

**Early
Detection
Research
Network**



**EDRN
Informatics
Infrastructure**

Dan Crichton

PI, JPL Informatics Center

Co-Director, JPL/Caltech Center for Data Science and Technology

NASA Jet Propulsion Laboratory, Caltech

With Heather Kincaid, NASA/JPL & Kristen Anton, Geisel School of
Medicine, Dartmouth

*(in collaboration with NCI, Dartmouth, Fred Hutchinson Cancer Center
and many others ...)*



JPL

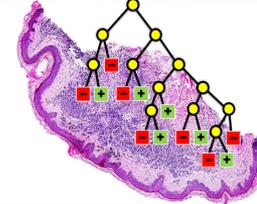
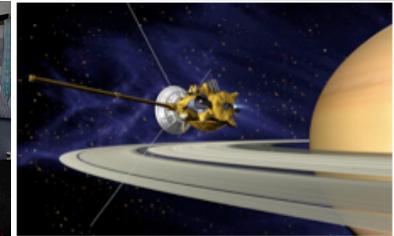
Jet Propulsion Laboratory
California Institute of Technology



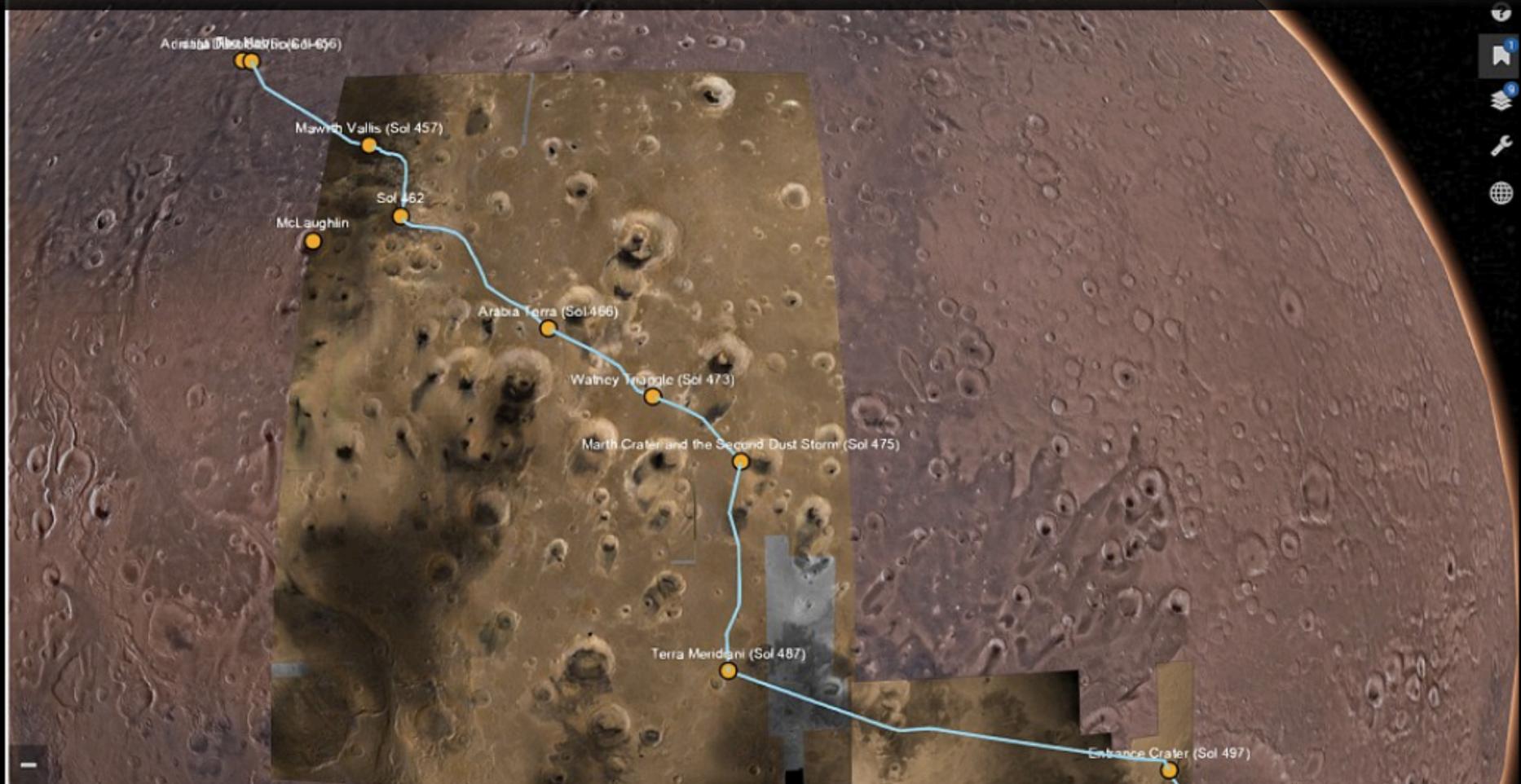
October 15, 2015

NASA/JPL Informatics Center

- Development of an advanced Knowledge System for EDRN to *capture, share* and support *reproducible analysis* from the biomarker data results
- 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*”



