The Cancer Genome Atlas (TCGA) is a project of the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) to develop a comprehensive molecular atlas of cancer.

TCGA’s integrated database of molecular and clinical annotation will provide scientists with accelerated access to the discovery with the end goal being better diagnostics, therapies, and preventive measures for cancer.
TCGA General Eligibility Criteria

- **General Biospecimen Criteria**
  - Primary, *de novo* tumor
  - Invasive cancer, not *in situ*
  - Metastatic and/or recurrences encouraged in addition to primary tumor
  - Diagnosis consistent with TCGA list of eligible tumors and subtypes
  - Prior malignancies may not have had prior systemic chemotherapy or local radiation in same field
  - Tumor Nuclei ≥60%; Tumor necrosis ≤20%
  - Blood as germline control for most cases

- **Additional Requirements**
  - Complete, de-identified pathology report in English
  - Completion of required clinical data elements
  - Submission of pre-screen frozen sections and diagnostic slides
Eligibility Exceptions

- Metastatic melanoma tumors accepted if no primary tumor

- ‘Acceptable’ prior treatments
  - BCG for bladder
  - Interferon for melanoma
  - Hydroxyurea for acute myeloid leukemia
TCGA Tumor Studies – Open

- Acute Myeloid Leukemia
- Bladder Urothelial Carcinoma
- Breast Invasive Carcinoma*
- Cervical Carcinoma
- Diffuse Large B-Cell Lymphoma
- Esophageal Carcinoma
- Head and Neck Carcinoma

- Hepatocellular Carcinoma
- Kidney Papillary Carcinoma
- Lower Grade Glioma
- Melanoma
- Pancreatic Adenocarcinoma
- Prostate Adenocarcinoma
- Stomach Adenocarcinoma
- Sarcoma

* Open only to non-ductal histologies
TCGA Tumor Studies – Closed

- Colon Adenocarcinoma
- Endometrial Carcinoma
- Glioblastoma Multiforme
- Kidney Clear Cell Carcinoma
- Lung Adenocarcinoma
- Lung Squamous Cell Carcinoma
- Ovarian Serous Cystadenocarcinoma
- Rectal Adenocarcinoma
- Thyroid Papillary Carcinoma

NOTE - All studies are open to cases from patients of African descent
TCGA Rare Tumor Studies

- Adrenocortical Carcinoma
- Chromophobe Kidney*
- Mesothelioma
- Pheochromocytoma/Paraganglioma
- Uterine Carcinoma*
- Acute Lymphoblastic Leukemia
- Anaplastic Thyroid
- Cholangiocarcinoma
- Malignant Peripheral Nerve Sheath Tumors
- Myelodysplastic Syndrome (MDS)
- Small Cell Lung Cancer
- Testicular Germ Cell Tumor

* Closed to new accrual
TCGA: “No Platform Left Behind”

36 cancer types

- glioblastoma multiforme (brain)
- squamous carcinoma (lung)
- serous cystadenocarcinoma (ovarian)

Biospecimen Core Resource with more than 150 Tissue Source Sites

- 5 Cancer Genomic Characterization Centers
- 3 Genome Sequencing Centers
- 7 Genome Data Analysis Centers

Data Coordinating Center

Multiple data types

- Clinical diagnosis
- Treatment history
- Histologic diagnosis
- Pathologic report/images
- Tissue anatomic site
- Surgical history
- Gene expression/RNA sequence
- Chromosomal copy number
- Loss of heterozygosity
- Methylation patterns
- miRNA expression
- DNA sequence
- RPPA (protein)
- Subset for Mass Spec

The Cancer Genome Atlas
TCGA: “No Platform Left Behind”

36 cancer types

glioblastoma multiforme (brain)

squamous carcinoma (lung)

serous cystadenocarcinoma (ovarian)

Etc. Etc. Etc.
TCGA: “No Platform Left Behind”

36 cancer types

- Glioblastoma multiforme (brain)
- Squamous carcinoma (lung)
- Serous cystadenocarcinoma (ovarian)

