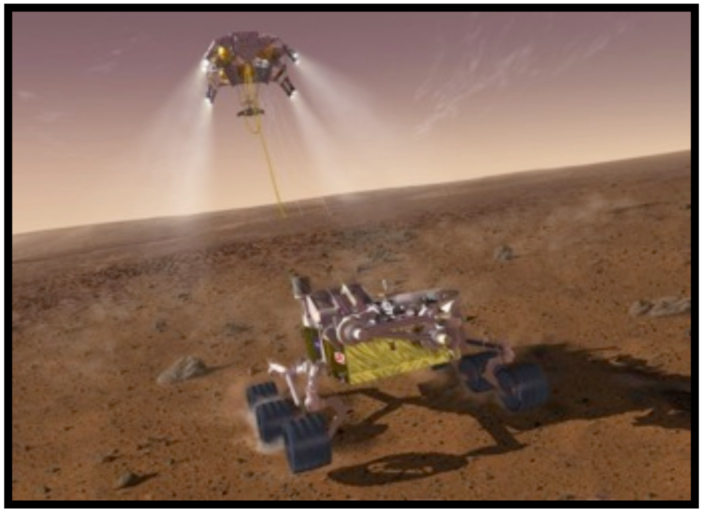
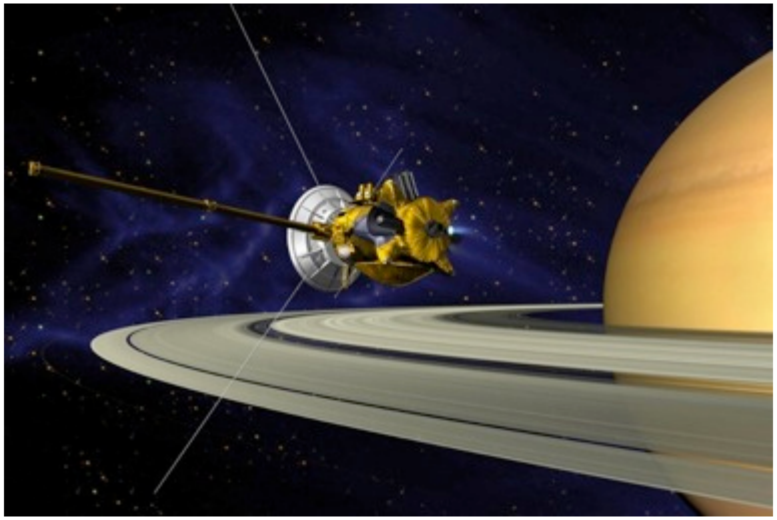
An Extensible and Scalable Knowledge System Architecture for Cancer Research

Daniel Crichton
Dan.Crichton@jpl.nasa.gov
Principal Computer Scientist and Program Manager
Director, Center for Data Science and Technology
Principal Investigator, JPL Informatics Center
NASA Jet Propulsion Laboratory, California Institute of Technology

Thanks to Kristen Anton (Dartmouth), Heather Kincaid (JPL)

NASA Big Data Challenges



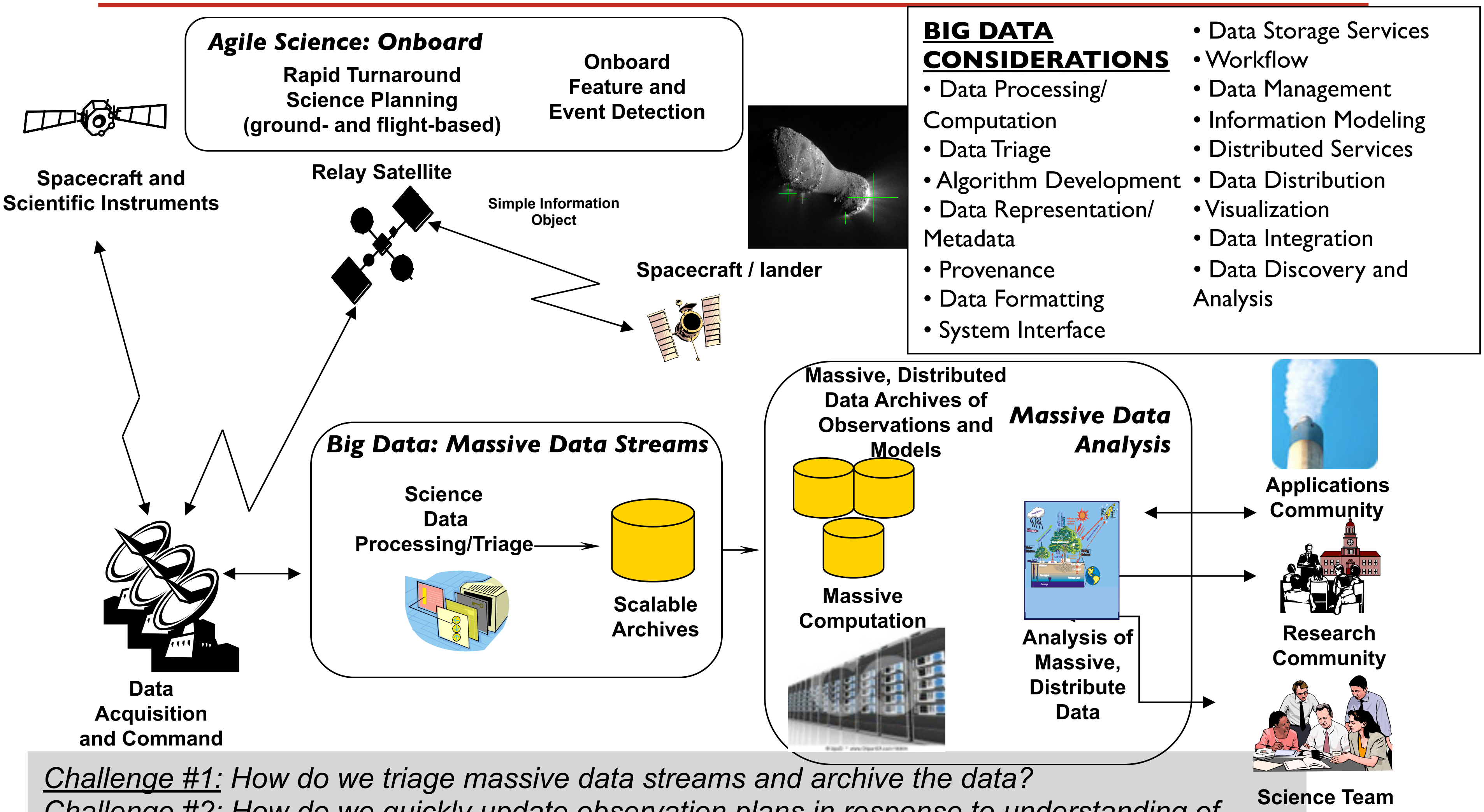


- Scientific missions that explore the Earth and solar system return a wealth of data

- Capturing and archiving these data are essential for preserving these data and supporting research and analysis

- Federal research analysis programs at NASA require data be from public archives

NASA/JPL Big Data Research Areas



Agile Science: Onboard

Rapid Turnaround
Science Planning
(ground- and flight-based)

Onboard
Feature and
Event Detection

Spacecraft and
Scientific Instruments

Relay Satellite

Simple Information
Object

Spacecraft / lander

Data
Acquisition
and Command

Big Data: Massive Data Streams

Science
Data
Processing/Triage

Scalable
Archives

Massive, Distributed
Data Archives of
Observations and
Models

Massive
Computation

***Massive Data
Analysis***

Analysis of
Massive,
Distribute
Data

Applications
Community

Research
Community

Science Team

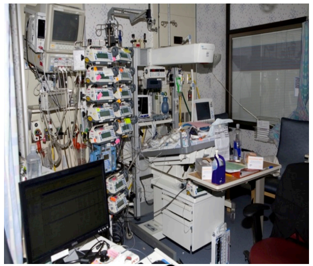
BIG DATA CONSIDERATIONS

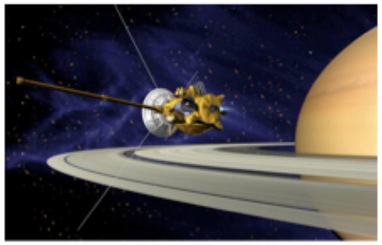
- Data Processing/ Computation
- Data Triage
- Algorithm Development
- Data Representation/ Metadata
- Provenance
- Data Formatting
- System Interface
- Data Storage Services
- Workflow
- Data Management
- Information Modeling
- Distributed Services
- Data Distribution
- Visualization
- Data Integration
- Data Discovery and Analysis

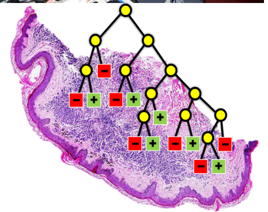
<u>Challenge #1:</u> How do we triage massive data streams and archive the data?
<u>Challenge #2:</u> How do we quickly update observation plans in response to understanding of newly acquired science data, especially for time-limited missions?
<u>Challenge #3:</u> How can we use advanced data science methods to systematically derive scientific inferences from massive, distributed science measurements and models?

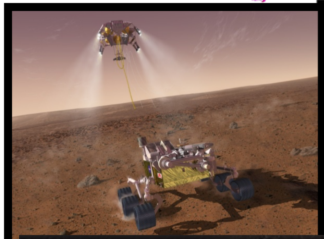
NASA/JPL Informatics Center:
Crossing Disciplines to Support Scientific Research

- Development of an advanced Knowledge System to *capture*, *share* and support *reproducible analysis* from the biomarker data results
  - Supporting EDRN program since 2001
  - Supporting MCL program since 2015
- NASA-NCI partnership, leveraging informatics and data science technologies from planetary and Earth science
  - Reproducible, Big Data Systems for exploring the universe
  - 2011 NASA Group Award for "*innovative use of NASA software technologies to support cancer research*"



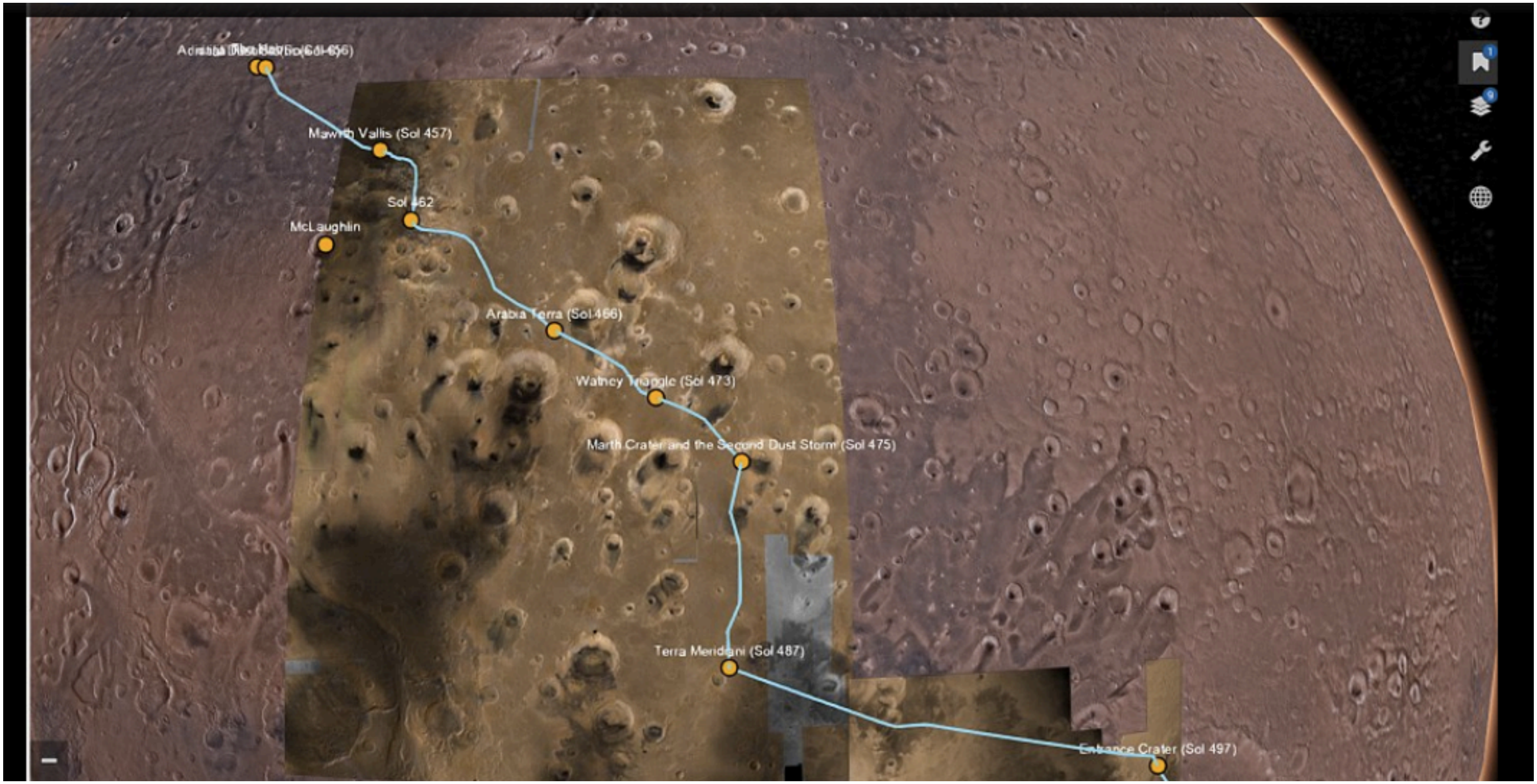






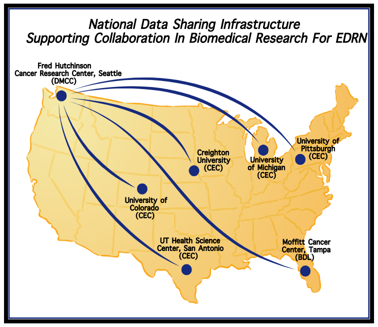
Mawrth Vallis (Sol 457)
Sol 462
McLaughlin
Arabia Terra (Sol 466)
Wahney Triangle (Sol 473)
Marth Crater and the Second Dust Storm (Sol 475)
Terra Meridiani (Sol 487)
Entrance Crater (Sol 497)

