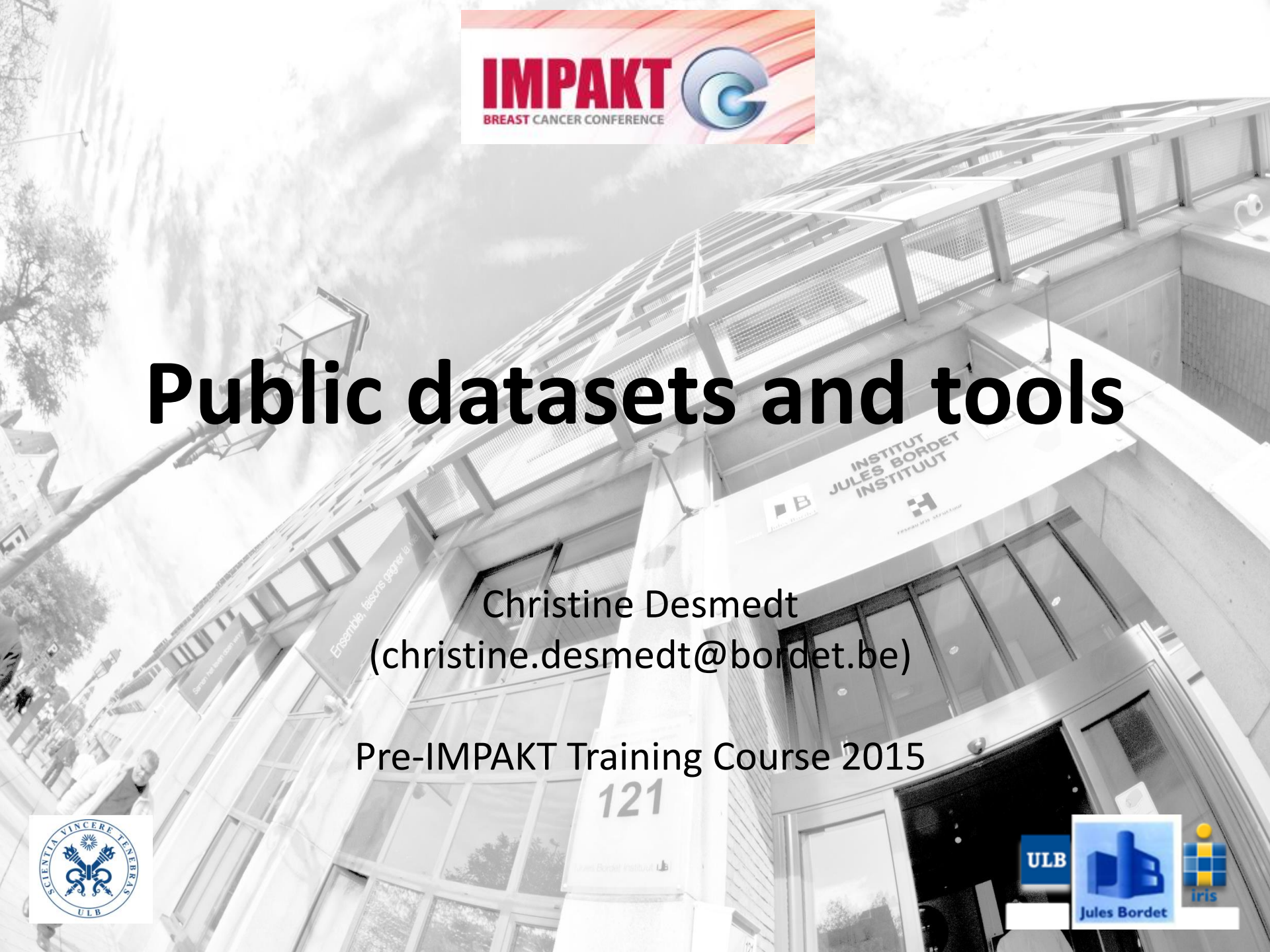


Public datasets and tools

Christine Desmedt
(christine.desmedt@bordet.be)

Pre-IMPAKT Training Course 2015

(FREE) Public datasets and tools

Christine Desmedt
(christine.desmedt@bordet.be)

Pre-IMPAKT Training Course 2014

Outline

1. Clinical tools
2. Research databases
3. Research tools

1. Clinical Tools

Clinical tools

- General info:
  - Cancer.gov
- Clinical trials:
  - Clinicaltrials.gov
- Adjuvant treatment:
  - Adjuvant! Online

Cancer.gov

National Cancer Institute
at the National Institutes of Health

We Can Answer Your Questions
1-800-4-CANCER

SEARCH

NCI Home | Cancer Topics | Clinical Trials | Cancer Statistics | Research & Funding | News | About NCI

Find a Cancer Type

A B C D E F G H I
J K L M N O P Q R
S T U V W X Y Z

NCI Vision & Priorities

NCI Director Harold E. Varmus, M.D.

Provocative Questions Project
Explore provocative research questions with the research community

The National Cancer Program
2013 Annual Plan and Budget Proposal

NCI Budget Overview
Explore how NCI allocates appropriated funds

Popular Resources

NCI Dictionary of Cancer Terms
Cancer-related terms

NCI Drug Dictionary
Definitions, names, and links

NCI Publications
Order/download free booklets

Funding Opportunities
Research and training

Better Research Through Team Tools

Researchers are using team-based approaches to address complex scientific challenges like cancer. If you manage, support, or conduct team-based research, NCI's Team Science Toolkit offers a growing collection of more than 900 resources to help maximize your work.

- What is team science?
- Team Science Toolkit


Changes to NCI's Clinical Trials Programs


Reducing Cancer Health Disparities


Better Research Through Team Tools

Types of Cancer

Common Cancer Types
Bladder Cancer
Breast Cancer
Colon and Rectal Cancer
Endometrial Cancer
Kidney (Renal Cell) Cancer
Leukemia
Lung Cancer
Melanoma
Non-Hodgkin Lymphoma
Pancreatic Cancer
Prostate Cancer
Thyroid Cancer

All Cancer Types
Cancers by Body Location/System
Childhood Cancers
Adolescents and Young Adults
Women's Cancers

Clinical Trials

Search for Clinical Trials
What Is a Clinical Trial?
Clinical Trial Results
Learn About Clinical Trials

Research & Funding

Funding Opportunities
Funding Patterns for R01 and R21 Grants in FY2013
Training Opportunities

Cancer Topics

What Is Cancer?
Treatment
Prevention, Genetics, Causes
Screening and Testing
Coping with Cancer
Smoking
Cancer Health Disparities
NCI Fact Sheets
Physician Data Query (PDQ®)

Cancer Research News

Cold Spring Harbor researchers find immunogenic mutations in tumor genomes correlate with increased patient survival

Northwestern U researchers find that vitamin D deficiency could be linked to aggressive prostate cancer

Hopkins researchers find that brain tumor cells penetrated by tiny, degradable particles can carry genetic instructions

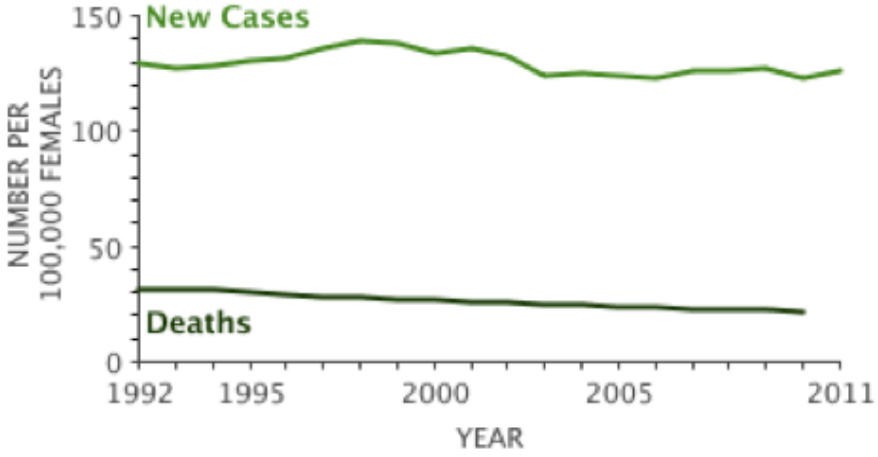
Visit the NCI News Center

Statistics at a Glance

At a Glance

Estimated New Cases in 2014	232,670
% of All New Cancer Cases	14.0%

Estimated Deaths in 2014	40,000
% of All Cancer Deaths	6.8%



Percent Surviving 5 Years
89.2%
2004-2010

Number of New Cases and Deaths per 100,000: The number of new cases of breast cancer was 124.5 per 100,000 women per year. The number of deaths was 22.6 per 100,000 women per year. These rates are age-adjusted and based on 2007-2011 cases and 2006-2010 deaths.

Lifetime Risk of Developing Cancer: Approximately 12.3 percent of women will be diagnosed with breast cancer at some point during their lifetime, based on 2008-2010 data.

Prevalence of this cancer: In 2011, there were an estimated 2,899,726 women living with breast cancer in the United States.

Many dictionaries

- NCI dictionary of Cancer Terms
- NCI Drug dictionary
- Glossary of statistical terms
- Glossary of Genetic terms

ClinicalTrials.gov

ClinicalTrials.gov
A service of the U.S. National Institutes of Health

ClinicalTrials.gov is a registry and results database of publicly and privately supported clinical studies of human participants conducted around the world. Learn more about clinical studies and about this site, including relevant history, policies, and laws.

