# R2 Introduction Workshop PRESSURE Part 2

## 21-May-2024

**Computer login:**
**func_user**
**Welkom01**


Lieke Hoyng , on behalf of the R2 Team

Dept. CEMM | Location AMC

Amsterdam University Medical Centers (AUMC)

University of Amsterdam, the Netherlands

R2 Support: **r2-support@amsterdamumc.nl**

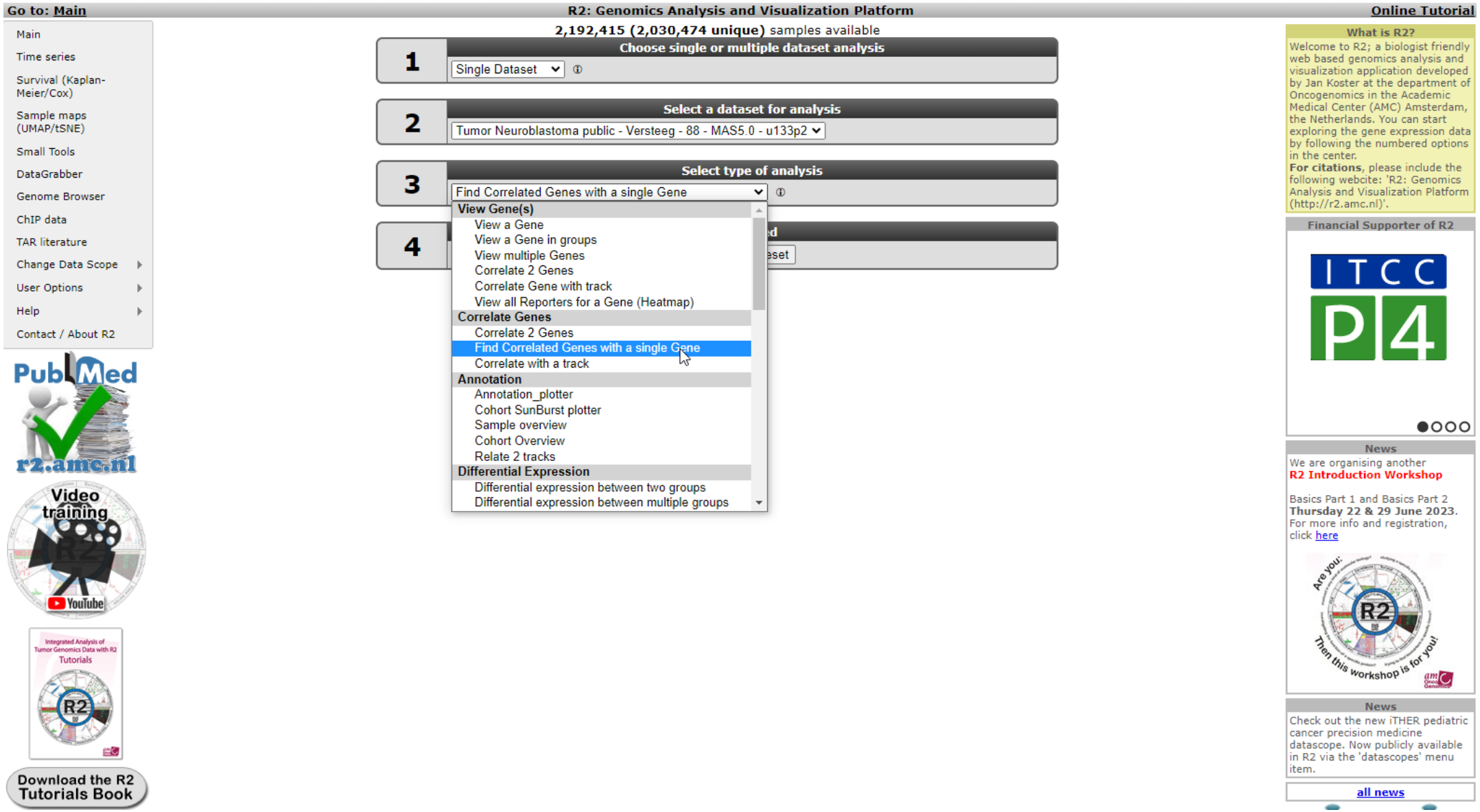
Jan Koster: **jankoster@amsterdamumc.nl**


