A web resource for exploring integrated cancer biomarker data in the context of mutation, differential expression, evolutionarily conserved expression patterns, and automatically mined literature evidence

Presented by:
Hayley Dingerdissen – GW Project Lead

Multi-PIs:
Dan Crichton – NASA JPL
Raja Mazumder – GW

Sub-awardees:
Frédéric Bastian – SIB; Vijay Shanker – UD
A semantic web of integrated cancer biomarker data

AUTOMATICALLY MINED LITERATURE

BIOMARKERS

DIFFERENTIAL EXPRESSION IN CANCER

CANCER EXPRESSION SPECIFICITY

NCBI SRA

DATA INTEGRATION AND ONTOLOGY UNIFICATION

COMMUNITY ENGAGEMENT AND VISUALIZATION

DATA PROVENANCE INFRASTRUCTURE

MORE

EVOLUTIONARY CONSERVATION OF NORMAL EXPRESSION

BIOMARKERS

MUTATION

TARGET

BIOMARKERS

NASS JPL

Early Detection Research Network

UNIVERSITY OF DELAWARE

reactome

PATHWAYS

IMMUNE EPITOME DATABASE AND ANALYSIS RESOURCE

Custom

FDA Approved Breast Cancer Biomarkers

RECOUNT2

GTEX Portal

Bgee

Evolutionary conservation of normal expression

4.0
Resulting web portal – a PCA3 case study

Details on this gene and all the associated data sources.

<table>
<thead>
<tr>
<th>Anatomical Entity ID</th>
<th>Anatomical Entity Name</th>
<th>Developmental Stage ID</th>
<th>Developmental Stage Name</th>
<th>Expression Lvl - Gene</th>
<th>Expression Lvl - Anatomical</th>
<th>Call Quality</th>
<th>Expression Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>UBERON:0002367</td>
<td>prostate gland</td>
<td>HsapDv:0000094</td>
<td>65-79 year-old human stage (human)</td>
<td>HIGH</td>
<td>HIGH</td>
<td>GOLD</td>
<td>1.19e4</td>
</tr>
<tr>
<td>UBERON:0002367</td>
<td>prostate gland</td>
<td>HsapDv:0000091</td>
<td>human late adulthood stage (human)</td>
<td>MEDIUM</td>
<td>MEDIUM</td>
<td>GOLD</td>
<td>2.30e4</td>
</tr>
<tr>
<td>UBERON:0002367</td>
<td>prostate gland</td>
<td>HsapDv:0000087</td>
<td>human adult stage (human)</td>
<td>MEDIUM</td>
<td>MEDIUM</td>
<td>GOLD</td>
<td>2.36e4</td>
</tr>
<tr>
<td>UBERON:0002245</td>
<td>cerebellar hemisphere</td>
<td>HsapDv:0000094</td>
<td>65-79 year-old human stage (human)</td>
<td>ABSENT</td>
<td>ABSENT</td>
<td>GOLD</td>
<td>2.52e4</td>
</tr>
<tr>
<td>UBERON:0000955</td>
<td>brain</td>
<td>HsapDv:0000087</td>
<td>human adult stage (human)</td>
<td>MEDIUM</td>
<td>MEDIUM</td>
<td>GOLD</td>
<td>2.54e4</td>
</tr>
<tr>
<td>UBERON:0014890</td>
<td>right hemisphere of cerebellum</td>
<td>HsapDv:0000094</td>
<td>65-79 year-old human stage (human)</td>
<td>ABSENT</td>
<td>ABSENT</td>
<td>GOLD</td>
<td>2.55e4</td>
</tr>
</tbody>
</table>
Future directions

Development plans

- Refining biomarker data model
- Integrated views
- Increase external linkages
- Increase cancer biomarker (and other) datasets
- Explore linkages to imaging projects/databases

Collaboration

- SingleCellITK
- INODE (use-case for EU grant with SIB)
- IEDB
- IMAT (planning to submit LOI)
- Looking for additional ITCR collaborators

Upcoming outreach

- ISMB (at ITCR booth)
- JCO manuscript submission (planned)
- AMIA (submitted)
- EDRN (planning to submit)
- OncoMX user community workshop (tentatively September 16, 2019)
Acknowledgements

• **Raja Mazumder (Multi-PI)**
  • Robel Kahsay
  • Amanda Bell
  • Evan Holmes
  • Stephanie Singleton
  • Reza Mousavi
  • Charles Hadley King
  • Jonathon Keeney
  • Nicole Ferrer
  • Olga Reghay
  • Dacian Stremtan

• **Jeet Vora**
• **Rahi Navelkar**
• **Yao Ren**
• **Janisha Patel**
• **Naila Gulzar**

• **Daniel Crichton (Multi-PI)**
  • Heather Kincaid
  • David Liu

**Previous contributions**

• Ashique Mahmood
• Lama ElZohary
• Yu Hu
• Ting-Chia Chang
• Prince Birring
• Julien Wollbrett

• **Frédéric Bastian (Sub-Awardee)**
• Marc Robinson-Rechavi

• **K. Vijay-Shanker (Sub-Awardee)**
• Samir Gupta

Funding: NCI U01CA2015010; NCI EDRN award 156620
Questions?