Creating Knowledge from Big Data: Exploring Mars Imaging Data with Mars Trek


Mawrth Vallis (Sol 457)
Sol 462
McLaughlin
Arabia Terra (Sol 466)
Watney Triangle (Sol 473)
Marth Crater and the Second Dust Storm (Sol 475)
Terra Meridiani (Sol 487)
Entrance Crater (Sol 497)

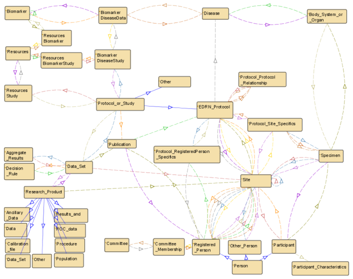
Derived from about 800 TBs of Imaging Data in the NASA Planetary Data System

Informatics and Data Science Needs and Capabilities

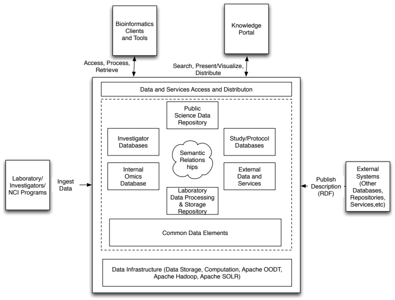


National Data Sharing Infrastructure
Supporting Collaboration In Biomedical Research For EDRN
Fred Hutchinson Cancer Research Center, Seattle (DMCC)
Creighton University (CEC)
University of Michigan (CEC)
University of Pittsburgh (CEC)
University of Colorado (CEC)
UT Health Science Center, San Antonio (CEC)
Moffitt Cancer Center, Tampa (BDL)

National Data Sharing Architectures

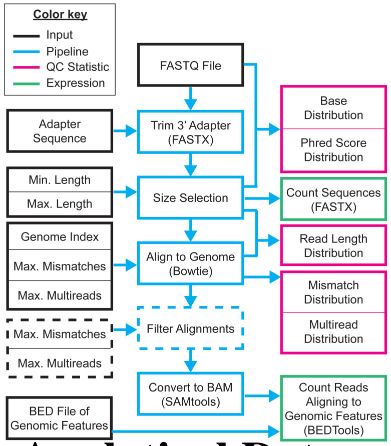


Common Data Elements & Models

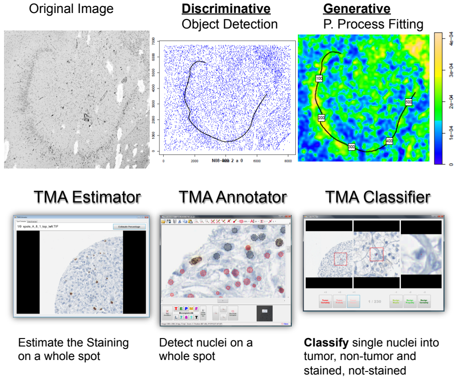


Bioinformatics Clients and Tools
Knowledge Portal
Access, Process, Retrieve
Search, Present/Visualize, Distribute
Data and Services Access and Distributon
Public Science Data Repository
Investigator Databases
Study/Protocol Databases
Semantic Relations hips
Internal Omics Database
External Data and Services
Laboratory Data Processing & Storage Repository
Common Data Elements
Data Infrastructure (Data Storage, Computation, Apache OODT, Apache Hadoop, Apache SOLR)
Laboratory/ Investigators/ NCI Programs
Ingest Data
Publish Description (RDF)
External Systems (Other Databases, Repositories, Services,etc)

Big Data Infrastructure



Analytical Data Pipelines



Estimate the Staining on a whole spot

Detect nuclei on a whole spot

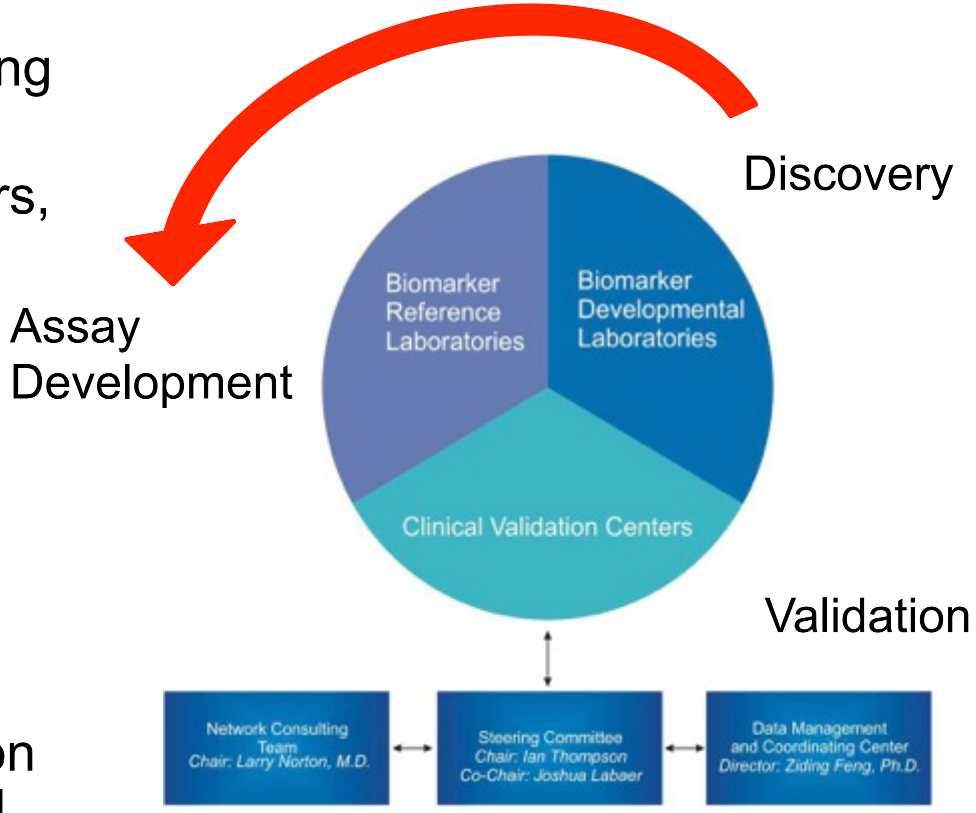
Classify single nuclei into tumor, non-tumor and stained, not-stained

Intelligent Data Algorithms

Visualization Techniques

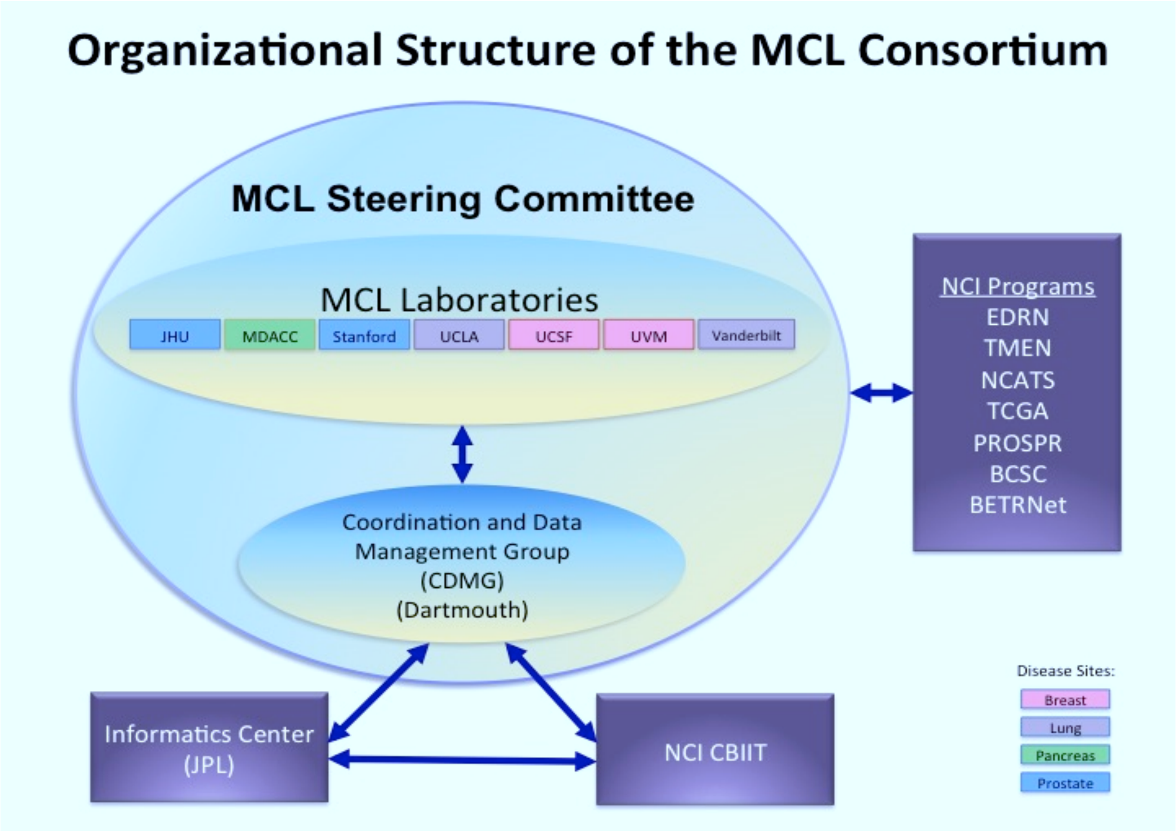
EDRN – Early Detection Research Network

- Early Detection Research Network (EDRN) is a network of 40+ institutions all performing research geared towards the discovery of cancer biomarkers, which are early indicators of onset of disease
- NCI/NIH funded program
  - Started in ~2000
  - NCI's flagship program
- Informatics efforts cited as a model for biomarker research
- Cross-disciplinary collaboration (FHCRC, JPL, Dartmouth and NCI)


Discovery
Assay Development
Validation
Biomarker Reference Laboratories
Biomarker Developmental Laboratories
Clinical Validation Centers
Network Consulting Team
Chair: Larry Norton, M.D.
Steering Committee
Chair: Ian Thompson
Co-Chair: Joshua Labaer
Data Management and Coordinating Center
Director: Ziding Feng, Ph.D.