Example: Exploring Mars Imaging Data with Mars Trek

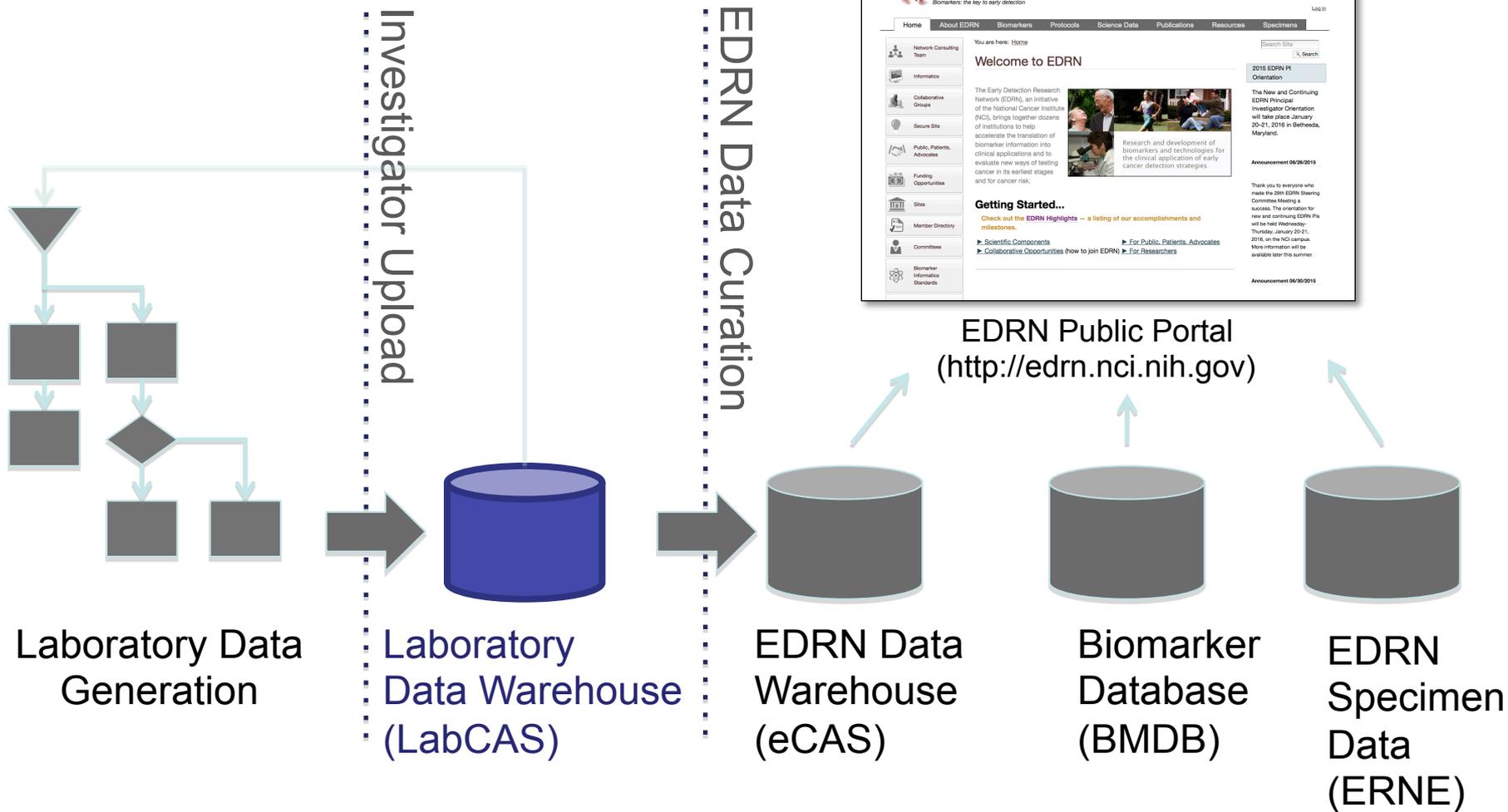


Derived from about 800 TBs of Imaging Data in the NASA Planetary Data System

In 2000, EDRN investigators identified the following needs:

- Ability to develop informatics standards such as common data elements (CDEs) for biomarker research to allow for data to be linked together and analyzed
- Informatics to support multi-institutional validation studies
- Create tools to facilitate the use of the informatics standards for data collection (e.g. CDE Form Tool, EDRN Data Model, Public Portal)
- Ability to access specimens across EDRN Clinical validation Centers
- Ability to support secure data transfer, data analysis and communication
- Ability to support EDRN-wide data storage, curation and retrieval of multidimensional, multi-format data

Capturing and Sharing Biomarker Data Results



EDRN Informatics Projects

EDRN has a number of focused projects that have been started in informatics to support the goals of capturing and sharing the state of biomarker research

Virtual Specimen Repository (ERNE): Access to information on specimens across EDRN;

EDRN-wide Portal: Access to EDRN-wide biomarker study results;

Science Data Warehouse (eCAS): Capture of biomarker data results from EDRN studies into a central repository. Security integrated.

Biomarker Database: Capture of biomarkers under study within the EDRN

Common Data Elements: Common sets of terms used to construct databases and forms

Laboratory Data Sharing (LabCAS): Enhance EDRN capabilities to automated the processing, capture and sharing of data from EDRN studies

Key Informatics Accomplishments

- Developed a **national, biomarker knowledge system** using advanced informatics technology
- Pioneered the concept 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 EDRN biomarkers; captured over 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 Laboratory's 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**.

LabCAS: Laboratory Catalog and Archive Service

- LabCAS is a new capability under development to provide investigators with a secure, reliable means to capture their **pre-publication** research datasets
- LabCAS also provides integrated data processing
- Enable investigators and collaborative groups/projects to share data in a secure manner as early as possible
- Scale to support data intensive projects
 - Repeatable data processing pipelines



The screenshot shows the LabCAS website interface. At the top, there is a red header with the National Cancer Institute logo and the text "National Cancer Institute" and "U.S. National Institutes of Health | www.cancer.gov". Below the header, the "Early Detection Research Network" logo is displayed, along with the tagline "Biomarkers: The Key to Early Detection". The "DCP Division of Cancer Prevention" logo is also visible. A "Log In" link and the text "LABCAS" are in the top right corner. The main content area features a large image of green, rod-shaped bacteria with the text "LabCAS Laboratory Catalog and Archive Service" overlaid. Below the image, there are two columns of text: "What is LabCAS" and "More Information".