r2platform.com

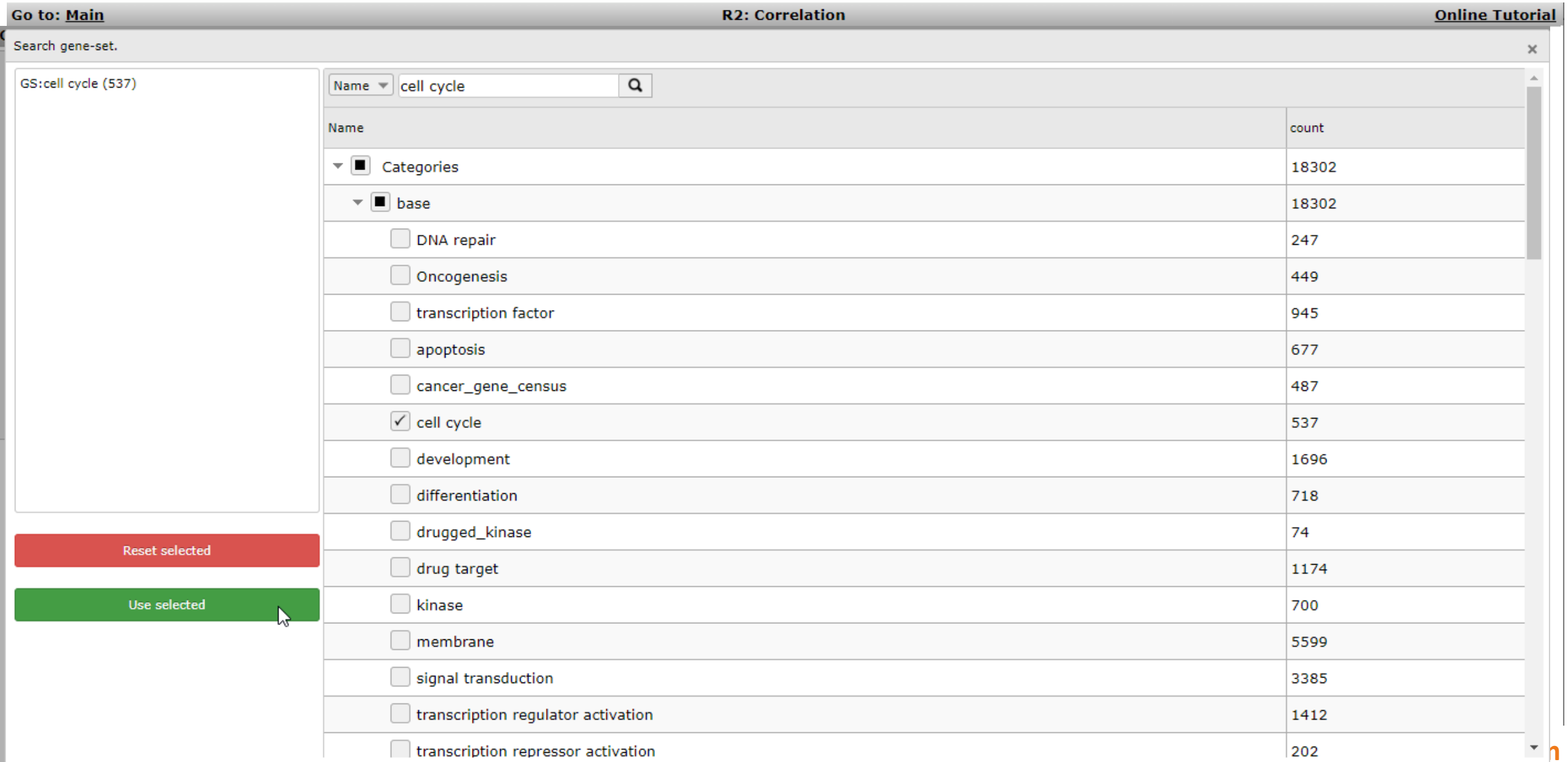
R2-Platform

# Correlations with a gene

# Correlate with a gene



# Correlate with a gene



# Correlate with a gene

Go to: I

| View | Gene | R | P | Presence |
|---|---|---|---|---|
| | KATNB1 | 0.63 | 2.35e-8 | 88/88 |
| | H2AFX | 0.575 | 7.36e-7 | 88/88 |
| | CHEK1 | 0.565 | 0.00000115 | 83/88 |
| | CHAF1A | 0.561 | 0.00000119 | 80/88 |
| | RUVBL1 | 0.56 | 0.00000106 | 86/88 |
| | MCM2 | 0.557 | 0.00000117 | 88/88 |
| | | | | 81/88 |
| | | | | 88/88 |
| | | | | 73/88 |
| | | | | 80/88 |
| | | | | 88/88 |
| | | | | 88/88 |
| | | | | 79/88 |
| | | | | 88/88 |
| | | | | 88/88 |
| | | | | 68/88 |
| | | | | 88/88 |

**Gene details**

**Gene symbol:**
MCM2 (GeneID: 4171)

**Description:**
minichromosome maintenance complex component 2

**Alternative names:**
- MCM2
- BM28
- CCNL1
- CDCL1
- D3S3194
- MITOTIN
- cdc19

**Other designations:**
- DNA replication licensing factor MCM2
- MCM2 minichromosome maintenance deficient 2, mitotin
- cell devision cycle-like 1
- cyclin-like 1
- minichromosome maintenance deficient 2 (mitotin)
- minichromosome maintenance protein 2 homolog
- nuclear protein BM28

**Ontology data:**
- cell cycle

| View | Gene | R | P | Presence |
|---|---|---|---|---|
| | CDC42 | -0.621 | 2.57e-8 | 87/88 |
| | SEPT4 | -0.524 | 0.00000828 | 86/88 |
| | SEPT8 | -0.507 | 0.0000197 | 88/88 |
| | SEPT7 | -0.494 | 0.0000337 | 88/88 |
| | CDKN2D | -0.483 | 0.0000563 | 81/88 |
| | MAPK1 | -0.482 | 0.0000562 | 88/88 |
| | PTP4A1 | -0.418 | 0.000516 | 88/88 |
| | CLIP1 | -0.416 | 0.000539 | 88/88 |
| | LIG4 | -0.403 | 0.00085 | 88/88 |
| | TACC1 | -0.395 | 0.00104 | 88/88 |
| | AHR | -0.388 | 0.00132 | 88/88 |
| | PAFAH1B1 | -0.379 | 0.00165 | 88/88 |
| | DMTF1 | -0.374 | 0.00192 | 88/88 |
| | ING4 | -0.364 | 0.00262 | 88/88 |
| | CYLD | -0.359 | 0.00301 | 88/88 |
| | RASSF4 | -0.358 | 0.00302 | 88/88 |
| | MAD2L2 | -0.33 | 0.00736 | 88/88 |

