

# Package ‘LungCancerACvsSCCGEO’

December 18, 2025

**Version** 1.46.0

**Date** 2013-7-13

**Title** A lung cancer dataset that can be used with maPredictDSC package for developing outcome prediction models from Affymetrix CEL files.

**Author** Adi Laurentiu Tarca <atarca@med.wayne.edu>

**Depends** R (>= 2.15.0)

**Maintainer** Adi Laurentiu Tarca <atarca@med.wayne.edu>

**Description** This package contains 30 Affymetrix CEL files for 7 Adenocarcinoma (AC) and 8 Squamous cell carcinoma (SCC) lung cancer samples taken at random from 3 GEO datasets (GSE10245, GSE18842 and GSE2109) and other 15 samples from a dataset produced by the organizers of the IMPROVER Diagnostic Signature Challenge available from GEO (GSE43580).

**License** GPL-2

**URL** <http://bioinformaticsprb.med.wayne.edu/>

**biocViews** CancerData, LungCancerData, MicroarrayData, GEO

**LazyLoad** yes

**git\_url** <https://git.bioconductor.org/packages/LungCancerACvsSCCGEO>

**git\_branch** RELEASE\_3\_22

**git\_last\_commit** 574147c

**git\_last\_commit\_date** 2025-10-29

**Repository** Bioconductor 3.22

**Date/Publication** 2025-12-18

## Contents

|                                |   |
|--------------------------------|---|
| LungCancerACvsSCCGEO . . . . . | 2 |
| Index                          | 3 |

---

|                      |                                                                                                                                                      |
|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| LungCancerACvsSCCGEO | <i>Annotation of a small set of training and test set samples (30 total) used by team 221 in the IMPROVER DSC for the lung cancer sub-challenge.</i> |
|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|

---

### Description

The LungCancerACvsSCCGEO dataset consists: i) a data frame `anoLC` giving the file names of the affy cel files used in the training phase and their corresponding phenotype (AC or SCC) and ii) `gsLC` the gold standard, i.e. the class membership of each test sample appearing in `anoLC`.

### Usage

```
data(LC)
```

### Source

GEO for the training data, while the test data comes from the citation below: Adi L. Tarca, Mario Lauria, Michael Unger, Erhan Bilal, Stephanie Boue, Kushal Kumar Dey, Julia Hoeng, Heinz Koeppl, Florian Martin, Pablo Meyer, Preetam Nandy, Raquel Norel, Manuel Peitsch, Jeremy J Rice, Roberto Romero, Gustavo Stolovitzky, Marja Talikka, Yang Xiang, Christoph Zechner, and IMPROVER DSC Collaborators, Strengths and limitations of microarray-based phenotype prediction: Lessons learned from the IMPROVER Diagnostic Signature Challenge. Bioinformatics, submitted 2013.

# Index

## \* **datasets**

LungCancerACvsSCCGEO, [2](#)

anoLC (LungCancerACvsSCCGEO), [2](#)

gsLC (LungCancerACvsSCCGEO), [2](#)

LungCancerACvsSCCGEO, [2](#)