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 LABCAS

Home

LabCAS
Laboratory Catalog and Archive Service

What is LabCAS

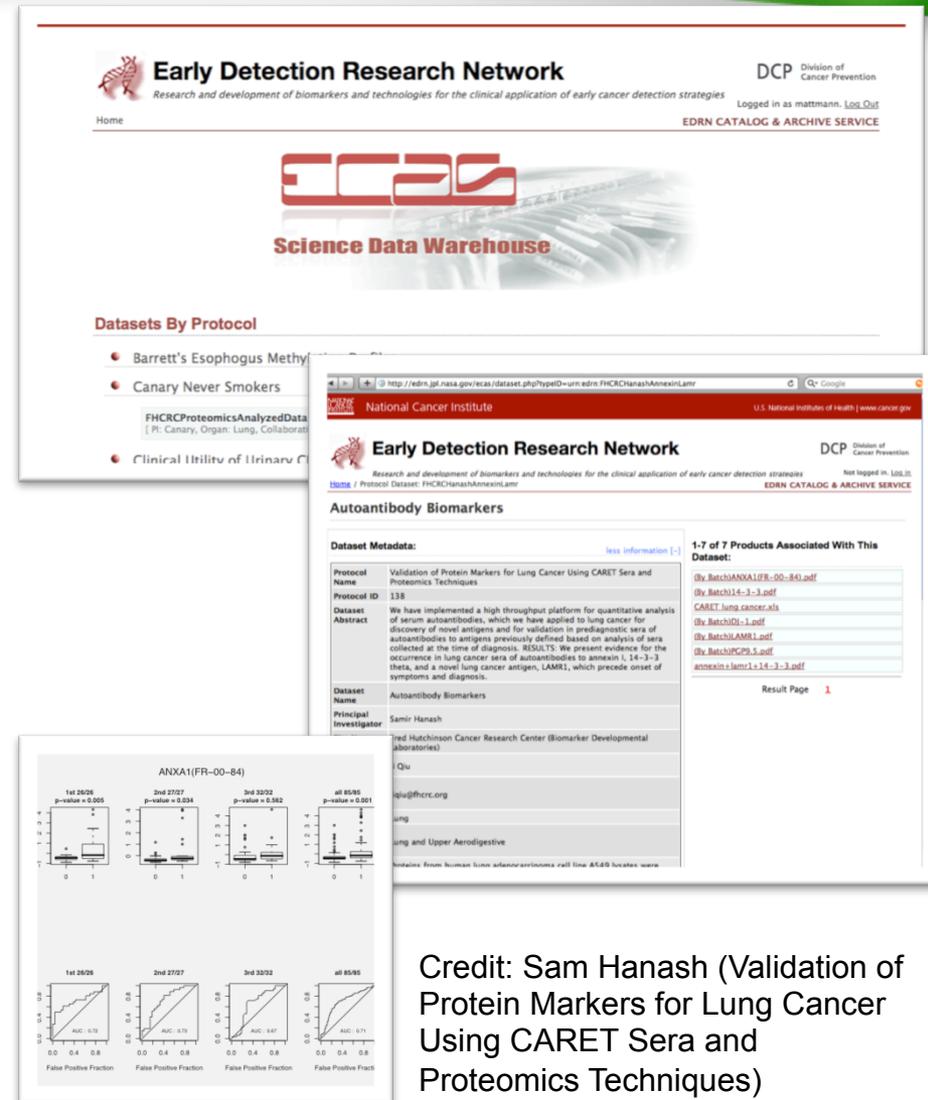
EDRN LabCAS is a better way to catalog and archive laboratory data files in a way that facilitates later retrieval and dissemination among collaborators.

More Information

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

eCAS: Capture and Sharing of Public Data Sets

- **EDRN has a warehouse of public biomarker data for use today!**
 - Uses the EDRN 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 EDRN systems
- **Provides a long term and central capture of EDRN study results for the broad community**



Early Detection Research Network
Research and development of biomarkers and technologies for the clinical application of early cancer detection strategies
Home | DCP Division of Cancer Prevention | Logged in as mattmann... | EDRN CATALOG & ARCHIVE SERVICE

eCAS Science Data Warehouse

Datasets By Protocol

- Barrett's Esophagus Methy...
- Canary Never Smokers
- FHCRCProteomicsAnalyzedData [P: Canary, Organ: Lung, Collaborat...
- Clinical Utility of Urinary C...

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
Fred Hutchinson Cancer Research Center (Biomarker Developmental laboratories)
J Qi
jq@fhcr.org
Lung
Lung and Upper Aerodigestive
Contains from human liver adenocarcinoma cell line A549. Excludes sera

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](#)
- [\(by Batch\)00-1.pdf](#)
- [\(by Batch\)LAMR1.pdf](#)
- [\(by Batch\)FCPS-1.pdf](#)
- [annexin1.html+14-3-3.pdf](#)

Result Page 1

ANXA1(FR-00-84)

Stage	n	p-value	AUC
1st	2626	0.005	0.73
2nd	2727	0.034	0.78
3rd	3202	0.042	0.87
all	8585	0.001	0.71

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

Biomarker Database: Capture of EDRN Biomarkers Under Research

- **A database of annotated biomarkers that are either under development or reported in publications**
 - Over 900 biomarkers captured
 - Based on EDRN research

- **A national biomarker registry to track EDRN progress**
 - Similar to gene and protein type registries
 - Essential to tracking progress and supporting national collaboration of biomarker discovery and validation

- **Share results with the broader research community**

- **Integrate with existing databases (e.g., genomic, publication, etc. databases)**

Biomarker Database Content

928 individual biomarkers

Ten organs represented

- Bladder (2)
- Breast (193)
- Colon (13)
- Esophagus (11)
- Head and Neck (8)
- Liver (9)
- Lung (182)
- Ovary (205)
- Pancreas (7)
- Prostate (387)

- 27 biomarker “panels” or “signatures” included
- 79 biomarkers associated with multiple organs (e.g. information detailing *p16* activity in Breast, Esophagus, Lung and Prostate)
- Expert review completed in each Collaborative Group
- Markers from all four Collaborative Groups “Accepted” and released to public view