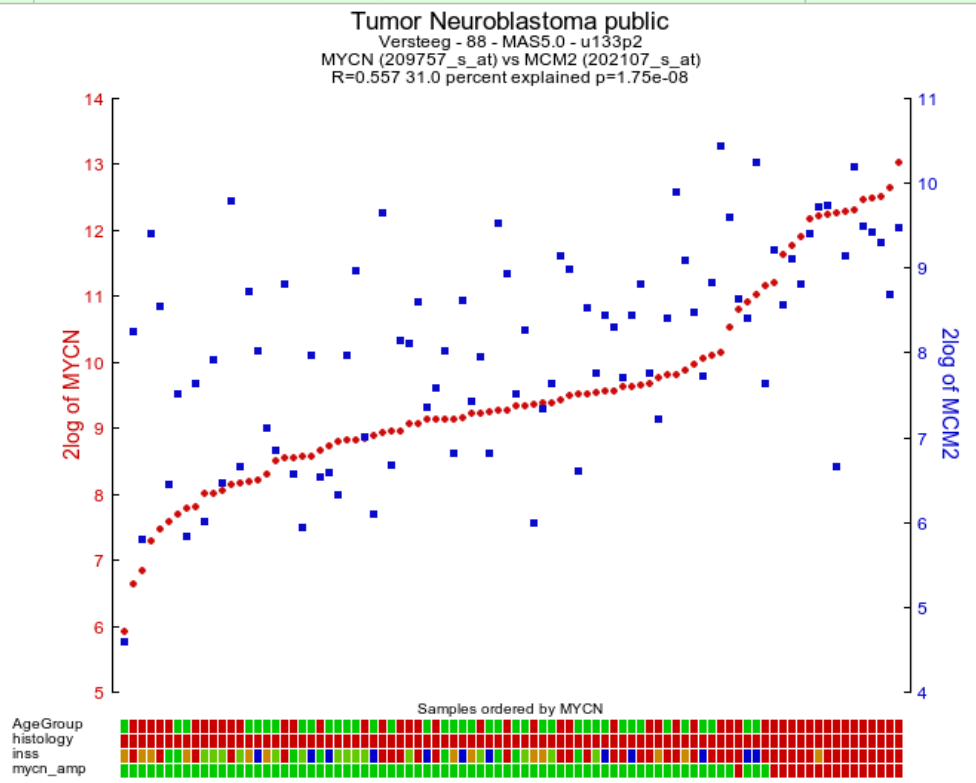
# Two Gene View for MYCN with MCM2

Go to: Main | R2: Two Gene View for MYCN with MCM2 | Online Tutorial

## Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2 public

Navigate through the result from the previous analysis

BIRC5 ⇦ | MCM2 (202107_s_at) pvalue=NA | ⇨ ZMYND11




**Significance of correlation:**
**r-value**=0.557 **p-value**=1.75e-08 **T-value**=6.218 **degrees of freedom**=86

| GeneID | Hugo | Description | R2 gene categories |
|---|---|---|---|
| 4171 | MCM2 | minichromosome maintenance complex component 2 | cell cycle |
| 4613 | MYCN | v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog | development, transcription factor |

▸ View additional details

**Pubsniffer**
MYCN vs MCM2

MCM2
MYCN

**KaplanScanner**
MCM2 (overall-2011/03)
MYCN (overall-2011/03)
MCM2 (relapsefree-2011/03)
MYCN (relapsefree-2011/03)

**Time Series**
Check

**ProbePlus**
209757_s_at
202107_s_at

**CliniSnitch**
MCM2
MYCN

**Across datasets**
2D Distribution

**Data set**
Sample overview

**Sample Map(s)**
map 2021-01-06
map 2022-03-01

# Pub sniffer

Go to: Main

- Main
- Time series
- AmpliconView
- Kaplan-Meier
- Sample maps
- Small Tools
- DataGrabber
- Genome Browser
- ChIP data
- TAR literature
- Change Data Scope
- User Options
- Help
- Contact / About R2