EDRN Organizational Structure

MCL - Consortium for the Characterization of Screen Detected Lesions



Organizational Structure of the MCL Consortium

MCL Steering Committee

MCL Laboratories

JHU

MDACC

Stanford

UCLA

UCSF

UVM

Vanderbilt

Coordination and Data Management Group (CDMG) (Dartmouth)

NCI Programs

EDRN

TMEN

NCATS

TCGA

PROSPR

BCSC

BETRNet

Informatics Center (JPL)

NCI CBIIT

Disease Sites:

Breast

Lung

Pancreas

Prostate

MCL Consortium Working Groups

- Methodology
- Genomics and Genotyping
- Imaging
- Pathology
- Data, Informatics and Resource Sharing

To conduct a comprehensive molecular and cellular characterization of tumor tissue, cell, and microenvironment components to distinguish screen-detected early lesions from interval and symptom-detected cancers. (Funded 2015)

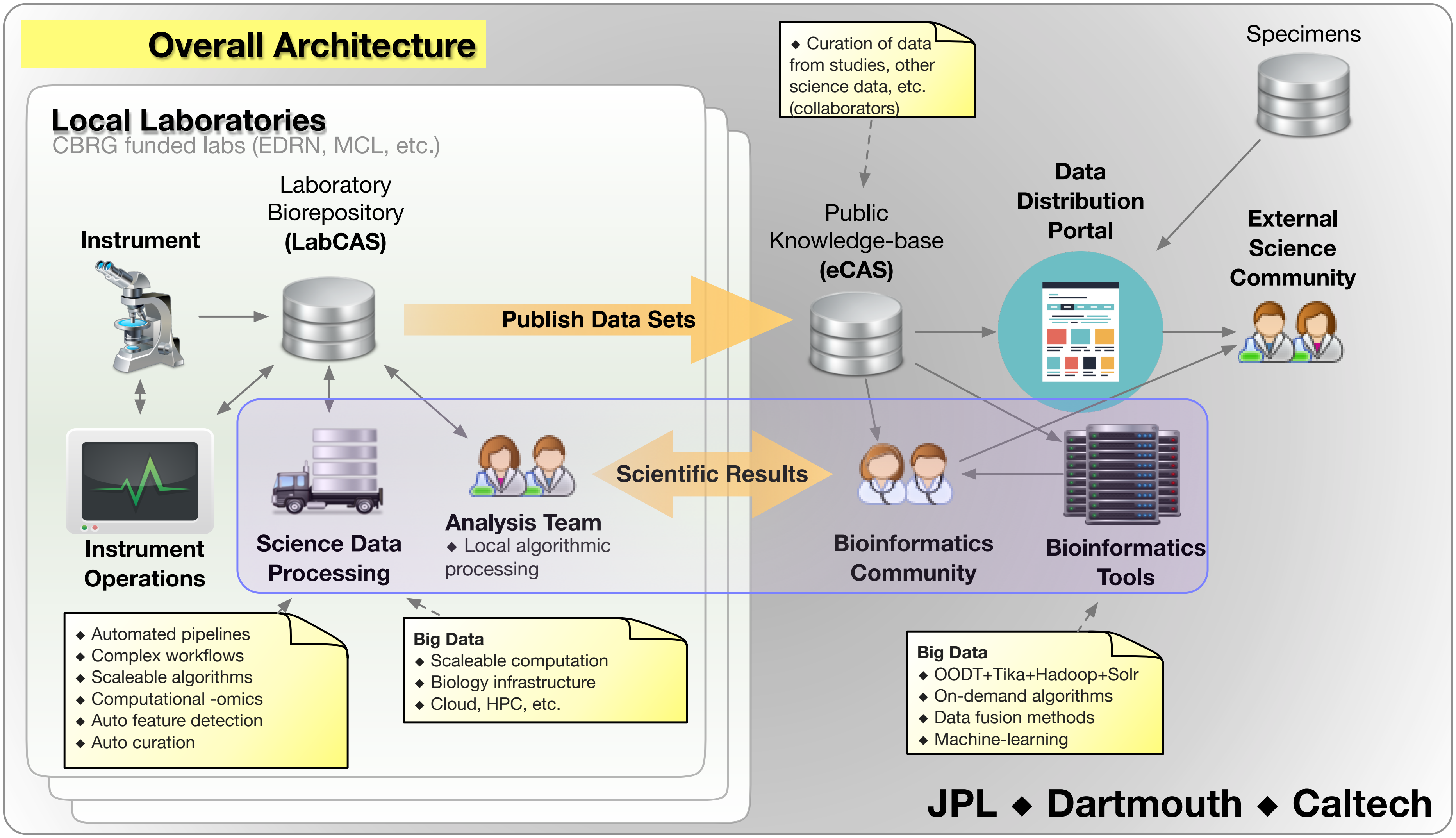
Key Informatics Accomplishments in Life Sciences

- Implemented a national, biomarker knowledge system re-using advanced informatics technology developed for planetary data science
- Pioneered the concept of providing access to information about biospecimens across EDRN at a national level (2001)
- Developed a repository for capturing scientific data sets; captured 90 data sets; integrated with the Canary Foundation infrastructure.
- Developed a biomarker database for capturing and annotating EDRN biomarkers; high-quality curation on more than 900 biomarkers
- Developed a public portal that provides dissemination of EDRN information as well as scientific data and results; over 2400 unique visitors a month
- Developing new tools for the Laboratories to support the processing, capture, curation and sharing of data before publications
- Received NASA Award in 2011 for the "innovative use of NASA software technologies to support cancer research" due to significant reuse of capability
- Began leveraging the architecture across multiple NCI programs

Cancer Biomarker Bioinformatics Workshop

- The NCI and NASA Jet Propulsion Laboratory held a workshop in May 2013 at Caltech to address informatics and data-driven research in cancer biomarkers
  - http://edrn.nci.nih.gov/cancer-bioinformatics-workshop/cancer-biomarker-bioinformatics-workshop-report-may-2013
  - A major outcome focused on data usability, reproducibility of results, methods and algorithms to systematize data analysis, and scalable computing infrastructures.
- Key Recommendations
  - Systematic approaches to the generation, capture, management of data to enable reproducibility
  - Increased emphasis on data curation to promote data reuse
  - Automation of data process/analytics software pipelines
  - Data integration and fusion of data from multiple platforms, studies
  - Scalable data infrastructures and repositories
  - Use of big data tools and bioinformatics techniques to scale data analysis
  - Increased training of scientists in the use of computational tools/methods

Data and Computational Flow



Overall Architecture

Local Laboratories

CBRG funded labs (EDRN, MCL, etc.)

- ◆ Curation of data from studies, other science data, etc. (collaborators)

- ◆ Automated pipelines
- ◆ Complex workflows
- ◆ Scaleable algorithms
- ◆ Computational -omics
- ◆ Auto feature detection
- ◆ Auto curation

Big Data

- ◆ Scaleable computation
- ◆ Biology infrastructure
- ◆ Cloud, HPC, etc.

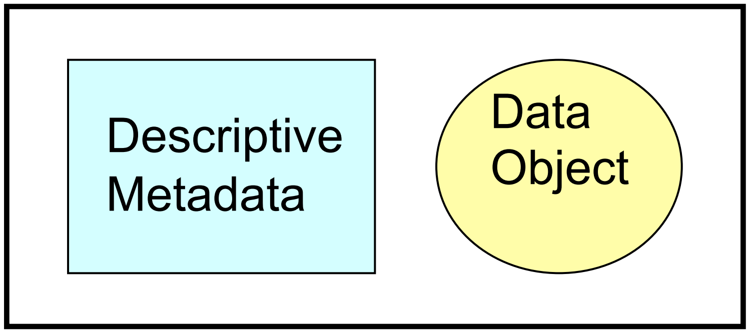
Big Data

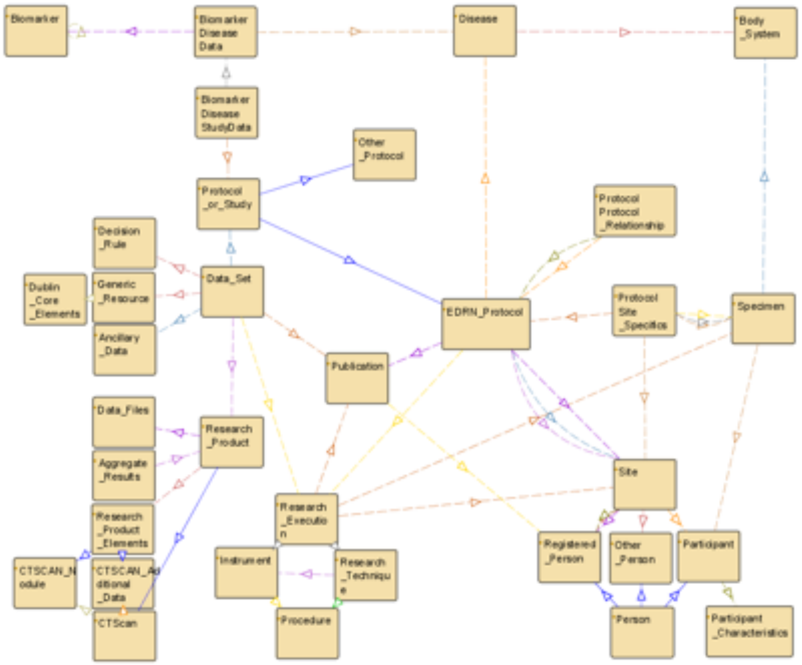
- ◆ OODT+Tika+Hadoop+Solr
- ◆ On-demand algorithms
- ◆ Data fusion methods
- ◆ Machine-learning

JPL ◆ Dartmouth ◆ Caltech

Common Data Elements and Information Models

- CDEs provide a common set of data semantics to capture and share data
    - EDRN CDEs
    - MCL CDEs
- Work with CBIIT to reuse standard elements in caDSR for consistency across NCI
- All archived data are captured as information objects (same as planetary)
    - Metadata described using CDEs
    - Data captured and stored in a data repository



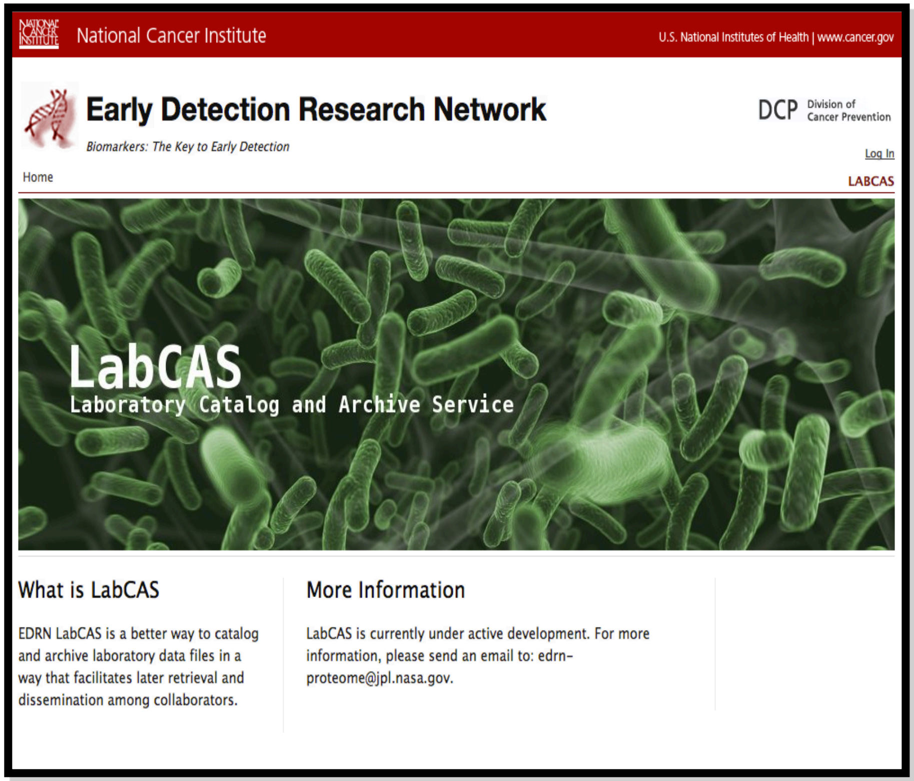


Biomarker Ontology Information Model

LabCAS: Laboratory Catalog and Archive Service

- Provides investigators with a secure, reliable means to capture their **pre-publication** research datasets
- Provides integrated data processing
- Enables investigators and collaborative groups/projects to share data in a secure manner as early as possible
- Scales to support data intensive projects
- Facilitates repeatable data processing pipelines

LabCAS is a new capability under development


National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov
Early Detection Research Network
Biomarkers: The Key to Early Detection
DCP Division of Cancer Prevention
Log In
Home
LABCAS
LabCAS
Laboratory Catalog and Archive Service