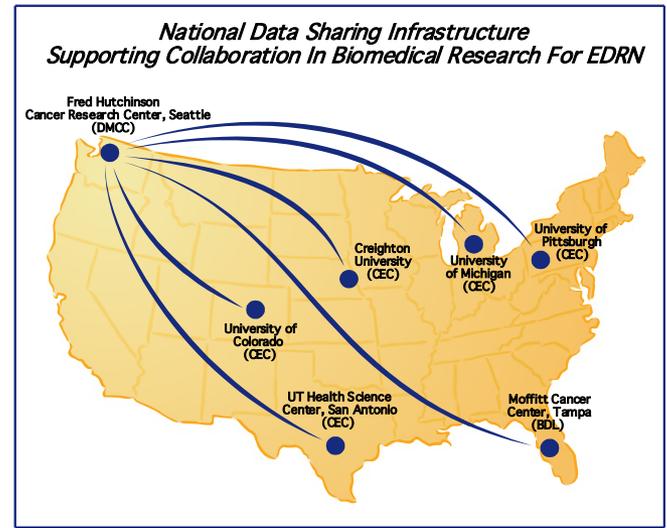
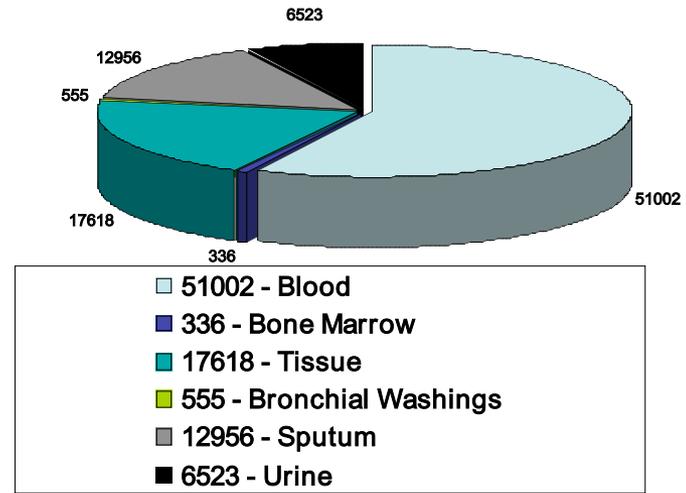
Research, Communicate, Curate, Annotate

Ability to locate specimens across EDRN Clinical Centers: ERNE

- **Specimen Locator System, dubbed ERNE, EDRN Resource Network Exchange was developed to query data across EDRN's Clinical Validation Centers (CVC)**
- **The system is based on NASA JPL's Object-oriented Data module which can be easily tailored to the CVC's institutional informatics system**
 - Same software module used to share earth and planetary science data
- **ERNE allows the user to query the availability of specimens in real-time**
- **This is a first-ever system developed to query specimen on disparately distributed specimens across the country**
 - "ERNE" has been a model studied by many groups

Participants in Virtual Specimen Bank

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



Public Portal

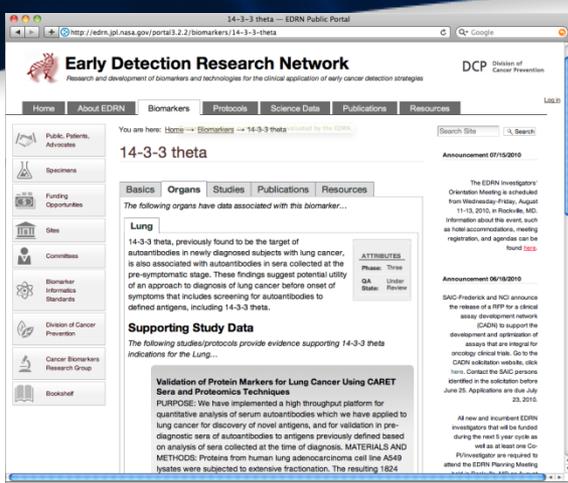
- **A one stop shop to access EDRN programmatic and science information**
 - Operated by NCI on the cancer.gov network
 - Used by NCI and EDRN for disseminating program information
 - Information spans from biomarker data all the way to member information
 - Integrates information from the DMCC
 - **Google-like search feature**

- **We have approximately 900 registered users and we get about 2400 unique visits a month**
 - A lot of ad hoc requests for data from outside EDRN



EDRN Public Portal
(<http://edrn.nci.nih.gov>)

Accessing EDNRN Data: Data may come from any of these sources

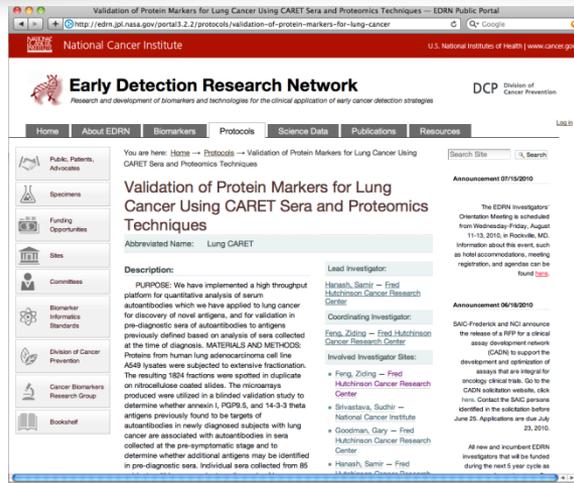


Early Detection Research Network

14-3-3 theta

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



Early Detection Research Network

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

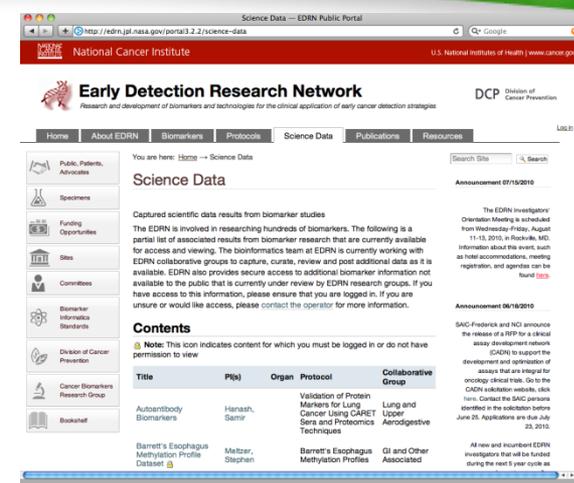
Description: The EDNRN Investigators' Orientation Meeting is scheduled from Wednesday-Friday, August 11-13, 2010, in Rockville, MD. Information about the event, such as hotel accommodations, meeting registration, and agendas can be found [here](#).