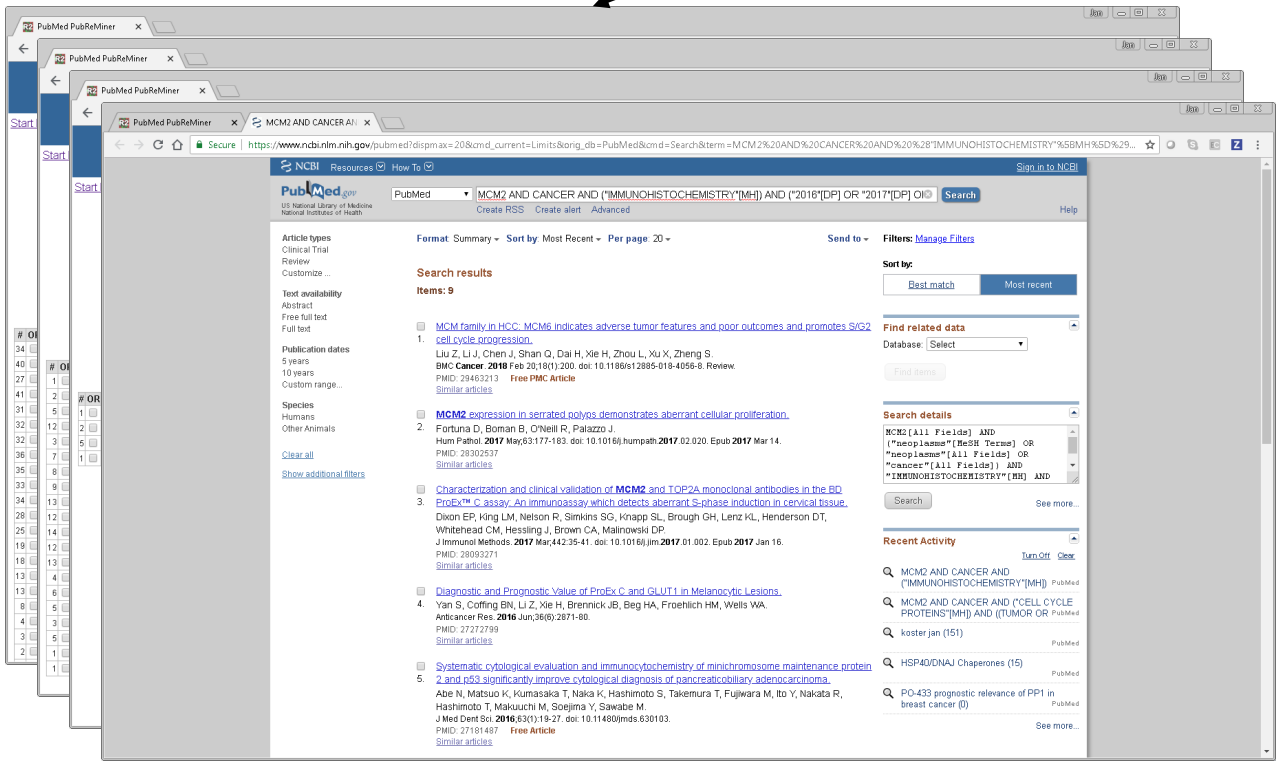
**R2: PubSniffer** — Online Tutorial

Checking NCBI PubMed for MYCN, MCM2

| Query | # Articles | Outlink |
|---|---|---|
| 'MYCN AND MCM2' | 1 | Pubreminer |
| 'MYCN' | 2934 | Pubreminer |
| 'MCM2' | 1336 | Pubreminer |
| 'MYCN AND MCM2 AND cancer' | 1 | Pubreminer |
| 'MYCN AND MCM2 AND development' | 0 | Pubreminer |
| 'MYCN AND MCM2 AND neuroblastom*' | 1 | Pubreminer |
| 'MCM2 AND cancer' | 665 | Pubreminer |
| 'MCM2 AND development' | 199 | Pubreminer |
| 'MCM2 AND neuroblastom*' | 7 | Pubreminer |

**Co-ocurrences of "MYCN AND MCM2" within sentences of abstracts**

| | |
|---|---|
| PubMed(7) | **MYCN** was bound to the proximal promoters of the **MCM2** to -8 genes |



