



---

## Webinar 3: Expression Data Analysis & OQL

May 14, 2020



- April 30: Introduction to cBioPortal (recording on [cBioPortal.org](https://cBioPortal.org))
- May 7: Mutation Details & Patient View (recording on [cBioPortal.org](https://cBioPortal.org))
- May 14: Expression Data Analysis & OQL
- May 21: Group Comparison
- May 28: API & R Client

All webinars are on Thursdays 11am-12pm EDT

# Acknowledgements



Nikolaus Schultz  
Jianjiong Gao  
Benjamin Gross  
S. Onur Sumer  
Yichao Sun  
Hongxin Zhang  
Adam Abeshouse  
Ritika Kundra  
Ino de Bruijn  
Robert Sheridan  
Angelica Ochoa  
Aaron Lisman  
Manda Wilson  
Avery Wang  
Jing Su  
Ramyasree Madupuri  
Gaofei Zhao  
Xiang Li



Ethan Cerami  
Chris Sander  
Tali Mazor  
Luke Sikina  
Pieter Lukasse  
Priti Kumari  
Augustin Luna  
James Lindsay



The Children's Hospital of Philadelphia

Adam Resnick  
Allison Heath  
Karthik Kalletla  
John Maris



Olivier Elemento  
Alexandros Sigaras



Fedde Schaeffer  
Oleguer Plantalech  
Pim van Nierop  
Paul van Dijk  
Sander Rodenburg  
Sjoerd van Hagen  
Kees van Bochove



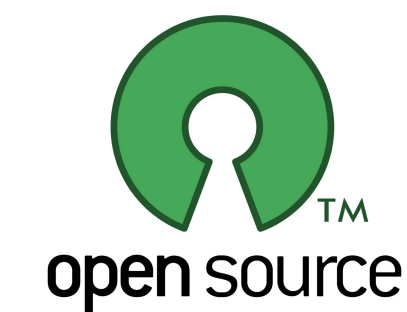
Trevor Pugh  
Kelsey Zhu



Ugur Dogrusoz  
Ziya Erkoc

## Alumni

B. Arman Aksoy  
Istemi Bahceci  
Caitlin Byrne  
Hsiao-Wei Chen  
Ersin Ciftci  
Fred Criscuolo  
Leonard Dervishi  
Gideon Dresdner  
Andy Duffile  
Catherine Del Vecchio Fitz  
Arthur Goldberg  
Zachary Heins  
Michael Heuer  
Anders Jacobsen  
Ngoc Nguyen



Erik Larsson  
Dong Li  
Tamba Monroe  
Peter Kok  
Irina Pulyakhina  
Pichai Raman  
M. Furkan Sahin  
Kaan Sancak  
Sander de Ridder  
Sander Tan  
Jiaojiao Wang  
Stuart Watt  
James Xu  
Dionne Zaal  
Riza Nugraha

## Funding: Present & past



Email: [cbioportal@googlegroups.com](mailto:cbioportal@googlegroups.com)

 [@cBioPortal](https://twitter.com/cBioPortal)

 [github.com/cBioPortal](https://github.com/cBioPortal)



# Questions?

---

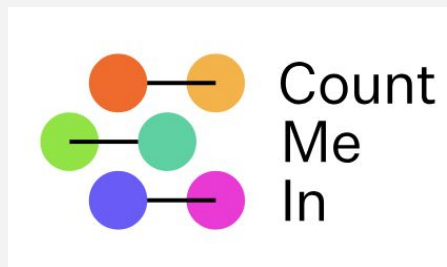
- During this webinar, please ask questions using the Zoom Q&A feature
- We will try to answer some live, some directly, and some on future webinars
- The recording of all webinars will be posted on <https://cbioportal.org/tutorials/>
- If you still have questions after the webinar, please contact us at [cbioportal@googlegroups.com](mailto:cbioportal@googlegroups.com)