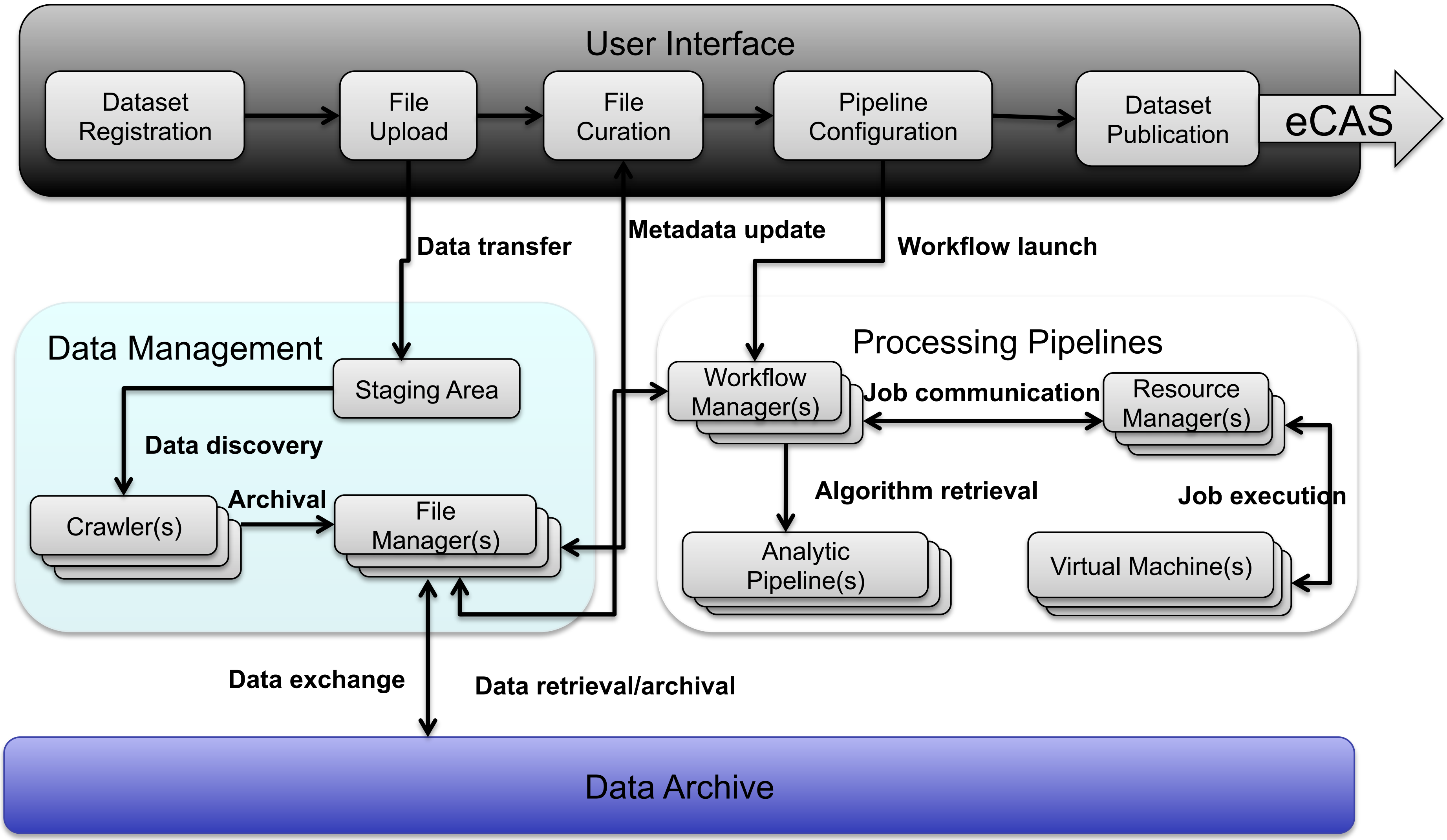
What is LabCAS

EDRN LabCAS is a better way to catalog and archive laboratory data files in a way that facilitates later retrieval and dissemination among collaborators.

More Information

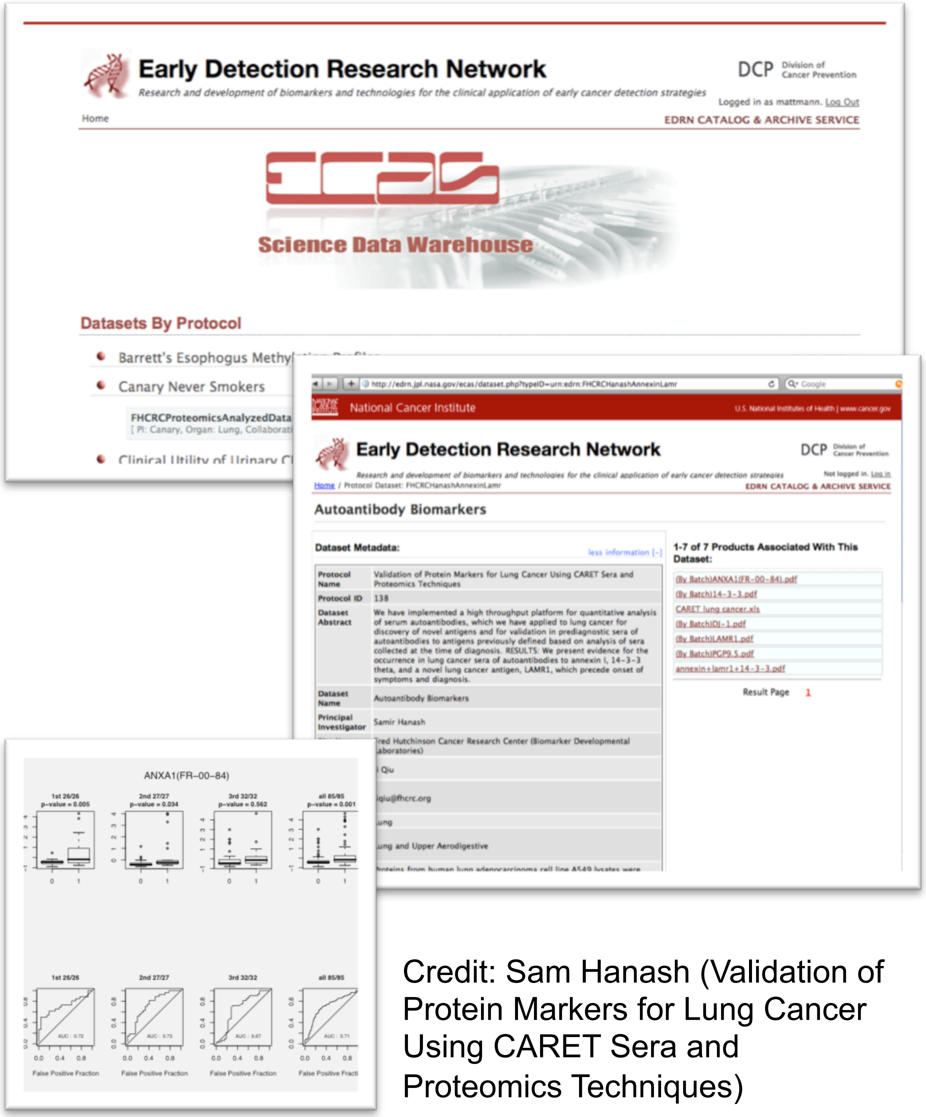
LabCAS is currently under active development. For more information, please send an email to: edrn-proteome@jpl.nasa.gov.

Data Capture , Processing and Ingestion



eCAS: Capture and Sharing of Public Data Sets

- **EDRN has a warehouse of public biomarker data**
  - Uses the CDEs to populate a catalog describing the data sets
  - Supports public release/access to the data
  - Supports peer review of the data by collaborative groups prior to public release
  - Integrated with the rest of the knowledge system
  - Supports reproducibility studies
- **Provides a long term and central capture of biomarker study results for the broad community**
- **Being extended to MCL**



Early Detection Research Network

Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies

DCP Division of Cancer Prevention

Logged in as mattmann. Log Out

Home

EDRN CATALOG & ARCHIVE SERVICE

Science Data Warehouse

Datasets By Protocol

- Barrett's Esophogus Methyl
- Canary Never Smokers
  - FHCRCProteomicsAnalyzedData
    [PI: Canary, Organ: Lung, Collaborati
- Clinical Utility of Urinary C

http://edrn.jpl.nasa.gov/ecas/dataset.php?typeID=urn:edrn:FHCRCHanashAnnexinLamr

National Cancer Institute

U.S. National Institutes of Health | www.cancer.gov

Early Detection Research Network

Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies

DCP Division of Cancer Prevention

Not logged in. Log in

Home / Protocol Dataset: FHCRCHanashAnnexinLamr

EDRN CATALOG & ARCHIVE SERVICE

Autoantibody Biomarkers

Dataset Metadata:

less information [-]

Protocol Name	Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques
Protocol ID	138
Dataset Abstract	We have implemented a high throughput platform for quantitative analysis of serum autoantibodies, which we have applied to lung cancer for discovery of novel antigens and for validation in prediagnostic sera of autoantibodies to antigens previously defined based on analysis of sera collected at the time of diagnosis. RESULTS: We present evidence for the occurrence in lung cancer sera of autoantibodies to annexin I, 14-3-3 theta, and a novel lung cancer antigen, LAMR1, which precede onset of symptoms and diagnosis.
Dataset Name	Autoantibody Biomarkers
Principal Investigator	Samir Hanash
	red Hutchinson Cancer Research Center (Biomarker Developmental aboratories)
	Qiu
	iqiu@fhcrc.org
	ung
	ung and Upper Aerodigestive

1-7 of 7 Products Associated With This Dataset:

- (By Batch)ANXA1(FR-00-84).pdf
- (By Batch)14-3-3.pdf
- CARET lung cancer.xls
- (By Batch)DJ-1.pdf
- (By Batch)LAMR1.pdf
- (By Batch)PGP9.5.pdf
- annexin+lamr1+14-3-3.pdf

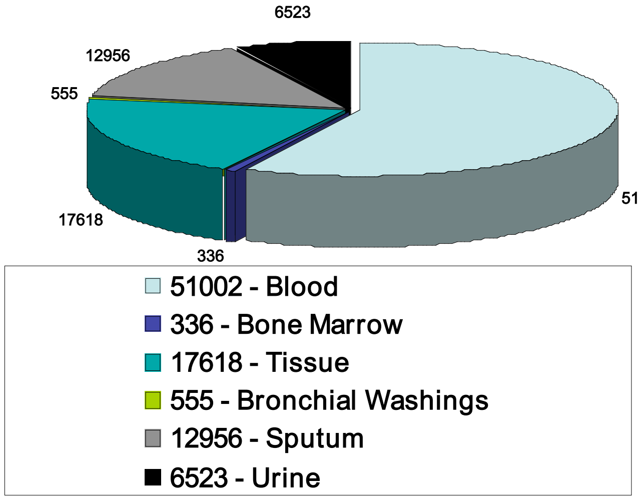
Result Page 1

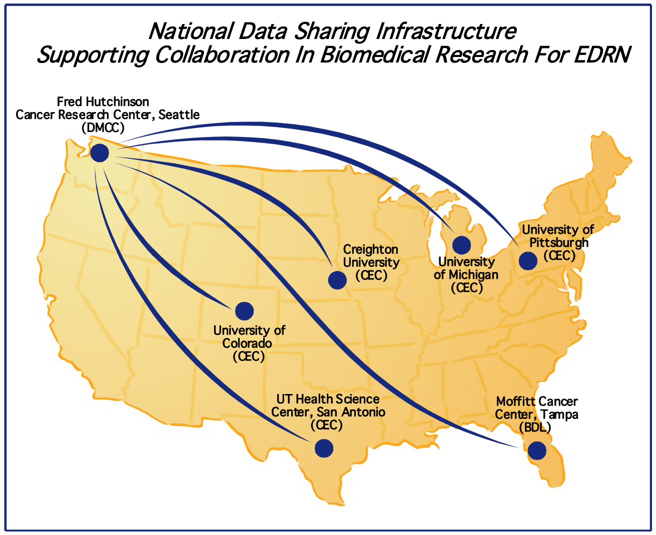
ANXA1(FR–00–84)
1st 26/26 p–value = 0.005
2nd 27/27 p–value = 0.034
3rd 32/32 p–value = 0.562
all 85/85 p–value = 0.001
AUC : 0.72
AUC : 0.73
AUC : 0.67
AUC : 0.71
False Positive Fraction