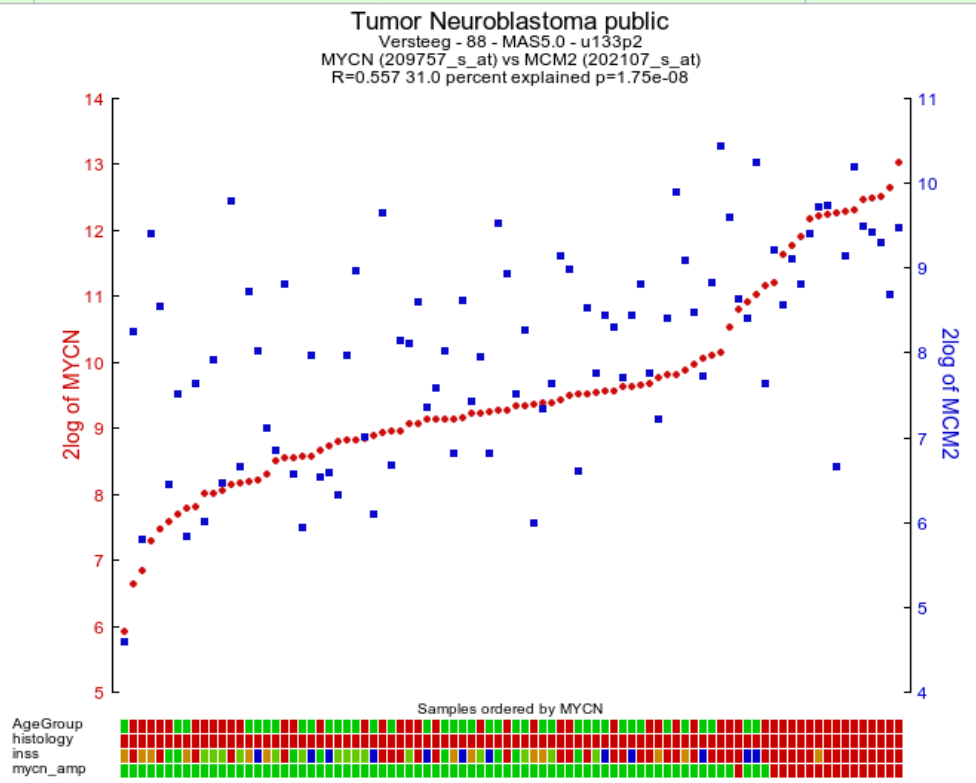
# Two gene View for MYCN with MCM2

Go to: Main | R2: Two Gene View for MYCN with MCM2 | Online Tutorial

## Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2 public

Navigate through the result from the previous analysis

BIRC5 ⇦ | MCM2 (202107_s_at) pvalue=NA | ⇨ ZMYND11



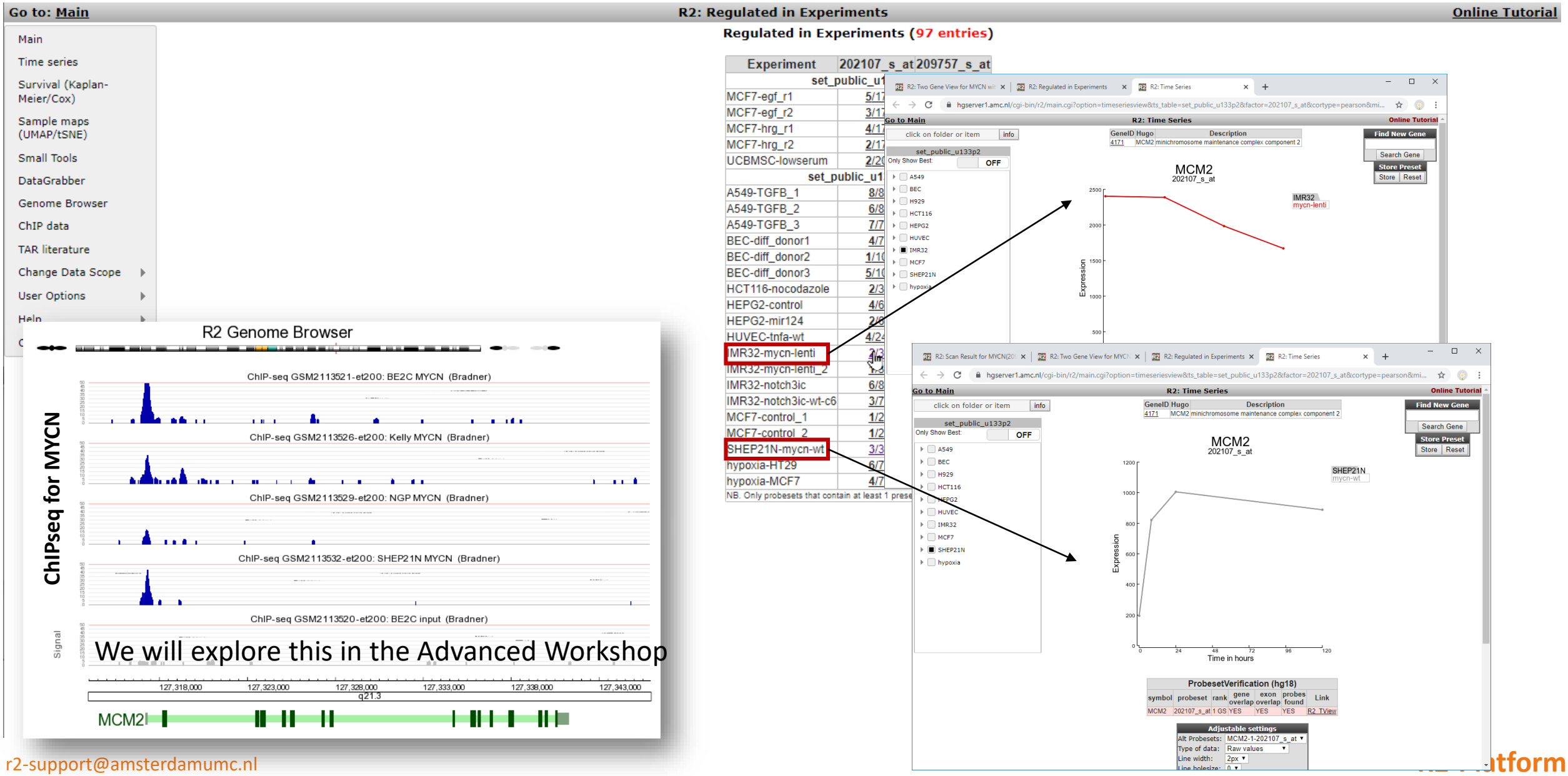

**Significance of correlation:**
**r-value**=0.557 **p-value**=1.75e-08 **T-value**=6.218 **degrees of freedom**=86

| GeneID | Hugo | Description | R2 gene categories |
|---|---|---|---|
| 4171 | MCM2 | minichromosome maintenance complex component 2 | cell cycle |
| 4613 | MYCN | v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog | development, transcription factor |