Lead Investigator: Hanah, 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, Sudhi - National Cancer Institute
- Goodman, Gary - Fred Hutchinson Cancer Research Center
- Hanah, Samir - Fred Hutchinson Cancer Research Center



Early Detection Research Network

Science Data

Captured scientific data results from biomarker studies

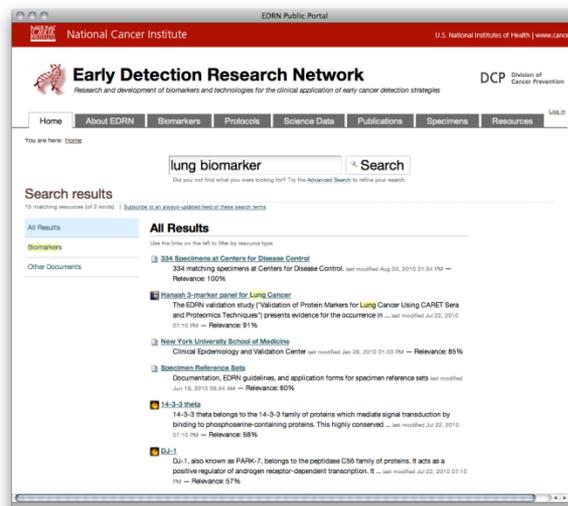
The EDNRN is involved in researching hundreds of biomarkers. The following is a partial list of associated results from biomarker research that are currently available for access and viewing. The bioinformatics team at EDNRN is currently working with EDNRN collaborative groups to capture, curate, review and post additional data as it is available. EDNRN also provides secure access to additional biomarker information not available to the public that is currently under review by EDNRN 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.

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

Biomarker Annotations

Protocols

Biomarker Data Results

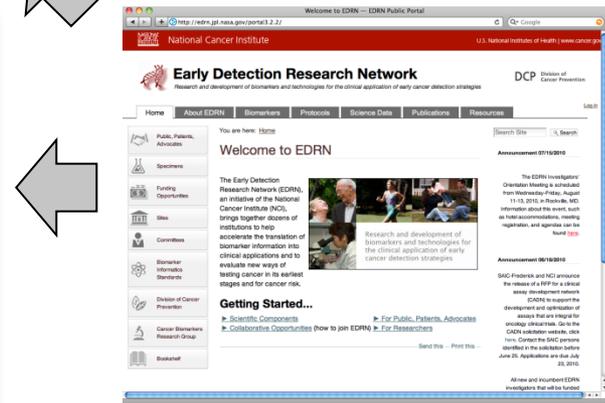


Early Detection Research Network

Search results for **lung biomarker**

All Results

- 334 Specimens at Centers for Disease Control
- 334 matching specimens at Centers for Disease Control, last modified Aug 20, 2010 01:34 PM - Relevance: 100%
- Hanah 3-marker panel for Lung Cancer
- The EDNRN validation study ("Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques") presents evidence for the occurrence of... last modified Jul 22, 2010 07:16 PM - Relevance: 91%
- New York University School of Medicine
- Clinical Epidemiology and Validation Center last modified Jun 20, 2010 01:33 PM - Relevance: 85%
- Specimen Reference Sets
- Documentation, EDNRN guidelines, and application forms for specimen reference sets last modified Jun 10, 2010 03:44 AM - Relevance: 80%
- 14-3-3 theta
- The EDNRN validation study ("Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques") presents evidence for the occurrence of... last modified Jul 22, 2010 07:16 PM - Relevance: 58%
- Du-1
- Du-1, also known as PARC-7, belongs to the peptidase C16 family of proteins. It acts as a positive regulator of androgen receptor-dependent transcription. It... last modified Jul 22, 2010 07:16 PM - Relevance: 57%

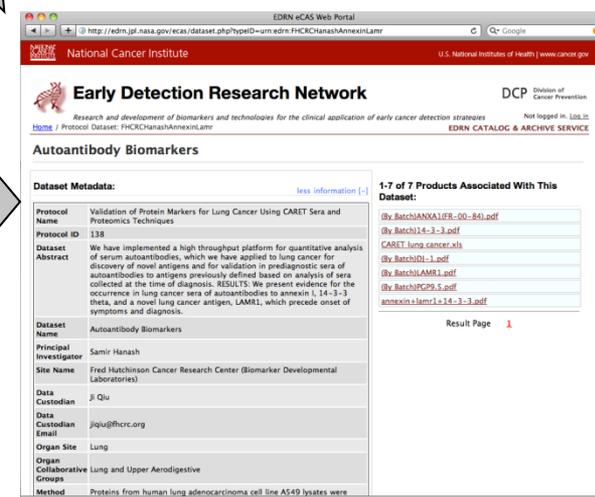


Early Detection Research Network

Welcome to EDNRN

Getting Started...