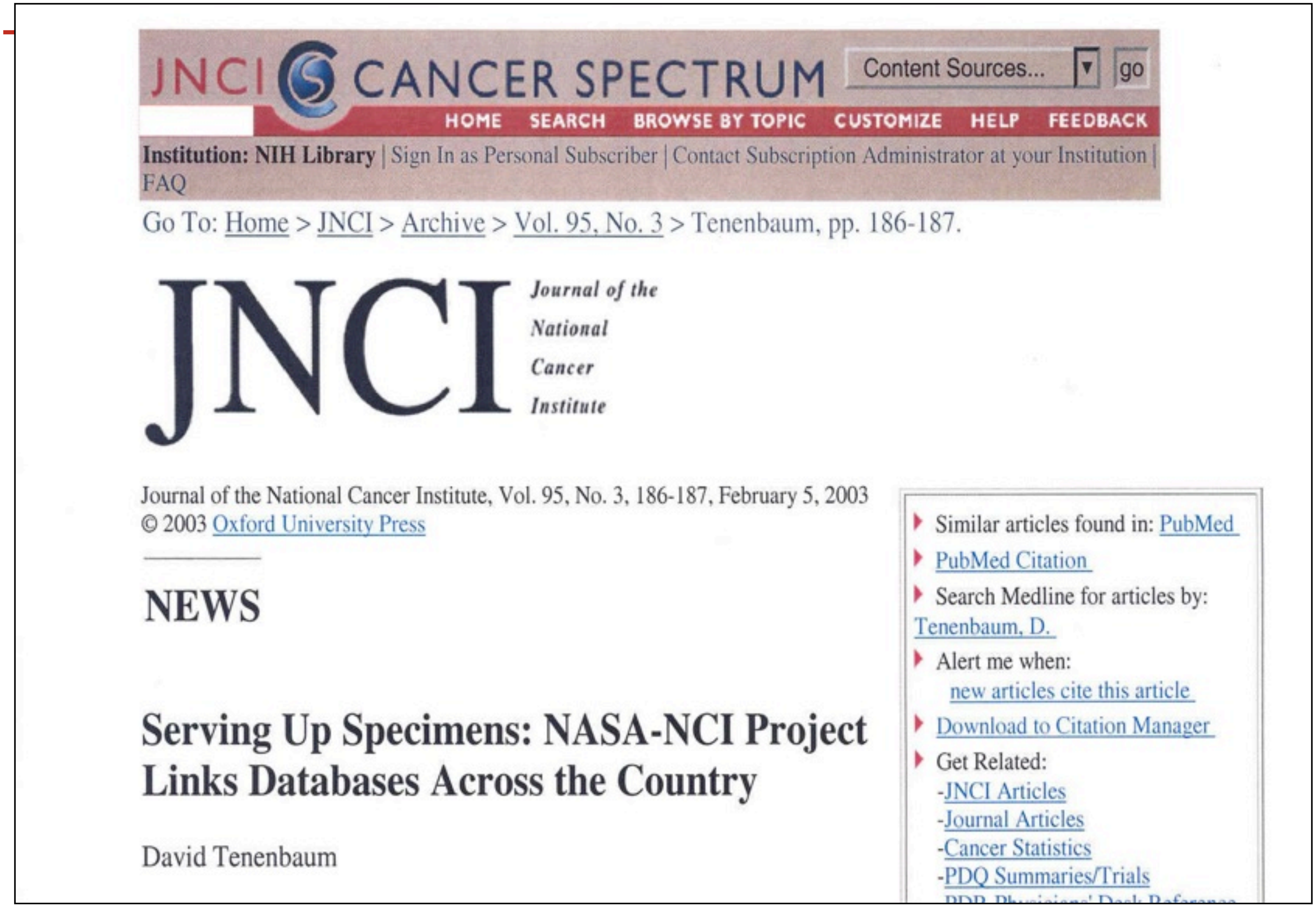
Credit: Sam Hanash (Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques)

Virtual, Distributed Specimen System

H. Lee Moffitt Cancer Center
University of Texas, San Antonio
Creighton University
University of Colorado
University of Pittsburgh
University of Michigan/Dartmouth College
(Great Lakes New England Consortium)
Brigham and Womens Hospital
MD Anderson
New York University
UCSD
Center for Disease Control
Johns Hopkins
Duke University
Fred Hutchinson Cancer Research Center
Fox Chase Cancer Research Center




National Data Sharing Infrastructure
Supporting Collaboration In Biomedical Research For EDRN
Fred Hutchinson Cancer Research Center, Seattle (DMCC)
Creighton University (CEC)
University of Michigan (CEC)
University of Pittsburgh (CEC)
University of Colorado (CEC)
UT Health Science Center, San Antonio (CEC)
Moffitt Cancer Center, Tampa (BDL)



Institution: NIH Library | Sign In as Personal Subscriber | Contact Subscription Administrator at your Institution | FAQ

Go To: Home > JNCI > Archive > Vol. 95, No. 3 > Tenenbaum, pp. 186-187.

JNCI *Journal of the National Cancer Institute*

Journal of the National Cancer Institute, Vol. 95, No. 3, 186-187, February 5, 2003

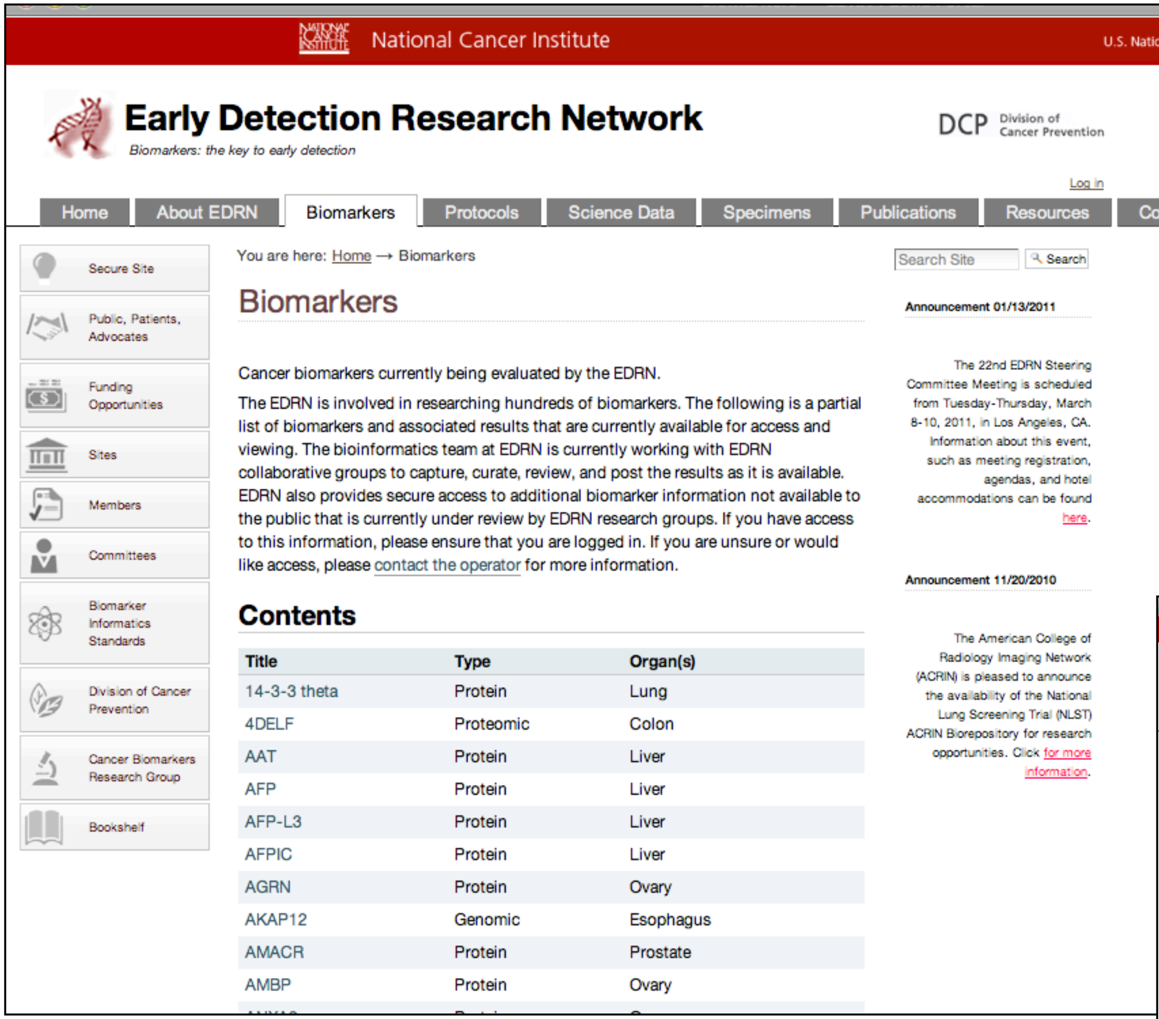
NEWS

Serving Up Specimens: NASA-NCI Project Links Databases Across the Country

David Tenenbaum

- Similar articles found in: PubMed
- PubMed Citation
- Search Medline for articles by: Tenenbaum, D.
- Alert me when: new articles cite this article
- Download to Citation Manager
- Get Related:
  -JNCI Articles
  -Journal Articles
  -Cancer Statistics
  -PDQ Summaries/Trials
  -PDR Physicians' Desk Reference

Enrichment with Biological Database References and Information


National Cancer Institute
U.S. National Institutes of H
DCP Division of Cancer Prevention
Log in
Home
About EDRN
Biomarkers
Protocols
Science Data
Specimens
Publications
Resources
Committees
Secure Site
Public, Patients, Advocates
Funding Opportunities
Sites
Members
Committees
Biomarker Informatics Standards
Division of Cancer Prevention
Cancer Biomarkers Research Group
Bookshelf
Search Site
Search

Early Detection Research Network

Biomarkers: the key to early detection

You are here: Home → Biomarkers

Biomarkers

Cancer biomarkers currently being evaluated by the EDRN.

The EDRN is involved in researching hundreds of biomarkers. The following is a partial list of biomarkers and associated results that are currently available for access and viewing. The bioinformatics team at EDRN is currently working with EDRN collaborative groups to capture, curate, review, and post the results as it is available. EDRN also provides secure access to additional biomarker information not available to the public that is currently under review by EDRN research groups. If you have access to this information, please ensure that you are logged in. If you are unsure or would like access, please contact the operator for more information.

Contents

Title	Type	Organ(s)
14-3-3 theta	Protein	Lung
4DELF	Proteomic	Colon
AAT	Protein	Liver
AFP	Protein	Liver
AFP-L3	Protein	Liver
AFPIC	Protein	Liver
AGRN	Protein	Ovary
AKAP12	Genomic	Esophagus
AMACR	Protein	Prostate
AMBP	Protein	Ovary

Announcement 01/13/2011

The 22nd EDRN Steering Committee Meeting is scheduled from Tuesday-Thursday, March 8-10, 2011, in Los Angeles, CA. Information about this event, such as meeting registration, agendas, and hotel accommodations can be found here.

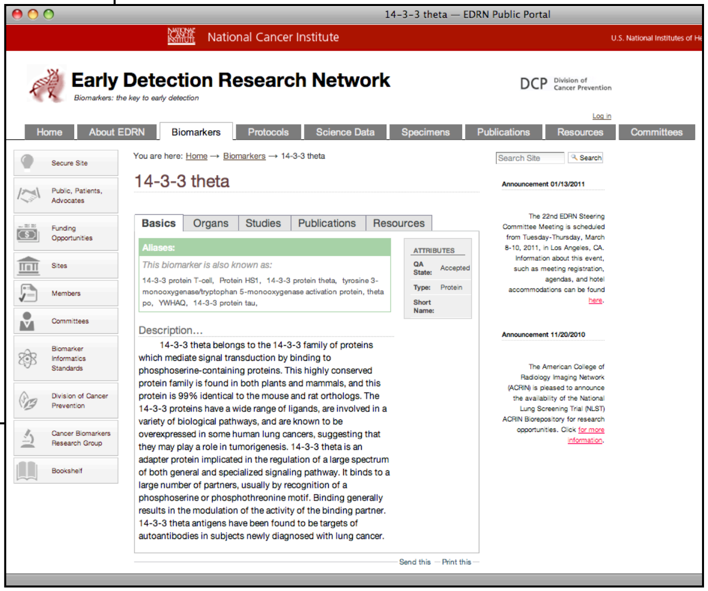
Announcement 11/20/2010

The American College of Radiology Imaging Network (ACRIN) is pleased to announce the availability of the National Lung Screening Trial (NLST) ACRIN Biorepository for research opportunities. Click for more information.

Facilitates capture and sharing of 'omics annotations

Provides connection to the following:

- Protocol
- Scientific Data
- Publications
- Additional Biomarker Resources


14-3-3 theta — EDRN Public Portal
National Cancer Institute
U.S. National Institutes of He
DCP Division of Cancer Prevention
Log in
Home
About EDRN
Biomarkers
Protocols
Science Data
Specimens
Publications
Resources
Committees
Secure Site
Public, Patients, Advocates
Funding Opportunities
Sites
Members
Committees
Biomarker Informatics Standards
Division of Cancer Prevention
Cancer Biomarkers Research Group
Bookshelf
Search Site
Search
Basics
Organs
Studies
Publications
Resources
Send this
Print this

Early Detection Research Network

Biomarkers: the key to early detection

You are here: Home → Biomarkers → 14-3-3 theta

14-3-3 theta

Aliases:

This biomarker is also known as:

14-3-3 protein T-cell, Protein HS1, 14-3-3 protein theta, tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta po, YWHAQ, 14-3-3 protein tau,

ATTRIBUTES

QA State:	Accepted
Type:	Protein
Short Name:	

Description...

14-3-3 theta belongs to the 14-3-3 family of proteins which mediate signal transduction by binding to phosphoserine-containing proteins. This highly conserved protein family is found in both plants and mammals, and this protein is 99% identical to the mouse and rat orthologs. The 14-3-3 proteins have a wide range of ligands, are involved in a variety of biological pathways, and are known to be overexpressed in some human lung cancers, suggesting that they may play a role in tumorigenesis. 14-3-3 theta is an adapter protein implicated in the regulation of a large spectrum of both general and specialized signaling pathway. It binds to a large number of partners, usually by recognition of a phosphoserine or phosphothreonine motif. Binding generally results in the modulation of the activity of the binding partner. 14-3-3 theta antigens have been found to be targets of autoantibodies in subjects newly diagnosed with lung cancer.