- Platform for **exploratory and interactive visualization, analysis and download** of large-scale cancer genomics data sets
- **Open source** software jointly developed by Dana-Farber Cancer Institute, Memorial Sloan Kettering Cancer Center, Princess Margaret Cancer Centre, Children's Hospital of Philadelphia, and The Hyve
- **Public website** (cBioPortal.org) with public data (TCGA, ICGC, published sequencing studies)
- **Private instances** are installed at academic and commercial institutions world-wide

# What data is in cBioPortal.org?

---

## Data sources



## Clinical data:

- Demographic
- Disease
- Treatments
- Survival
- etc



## Molecular data:

- Mutations
- Fusions
- Copy number
- DNA Methylation
- **RNA expression**
- **Protein levels**

# FAQs related to expression data (from the two previous webinars)

---

Are there genomics data for non-coding genes in cBioPortal?

- **No for mutations**
- **Yes for copy number and RNA expression** (TCGA)

Are there transcript-specific expression data?

- **No.**

What types of expression data are available in cBioPortal?

- **RNA** (mRNA, miRNA, lncRNA):
  - Microarray
  - RNAseq
- **Protein**
  - Reverse phase protein array (RPPA)-based
  - Mass spectrometry (MS)-based

# FAQs related to expression data (from the two previous webinars)

---

## What are the different versions of TCGA datasets?

- **TCGA Firehose Legacy**

- Clinical data, mutations, copy number, RNA, methylation, protein level (RPPA)
- Last updated: 2016

- **TCGA Published Studies**

- Study-specific
- miRNA expression data

- **TCGA PanCancer Atlas**

- Normalized across cancer types
- Currently clinical data, mutations, copy number, RNAseq
- More (e.g. DNA methylation, miRNA, protein levels, etc) will be added

### **Breast**

---

#### **Invasive Breast Carcinoma**

- ☐ Breast Invasive Carcinoma (TCGA, Cell 2015)
- ☐ Breast Invasive Carcinoma (TCGA, Firehose Legacy)
- ☐ Breast Invasive Carcinoma (TCGA, Nature 2012)
- ☐ Breast Invasive Carcinoma (TCGA, PanCancer Atlas)



# FAQs related to expression data (from the two previous webinars)


---

- What are expression z-scores?
  - Z-scores are calculated for all expression data and they are available for query by gene.
  - One of the following reference populations is used when calculating z-scores:
    - All samples in the same study
    - Samples that are diploid of the same gene (based on GISTIC) in the same study
    - Normal samples - only available in one study (Prostate, MSK 2010)
  - Z-scores  $\pm 2$  are the default thresholds for stratifying samples for downstream analyses such as survival analysis and other comparisons.
  - Statements such as “Gene X is over expressed in 10% of tumors of study Z in cBioPortal” is NOT an accurate description.



<https://www.cbioportal.org/>

# Getting help



[Data Sets](#) [Web API](#) [R/MATLAB](#) [Tutorials](#) [FAQ](#) [News](#)

## Tutorials

Tutorial #1: Single Study Exploration

Tutorial Objectives

- Introduction
- Show the data
  - From the portal
  - From the database
- Walk through the data
  - Study
  - Clinical
  - Mutations
  - Copy number
- Show how to use the data

Tutorial #2: Single Study Query

Tutorial Objectives

- Show how to use the query interface
- Walk through the query
  - Oncology
  - Cancer
  - Mutations
  - Plots
  - Mutations
  - Co-expression
- Show how to use the query

Tutorial #3: Patient View

Tutorial Objectives

- Show how to use the patient view
- Walk through the patient view
  - Summary
  - Clinical
  - Pathology
  - Tissues
- Highlight potential use cases for patient view
- Show how to use the patient view

Tutorial #4: Virtual Studies

Tutorial Objectives

- Show how to use the virtual studies
- Walk through the virtual studies
  - Summary
  - Clinical
  - Pathology
  - Tissues
- Highlight potential use cases for virtual studies
- Show how to use the virtual studies