Find Studies | About Clinical Studies | Submit Studies | Resources | About This Site

ClinicalTrials.gov currently lists 188,780 studies with locations in all 50 states and in 190 countries.

Text Size

Search for Studies
Example: "Heart attack" AND "Los Angeles"
HER2 mutations breast cancer [Search]
Advanced Search | See Studies by Topic
See Studies on a Map

Search Help
- How to search
- How to find results of studies
- How to read a study record

Locations of Recruiting Studies

- Non-U.S. Only (53%)
- U.S. Only (41%)
- Both U.S. and Non-U.S. (6%)

Total N = 35,128 studies
Data as of April 21, 2015

- See more trends, charts, and maps

For Patients & Families
- How to find studies
- See studies by topic
- Learn about clinical studies
- Learn more…

For Researchers
- How to submit studies
- Download content for analysis
- About the results database
- Learn more…

For Study Record Managers
- Why register?
- How to register study records
- FDAAA 801 Requirements
- Learn more…

Learn More
- ClinicalTrials.gov Online Training
- Glossary of common site terms

For the Press
Using our RSS Feeds

Neratinib in Metastatic HER2 Non-amplified But HER2 Mutant Breast Cancer

This study is currently recruiting participants. (see Contacts and Locations)

Verified April 2015 by Washington University School of Medicine

Sponsor:
Washington University School of Medicine

Information provided by (Responsible Party):
Washington University School of Medicine

ClinicalTrials.gov Identifier:
NCT01670877

First received: August 17, 2012
Last updated: April 7, 2015
Last verified: April 2015
History of Changes

Purpose

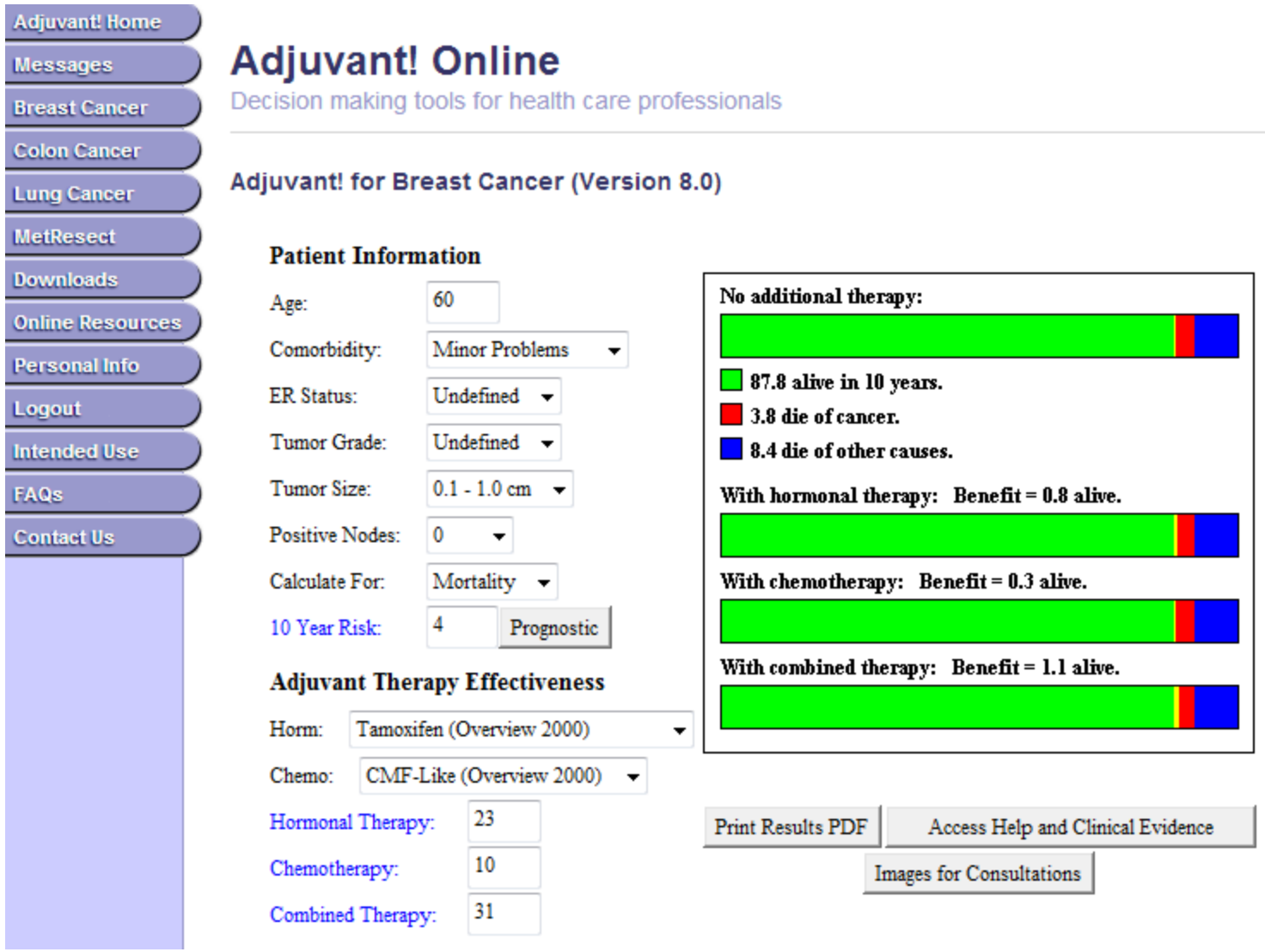
This phase II study will test **cancer** to see if it has a **HER2 mutation** and, if so, see how **HER2 mutated cancer** responds to treatment with neratinib.

Condition	Intervention	Phase
Breast Neoplasms	Drug: Neratinib	Phase 2

Study Type: Interventional
Study Design: Endpoint Classification: Efficacy Study
Intervention Model: Single Group Assignment
Masking: Open Label
Primary Purpose: Treatment

Official Title: A Phase II Study of Neratinib in Metastatic **HER2** Non-amplified But **HER2** Mutant **Breast Cancer**

Adjuvant! Online



Note of caution!

Validity of Adjuvant! Online program in older patients with breast cancer: a population-based study

Nienke A de Glas, Willemien van de Water, Ellen G Engelhardt, Esther Bastiaannet, Anton J M de Craen, Judith R Kroep, Hein Putter, Anne M Stiggelbout, Nir I Weijl, Cornelis J H van de Velde, Johanneke E A Portielje, Gerrit-Jan Liefers

Lancet Oncol 2014; 15: 722–29
Published Online May 14, 2014 http://dx.doi.org/10.1016/S1470-2045(14)70200-1
See Comment page 672

Department of Surgery (N A de Glas MD, W van de Water MD, E Bastiaannet PhD, Prof C J H van de Velde PhD, G-J Liefers PhD), Department of Gerontology and Geriatrics (N A de Glas, W van de Water, E Bastiaannet, A J M de Craen PhD), Department of Medical Decision Making (E G Engelhardt MSc, Prof A M Stiggelbout PhD), Department of Medical Oncology (J R Kroep PhD), Department of Medical Statistics (Prof Hein Putter PhD), Leiden University Medical Centre, Leiden, Netherlands; Department of Medical Oncology, Bronovo Hospital The Hague, The Hague, Netherlands (N I Weijl PhD); and Department of Medical

Summary

Background Adjuvant! Online is a prediction tool that can be used to aid clinical decision making in patients with breast cancer. It was developed in a patient population aged 69 years or younger, and subsequent validation studies included small numbers of older patients. Since older patients with breast cancer differ from younger patients in many aspects, the aim of this study was to investigate the validity of Adjuvant! Online in a large cohort of unselected older patients.

Methods We included patients from the population-based FOCUS cohort, which included all consecutive patients aged 65 years or older who were diagnosed with invasive or in-situ breast cancer between Jan 1, 1997, and Dec 31, 2004, in the southwestern part of the Netherlands. We included all patients who fulfilled the criteria as stated by Adjuvant! Online: patients with unilateral, unicentric, invasive adenocarcinoma; no evidence of metastatic or residual disease; no evidence of T4 features; and no evidence of inflammatory breast cancer. We entered data from all patients with the "average for age" comorbidity status (model 1) and with an individualised comorbidity status (model 2).

Findings We included 2012 patients. Median age of patients in the cohort was 74·0 years (IQR 69·0–79·0). 904 (45%) of 2012 patients died during follow-up, whereas 326 (16%) patients had recurrence. Median follow-up for overall survival was 9·0 years (IQR 7·4–10·7), and 6·6 years (4·4–6·6) for patients without recurrence. Using model 1, Adjuvant! Online overestimated 10-year overall survival by 9·8% ([95% CI 5·9–13·7], p<0·0001) and 10-year cumulative recurrence survival by 8·7% ([6·7–10·7], p<0·0001). By contrast, when using model 2, Adjuvant! Online underestimated the 10-year overall survival by –17·1% ([95% CI –21·0 to –13·2], p<0·0001). However, when using model 2, Adjuvant! Online predicted cumulative recurrence accurately in all patients (–0·7% [95% CI –2·7–1·3], p=0·48).