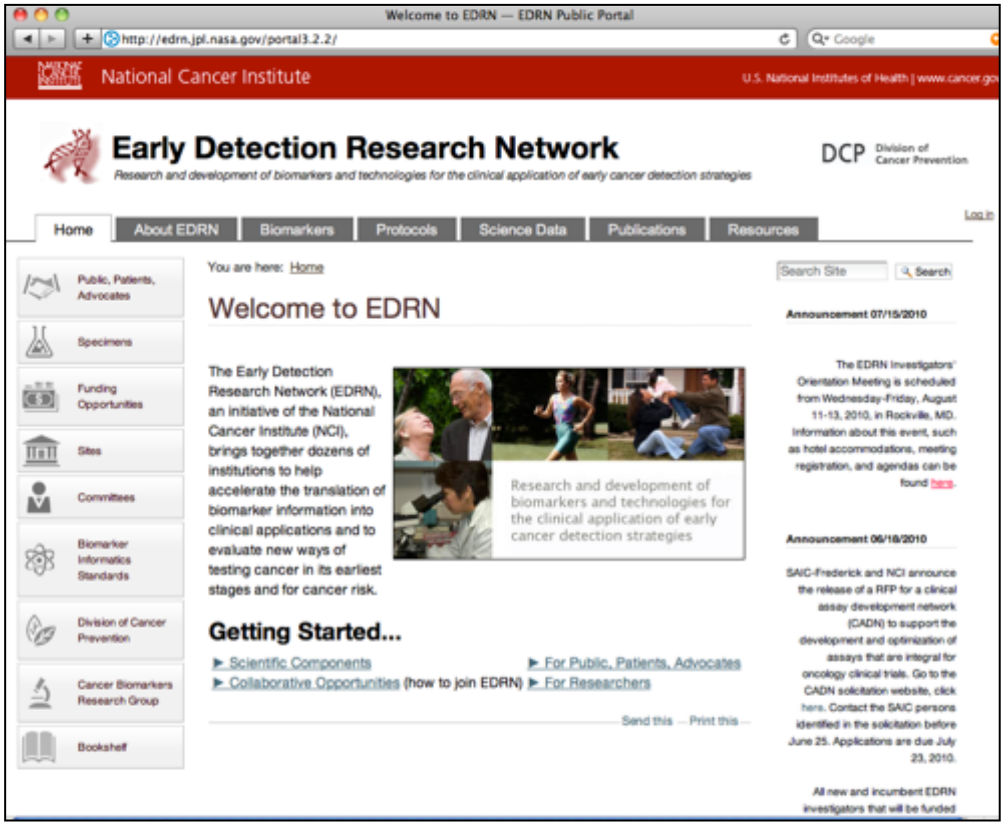
Announcement 01/13/2011

The 22nd EDRN Steering Committee Meeting is scheduled from Tuesday-Thursday, March 8-10, 2011, in Los Angeles, CA. Information about this event, such as meeting registration, agendas, and hotel accommodations can be found here.

Announcement 11/20/2010

The American College of Radiology Imaging Network (ACRIN) is pleased to announce the availability of the National Lung Screening Trial (NLST) ACRIN Biorepository for research opportunities. Click for more information.

Portal: Dissemination and Access to Knowledge System Data


Welcome to EDRN — EDRN Public Portal
http://edrn.jpl.nasa.gov/portal3.2.2/
Google
National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov
DCP Division of Cancer Prevention
Home
About EDRN
Biomarkers
Protocols
Science Data
Publications
Resources
Log in
Search Site
Search
Public, Patients, Advocates
Specimens
Funding Opportunities
Sites
Committees
Biomarker Informatics Standards
Division of Cancer Prevention
Cancer Biomarkers Research Group
Bookshelf
Send this
Print this

Early Detection Research Network

Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies

You are here: Home

Welcome to EDRN

The Early Detection Research Network (EDRN), an initiative of the National Cancer Institute (NCI), brings together dozens of institutions to help accelerate the translation of biomarker information into clinical applications and to evaluate new ways of testing cancer in its earliest stages and for cancer risk.

Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies

Getting Started...

- ► Scientific Components
- ► Collaborative Opportunities (how to join EDRN)
- ► For Public, Patients, Advocates
- ► For Researchers

Announcement 07/15/2010

The EDRN Investigators' Orientation Meeting is scheduled from Wednesday-Friday, August 11-13, 2010, in Rockville, MD. Information about this event, such as hotel accommodations, meeting registration, and agendas can be found here.

Announcement 06/18/2010

SAIC-Frederick and NCI announce the release of a RFP for a clinical assay development network (CADN) to support the development and optimization of assays that are integral for oncology clinical trials. Go to the CADN solicitation website, click here. Contact the SAIC persons identified in the solicitation before June 25. Applications are due July 23, 2010.

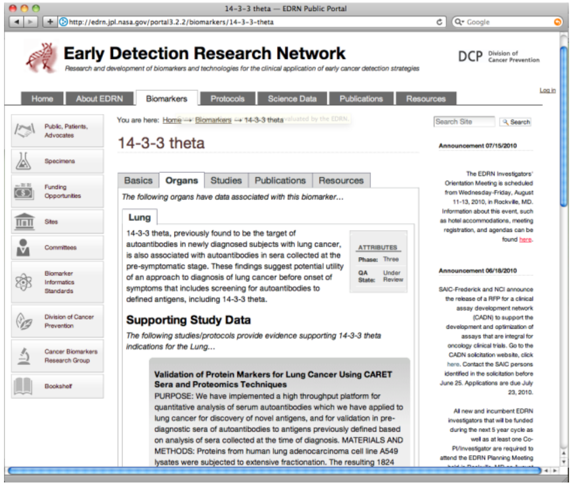
All new and incumbent EDRN investigators that will be funded

http://cancer.gov/edrn
http://mcl.jpl.nasa.gov

- Gateway to information
- Information managed both within and outside the knowledge system
- Initial starting point for community to access research data
- Google-like search to access the wealth of data
- Multi-level Security protects pre-publication and sensitive data

Navigating the Knowledge System: Data Semantically Linked

Jet Propulsion Laboratory
California Institute of Technology


14-3-3 theta — EDRN Public Portal
http://edrn.jpl.nasa.gov/portal3.2.2/biomarkers/14-3-3-theta
Google
Early Detection Research Network
Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies
DCP Division of Cancer Prevention
Log in
Home
About EDRN
Biomarkers
Protocols
Science Data
Publications
Resources
Public, Patients, Advocates
Specimens
Funding Opportunities
Sites
Committees
Biomarker Informatics Standards
Division of Cancer Prevention
Cancer Biomarkers Research Group
Bookshelf
Search Site
Search
Basics
Organs
Studies
Publications
Resources

You are here: Home → Biomarkers → 14-3-3 theta

14-3-3 theta

The following organs have data associated with this biomarker...

Lung

14-3-3 theta, previously found to be the target of autoantibodies in newly diagnosed subjects with lung cancer, is also associated with autoantibodies in sera collected at the pre-symptomatic stage. These findings suggest potential utility of an approach to diagnosis of lung cancer before onset of symptoms that includes screening for autoantibodies to defined antigens, including 14-3-3 theta.

ATTRIBUTES

Phase:	Three
QA State:	Under Review

Supporting Study Data

The following studies/protocols provide evidence supporting 14-3-3 theta indications for the Lung...

Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques
PURPOSE: We have implemented a high throughput platform for quantitative analysis of serum autoantibodies which we have applied to lung cancer for discovery of novel antigens, and for validation in pre-diagnostic sera of autoantibodies to antigens previously defined based on analysis of sera collected at the time of diagnosis. MATERIALS AND METHODS: Proteins from human lung adenocarcinoma cell line A549 lysates were subjected to extensive fractionation. The resulting 1824

Announcement 07/15/2010

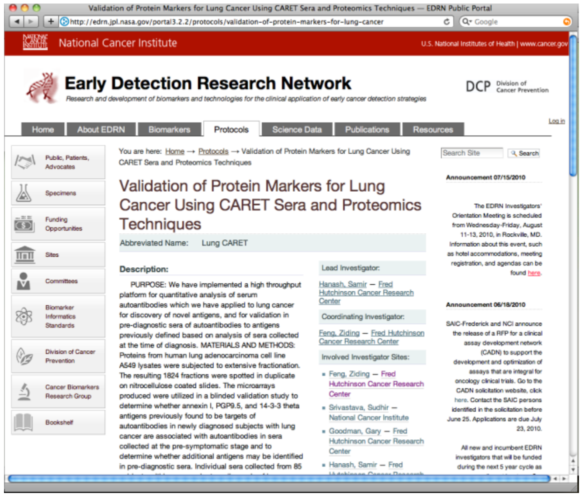
The EDRN Investigators' Orientation Meeting is scheduled from Wednesday-Friday, August 11-13, 2010, in Rockville, MD. Information about this event, such as hotel accommodations, meeting registration, and agendas can be found here.

Announcement 06/18/2010

SAIC-Frederick and NCI announce the release of a RFP for a clinical assay development network (CADN) to support the development and optimization of assays that are integral for oncology clinical trials. Go to the CADN solicitation website, click here. Contact the SAIC persons identified in the solicitation before June 25. Applications are due July 23, 2010.

All new and incumbent EDRN investigators that will be funded during the next 5 year cycle as well as at least one Co-PI/investigator are required to attend the EDRN Planning Meeting

Biomarker Annotations


Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques — EDRN Public Portal
http://edrn.jpl.nasa.gov/portal3.2.2/protocols/validation-of-protein-markers-for-lung-cancer
Google
National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov
Early Detection Research Network
Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies
DCP Division of Cancer Prevention
Log in
Home
About EDRN
Biomarkers
Protocols
Science Data
Publications
Resources
Public, Patients, Advocates
Specimens
Funding Opportunities
Sites
Committees
Biomarker Informatics Standards
Division of Cancer Prevention
Cancer Biomarkers Research Group
Bookshelf
Search Site
Search

You are here: Home → Protocols → Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques

Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques

Abbreviated Name: Lung CARET

Description:

PURPOSE: We have implemented a high throughput platform for quantitative analysis of serum autoantibodies which we have applied to lung cancer for discovery of novel antigens, and for validation in pre-diagnostic sera of autoantibodies to antigens previously defined based on analysis of sera collected at the time of diagnosis. MATERIALS AND METHODS: Proteins from human lung adenocarcinoma cell line A549 lysates were subjected to extensive fractionation. The resulting 1824 fractions were spotted in duplicate on nitrocellulose coated slides. The microarrays produced were utilized in a blinded validation study to determine whether annexin I, PGP9.5, and 14-3-3 theta antigens previously found to be targets of autoantibodies in newly diagnosed subjects with lung cancer are associated with autoantibodies in sera collected at the pre-symptomatic stage and to determine whether additional antigens may be identified in pre-diagnostic sera. Individual sera collected from 85

Lead Investigator:

Hanash, Samir — Fred Hutchinson Cancer Research Center

Coordinating Investigator:

Feng, Ziding — Fred Hutchinson Cancer Research Center

Involved Investigator Sites:

- Feng, Ziding — Fred Hutchinson Cancer Research Center
- Srivastava, Sudhir — National Cancer Institute
- Goodman, Gary — Fred Hutchinson Cancer Research Center
- Hanash, Samir — Fred Hutchinson Cancer Research

Announcement 07/15/2010

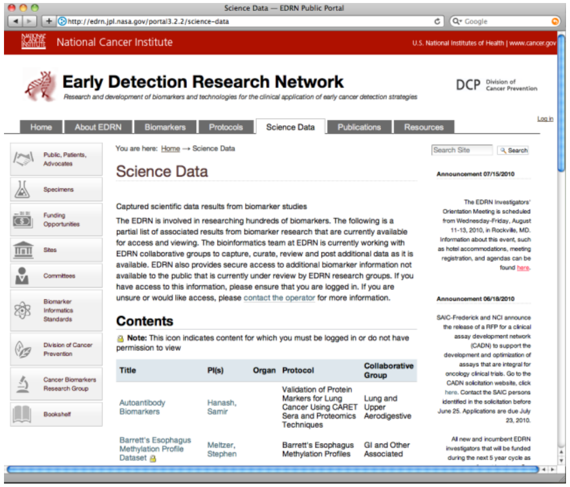
The EDRN Investigators' Orientation Meeting is scheduled from Wednesday-Friday, August 11-13, 2010, in Rockville, MD. Information about this event, such as hotel accommodations, meeting registration, and agendas can be found here.

Announcement 06/18/2010

SAIC-Frederick and NCI announce the release of a RFP for a clinical assay development network (CADN) to support the development and optimization of assays that are integral for oncology clinical trials. Go to the CADN solicitation website, click here. Contact the SAIC persons identified in the solicitation before June 25. Applications are due July 23, 2010.