▸ View additional details

**Pubsniffer**
MYCN vs MCM2
**GeneCards**
MCM2
MYCN
**KaplanScanner**
MCM2 (overall-2011/03)
MYCN (overall-2011/03)
MCM2 (relapsefree-2011/03)
MYCN (relapsefree-2011/03)
**Time Series**
Check
209757_s_at
202107_s_at
**CliniSnitch**
MCM2
MYCN
**Across datasets**
2D Distribution
**Data set**
Sample overview
**Sample Map(s)**
map 2021-01-06
map 2022-03-01

# Regulated in timeseries experiments



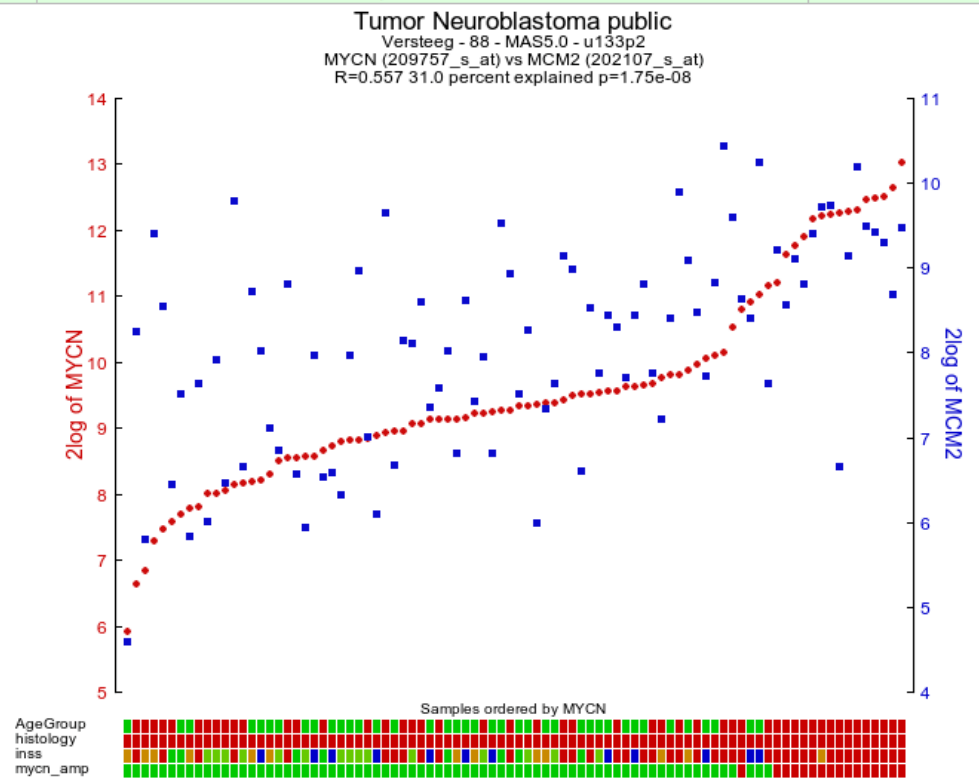
# 2d distribution

Go to: Main | R2: Two Gene View for MYCN with MCM2 | Online Tutorial

## Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2 public

Navigate through the result from the previous analysis

BIRC5 ⇦ | MCM2 (202107_s_at) pvalue=NA | ⇨ ZMYND11



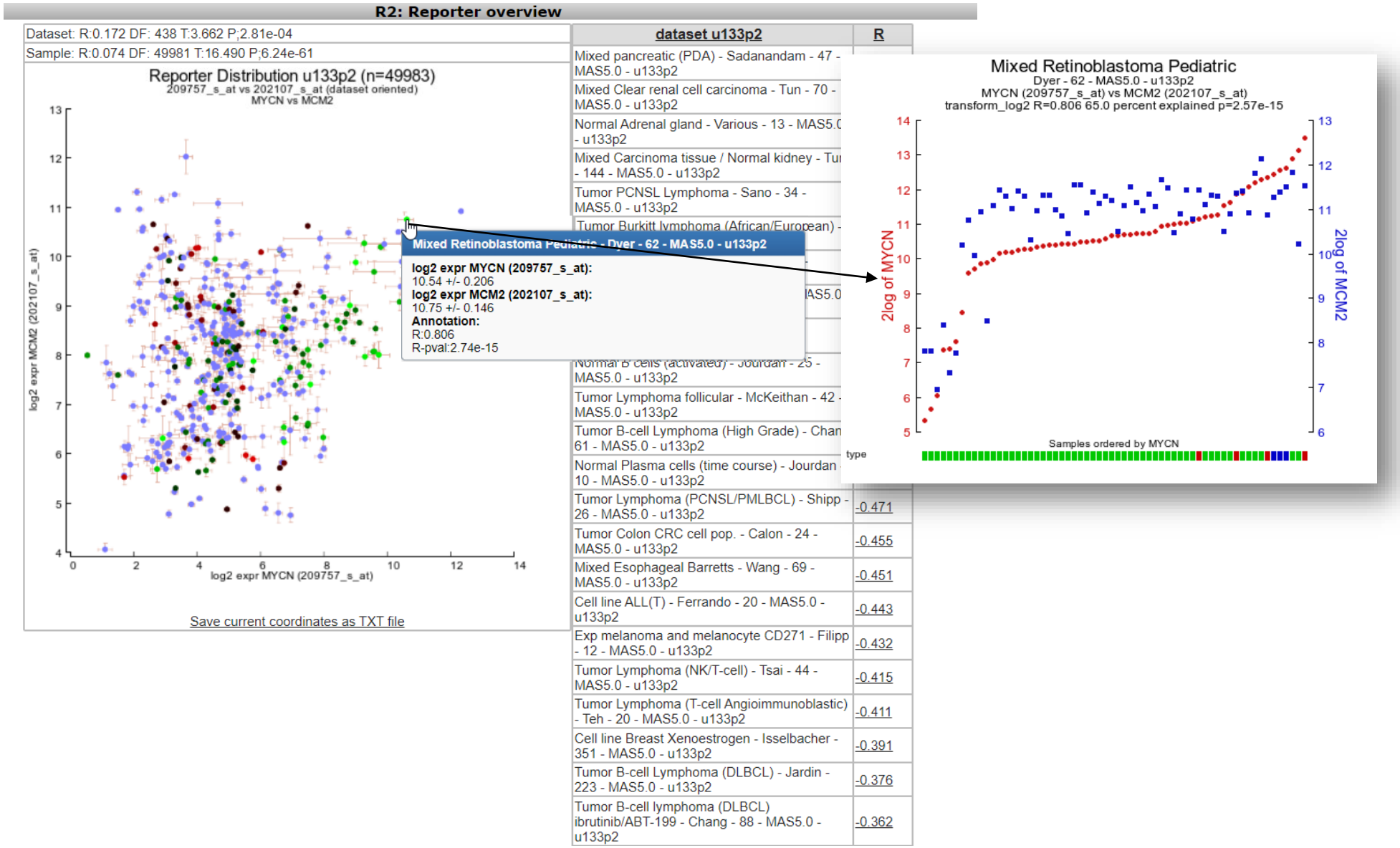

**Significance of correlation:**
**r-value**=0.557 **p-value**=1.75e-08 **T-value**=6.218 **degrees of freedom**=86

| GeneID | Hugo | Description | R2 gene categories |
|---|---|---|---|
| 4171 | MCM2 | minichromosome maintenance complex component 2 | cell cycle |
| 4613 | MYCN | v-myc avian myelocytomatosis viral oncogene neuroblastoma derived homolog | development, transcription factor |