Interpretation Adjuvant! Online does not accurately predict overall survival and recurrence in older patients with early breast cancer.

Funding Dutch Cancer Foundation.

European Journal of Cancer (2012) **48**, 982–989

ELSEVIER

Available at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.ejconline.com

EJC

Adjuvant! Online is overoptimistic in predicting survival of Asian breast cancer patients

Nirmala Bhoo-Pathy [a,b], Cheng-Har Yip [c], Mikael Hartman [d,e], Nakul Saxena [d], Nur Aishah Taib [c], Gwo-Fuang Ho [f], Lai-Meng Looi [g], Awang M. Bulgiba [h], Yolanda van der Graaf [b], Helena M. Verkooijen [d,i,*]

[a] *National Clinical Research Centre, Level 3, Dermatology Block, Kuala Lumpur Hospital, 50586 Jalan Pahang, Kuala Lumpur, Malaysia*
[b] *Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, P.O. Box 85500, 3508 GA Utrecht, The Netherlands*
[c] *Department of Surgery, Faculty of Medicine, University of Malaya, 50603 Lembah Pantai, Kuala Lumpur, Malaysia*
[d] *Saw Swee Hock School of Public Health, National University of Singapore, Block MD3, 16 Medical Drive, Singapore 117597, Singapore*
[e] *Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Block MD3, 16 Medical Drive, Singapore 117597, Singapore*
[f] *Department of Oncology, Faculty of Medicine, University of Malaya, 50603 Lembah Pantai, Kuala Lumpur, Malaysia*
[g] *Department of Pathology, Faculty of Medicine, University of Malaya, 50603 Lembah Pantai, Kuala Lumpur, Malaysia*
[h] *Julius Centre University of Malaya, Faculty of Medicine, University of Malaya, 50603 Lembah Pantai, Kuala Lumpur, Malaysia*
[i] *Imaging Division, University Medical Center Utrecht, P.O. Box 85500, 3508 GA Utrecht, The Netherlands*

2. Research databases

Research databases

- Gene Expression Omnibus: www.ncbi.nlm.nih.gov/geo/
- Array Express: www.ebi.ac.uk/arrayexpress/
- EGA (European Genome-phenome Archive)

- METABRIC
- TCGA
- ICGC
- The Cancer Cell Line Encyclopedia

METABRIC

(Molecular Taxonomy of Breast Cancer International Consortium)

ARTICLE

doi:10.1038/nature10983

The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups

Christina Curtis[1,2]†*, Sohrab P. Shah[3,4]*, Suet-Feung Chin[1,2]*, Gulisa Turashvili[3,4]*, Oscar M. Rueda[1,2], Mark J. Dunning[2], Doug Speed[2,5]†, Andy G. Lynch[1,2], Shamith Samarajiwa[1,2], Yinyin Yuan[1,2], Stefan Gräf[1,2], Gavin Ha[3], Gholamreza Haffari[3], Ali Bashashati[3], Roslin Russell[2], Steven McKinney[3,4], METABRIC Group‡, Anita Langerød[6], Andrew Green[7], Elena Provenzano[8], Gordon Wishart[8], Sarah Pinder[9], Peter Watson[3,4,10], Florian Markowetz[1,2], Leigh Murphy[10], Ian Ellis[7], Arnie Purushotham[9,11], Anne-Lise Børresen-Dale[6,12], James D. Brenton[2,13], Simon Tavaré[1,2,5,14], Carlos Caldas[1,2,8,13] & Samuel Aparicio[3,4]

The elucidation of breast cancer subgroups and their molecular drivers requires integrated views of the genome and transcriptome from representative numbers of patients. We present an integrated analysis of copy number and gene expression in a discovery and validation set of 997 and 995 primary breast tumours, respectively, with long-term clinical follow-up. Inherited variants (copy number variants and single nucleotide polymorphisms) and acquired somatic copy number aberrations (CNAs) were associated with expression in ~40% of genes, with the landscape dominated by ***cis***- and ***trans***-acting CNAs. By delineating expression outlier genes driven in ***cis*** by CNAs, we identified putative cancer genes, including deletions in ***PPP2R2A***, ***MTAP*** and ***MAP2K4***. Unsupervised analysis of paired DNA–RNA profiles revealed novel subgroups with distinct clinical outcomes, which reproduced in the validation cohort. These include a high-risk, oestrogen-receptor-positive 11q13/14 ***cis***-acting subgroup and a favourable prognosis subgroup devoid of CNAs. ***Trans***-acting aberration hotspots were found to modulate subgroup-specific gene networks, including a TCR deletion-mediated adaptive immune response in the 'CNA-devoid' subgroup and a basal-specific chromosome 5 deletion-associated mitotic network. Our results provide a novel molecular stratification of the breast cancer population, derived from the impact of somatic CNAs on the transcriptome.

Inherited genetic variation and acquired genomic aberrations contribute to breast cancer initiation and progression. Although somatically acquired CNAs are the dominant feature of sporadic breast cancers, the driver events that are selected for during tumorigenesis are difficult to elucidate as they co-occur alongside a much larger landscape of random non-pathogenic passenger alterations[1,2] and germline copy number variants (CNVs). Attempts to define subtypes of breast cancer and to discern possible somatic drivers are still in their relative infancy[3–6], in part because breast cancer represents multiple diseases, implying that large numbers (many hundreds or thousands) of patients must be studied. Here we describe an integrated genomic/transcriptomic analysis of breast cancers with long-term clinical outcomes composed of a discovery set of 997 primary tumours and a validation set of 995 tumours from METABRIC (Molecular Taxonomy of Breast Cancer International Consortium).

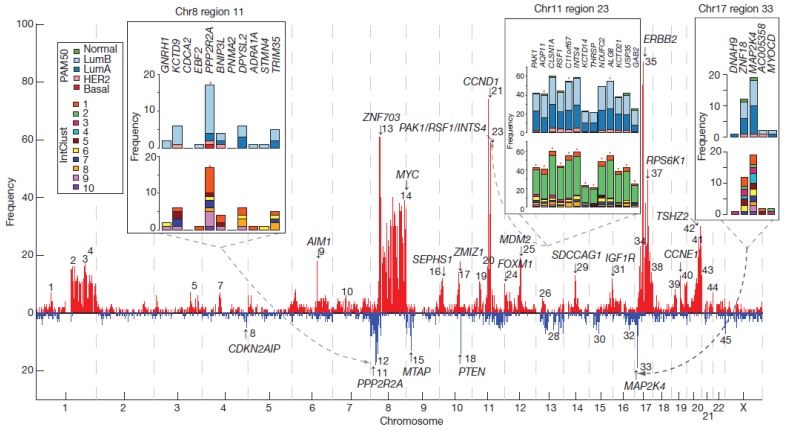
A breast cancer population genomic resource

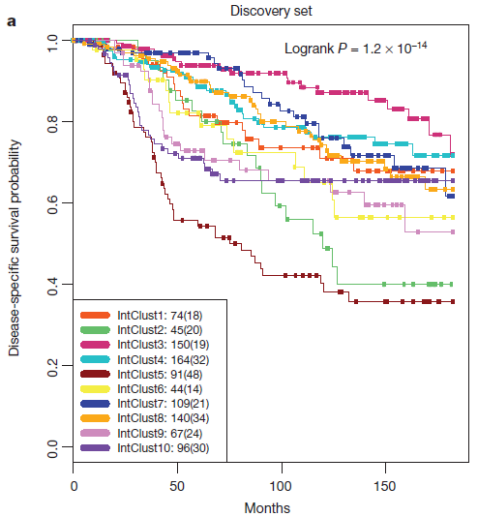
We assembled a collection of over 2,000 clinically annotated primary fresh-frozen breast cancer specimens from tumour banks in the UK and Canada (Supplementary Tables 1–3). Nearly all oestrogen receptor (ER)-positive and/or lymph node (LN)-negative patients did not receive chemotherapy, whereas ER-negative and LN-positive patients did. Additionally, none of the $HER2^+$ patients received trastuzumab. As such, the treatments were homogeneous with respect to clinically relevant groupings. An initial set of 997 tumours was analysed as a discovery group and a further set of 995 tumours, for which complete data later became available, was used to test the reproducibility of the integrative clusters (described below). An overview of the main analytical approaches is provided in Supplementary Fig. 1. Details concerning expression and copy number profiling, including sample assignment to the PAM50 intrinsic subtypes[3,4,7] (Supplementary Fig. 2), copy number analysis (Supplementary Tables 4–8) and validation (Supplementary Figs 3 and 4 and Supplementary Tables 9–11), and *TP53* mutational profiling (Supplementary Fig. 5) are described in the Supplementary Information.