Tutorial #5: Onco Query Language (OQL)


Tutorial Objectives

- Explain the OQL
- Delineate the different ways to define groups and enter the group comparison view
- Show how to use the OQL

Tutorial #6: Group Comparison

Tutorial Objectives

- Introduction
- Explain the group comparison functionality
  - Explain what the new group comparison functionality enables
  - Delineate the different ways to define groups and enter the group comparison view
  - Highlight potential use cases for group comparison functionality
- Show how to use the group comparison






[Data Sets](#) [Web API](#) [R/MATLAB](#) [Tutorials](#) [FAQ](#) [News](#)

## FAQs

- General Questions
  - What is the cBioPortal for Cancer Genomics?
  - How do I get started?
  - What data types are in the portal?
  - What is the process of data curation?
  - How do I get updates on new portal developments and new data sets?
  - Does the portal work on all browsers and operating systems?
  - How do I cite the cBioPortal?
  - Can I use figures from the cBioPortal in my publications or presentations?
  - Can I save or bookmark my results in cBioPortal?

### cBioPortal for Cancer Genomics Discussion Group

60 of 2034 topics (99+ unread)  



PROTOCOL | CANCER

## Integrative Analysis of Complex Cancer Genomics and Clinical Profiles Using the cBioPortal

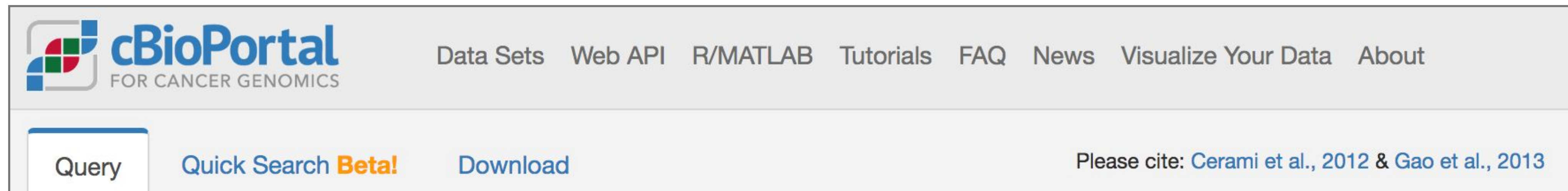
Jianjiong Gao<sup>1</sup>, Bülent Arman Aksoy<sup>1</sup>, Ugur Dogrusoz<sup>2</sup>, Gideon Dresdner<sup>1</sup>, Benjamin Gross<sup>1</sup>, S. Onur Sumer<sup>1</sup>, Yichao Sun<sup>1</sup>, Anders Jacobsen<sup>1</sup>, Rileen Sinha<sup>1</sup>, Erik Larsson<sup>3</sup>, Ethan Cerami<sup>1,4</sup>, Chris Sander<sup>1</sup>, and Nikolaus Schultz<sup>1</sup>

*Sci. Signal.* 02 Apr 2013:  
Vol. 6, Issue 269, pp. p11  
DOI: 10.1126/scisignal.2004088



# Using cBioPortal for publication

- **Please use cBioPortal in your publications!**
  - Figures are downloadable as PDF/SVG so you can customize them
- Cite the studies that generated the data you are using (if applicable)
- Cite cBioPortal





# Questions?

---

- Join us for the remaining webinars
- The recording of all webinars will be posted on <https://cbioportal.org/tutorials/>
- If you still have questions, please ask them via our Google Group or email directly [cbioportal@googlegroups.org](mailto:cbioportal@googlegroups.org)

- April 30: Introduction to cBioPortal (recording on cBioPortal.org)
- May 7: Mutation Details & Patient View (recording on cBioPortal.org)
- May 14: Expression Data Analysis & OQL (recording on cBioPortal.org soon)
- **May 21: Group Comparison**
- **May 28: API & R Client**

*All webinars are on Thursdays 11am-12pm EDT*