- Scientific Components
- Collaborative Opportunities (How to join EDNRN)
- For Public, Patients, Advocates
- For Researchers



Early Detection Research Network

Autoantibody Biomarkers

Dataset Metadata:

1-7 of 7 Products Associated With This Dataset:

Protocol Name	Principal Investigator	Site Name	Data Custodian	Organ Site	Organ Collaborative Groups	Method
Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques	Samir Hanah	Fred Hutchinson Cancer Research Center (Biomarker Developmental Laboratories)	jiqiu@fhcr.org	Lung	Lung and Upper Aerodigestive	Proteins from human lung adenocarcinoma cell line A549 lysates were

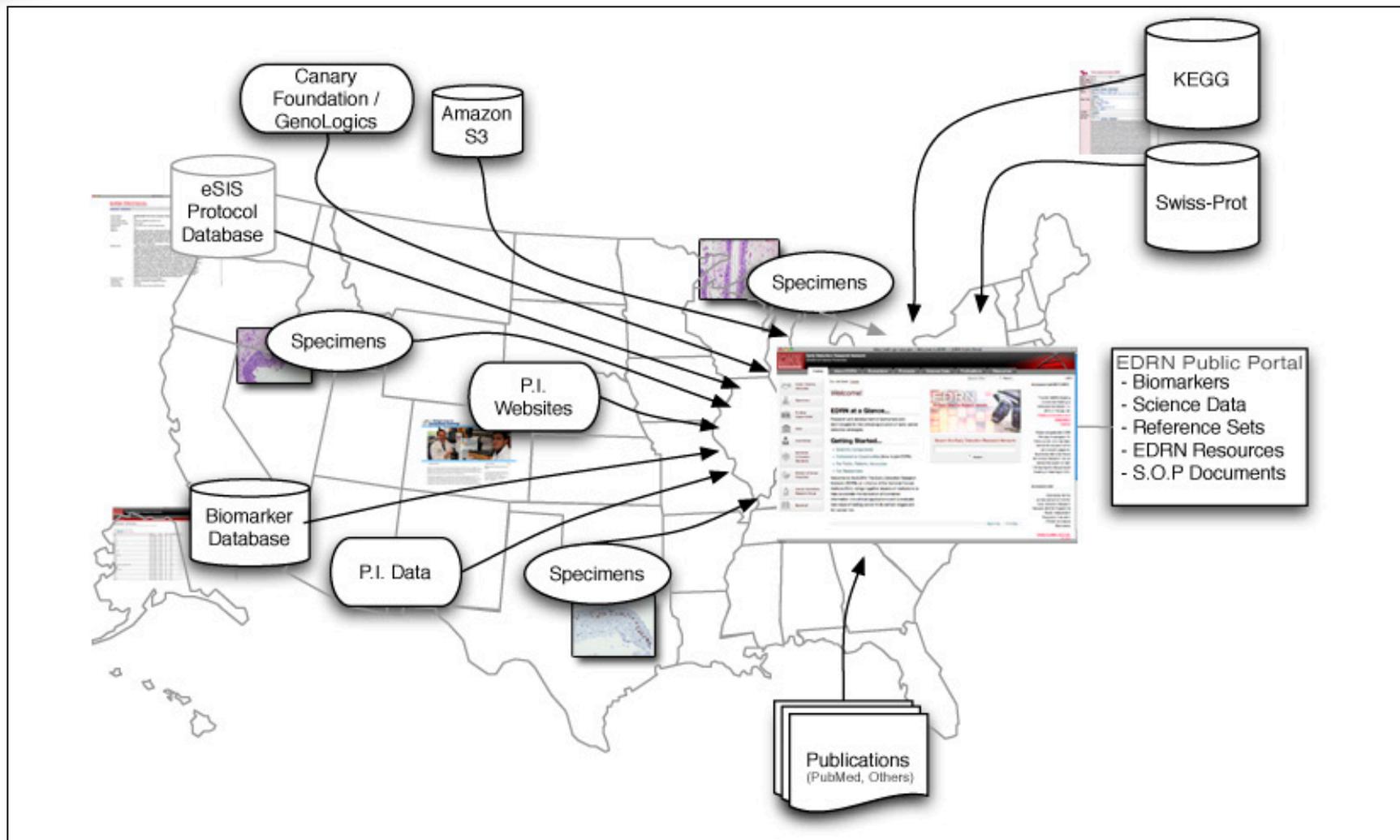
Specimens

Linked through Public Portal

Access to download data

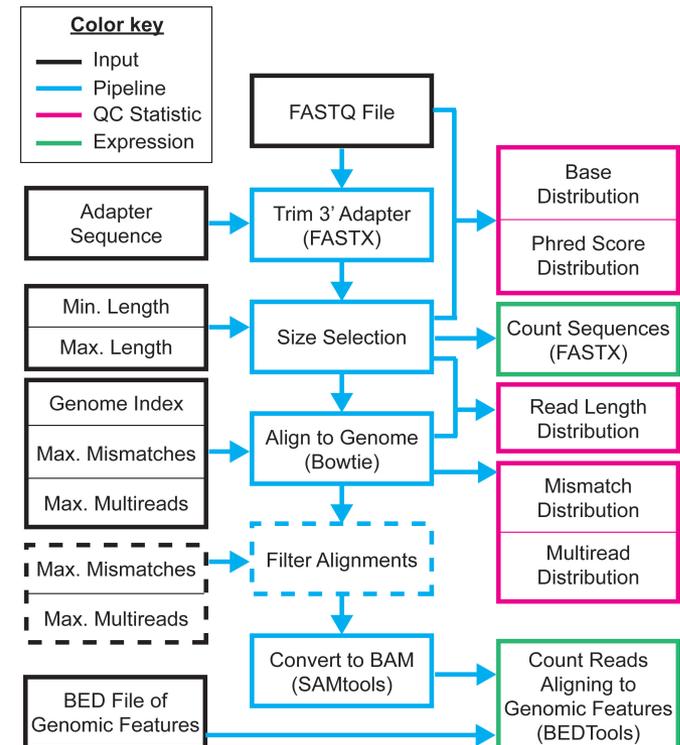
What's Emerged...

The EDRN Knowledge System



Numerous Collaborations

- **Reproducibility experiments**
 - EDRN PLCO
- **Analytic Data Pipelines**
 - Boston University
 - Vanderbilt
 - PNNL
 - NIST/Stanford
- **Bioinformatics Tool Integration**
 - Secretome
- **Data citation in eCAS**
- **Biomarker Database/Results**



*Courtesy of Josh Campbell and
Teresa Wang*

MCL Knowledge Environment



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

Home About ED RN Biomarkers Protocols Science Data Publications Resources

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

14-3-3 theta

Announcement 07/19/2010

The ED RN investigators Overview Meeting is scheduled from Wednesday-Friday, August 11-13, 2010, in Rockville, MD. Information about the event, such as hotel accommodations, meeting registration, and agendas can be found [here](#).