Genome variation affects tumour expression architecture

Genomic variants are considered to act in *cis* when a variant at a locus has an impact on its own expression, or in *trans* when it is associated

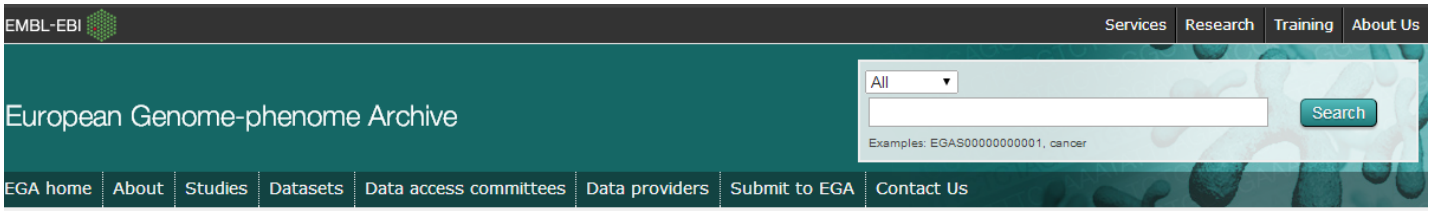
[1]Department of Oncology, University of Cambridge, Hills Road, Cambridge CB2 2XZ, UK. [2]Cancer Research UK, Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK. [3]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia V6T 2B5, Canada. [4]Molecular Oncology, British Columbia Cancer Research Centre, Vancouver, British Columbia V5Z 1L3, Canada. [5]Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Cambridge CB3 0WA, UK. [6]Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Montebello, 0310 Oslo, Norway. [7]Department of Histopathology, School of Molecular Medical Sciences, University of Nottingham, Nottingham NG5 1PB, UK. [8]Cambridge Breast Unit, Addenbrooke's Hospital, Cambridge University Hospital NHS Foundation Trust and NIHR Cambridge Biomedical Research Centre, Cambridge CB2 2QQ, UK. [9]King's College London, Breakthrough Breast Cancer Research Unit, London WC2R 2LS, UK. [10]Manitoba Institute of Cell Biology, University of Manitoba, Manitoba R3E 0V9, Canada. [11]NIHR Comprehensive Biomedical Research Centre at Guy's and St Thomas' NHS Foundation Trust and King's College London, London WC2R 2LS, UK. [12]Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, 0316 Oslo, Norway. [13]Cambridge Experimental Cancer Medicine Centre, Cambridge CB2 0RE, UK. [14]Molecular and Computational Biology Program, University of Southern California, Los Angeles, California 90089, USA. †Present addresses: Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California 90033, USA (Ch.C.); University College London, Genetics Institute, WC1E 6BT, UK (D.S.).
*These authors contributed equally to this work.
‡Lists of participants and affiliations appear at the end of the paper.





METABRIC

(www.ebi.ac.uk/ega/studies/EGAS00000000083)



STUDY: METABRIC

Study Description

The elucidation of breast cancer subgroups and their molecular drivers requires integrated views of the genome and transcriptome from... Show More

Study ID	Alternative Stable ID	type
EGAS00000000083		Genotype/Expression

Data provider(s)

- University of Cambridge & Cancer Research UK Cambridge Research Institute
- The Cancer Research UK Cambridge Research Institute

Who archives the data?

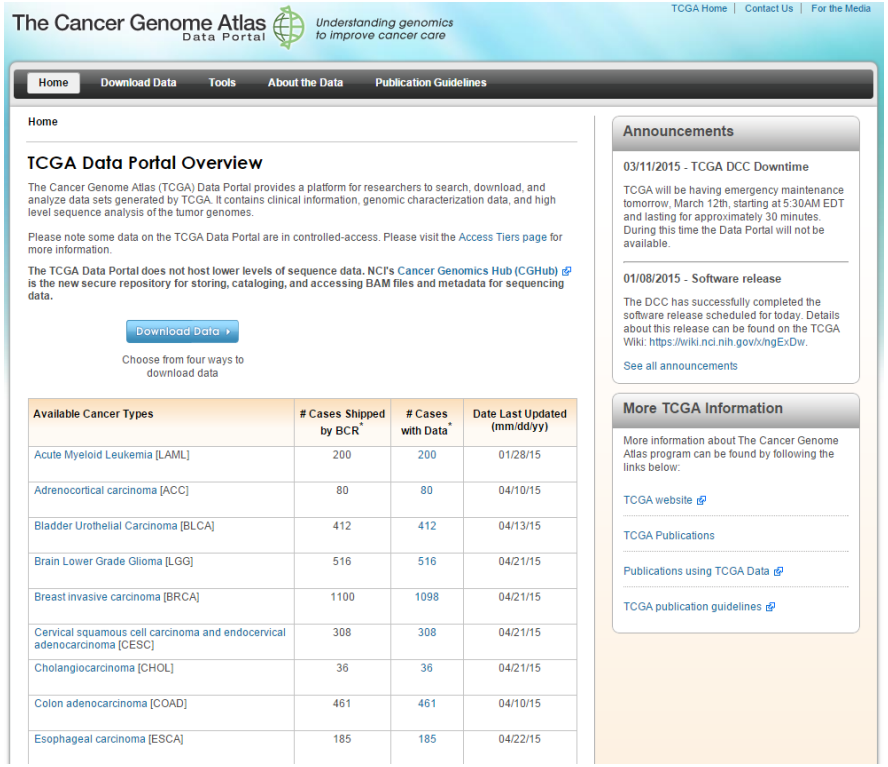
This study includes 10 datasets:

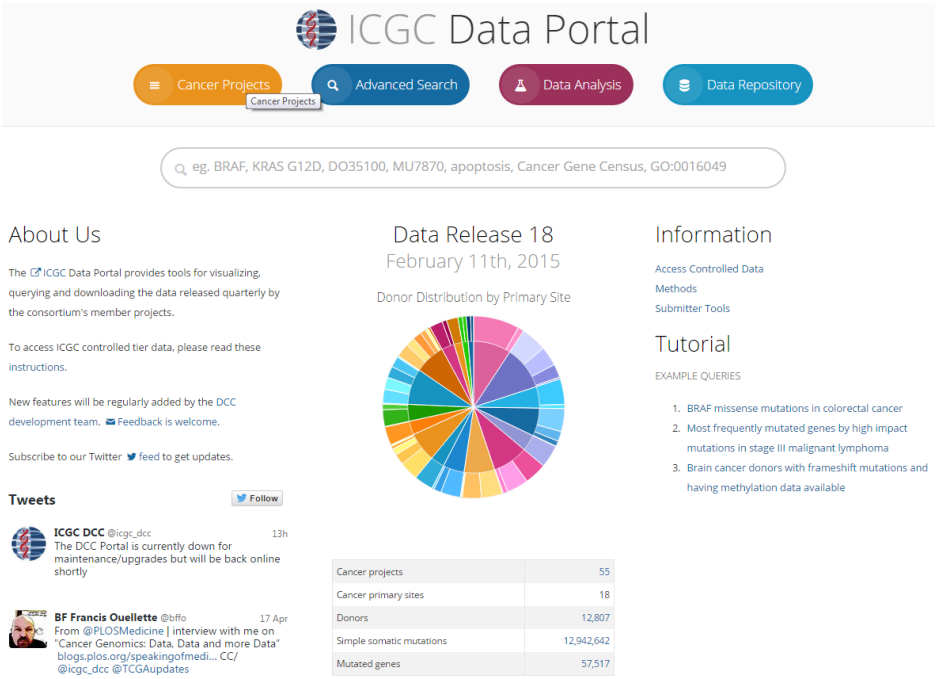
Click on a Dataset ID in the table below to learn more, and to find out who to contact about access to these data

Dataset ID	Technology	Type	Samples	Description
EGAD00010000162	Illumina HT 12		2136	Illumina HT 12 IDATS
EGAD00010000164	Affymetrix SNP 6.0		1992	Affymetrix 6.0 CEL files
EGAD00010000210	Illumina HT 12		997	Normalized expression data; discovery set
EGAD00010000211	Illumina HT 12		995	Normalized expression data; validation set
EGAD00010000212	Illumina HT 12		144	Normalized expression data; normals
EGAD00010000213	Affymetrix SNP 6.0		997	Segmented (CBS) copy number aberrations (CNA); discovery set
EGAD00010000214	Affymetrix SNP 6.0		997	Segmented (CBS) copy number variants (CNV); discovery set
EGAD00010000215	Affymetrix SNP 6.0		995	Segmented (CBS) copy number aberrations (CNA); validation set

TCGA/ICGC

(The Cancer Genome Atlas/ International Cancer Genomics Consortium)