Poster #22 Tonight

We are looking for biomarker data (preferably mapped to gene level) – please let us know if you have data and would like to collaborate!

Join us at our next workshop!
Tentatively September 16, 2019

Contact us:
Hayley Dingerdissen – hmhamilt@gwmail.gwu.edu
Dan Crichton – daniel.j.crichton@jpl.nasa.gov
Raja Mazumder – mazumder@gwu.edu

@OncoMX_KB
A web resource for exploring integrated cancer biomarker data in the context of mutation, differential expression, evolutionarily conserved expression patterns, and automatically mined literature evidence

Presented by:
Hayley Dingerdissen – GW Project Lead

Multi-PIs:
Dan Crichton – NASA JPL
Raja Mazumder – GW

Sub-awardees:
Frédéric Bastian – SIB; Vijay Shanker - UD
Resulting web portal – a PCA3 case study

https://www.oncomx.org/
(https://beta.oncomx.org/)

**PCA3  prostate cancer associated 3 [ Homo sapiens (human) ]**

Gene ID: 50652, updated on 13-Feb-2019

**Summary**

- **Official Symbol**: PCA3 provided by HGNC
- **Official Full Name**: prostate cancer associated 3 provided by HGNC
- **Primary source**: HGNC:HGNC:8637
- **See related**: Ensembl:ENSG00000225937, MIM:604845
- **Gene type**: ncRNA
- **RefSeq status**: REVIEWED
- **Organism**: Homo sapiens
- **Lineage**: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarhini; Hominidae; Homo
- **Also known as**: DD3; PCAT3; NCRNA00019; PRUNE2-AS1

This gene produces a spliced, long non-coding RNA that is highly overexpressed in most types of prostate cancer cells and is used as a specific biomarker for this type of cancer. This gene is embedded in an intronic region of the prune2 gene on the opposite DNA strand. The transcript regulates prune2 levels through formation of a double-stranded RNA that undergoes adenosine deaminase acting on RNA-dependent adenosine-to-inosine RNA editing. In prostate cancer derived cells, overexpression of PCA induced downregulation of prune2, leading to decreased cell proliferation. Conversely, silencing in prostate cancer cells resulted in increased proliferation. Regulation of this gene appears to be sensitive to androgen-receptor activation, a molecular signature of prostate cancer. Alternative splicing results in multiple transcript variants. [provided by RefSeq, Jan 2017]

**Expression**

Restricted expression toward prostate (RPKM 4.9) [See more]

Resulting web portal – a PCA3 case study

Explore gene specific differential expression data

Try: AFP YWHAE BRCA1
Resulting web portal – a PCA3 case study

<table>
<thead>
<tr>
<th>Gene Symbol/Panel</th>
<th>Type</th>
<th>Associated Dataset</th>
<th>Is Panel</th>
<th>Phase</th>
<th>QA State</th>
<th>Organ</th>
<th>HGNC Symbol</th>
<th>Reference Resource</th>
<th>UniProtKB/SwissProt AC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laxman 7-marker panel for prostate cancer</td>
<td>Gene</td>
<td></td>
<td>Panel</td>
<td></td>
<td>Curated</td>
<td>Prostate</td>
<td></td>
<td>Reference 1</td>
<td>Reference 2</td>
</tr>
<tr>
<td>MIPS (MI Prostate Score Urine test)</td>
<td>Gene</td>
<td></td>
<td>Panel</td>
<td></td>
<td>Curated</td>
<td>Prostate</td>
<td></td>
<td>Reference 1</td>
<td>Reference 2</td>
</tr>
<tr>
<td>PCA3</td>
<td>Gene</td>
<td></td>
<td>Biomarker</td>
<td></td>
<td>Curated</td>
<td>Prostate</td>
<td>PCA3</td>
<td>Reference 1</td>
<td>Reference 2 Reference 3</td>
</tr>
</tbody>
</table>

Showing 1 to 3 of 3 entries (filtered from 939 total entries)
Resulting web portal – a PCA3 case study
Resulting web portal – a PCA3 case study

OVERVIEW SEARCH RESULTS FOR GENE: PCA3 /

Gene specific overview of data in OncomiX. Access to gene specific details here.