Multiple data types

- Clinical diagnosis
- Treatment history
- Histologic diagnosis
- Pathologic report/images
- Tissue anatomic site
- Surgical history
- Gene expression/RNA sequence
- Chromosomal copy number
- Loss of heterozygosity
- Methylation patterns
- miRNA expression
- DNA sequence
- RPPA (protein)
- Subset for Mass Spec
TCGA: “No Platform Left Behind”

36 cancer types

- glioblastoma multiforme (brain)
- squamous carcinoma (lung)
- serous cystadenocarcinoma (ovarian)

Biospecimen Core Resource with more than 150 Tissue Source Sites

5 Cancer Genomic Characterization Centers

3 Genome Sequencing Centers

7 Genome Data Analysis Centers

Data Coordinating Center

Multiple data types

- Clinical diagnosis
- Treatment history
- Histologic diagnosis
- Pathologic report/images
- Tissue anatomic site
- Surgical history
- Gene expression/RNA sequence
- Chromosomal copy number
- Loss of heterozygosity
- Methylation patterns
- miRNA expression
- DNA sequence
- RPPA (protein)
- Subset for Mass Spec

The Cancer Genome Atlas
<table>
<thead>
<tr>
<th>Platform</th>
<th>Pilot (GMB, OV)</th>
<th>Expansion</th>
</tr>
</thead>
</table>
| SNP/copy number | Affymetrix SNP 6.0  
                   | Agilent CGH Array  
                   | Illumina 1M Duo  
                   | Affymetrix SNP 6.0  
                   | Low-pass sequencing |
| DNA methylation | Infinium 4500K array                               | Infinium 4500K array                                   |
| mRNA         | Agilent 244K Array  
                   | Affymetrix Human Exon Array  
                   | Affymetrix U133 Array  
                   | RNAseq                                             |
| miRNA        | Agilent 8X15K Array                                 | RNAseq                                                 |
| Mutation     | 600-1000 genes                                     | Illumina exome sequencing  
                   | Illumina whole genome sequencing (10%)            |
| Proteomics   | N/A                                                  | Reverse-phase protein array; mass spectrometry         |
TCGA Network

TCGA Network Key:
- Data Coordinating Center (DCC)
- National Cancer Institute (NCI)
- Biospecimen Core Resource (BCR)
- Genomics Characterization Center (GCC)
- Genomics Sequencing Center (GSC)

- British Columbia Cancer Agency (GCC) Vancouver, British Columbia
- University of Southern California (GCC) Los Angeles, California
- MD Anderson Cancer Center (GCC) Houston, Texas
- Baylor College of Medicine (GSC) Houston, Texas
- Washington University (GSC) St. Louis, Missouri
- Nationwide Children’s Hospital (BCR) Columbus, Ohio
- National Cancer Institute (NCI) Bethesda, Maryland
- The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins (GCC) Baltimore, Maryland
- SRA International, Inc. (DCC) Fairfax, Maryland
- Broad Institute (GCC and GSC) Cambridge, Massachusetts

The Cancer Genome Atlas
The BCR Role

- Onboarding and Training of Tissue Source Sites (TSS)
  - IRB/MTA coordination
  - OpenClinica training (web-based clinical data capture system)
  - On-site training and shipment coordination
- Pathology Validation
  - Confirmation of acceptable TCGA histology
  - Verification of tumor nuclei and necrosis metrics via BCR pathologists
- Molecular Processing
  - Standardized co-isolation of DNA and RNA
  - Quality control, including RNA integrity determination and identity panel
  - Distribution of nucleic acids to Genome Characterization Centers (GCCs) and Genome Sequencing Centers (GSCs)
- Clinical Data Collection
  - Development of clinical data forms for TCGA studies
  - Coordination, collection and quality control of clinical data
  - Electronic transfer to Data Coordinating Center (DCC)
  - Assistance with TCGA clinical data analysis
BCR Pipeline

Genome Characterization Centers (GCC) & Genome Sequencing Centers (GSC)

Tissue Source Sites

Initial Qualification

Clinical Outreach

Logistics

Molecular Qualification

Molecular

Informatics

Pathology Qualification

Pathology & Virtual Microscopy

Histology

DCC

The Cancer Genome Atlas
Clinical Form Development

- **Form Creation**
  - Generic forms are modified by Disease Working Groups for each tumor type