N=11,079

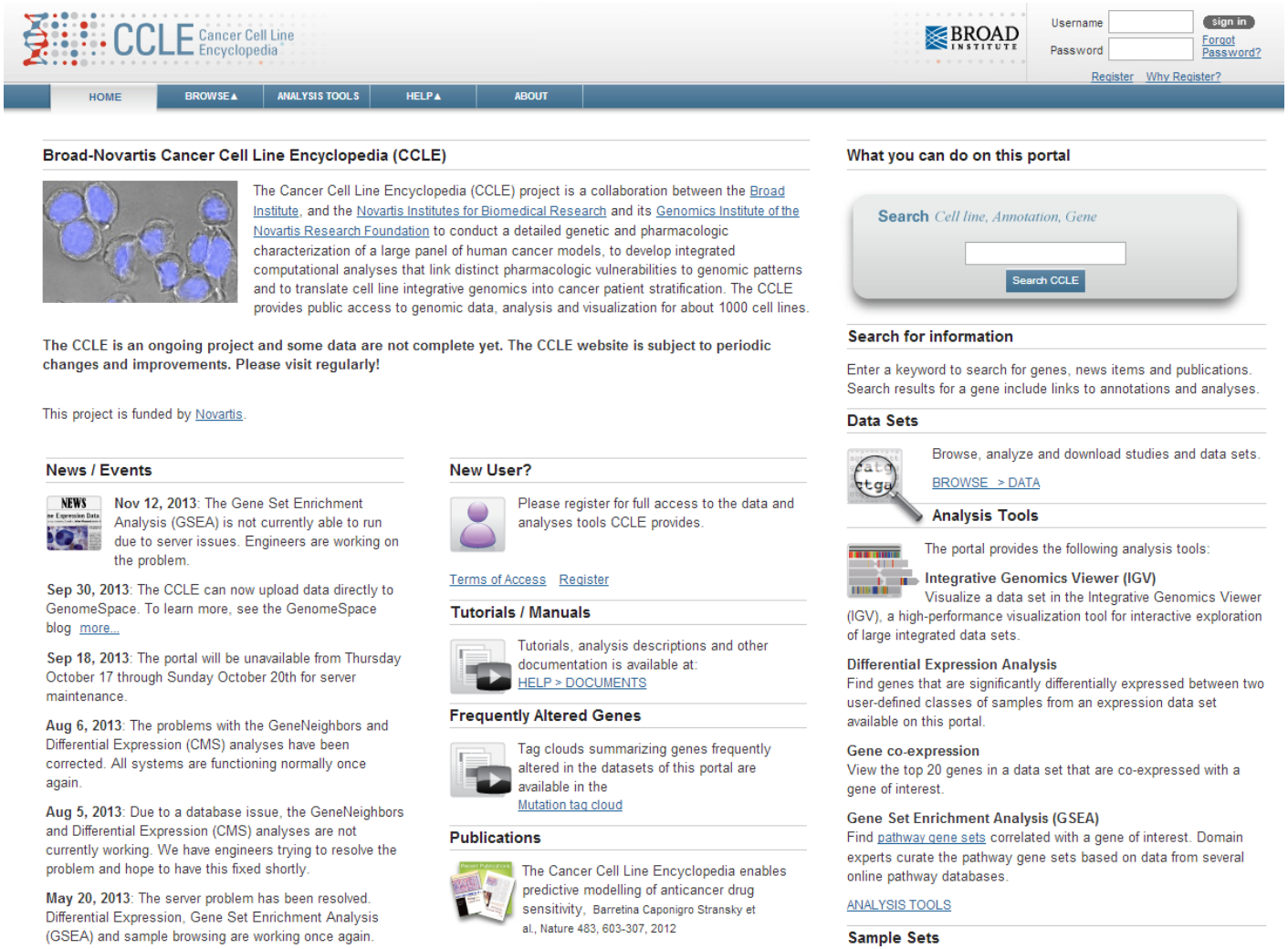
N BC= 1,098

N=12,807

N BC= 1,186

The Cancer Cell Line Encyclopedia

(www.broadinstitute.org/ccle/home)



3. Research Tools

Research tools

- R: a door to many research avenues…!

(www.bioconductor.org)

- GSEA (www.broadinstitute.org/gsea/)
- MSigDB (http://www.broadinstitute.org/gsea/msigdb/index.jsp)
- DAVID (http://david.abcc.ncifcrf.gov/)
- Oncomine (www.oncomine.org)

GSEA & MSigDB

- **Gene Set Enrichment Analysis** (GSEA) is a computational method that determines whether an a priori defined set of genes shows statistically significant, concordant differences between two biological states (e.g. phenotypes).
- The **Molecular Signatures Database** (MSigDB) is a collection of annotated gene sets for use with GSEA software.

Annotation of mutations

- COSMIC (http://cancer.sanger.ac.uk/cancergenome/projects/cosmic)
- Oncotator (www.broadinstitute.org/oncotator)
- Annovar (http://www.openbioinformatics.org/annovar)

- DGIdb (dgidb.genome.wustl.edu)

- …!

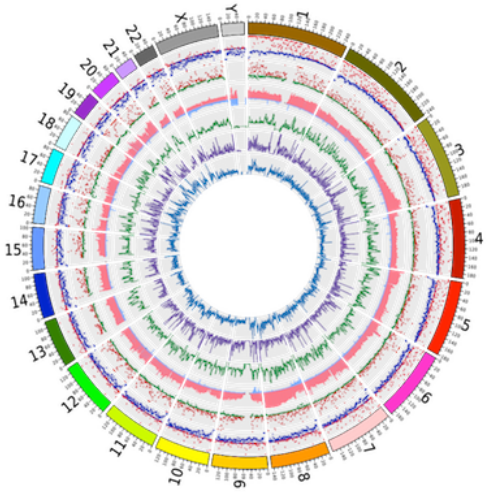
COSMIC

COSMIC
Catalogue of somatic mutations in cancer

Home ▼ | About ▼ | Licensing ▼ | Data Download ▼ | News ▼ | Help ▼ | Enter search here... | Login ▼

COSMIC v72

eg: Braf, COLO-829, Carcinoma, V600E, BRCA-UK, Campbell SEARCH



Genomic Landscape of Cancer

R Resources

Key COSMIC resources

Cell Lines Project
COSMIC Whole Genomes
Cancer Gene Census
Drug Sensitivity
Mutational Signatures New
GRCh37 Cancer Archive New

T Tools

Additional tools to explore COSMIC

Cancer Browser
Genome Browser
CONAN
COSMIC Mart

C Expert Curation

High quality curation by expert postdoctoral scientists

Cancer Gene Census
Curated Genes
Gene Fusions
Genome-Wide Screens

D Data

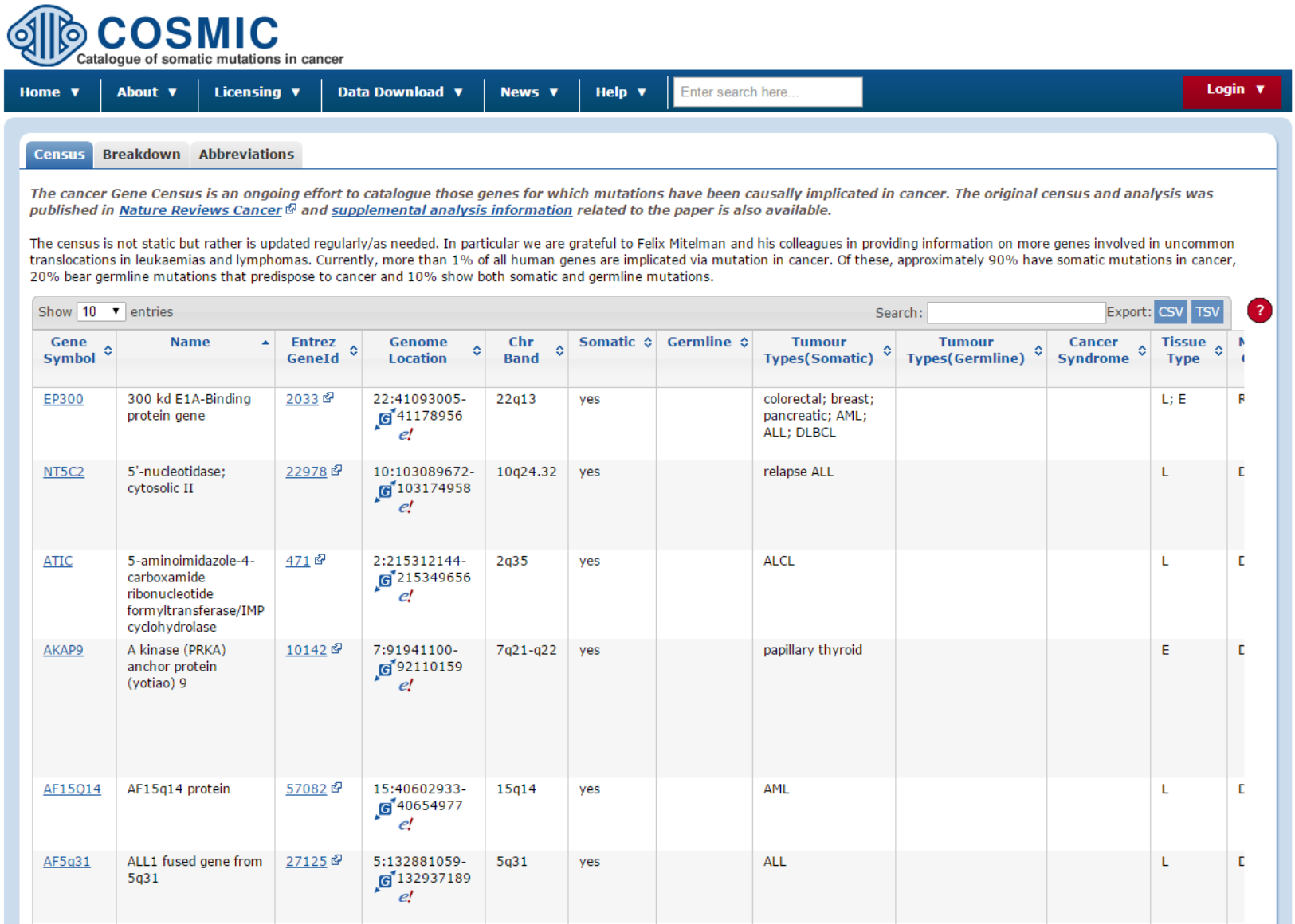
Further details on using COSMIC's content