Basics Organs Studies Publications Resources

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

Lung

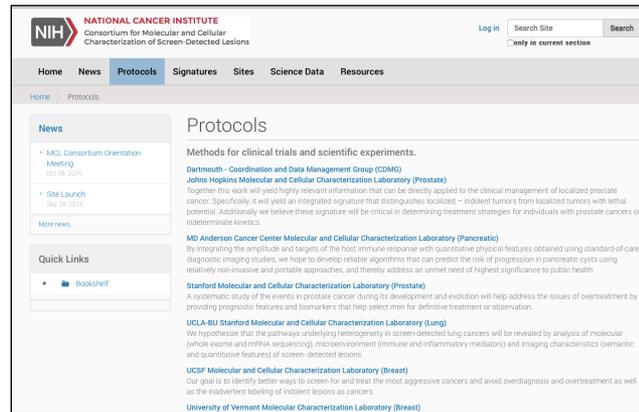
14-3-3 theta, previously found to be the target of autoantibodies in newly diagnosed subjects with lung cancer, is also associated with autoantibodies in sera collected at the pre-symptomatic stage. These findings suggest potential utility of an approach to diagnosis of lung cancer before onset of symptoms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

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



NATIONAL CANCER INSTITUTE
Consortium for Molecular and Cellular Characterization of Screen-Detected Lesions

Log in Search Site only in current section

Home News Protocols Signatures Sites Science Data Resources

Protocols

Methods for clinical trials and scientific experiments.

Dartmouth - Coordination and Data Management Group (CDMG)
Johns Hopkins Molecular and Cellular Characterization Laboratory (Prostate)

Together this work will yield highly relevant information that can be directly applied to the clinical management of localized prostate cancer. Specifically, it will yield an integrated signature that distinguishes localized - indolent tumors from localized tumors with lethal potential. Additionally we believe these signature will be critical in determining treatment strategies for individuals with prostate cancers of indeterminate kinetics.

MD Anderson Cancer Center Molecular and Cellular Characterization Laboratory (Pancreatic)

By integrating the amplitude and targets of the host immune response with quantitative physical features obtained using standard-of-care diagnostic imaging studies, we hope to develop reliable algorithms that can predict the risk of progression in pancreatic cystic lesion relatively noninvasive and portable approaches, and thereby address an unmet need of highest significance to public health.

Stanford Molecular and Cellular Characterization Laboratory (Prostate)

A systematic study of the events in prostate cancer during its development and evolution will help address the issues of overtreatment by providing prognostic features and biomarkers that help select men for definitive treatment or observation.

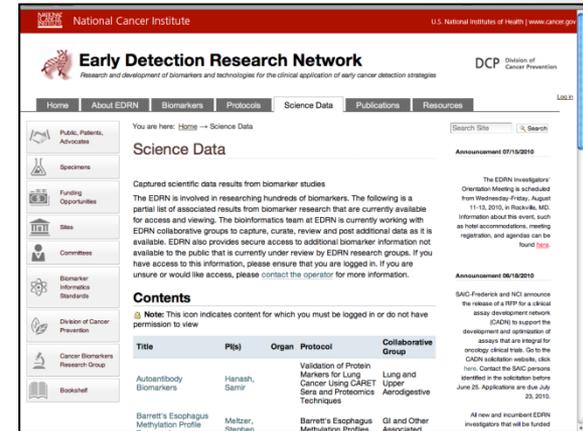
UCLA-BU Stanford Molecular and Cellular Characterization Laboratory (Lung)

We hypothesize that the pathways underlying heterogeneity in screen-detected lung cancers will be revealed by analysis of molecular (whole exome and mRNA sequencing), microenvironment (immune and inflammatory mediators) and imaging characteristics (semantic and quantitative features) of screen- detected lesions.

UCSF Molecular and Cellular Characterization Laboratory (Breast)

Our goal is to identify better ways to screen for and treat the most aggressive cancers and avoid overdagnosis and overtreatment as well as the inadvertent labeling of indolent lesions as cancer.

University of Vermont Molecular Characterization Laboratory (Breast)



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

Home About ED RN Biomarkers Protocols Science Data Publications Resources

You are here: Home → Science Data

Science Data

Announcement 07/19/2010

Captured scientific data results from biomarker studies

The ED RN 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 biominformatics team in ED RN is currently working with ED RN collaborative groups to capture, curate, review and post additional data as it is available. ED RN also provides secure access to additional biomarker information not available to the public that is currently under review by ED RN 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

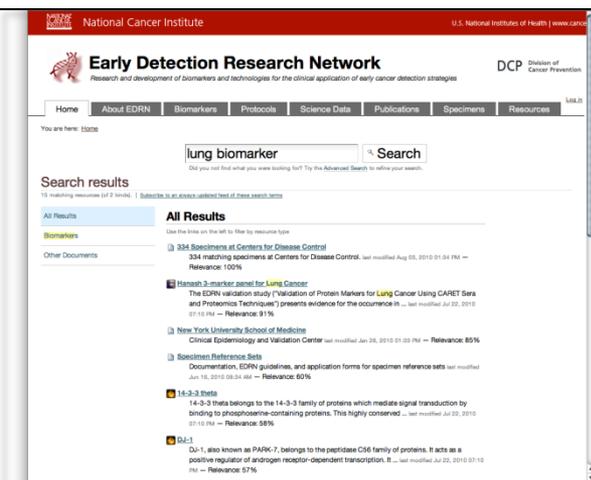
Title	File(s)	Organ	Protocol	Collaborative Group
Autoantibody Biomarkers	Hanah, Samir	Lung and Cancer Using CARET	Validation of Protein Markers for Lung and Cancer Using CARET	Upper Sire and Proteomics Aerodigestive
Barrett's Esophagus Methylation Profiles	Meltzer, Hanshan	GI and Other	Barrett's Esophagus Methylation Profiles	GI and Other