PCA3

Details on this gene and all the associated data sources.

Biomarker Details

EDRN Title/PCA3

Organ: Prostate

Phases: Curated

Alias:

prostate cancer associated transcript 3 (non-protein coding), NCRNA00019, D03, prostate cancer antigen 3 (non-protein coding), PCA3, non-protein coding RNA 19, prostate cancer antigen 3

Description:
The prostate cancer antigen 3 (PCA3) gene is a highly specific biomarker upregulated in prostate cancer. Because there is no established open reading frame, the gene is thought to express a non-coding RNA. The FDA has approved an assay for prostate cancer that uses the ratio of PCA3 RNA to PSA RNA in urine. The assay is used to help physicians determine the need for repeat prostate biopsies in men 56 years of age or older who have had a previous negative biopsy. When used in conjunction with other clinical information, the use of this test can reduce the need for unnecessary prostate biopsies.

EDRN Publications:

http://nci.nih.gov/ncipdpublications/view/96
http://nci.nih.gov/ncipdpublications/view/42
http://nci.nih.gov/ncipdpublications/view/40
http://nci.nih.gov/ncipdpublications/view/143
http://nci.nih.gov/ncipdpublications/view/12
http://nci.nih.gov/ncipdpublications/view/185
http://nci.nih.gov/ncipdpublications/view/247
http://nci.nih.gov/ncipdpublications/view/330

Expression Details

Log2 P.C. | Adj. P. value | Significant | Expression Trend | TCGA Study | Cancer Type | Cancer Variant | Patient Freq. | Source | PMID | Anatomical Entity ID

2.1 | 4.3e-10 | 6.32e-08 | Yes | Up | PROAD | D0312973 / Prostate cancer (PCa) | 49/52(94.23) | R14.96e+12 | UBERON.00002387

Showing 1 to 1 of 1 entries (Filtered from 19 total entries)

EDRN Details

Biomarker

Dyes Details

Show 10 * T-entries

Anatomical Entity ID | Anatomical Entity Name | Developmental Stage ID | Developmental Stage Name | Expression Level - Gene | Expression Level - Anatomical | Call Quality | Expression Rank

UBERON.00002367 | prostate gland | Hasp00016894 | 65.79 year-old human stage (human) | HIGH | HIGH | GOLD | 1.14e+4
UBERON.00002367 | prostate gland | Hasp0002551 | human late adulthood stage (human) | MEDIUM | MEDIUM | GOLD | 2.05e+4
UBERON.00002397 | human adult stage (human) | Hasp0002551 | human adult stage (human) | MEDIUM | MEDIUM | GOLD | 2.56e+4
UBERON.00002349 | cerebellum hemispheres | Hasp0002551 | human adult stage (human) | MEDIUM | MEDIUM | GOLD | 2.52e+4
UBERON.00002355 | brain | Hasp0002551 | human adult stage (human) | MEDIUM | MEDIUM | GOLD | 2.54e+4
Use-case driven interface development

<table>
<thead>
<tr>
<th>As a [user profile]</th>
<th>I want to [requirement]</th>
<th>So that [benefit]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research associate working on cancer informatics</td>
<td>Identify a list of genes expressed in lung cancer</td>
<td>I can identify common pathways implicated in lung cancer expression</td>
</tr>
<tr>
<td>Clinical researcher</td>
<td>Evaluate a list of biomarkers available in OncoMX and also in the EDRN portal</td>
<td>I can compare biomarkers available in OncoMX as well as EDRN portal</td>
</tr>
<tr>
<td>PhD student studying the role of mutation in breast cancer</td>
<td>Identify mutations in all forms of breast cancer</td>
<td>I can see which important mutations are conserved in breast cancer</td>
</tr>
<tr>
<td>Post-doc in bioinformatics</td>
<td>Observe differential expression reported in literature</td>
<td>I can compare experimentally identified genes with existing literature</td>
</tr>
</tbody>
</table>

**USER SEARCH**
- Mutations in breast cancer
- Expression in ovarian cancer
- Pathways implicated with marker across cancers

**DATA INTEGRATION**
- Biomarkers
- Cancer mutation & expression
- Normal expression across species
- Literature mining for disease associations
- Implicated pathways

**OUTPUT PERSPECTIVES**
- CANCER BIOMARKERS
- EVOLUTIONARY CONTEXT
- LITERATURE MINING
- BIOMARKERS IN PATHWAYS
Major features

Integrated, single, accessible, harmonized, online resource, built to support extensibility of different data types over time

- User community focus
- Provenance standards (BCO)
- Definition of a biomarker data model
- Ontology-guided data integration
- Leveraging existing literature
- Adherence to FAIR principles