Downloads
License
Submission
Genome Annotations
Datasheet V72
Help
FAQ

Statistics

Domain	Counts
Samples	1103964
Coding Mutations	3158657
Papers	21086
Fusions	10890
Genomic Rearrangements	61232
Whole Genomes	19672
Copy Number	842651
Gene Expression	8228797

Cancer Gene Census

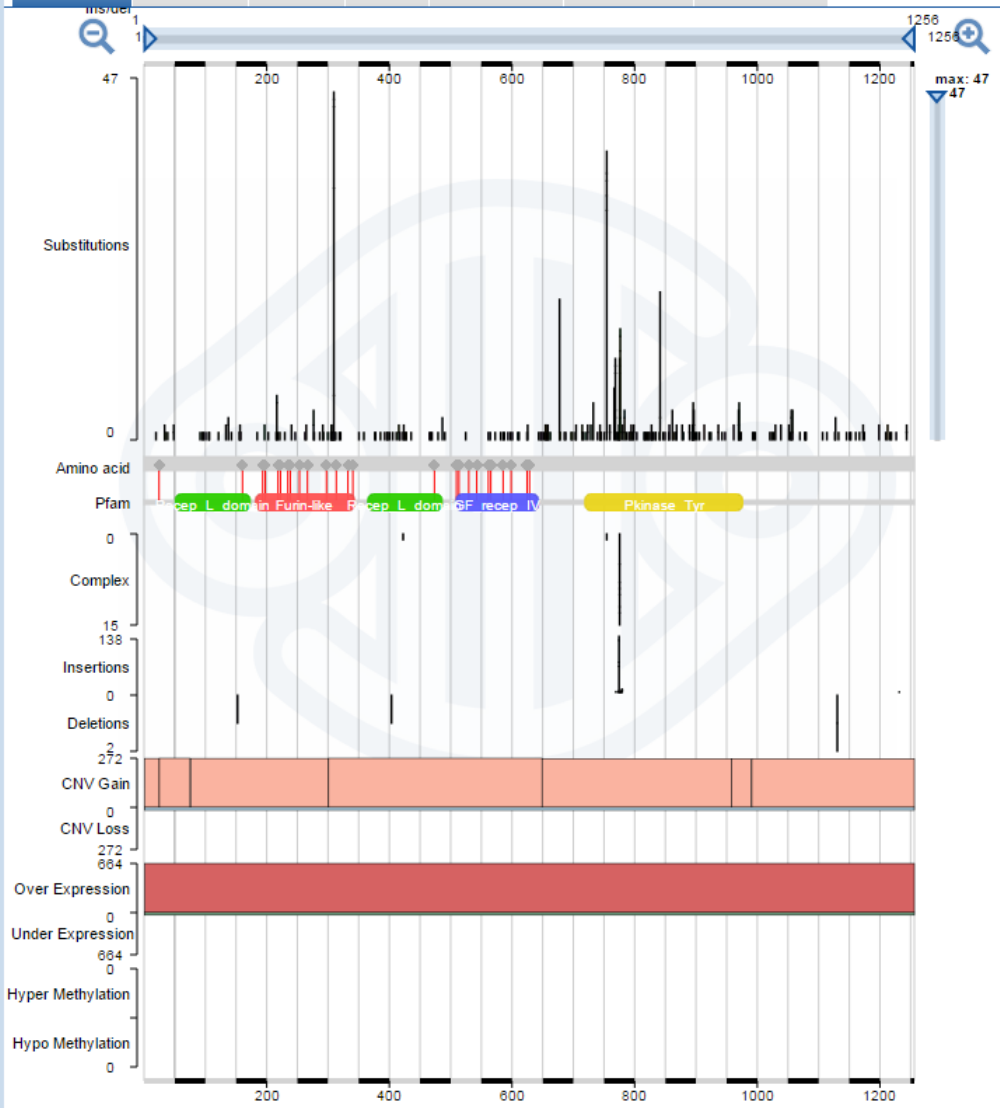


Gene Symbol	Name	Entrez GeneId	Genome Location	Chr Band	Somatic	Germline	Tumour Types(Somatic)	Tumour Types(Germline)	Cancer Syndrome	Tissue Type
EP300	300 kd E1A-Binding protein gene	2033	22:41093005-41178956	22q13	yes		colorectal; breast; pancreatic; AML; ALL; DLBCL			L; E
NT5C2	5'-nucleotidase; cytosolic II	22978	10:103089672-103174958	10q24.32	yes		relapse ALL			L
ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	471	2:215312144-215349656	2q35	yes		ALCL			L
AKAP9	A kinase (PRKA) anchor protein (yotiao) 9	10142	7:91941100-92110159	7q21-q22	yes		papillary thyroid			E
AF15Q14	AF15q14 protein	57082	15:40602933-40654977	15q14	yes		AML			L
AF5q31	ALL1 fused gene from 5q31	27125	5:132881059-132937189	5q31	yes		ALL			L

Cosmic » Gene » Analysis » ERBB2

View in GRCh37 Archive

Histogram | Mutations | Fusions | Tissue | Distribution | CNV & Expr | Methylation



Filters

Gene: ERBB2

Position

Start: 1

End: 1256

Sequence Type:

cDNA

Amino Acid

Cancer type — Select

- Systematic Screen
- Somatic Status
- Tumour Source
- Mutation Type
- Copy Number Variation
- Gene Expression
- Methylation

Apply | Reset

cBioPortal

cBioPortal for Cancer Genomics

Visualize, analyze, discover.

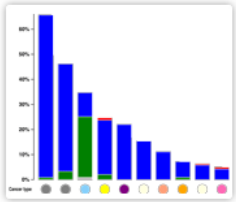
Memorial Sloan Kettering Cancer Center

HOME | DATA SETS | WEB API | R/MATLAB | TUTORIALS | FAQ | NEWS | TOOLS | ABOUT | VISUALIZE YOUR DATA

The cBioPortal for Cancer Genomics provides **visualization**, **analysis** and **download** of large-scale **cancer genomics** data sets.

Please adhere to the TCGA publication guidelines when using TCGA data in your publications.

Please cite Gao et al. *Sci. Signal.* 2013 & Cerami et al. *Cancer Discov.* 2012 when publishing results based on cBioPortal.



Query | **Download Data**

Select Cancer Study: All Cancer Studies

Select Data Type Priority: Mutation and CNA | Only Mutation | Only CNA

Enter Gene Set: Advanced: Onco Query Language (OQL)

User-defined List

Enter HUGO Gene Symbols or Gene Aliases

Submit

What's New

- New data and features released
- New tools released

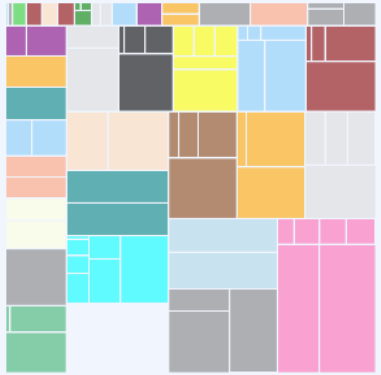
Sign up for low-volume email news alerts:

Subscribe

Or follow us @cbioportal on Twitter

Data Sets

The Portal contains data for **20958 tumor samples from 89 cancer studies.** [Details]



Example Queries

RAS/RAF alterations in colorectal cancer

BRCA1 and BRCA2 mutations in ovarian cancer

cBioPortal for Cancer Genomics

Visualize, analyze, discover.

Memorial Sloan Kettering Cancer Center

HOME | DATA SETS | WEB API | R/MATLAB | TUTORIALS | FAQ | NEWS | TOOLS | ABOUT | VISUALIZE YOUR DATA

Modify Query

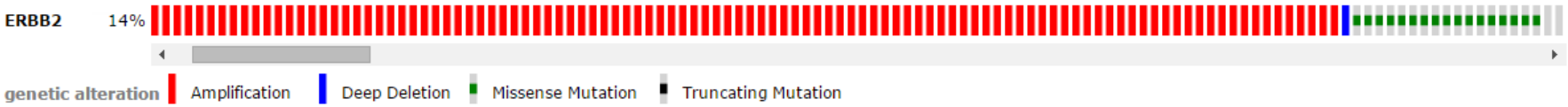
Breast Invasive Carcinoma (TCGA, Provisional)
Tumors with sequencing and CNA data (**962** samples) / **1** Gene

Gene Set / Pathway is altered in 138 (14.3%) of queried samples

OncoPrint | Plots | Mutations | Co-Expression | Protein Changes | Survival | Network | IGV | Download | Bookmark

Case Set: Tumors with sequencing and CNA data: All tumor samples that have CNA and sequencing data (962 samples)

Altered in 138 (14%) of cases



cBioPortal for Cancer Genomics

Visualize, analyze, discover.