All new and incumbent ED RN investigators that will be funded

Signature Annotations
(Biomarker Database)

Protocols

Signature Data Results
(eCAS)



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

Home About ED RN Biomarkers Protocols Science Data Publications Specimens Resources

You are here: Home

lung biomarker

Search results

All Results

Biomarkers

354 Specimens at Centers for Disease Control
334 matching specimens at Centers for Disease Control. last modified Aug 03, 2010 01:04 PM → Relevance: 100%

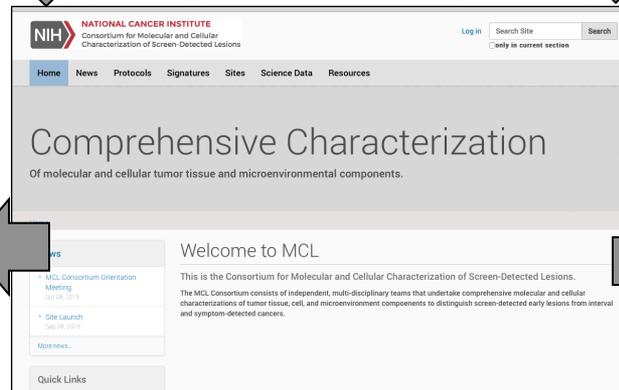
Hanah 3-marker panel for Lung Cancer
The ED RN validation study "Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques" presents evidence for the occurrence in... last modified Jul 20, 2010 07:10 PM → Relevance: 91%

New York University School of Medicine
Clinical Epidemiology and Variation Center last modified Jan 26, 2010 01:03 PM → Relevance: 85%

Specimen Reference Sets
Documentation, ED RN guidelines, and application forms for specimen reference sets last modified Jun 16, 2010 08:34 AM → Relevance: 80%

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 20, 2010 07:10 PM → Relevance: 58%

Du1
Du1, also known as PARC 7, belongs to the peptidase C16 family of proteins. It acts as a positive regulator of androgen receptor-dependent transcription. E... last modified Jul 23, 2010 01:10 PM → Relevance: 57%



NATIONAL CANCER INSTITUTE
Consortium for Molecular and Cellular Characterization of Screen-Detected Lesions

Log in Search Site only in current section

Home News Protocols Signatures Sites Science Data Resources

Comprehensive Characterization

Of molecular and cellular tumor tissue and microenvironmental components.

Welcome to MCL

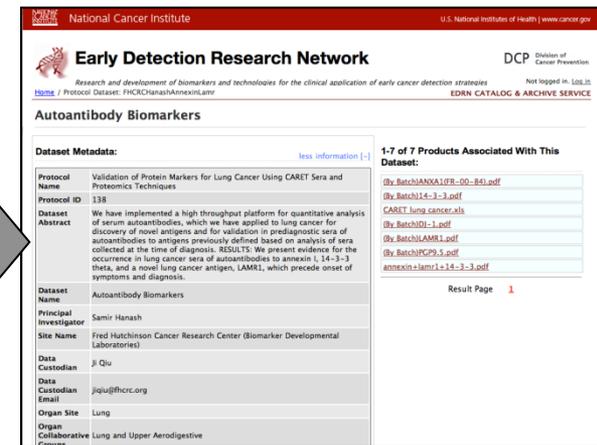
This is the Consortium for Molecular and Cellular Characterization of Screen-Detected Lesions. The MCL Consortium consists of independent, multi-disciplinary teams that undertake comprehensive molecular and cellular characterizations of tumor tissue, cell, and microenvironment components to distinguish screen-detected early lesions from interval and symptom-detected cancers.

MCL Consortium Orientation Meeting
04/08/2010

Site Launch
06/28/2010

More news...

Quick Links



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

Home / Protocol Dataset: FHCCR/HanahSamirLamr Not logged in. Log in ED RN CATALOG & ARCHIVE SERVICE

Autoantibody Biomarkers

Dataset Metadata: less information [-]

1-7 of 7 Products Associated With This Dataset:

Protocol Name	Protocol ID	Dataset Abstract	File(s)
Validation of Protein Markers for Lung Cancer Using CARET Sera and Proteomics Techniques	138	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. 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, LAM1, which precede onset of symptoms and diagnosis.	file:Batch14-3-3.pdf Batch14-3-3.pdf Batch14-3-3.pdf Batch14-3-3.pdf Batch14-3-3.pdf

Dataset Name: Autoantibody Biomarkers

Principal Investigator: Samir Hanah

Site Name: Fred Hutchinson Cancer Research Center (Biomarker Developmental Laboratories)

Data Custodian: Ji Qiu

Data Custodian Email: jqiu@fhccr.org

Organ Site: Lung and Upper Aerodigestive

Organ Collaborative Group: Lung and Upper Aerodigestive

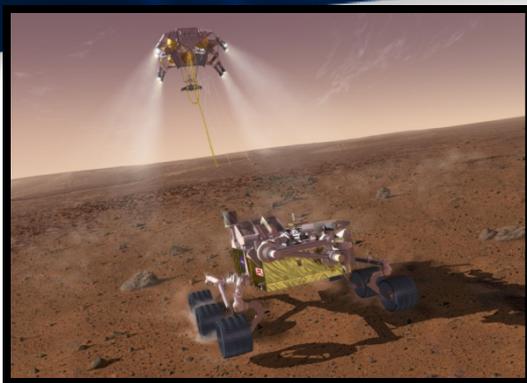
Result Page 1

Specimens
(ERNE)

Linked through
Public Portal

Access to download data
(eCAS)

Conclusion



Reproduced from *Wired* magazine



EDRN has a focused, leading-edge, informatics platform that can be leveraged today for sharing data.



http://twitter.com/edrn_ic

[http://www.facebook.com/
group.php?gid=56938589930](http://www.facebook.com/group.php?gid=56938589930)