- **Electronic Build**
  - Wire-frame forms are used as a tool to communicate and build studies.
  - Electronic forms are developed and launched in the OpenClinica system

- **Form Completion by TSSs**
  - Tissue Source Sites complete Case Report Forms (CRFs) using OpenClinica

- **Data Review and Quality Control**
  - All forms are reviewed for accuracy by the BCR, who works with TSSs to correct discrepancies.
Clinical Data Forms

- **Case Quality Control Form** *(required for all cases)*
  - Information about the submitted samples

- **Enrollment Form** *(required for all qualified cases)*
  - Disease-specific information and detailed patient information

- **Follow-up Form(s)** *(required for all living patients)*
  - Updated information not received at time of enrollment

- **Supplemental Treatment Form** *(optional data)*
  - Pharmaceutical Form: Chemotherapy, immunotherapy, targeted therapy, and hormone therapy given for the submitted malignancy
  - Radiation Form: Radiation treatment given for the submitted malignancy
  - Ablation/Embolization Form

- **Other Malignancy Form** *(if applicable)*
  - Captures diagnosis and treatment information related to prior malignancies and synchronous primaries

[Website Link](http://www.nationwidechildrens.org/biospecimen-core-resource-for-the-cancer-genome-atlas)
Minimum Clinical Data Set – Enrollment

1. TSS name/identifier/patient #
2. Primary site of disease
3. Histological type
4. Tumor grade
5. Gender
6. Date of birth (month/year)
7. Prior diagnosis of prior neoplasm (if yes, another form required)
8. Date of initial pathologic diagnosis (month/year)
9. Neoadjuvant therapy
10. Tumor stage
11. Date of last contact (month/year)
12. Vital status (month/year)
13. Date of death, if applicable (month/year)
14. New tumor events
15. Tumor-specific elements developed by relevant Disease Working Group

The Cancer Genome Atlas
Minimum Clinical Data Set – Follow-Up

1. Type of treatment incurred
2. Measure of success to treatment
3. Recurrences, progression since last form completed
4. New tumor events
5. Additional surgeries
6. Date of last contact (month/year)
7. Vital status (month/year)
8. Date of Death, if applicable
Examples of Tumor Specific Questions

- Smoking status (lung/bladder)
- ER/PR/Her2+ (breast)
- Biochemical recurrence (prostate)
- Residual tumor after surgery (ovarian/sarcoma)
- Prior melanoma diagnosis details (melanoma)
- International Prognostic Index (lymphoma)
TCGA Data Standardization

- All data elements registered by the Cancer Data Standards Registry (caDSR)
- Common Data Elements (CDEs) – caDSR identifiers are assigned to each TCGA data question
- Standardization allows cross comparison of data elements across TCGA and non-TCGA studies
- Disease site coding standardization according to International Classification of Diseases, version 10 (ICD-10 Coding)
- ICD-O-3 histology coding
Quality Control of Clinical Data

- OpenClinica Validations
- BCR Validation Engine
- Manual Review

As the BCR reviews cases, common issues lead to future improvements, which are always centered around standardizing data across forms and disease studies.
XML

- **EXtensible Markup Language**
- Common tool for data transmissions
- Designed to transport and store data
- Allows author to define structure
- Each data element is represented by a string of data tags containing information about the specific element
<shared:vital_status cde="2939553" display_order="22" owner="TSS" preferred_name="Vital_Status" procurement_status="Completed" restricted="false" source_system_identifier="674513" tier="1" xsd_ver="1.8">DECEASED</shared:vital_status>
Information Provided in the XML