Memorial Sloan Kettering Cancer Center

HOME | DATA SETS | WEB API | R/MATLAB | TUTORIALS | FAQ | NEWS | TOOLS | ABOUT | VISUALIZE YOUR DATA

Modify Query

Breast Invasive Carcinoma (TCGA, Provisional)
Tumors with sequencing and CNA data (**962** samples) / **1** Gene

Gene Set / Pathway is altered in 138 (14.3%) of queried samples

OncoPrint | Plots | Mutations | Co-Expression | Protein Changes | Survival | Network | IGV | Download | Bookmark

One Gene | Custom

Plot Parameters

Gene

ERBB2

Plot Type

mRNA vs. Copy Number

Data Type

- **mRNA** -

mRNA expression (RNA Seq V2 RSEM

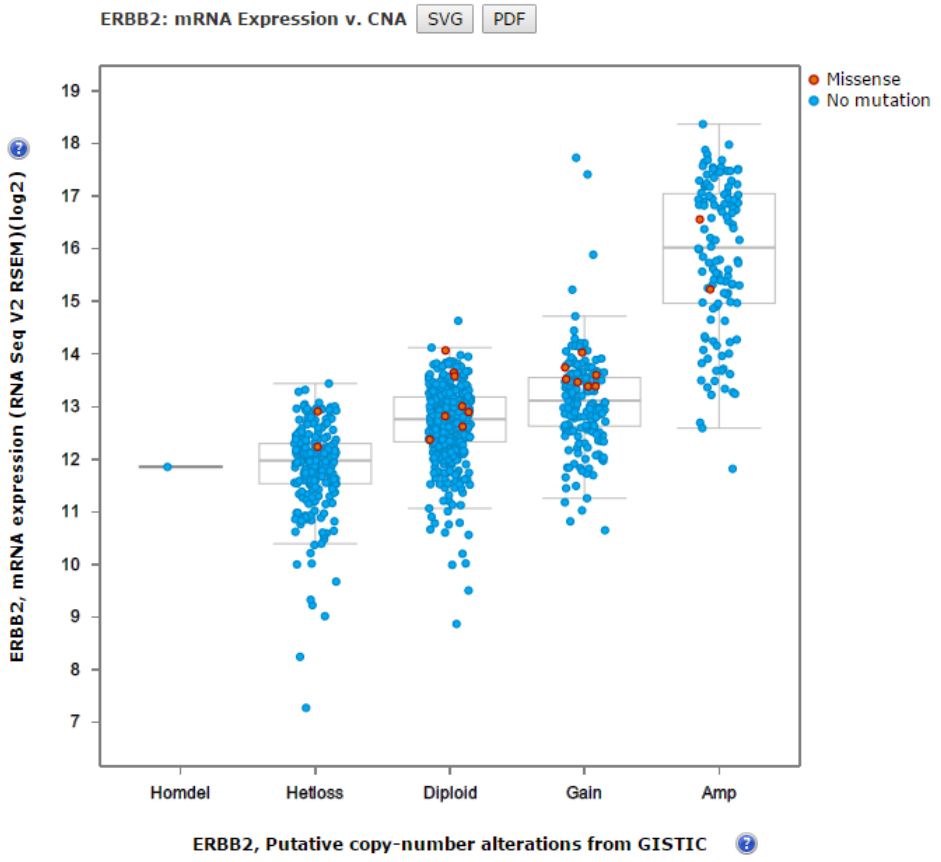
log scale

- **Copy Number** -

Putative copy-number alterations fror

Search case(s)

Case ID...



Modify Query

Breast Invasive Carcinoma (TCGA, Provisional)
Tumors with sequencing and CNA data (**962** samples) / **1** Gene

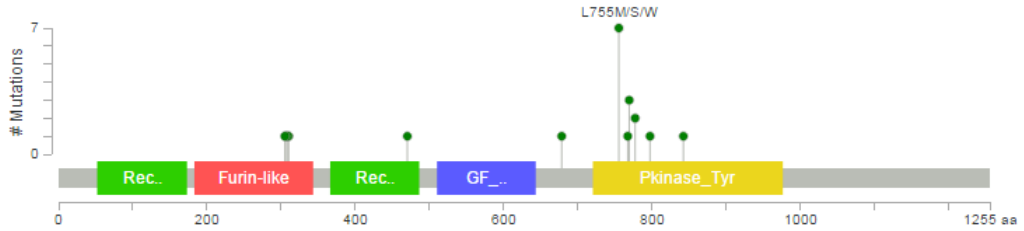
Gene Set / Pathway is altered in 138 (14.3%) of queried samples

OncoPrint | Plots | Mutations | Co-Expression | Protein Changes | Survival | Network | IGV | Download | Bookmark

ERBB2

ERBB2: [Somatic Mutation Rate: 2.0%]

ERBB2_HUMAN | PDF | SVG | Customize | Color Codes



3D Structure »

Show / hide columns | Showing 21 mutation(s) | Search:

Sample ID	AA change		Type	Copy #	COSMIC	Mutation Assessor	#Mut in Sample
TCGA-AO-A128-01	*V797A*	3D	Missense	Diploid	1	Medium	703
TCGA-A8-A08Z-01	*V842I*	3D	Missense	Diploid	12	Neutral	24
TCGA-AC-A3YI-01	*L755S*	3D	Missense	Diploid	35	Medium	11
TCGA-BH-A18P-01	*L755S*	3D	Missense	Gain	35	Medium	221
TCGA-A8-A0AB-01	*L755S*	3D	Missense	Gain	35	Medium	21
TCGA-A8-A0A6-01	*L755S*	3D	Missense	Gain	35	Medium	2052
TCGA-D8-A1XM-01	*L755S*	3D	Missense	Gain	35	Medium	42
TCGA-A2-A0T6-01	*L755W*	3D	Missense	Diploid	35	High	91
TCGA-C8-A135-01	*D769H*	3D	Missense	AMP	9	Neutral	37
TCGA-C8-A274-01	*I767M*	3D	Missense	Diploid	5	Neutral	136
TCGA-OL-A5D6-01	*V777L*	3D	Missense	Diploid	12	Neutral	30

OncoPrint | Plots | Mutations | Co-Expression | Protein Changes | Survival | Network | IGV | Download | Bookmark

Data Set: mRNA expression (RNA Seq V2 RSEM)

This table lists the genes with the highest expression correlation with the query genes. Click on a row to see the corresponding correlation plot.