All new and incumbent EDRN investigators that will be funded during the next 5 year cycle as

Protocols


Science Data — EDRN Public Portal
http://edrn.jpl.nasa.gov/portal3.2.2/science-data
Google
National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov
Early Detection Research Network
Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies
DCP Division of Cancer Prevention
Log in
Home
About EDRN
Biomarkers
Protocols
Science Data
Publications
Resources
Public, Patients, Advocates
Specimens
Funding Opportunities
Sites
Committees
Biomarker Informatics Standards
Division of Cancer Prevention
Cancer Biomarkers Research Group
Bookshelf
Search Site
Search

You are here: Home → Science Data

Science Data

Captured scientific data results from biomarker studies

The EDRN is involved in researching hundreds of biomarkers. The following is a partial list of associated results from biomarker research that are currently available for access and viewing. The bioinformatics team at EDRN is currently working with EDRN collaborative groups to capture, curate, review and post additional data as it is available. EDRN also provides secure access to additional biomarker information not available to the public that is currently under review by EDRN research groups. If you have access to this information, please ensure that you are logged in. If you are unsure or would like access, please contact the operator for more information.

Contents

🔒 **Note:** This icon indicates content for which you must be logged in or do not have permission to view

Title	PI(s)	Organ	Protocol	Collaborative Group
Autoantibody Biomarkers	Hanash, Samir		Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques	Lung and Upper Aerodigestive
Barrett's Esophagus Methylation Profile Dataset 🔒	Meltzer, Stephen		Barrett's Esophagus Methylation Profiles	GI and Other Associated

Announcement 07/15/2010

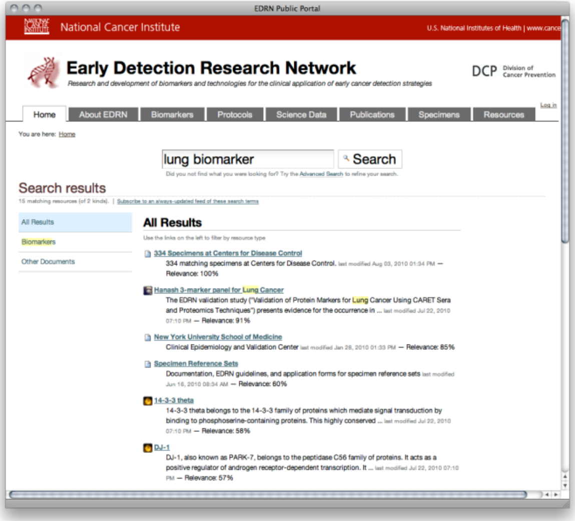
The EDRN Investigators' Orientation Meeting is scheduled from Wednesday-Friday, August 11-13, 2010, in Rockville, MD. Information about this event, such as hotel accommodations, meeting registration, and agendas can be found here.

Announcement 06/18/2010

SAIC-Frederick and NCI announce the release of a RFP for a clinical assay development network (CADN) to support the development and optimization of assays that are integral for oncology clinical trials. Go to the CADN solicitation website, click here. Contact the SAIC persons identified in the solicitation before June 25. Applications are due July 23, 2010.

All new and incumbent EDRN investigators that will be funded during the next 5 year cycle as

Biomarker Data Results


EDRN Public Portal
National Cancer Institute
U.S. National Institutes of Health | www.cance
Early Detection Research Network
Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies
DCP Division of Cancer Prevention
Log in
Home
About EDRN
Biomarkers
Protocols
Science Data
Publications
Specimens
Resources
lung biomarker
Search
All Results
Biomarkers
Other Documents

You are here: Home

Did you not find what you were looking for? Try the Advanced Search to refine your search.

Search results

15 matching resources (of 2 kinds). | Subscribe to an always-updated feed of these search terms

All Results

Use the links on the left to filter by resource type

334 Specimens at Centers for Disease Control
334 matching specimens at Centers for Disease Control. last modified Aug 03, 2010 01:34 PM — Relevance: 100%

Hanash 3-marker panel for Lung Cancer
The EDRN validation study ("Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques") presents evidence for the occurrence in ... last modified Jul 22, 2010 07:10 PM — Relevance: 91%

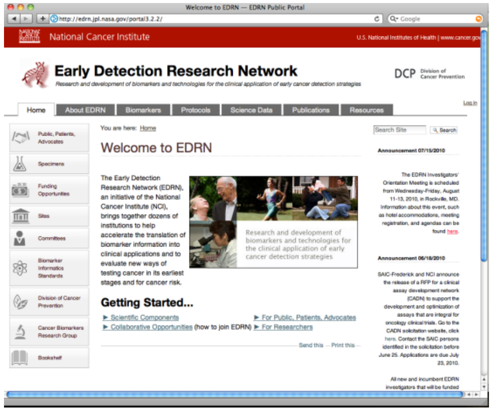
New York University School of Medicine
Clinical Epidemiology and Validation Center last modified Jan 28, 2010 01:33 PM — Relevance: 85%

Specimen Reference Sets
Documentation, EDRN guidelines, and application forms for specimen reference sets last modified Jun 18, 2010 08:34 AM — Relevance: 60%

14-3-3 theta
14-3-3 theta belongs to the 14-3-3 family of proteins which mediate signal transduction by binding to phosphoserine-containing proteins. This highly conserved ... last modified Jul 22, 2010 07:10 PM — Relevance: 58%

DJ-1
DJ-1, also known as PARK-7, belongs to the peptidase C56 family of proteins. It acts as a positive regulator of androgen receptor-dependent transcription. It ... last modified Jul 22, 2010 07:10 PM — Relevance: 57%

Specimens


Welcome to EDRN — EDRN Public Portal
http://edrn.jpl.nasa.gov/portal3.2.2/
Google
National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov
Early Detection Research Network
Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies
DCP Division of Cancer Prevention
Log in
Home
About EDRN
Biomarkers
Protocols
Science Data
Publications
Resources
Public, Patients, Advocates
Specimens
Funding Opportunities
Sites
Committees
Biomarker Informatics Standards
Division of Cancer Prevention
Cancer Biomarkers Research Group
Bookshelf
Search Site
Search
Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies
Send this — Print this

You are here: Home

Welcome to EDRN

The Early Detection Research Network (EDRN), an initiative of the National Cancer Institute (NCI), brings together dozens of institutions to help accelerate the translation of biomarker information into clinical applications and to evaluate new ways of testing cancer in its earliest stages and for cancer risk.

Getting Started...

► Scientific Components
► Collaborative Opportunities (how to join EDRN)
► For Public, Patients, Advocates
► For Researchers

Announcement 07/15/2010

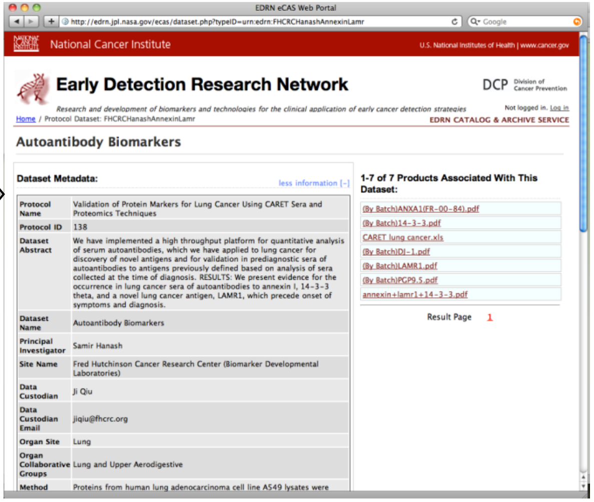
The EDRN Investigators' Orientation Meeting is scheduled from Wednesday-Friday, August 11-13, 2010, in Rockville, MD. Information about this event, such as hotel accommodations, meeting registration, and agendas can be found here.

Announcement 06/18/2010

SAIC-Frederick and NCI announce the release of a RFP for a clinical assay development network (CADN) to support the development and optimization of assays that are integral for oncology clinical trials. Go to the CADN solicitation website, click here. Contact the SAIC persons identified in the solicitation before June 25. Applications are due July 23, 2010.

All new and incumbent EDRN investigators that will be funded

Linked through
Public Portal


EDRN eCAS Web Portal
http://edrn.jpl.nasa.gov/ecas/dataset.php?typeID=urn:edrn:FHCRCHanashAnnexinLamr
Google
National Cancer Institute
U.S. National Institutes of Health | www.cancer.gov
Early Detection Research Network
Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies
DCP Division of Cancer Prevention
Not logged in. Log in
EDRN CATALOG & ARCHIVE SERVICE
less information [-]

Home / Protocol Dataset: FHCRCHanashAnnexinLamr

Autoantibody Biomarkers

Dataset Metadata:

Protocol Name	Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques
Protocol ID	138
Dataset Abstract	We have implemented a high throughput platform for quantitative analysis of serum autoantibodies, which we have applied to lung cancer for discovery of novel antigens and for validation in prediagnostic sera of autoantibodies to antigens previously defined based on analysis of sera collected at the time of diagnosis. RESULTS: We present evidence for the occurrence in lung cancer sera of autoantibodies to annexin I, 14-3-3 theta, and a novel lung cancer antigen, LAMR1, which precede onset of symptoms and diagnosis.
Dataset Name	Autoantibody Biomarkers
Principal Investigator	Samir Hanash
Site Name	Fred Hutchinson Cancer Research Center (Biomarker Developmental Laboratories)
Data Custodian	Ji Qiu
Data Custodian Email	jiqiu@fhcrc.org
Organ Site	Lung
Organ Collaborative Groups	Lung and Upper Aerodigestive
Method	Proteins from human lung adenocarcinoma cell line A549 lysates were

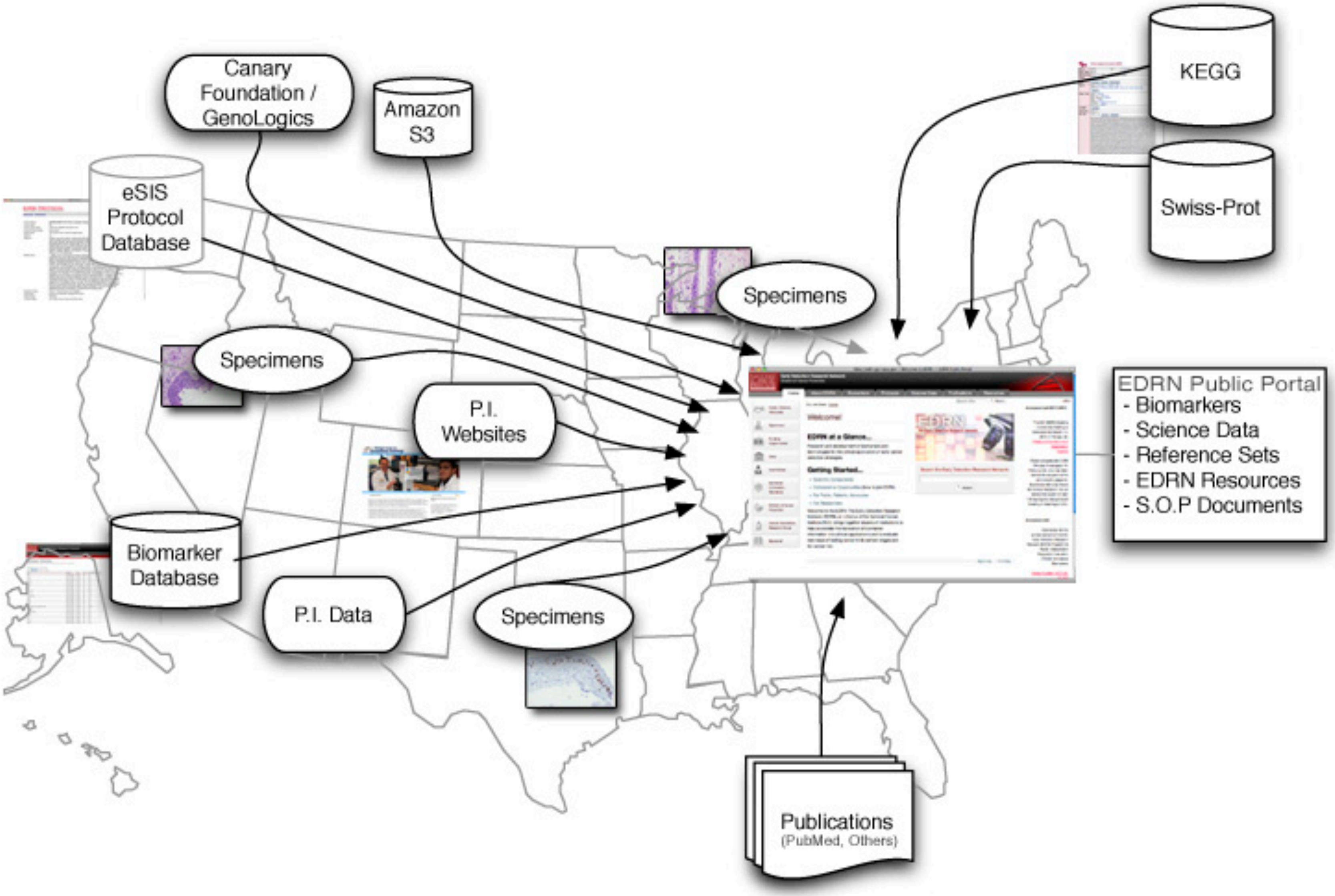
1-7 of 7 Products Associated With This Dataset:

- (By Batch)ANXA1(FR-00-84).pdf
- (By Batch)14-3-3.pdf
- CARET lung cancer.xls
- (By Batch)DJ-1.pdf
- (By Batch)LAMR1.pdf
- (By Batch)PGP9.5.pdf
- annexin+lamr1+14-3-3.pdf

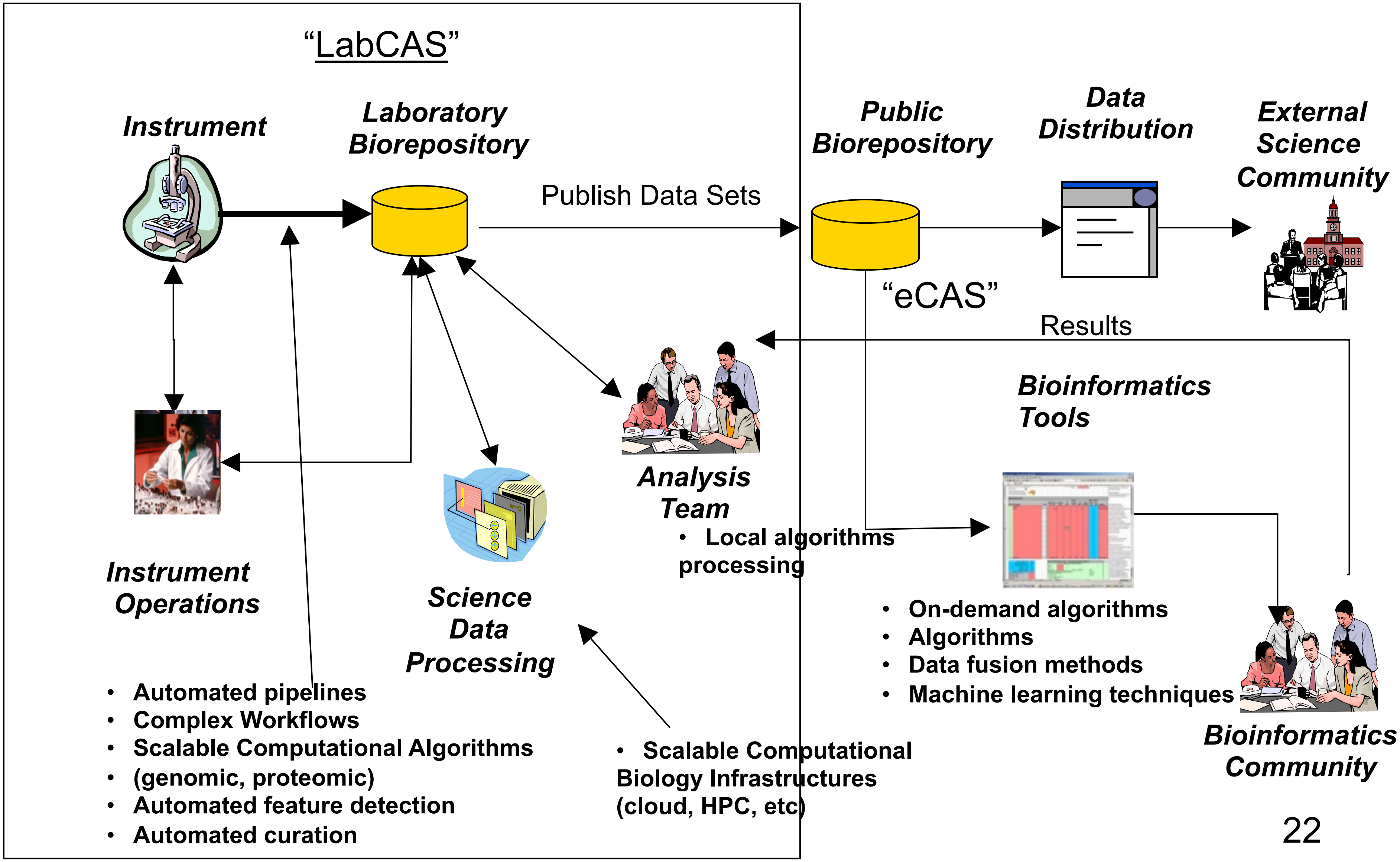
Result Page 1

Access to download data

A Virtual, National Integration Biomarkers Knowledge System



Moving towards data-driven discovery for cancer biomarkers



"LabCAS"

Instrument

Laboratory Biorepository

Publish Data Sets

Public Biorepository

Data Distribution

External Science Community

"eCAS"

Results

Bioinformatics Tools

Analysis Team

- **Local algorithms processing**

Instrument Operations

Science Data Processing

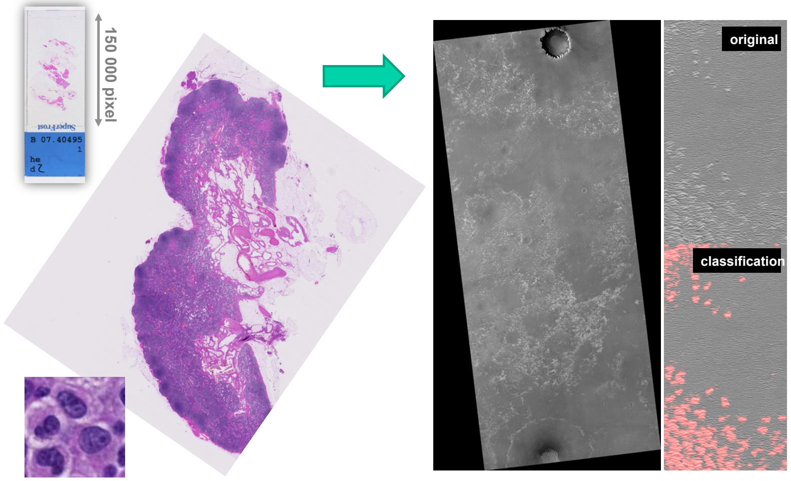
- **Automated pipelines**
- **Complex Workflows**
- **Scalable Computational Algorithms**
- **(genomic, proteomic)**
- **Automated feature detection**
- **Automated curation**

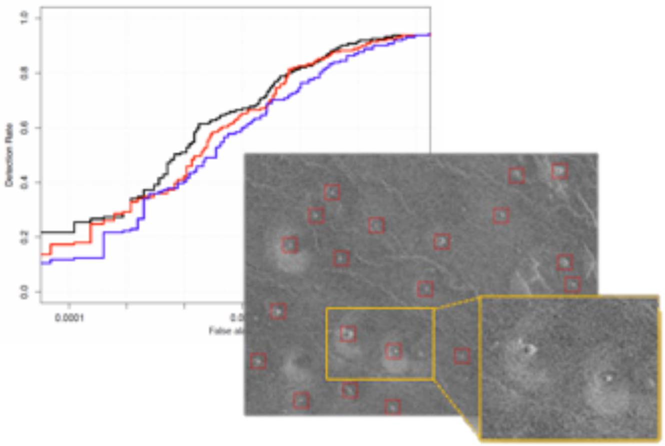
- **Scalable Computational Biology Infrastructures (cloud, HPC, etc)**

- **On-demand algorithms**
- **Algorithms**
- **Data fusion methods**
- **Machine learning techniques**

Bioinformatics Community

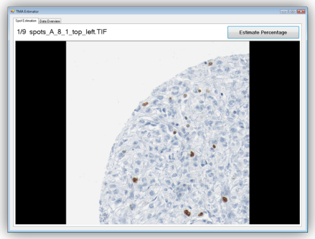
Application of Machine Learning Techniques





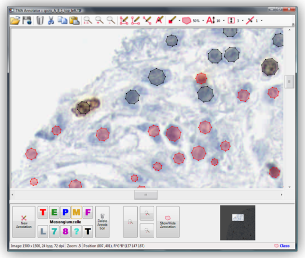
Volcanoes on Venus

TMA Estimator



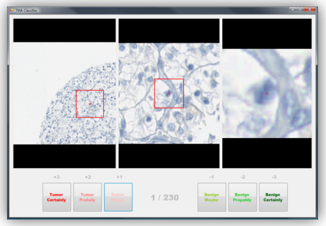
Estimate the Staining on a whole spot

TMA Annotator



Detect nuclei on a whole spot

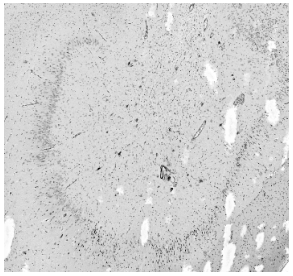
TMA Classifier



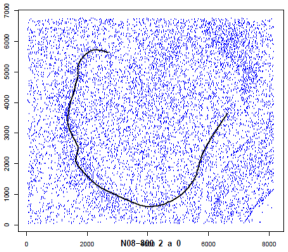
Classify single nuclei into tumor, non-tumor and stained, not-stained

Automated Classification

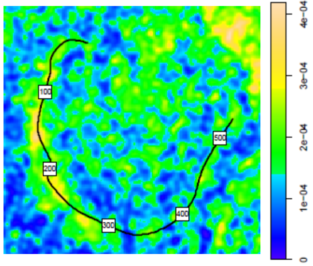
Original Image



Discriminative
Object Detection



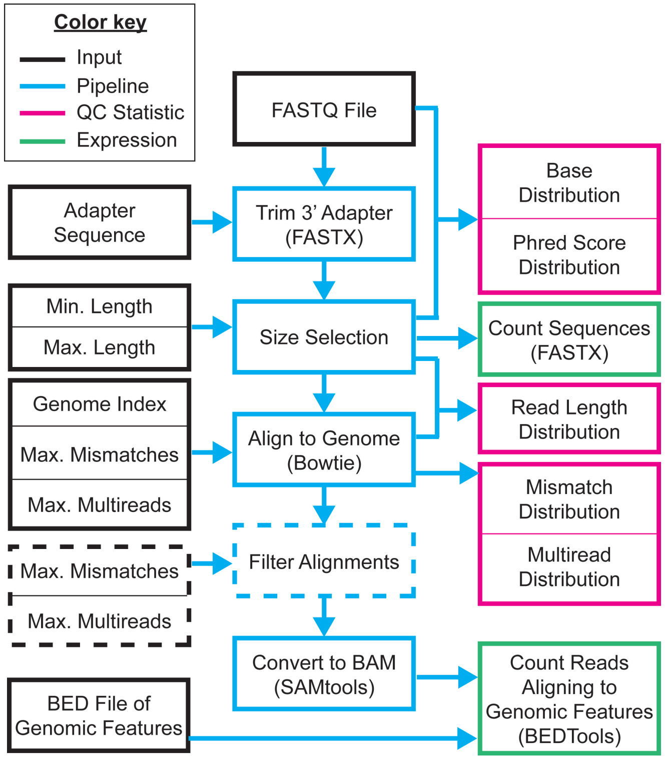
Generative
P. Process Fitting



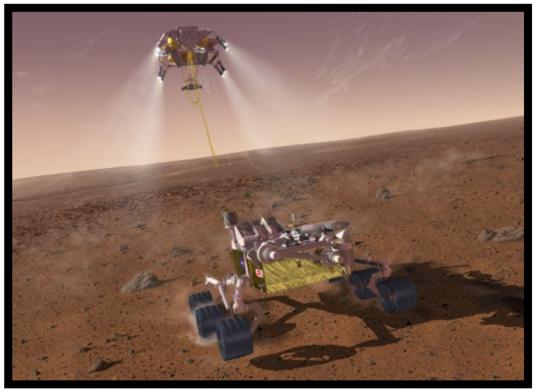
Feature/Object Detection

Potential Collaboration Areas

- Reproducibility experiments
- Analytic Data Pipelines/Computational Methods
- Bioinformatics Tool Integration
- Data Integration
- Data Capture, Curation and Sharing
- Image Archiving and Visualization
- Linking of Distributed Capabilities
- Ontologies, Common Data Elements, etc.



Courtesy of Josh Campbell and Teresa Wang



Early Detection Research Network
Informatics Center
at the Jet Propulsion Laboratory

Reproduced from *Wired* magazine

The JPL Informatics Center would like to collaborate with the ITCR Program to explore how data, tools, and methods can be shared to expand the knowledge system and support the EDRN and MCL Programs and other NCI research.

http://twitter.com/edrn_ic

http://www.facebook.com/group.php?gid=56938589930