<table>
<thead>
<tr>
<th>Information Tag</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Name</td>
<td>Description of the data element provided and whether it’s generic vs. disease specific</td>
<td><code>&lt;shared:vital_status&gt;</code></td>
</tr>
<tr>
<td>CDE Identifier</td>
<td>ID registered at the caDSR</td>
<td><code>cde=&quot;2939553&quot;</code></td>
</tr>
<tr>
<td>Display Order</td>
<td>Order of the element on the CRF</td>
<td><code>display_order=&quot;22&quot;</code></td>
</tr>
<tr>
<td>Owner</td>
<td>Entity that provided the information</td>
<td><code>owner=&quot;TSS&quot;</code></td>
</tr>
<tr>
<td>Preferred Name</td>
<td>Simplified version of “Data Name”</td>
<td><code>preferred_name=&quot;vital_status&quot;</code></td>
</tr>
<tr>
<td>Procurement Status</td>
<td>Collection status of the element</td>
<td><code>procurement_status=&quot;completed&quot;</code></td>
</tr>
<tr>
<td>Restricted</td>
<td>Indicates restricted information</td>
<td><code>restricted=&quot;false&quot;</code></td>
</tr>
<tr>
<td>Source System ID</td>
<td>ID used by OpenClinica</td>
<td><code>source_system_id=&quot;674513&quot;</code></td>
</tr>
<tr>
<td>Tier</td>
<td>Indicates whether the element was required</td>
<td><code>tier=&quot;1&quot;</code></td>
</tr>
<tr>
<td>XSD Version</td>
<td>Version of XSD used to generate XML</td>
<td><code>xsd_ver=&quot;1.8&quot;</code></td>
</tr>
<tr>
<td>Answer Provided by TSS</td>
<td>Information provided by the TSS for this case</td>
<td><code>&lt;DECEASED&gt;&lt;/shared:vital_status&gt;</code></td>
</tr>
</tbody>
</table>
Biospecimen XML vs. Clinical XML

- The BCR generates two XML files
- Biospecimen XML
  - Uploads within 24 hours of nucleic acids shipping to GCCs and GSCs
  - BCR Quality Metrics
  - Select Data from Case Quality Control Form
- Clinical XML
  - Uploaded once enrollment form data has passed BCR quality control
  - Select Data from Case Quality Control Form
  - Enrollment Form Data
  - Follow-up Form(s) Data
  - Treatment Form(s) Data
Biotabs

- DCC parses XML uploaded by the BCR
- Data extracted, ‘flattened,’ and put into .txt files
- Can be opened as a tab-delimited file in Excel
- Some data loss
- Some data not required for all tumor types [Not Available]

<table>
<thead>
<tr>
<th>bcr_patient_barcode</th>
<th>age_at_initial_pathologic_diagnosis</th>
<th>age_began_smoking_in_years</th>
<th>ajcc_cancer_metastasis_stage_code</th>
<th>ajcc_cancer_staging_handbook_edition</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCGA-BL-00C8</td>
<td>78</td>
<td>[Not Available] MX</td>
<td>[Not Available] [Not Available]</td>
<td>[Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-01ST</td>
<td>57</td>
<td>7th N0</td>
<td>Stage III</td>
<td>T3 Bladder, M0 ebc99a6-6c9a-4c0b-8ca7-270a5c37804a [Not Available] [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>65</td>
<td>[Not Available] N2</td>
<td>Stage IV</td>
<td>T4 Wall, Posterior 556f0b6-172e-4af1-8922-ae03688d8e8 [Not Available] [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>D2</td>
<td>7th N0</td>
<td>Stage III</td>
<td>T3 Wall, Posterior 5f4f205-813c-1a85-97b5-77942e960060 [Not Available] [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>62</td>
<td>[Not Available] N0</td>
<td>Stage IV</td>
<td>[Not Available] [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>75</td>
<td>[Not Available] N0</td>
<td>Stage III</td>
<td>T3a Bladder, M0 67b0786-9866-494e-b28a-5e2a9893626a [Not Available] [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>75</td>
<td>[Not Available] N0</td>
<td>Stage III</td>
<td>T3b Bladder, M0 22479a4-4ff3-974e-8485-6c109d5f15e0 [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>75</td>
<td>[Not Available] N0</td>
<td>Stage III</td>
<td>T3a Wall, Posterior 7ab1e7d-f2d0-4d66-8358-38b32d95f27e [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>73</td>
<td>[Not Available] N0</td>
<td>Stage III</td>
<td>T3a Trioge ebeda3-433a-459a-870a-6e059d183f99e [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>75</td>
<td>[Not Available] N0</td>
<td>Stage III</td>
<td>T3a Wall, Lateral 10e79b7-9cb8-4eb9-8b9b-f4d4-914b-3d7d5a1ce34 [Not Available] High Grade</td>
</tr>
<tr>
<td>TCGA-BL-015J</td>
<td>75</td>
<td>[Not Available] N0</td>
<td>Stage III</td>
<td>T3a [Not Available] 6a03d8b-6b6c-57a2-970a-f0e5d193f848 [Not Available] High Grade</td>
</tr>
</tbody>
</table>