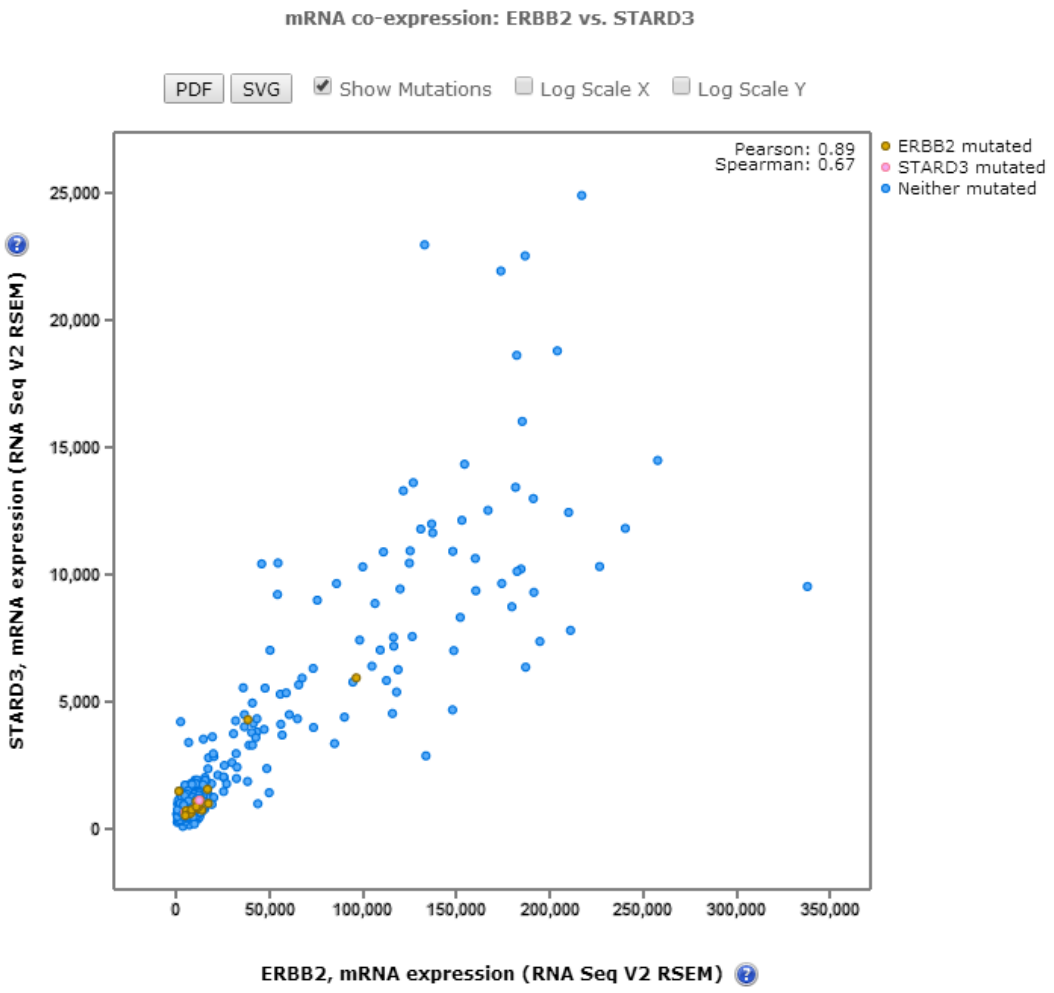
ERBB2

Search Gene

Show All

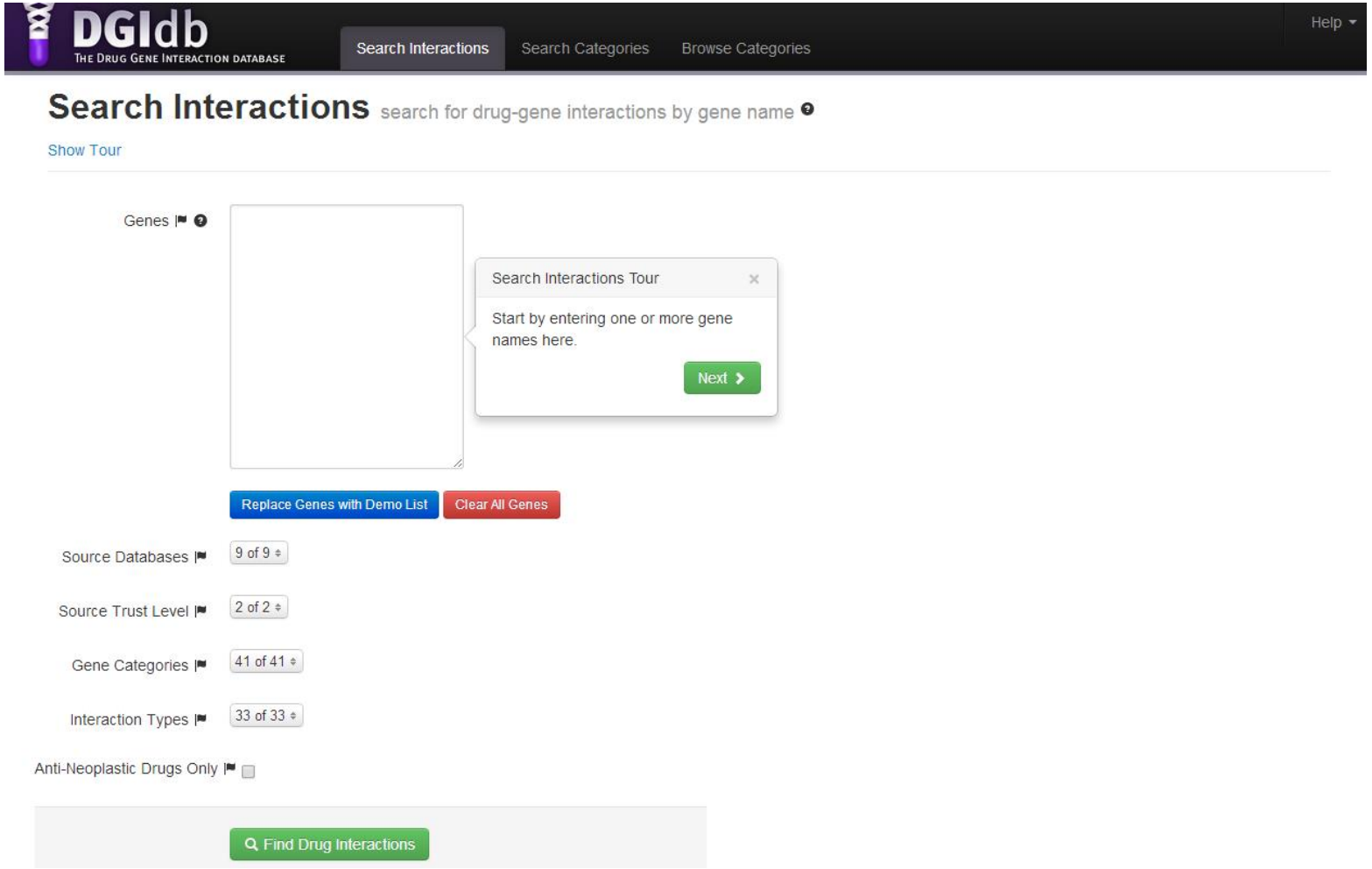
Correlated Gene	Pearson's Correlation	Spearman's Correlation
STARD3	0.89	0.67
GRB7	0.87	0.68
PGAP3	0.87	0.74
MIEN1	0.79	0.53
ORMDL3	0.78	0.67
PSMD3	0.71	0.46
CDK12	0.63	0.39
MED1	0.61	0.39
MED24	0.61	0.52
FBXL20	0.51	0.39
CASC3	0.50	0.43
MSL1	0.50	0.42
RPL19	0.45	0.32
WIPF2	0.45	0.43
PIP4K2B	0.43	0.46
CWC25	0.42	0.36
LASP1	0.40	0.48
TMEM86A	0.40	0.30
THRA	0.40	0.36
SRCIN1	0.40	0.48
MRPL45	0.39	0.38
PCGF2	0.38	0.46
MLLT6	0.36	0.35
CISD3	0.36	0.48
TNFAIP1	0.35	0.35
SEC16A	0.32	0.33
EPN3	0.32	0.41
RAB3D	0.31	0.32
ARHGAP8	0.30	0.37

1 to 29 of 29

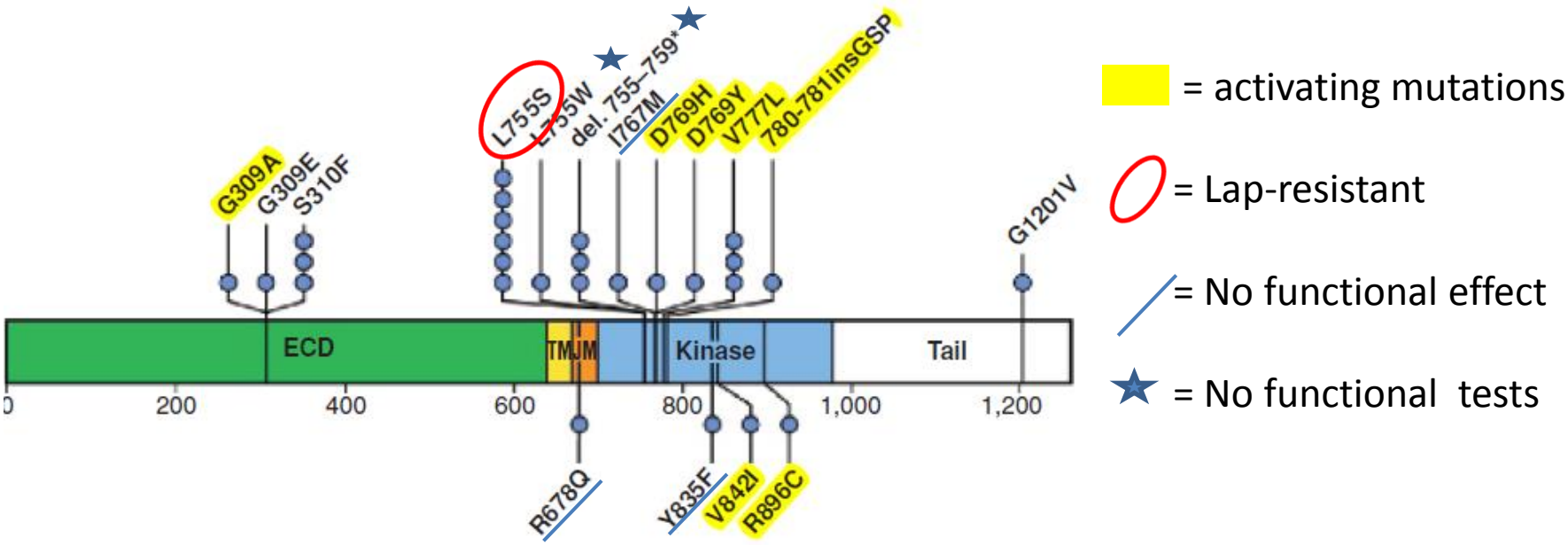


DGIdb

(http://dgidb.genome.wustl.edu/)



A note of caution however!



Adapted from Bose et al. Can Discovery 2012

ERBB2 mutations	Activity	Position	Mutation Assessor	Provean	SIFT	PolyPhen II
L755S	Lapatinib-resistance	chr17:g.37880220T>C	medium	Deleterious	Damaging	Probably damaging
D769H	activating	chr17:g.37880261G>C	neutral	Deleterious	Damaging	Probably damaging
V777L	activating	chr17:g.37881000G>T	neutral	Neutral	Tolerated	Benign
V842I	activating	chr17:g.37881332G>A	neutral	Neutral	Damaging	Probably damaging
G309A	activating	chr17:g.37868205G>C	low	Neutral	Damaging	Probably damaging
G309E	activating	chr17:g.37868205G>A	medium	Deleterious	Damaging	Probably damaging
R678Q	no functional effect	chr17:g.37879658G>A	low	Neutral	Tolerated	Possibly damaging
D769Y	activating	chr17:g.37880261G>T	low	Deleterious	Damaging	Probably damaging
R896C	activating	chr17:g.37881616C>T	low	Neutral	Tolerated	Benign
S310F	activating	chr17:g.37868208C>T	medium	Deleterious	Damaging	Probably damaging
I767M	no functional effect	chr17:g.37880257C>G	neutral	Neutral	Damaging	Probably damaging
Y835F	no functional effect	chr17:g.37881312A>T	low	Deleterious	Damaging	Probably damaging
P780-Y781insertionGSP	activating	chr17:g.37881011A>AGGGCTCCCC	not assessed*	Deleterious	not assessed*	not assessed*

Thank you for your attention!