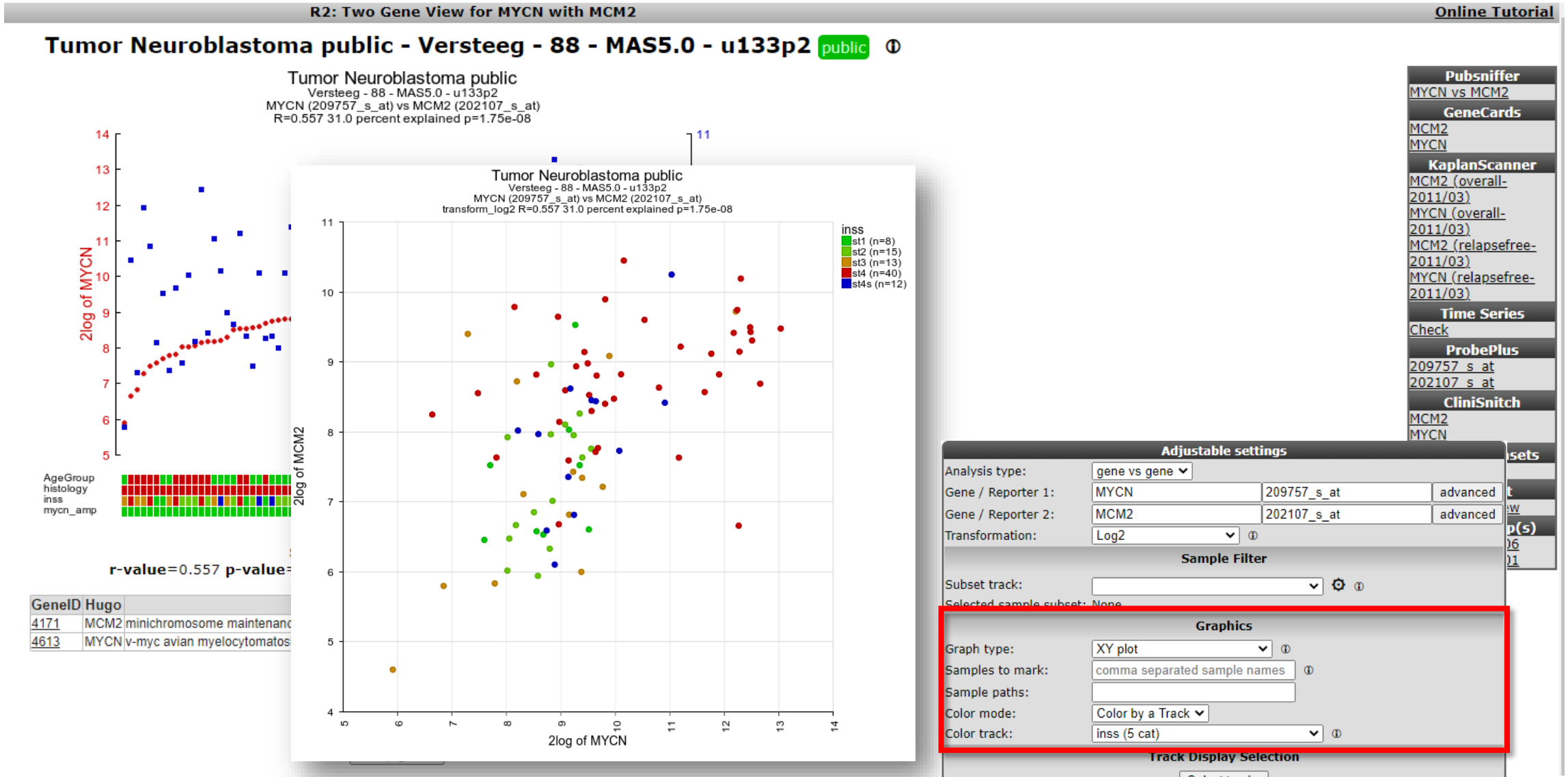
▸ View additional details

**Pubsniffer**
MYCN vs MCM2
**GeneCards**
MCM2
MYCN
**KaplanScanner**
MCM2 (overall-2011/03)
MYCN (overall-2011/03)
MCM2 (relapsefree-2011/03)
MYCN (relapsefree-2011/03)
**Time Series**
Check