NOTE: The BCR is currently working to improve the column headers on the TCGA Data Portal “biotab” files. The goal is to designate a single column header, across studies and forms, for all questions with the same meaning.
TCGA Public-Access Data

- Data are de-identified
- Molecular platforms:
  - SNP/copy number – level 3 SNP data
  - DNA methylation
  - mRNA
  - miRNA
  - Mutation – level 3 sequencing data
  - Proteomics
- Clinical data
TCGA Public-Access Data

- Redacted pathology reports
- Aperio whole slide images
  - Frozen section of tissue used for extractions
  - Formalin-fixed, paraffin-embedded H&E from a representative diagnostic block
- MSI testing (level 1, 2, and 3)
  - COAD and READ
  - STAD in the future
- HPV testing
  - Currently validating assay and will test HNSC cases
  - CESC in the future
TCGA Data Portal – Pathology Reports

- Synoptic Information:
  - Anatomic site
  - Histologic diagnosis
  - Laterality
  - Lymph node status
  - Tumor size(s)
  - Number of lesions
  - Tumor stage

- Additional Information:
  - ICD-O-3 histology + site coding
  - International reports translated to English
TCGA Public-Access Data

- Whole slide frozen sections 40X images
- Whole slide FFPE Diagnostic 40X images
- Aperio .svs file format
- JPEG compression
- Quality factor of 30
TCGA Controlled-Access Data

- Data Access Request is required
- Contains data that are unique to an individual
  - Individual germline variant data (SNP .cel files)
  - Primary sequence data - available at the Cancer Genomics Hub (https://cghub.ucsc.edu/)
  - Exon array data (applies to the Glioblastoma Multiforme and Ovarian Serous Cystadenocarcinoma projects only)
Annotations Manager

- For when things don’t exactly fit
- Example annotations include:
  - Neoadjuvant therapy
  - History of prior malignancy
  - Subject withdrew consent
  - Synchronous malignancy
  - Redactions
TCGA Number Decoded

TCGA-BL-A123-01A*-11D*-A123*-02

**PROJECT NAME**
The first part of the number represents the current project name. Currently, TCGA is the only project.

**TISSUE SOURCE SITE (TSS)**
This will be a numeric/alpha-numeric ID unique to each cancer-type at each TSS.

**PATIENT IDENTIFIER**
This ID is unique to each patient and is associated with a site ID.

**SAMPLE IDENTIFIER**
The sample ID represents both the sample type and the sample vial, as follows:

- **Sample type:**
  - Two digit # 01-29 (see details below)
  - #'s 01-09 = tumor types
  - #'s 10-19 = normal types
  - #'s 20-29 = control samples

- **Vial Identifier:**
  An alphabetic character represents the vial of a specific specimen. The values range from A-Z.

**PLATE IDENTIFIER**
Identifies individual plates associated with the sample.

**CENTER IDENTIFIER**
Identifies each of the Genomic Sequencing Centers (GSCs) and Cancer Genome Characterization Centers (CGCCs).

**PORTION IDENTIFIER**
The portion ID consists of a portion code and an analyte code.

- **Portion Code:** A two-digit number identifying the portion
- **Analyte Code:** Alphabetic code representing an analyte type (i.e., D for DNA). (see details below)

**Sample ID Details**
- 01 Solid Tumor
- 12 Buccal Smear
- 10 Normal Blood
- 20 Cell Line
- 11 Normal Tissue

**IMPORTANT NOTES**
* **SLIDES:** When the slides are cut, an additional number will be added to the Vial ID, to determine whether the slide is a top slide (TS1), bottom slide (BS1), or a middle slide (MS1).
* **ANALYTE CODE:** The analyte code will not be added until the specimen has passed the BCR pathology review.
* **PLATE ID:** The plate ID and the center ID will not be added until the specimen has been through the Molecular workflow.

