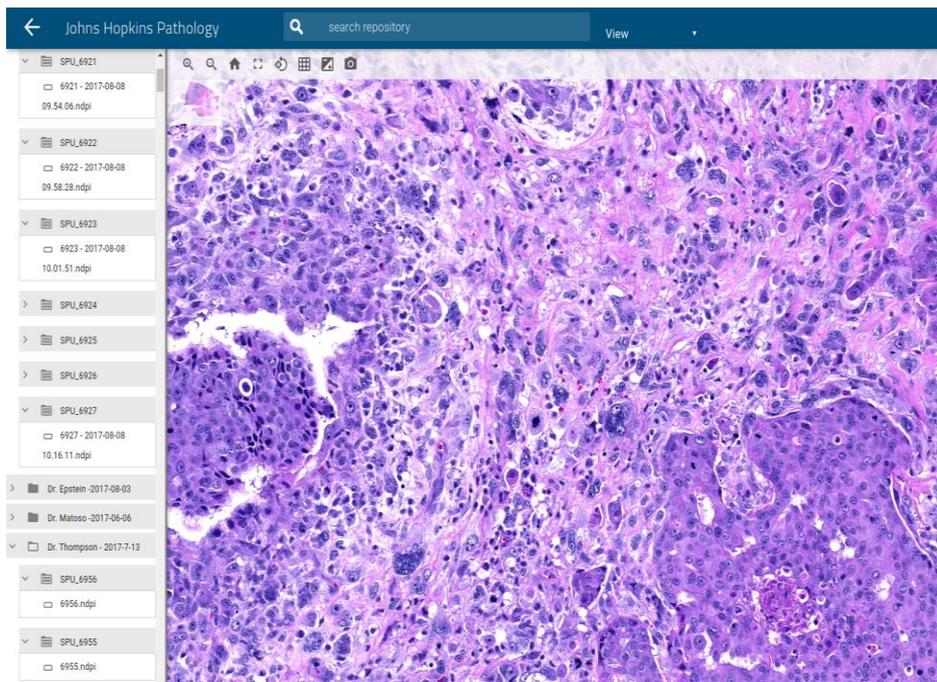
Platinum-based neoadjuvant chemotherapy
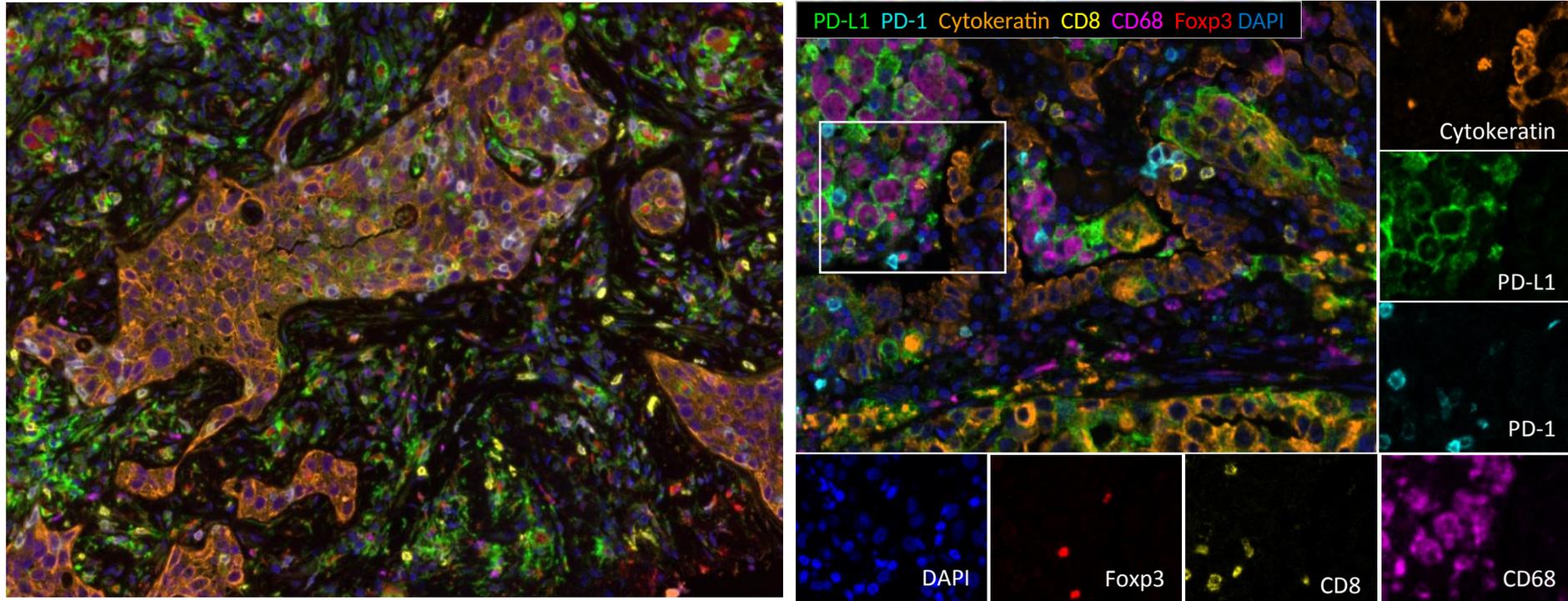
Baras et al. Oncoimmunology (PMC4910705)

*Courtesy of J. Taube and colleagues: Johns Hopkins*

AACR American Association for Cancer Research
**FINDING CURES TOGETHER℠**

- AACR Project GENIE is an international cancer registry formed through data sharing and contains data from 47,000+ sequenced tumors.

- Each sequenced tumor has an associated limited clinical data set.
  - Working to enhance the clinical data collected as part of the baseline.

- In addition to the genomic and clinical data, the BAM files; nucleic acid libraries; stained slides; and in many cases, tissue, can be used to drive further discovery.

- These data taken together with appropriate clinical and pathologic endpoints derived from patient EHRs and related clinical reports will improve patient treatment and outcomes.