**Analyte Code Details**
- R: RNA
- D: DNA
- T: Total RNA
- W: WGA DNA (Qiagen)
## Tumor Codes

<table>
<thead>
<tr>
<th>Tumor</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute Myeloid Leukemia</td>
<td>LAML</td>
</tr>
<tr>
<td>Bladder Urothelial Carcinoma</td>
<td>BLCA</td>
</tr>
<tr>
<td>Breast Invasive Carcinoma</td>
<td>BRCA</td>
</tr>
<tr>
<td>Cervical Carcinoma</td>
<td>CESC</td>
</tr>
<tr>
<td>Colon Adenocarcinoma</td>
<td>COAD</td>
</tr>
<tr>
<td>Rectal Adenocarcinoma</td>
<td>READ</td>
</tr>
<tr>
<td>Diffuse Large B-Cell Lymphoma</td>
<td>DLBC</td>
</tr>
<tr>
<td>Endometrial Carcinoma</td>
<td>UCEC</td>
</tr>
<tr>
<td>Esophageal Carcinoma</td>
<td>ESCA</td>
</tr>
<tr>
<td>Glioblastoma Multiforme</td>
<td>GBM</td>
</tr>
<tr>
<td>Head and Neck Carcinoma</td>
<td>HNSC</td>
</tr>
<tr>
<td>Hepatocellular Carcinoma</td>
<td>LIHC</td>
</tr>
</tbody>
</table>
## Tumor Codes

<table>
<thead>
<tr>
<th>Tumor</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kidney Clear Cell Carcinoma</td>
<td>KIRC</td>
</tr>
<tr>
<td>Kidney Papillary Carcinoma</td>
<td>KIRP</td>
</tr>
<tr>
<td>Lower Grade Glioma</td>
<td>LGG</td>
</tr>
<tr>
<td>Lung Adenocarcinoma</td>
<td>LUAD</td>
</tr>
<tr>
<td>Lung Squamous Cell Carcinoma</td>
<td>LUSC</td>
</tr>
<tr>
<td>Melanoma</td>
<td>SKCM</td>
</tr>
<tr>
<td>Ovarian Serous Carcinoma</td>
<td>OV</td>
</tr>
<tr>
<td>Pancreatic Adenocarcinoma</td>
<td>PAAD</td>
</tr>
<tr>
<td>Prostate Adenocarcinoma</td>
<td>PRAD</td>
</tr>
<tr>
<td>Stomach Adenocarcinoma</td>
<td>STAD</td>
</tr>
<tr>
<td>Sarcoma</td>
<td>SARC</td>
</tr>
<tr>
<td>Thyroid Papillary Carcinoma</td>
<td>THCA</td>
</tr>
</tbody>
</table>
The TCGA program has established the following policy to clarify freedom of TCGA and non-TCGA users to publish findings using TCGA data. There are no limitations on publications containing analyses using any TCGA data set if the data set meets one of the following three freedom-to-publish criteria: A marker paper has been published on that tumor type; or
- 18 months after 100 cases of a given tumor type have shipped from BCR to characterization and sequencing centers; or
- The author receives specific approval from the TCGA Publication Committee in consultation with appropriate disease-specific analysis group(s).
- Specifically, the status of each tumor dataset is available below. If you have questions, do not hesitate to contact tcca@mail.nih.gov.

TCGA Data Use Policy and Publication Guidelines promote the responsible use of TCGA data sets. All investigators, and their institutions, seeking access and use of TCGA data must acknowledge their agreement with TCGA policies and procedures, which are described in the TCGA Data Portal Data Use Certification.
Accessing TCGA Data

New data derived from TCGA analyses is deposited on a regular basis into the TCGA Data Portal. This process is controlled by the Data Coordinating Center (DCC).

http://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp
Participate in TCGA

We welcome additional collaborators at any time as we recognize the need to involve teams from around the world to acquire the samples and data necessary for a program of this size and scope.

The Cancer Genome Atlas Program Office
National Cancer Institute at NIH
31 Center Drive
Bldg. 31, Suite 3A20
Bethesda, MD 20892
Phone: (301) 594-9831
E-mail: tcga@mail.nih.gov