**ProbePlus**
209757_s_at
202107_s_at
**CliniSnitch**
MCM2
MYCN
**Across datasets**
2D Distribution
**Data set**
Sample overview
**Sample Map(s)**
map 2021-01-06
map 2022-03-01

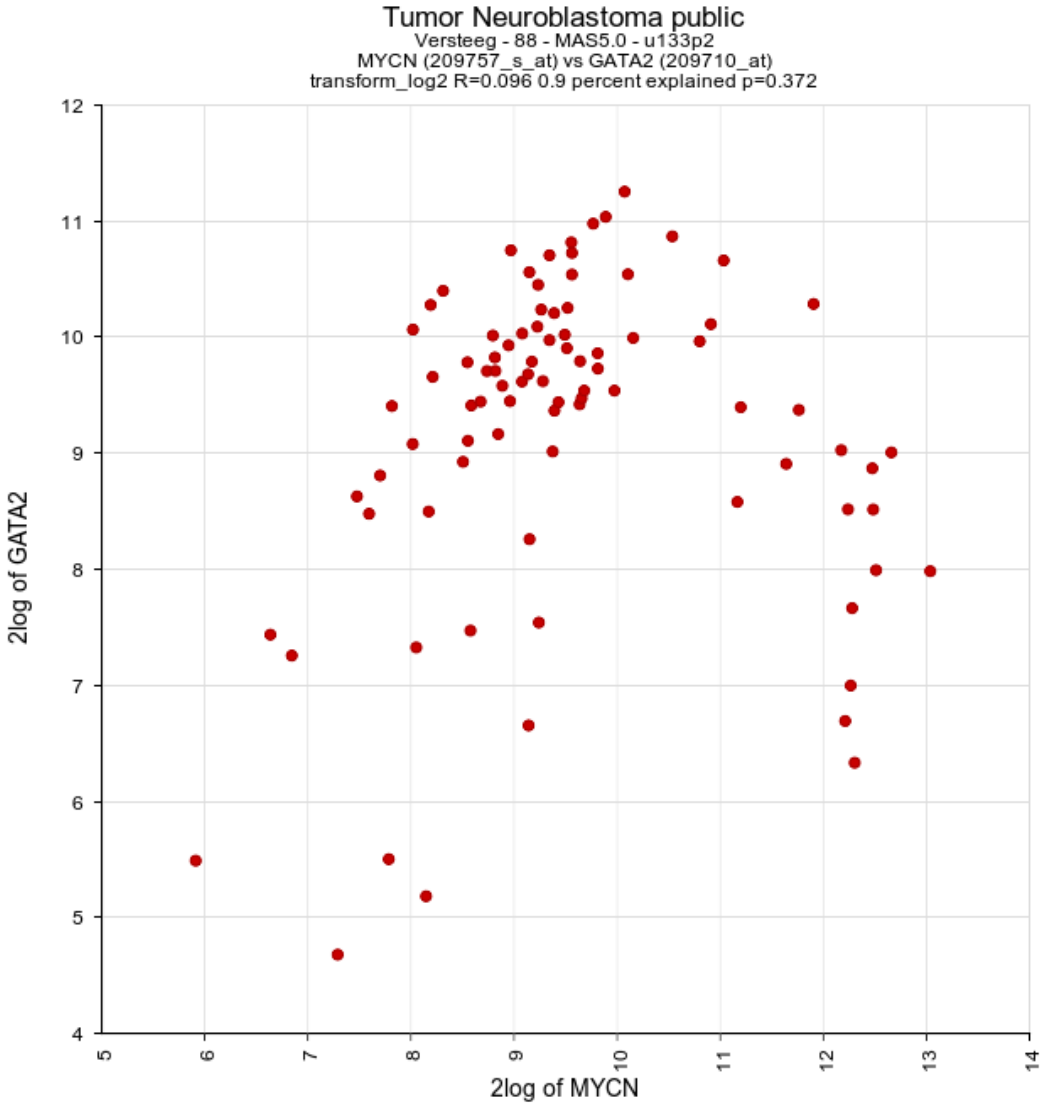
# 2d distribution



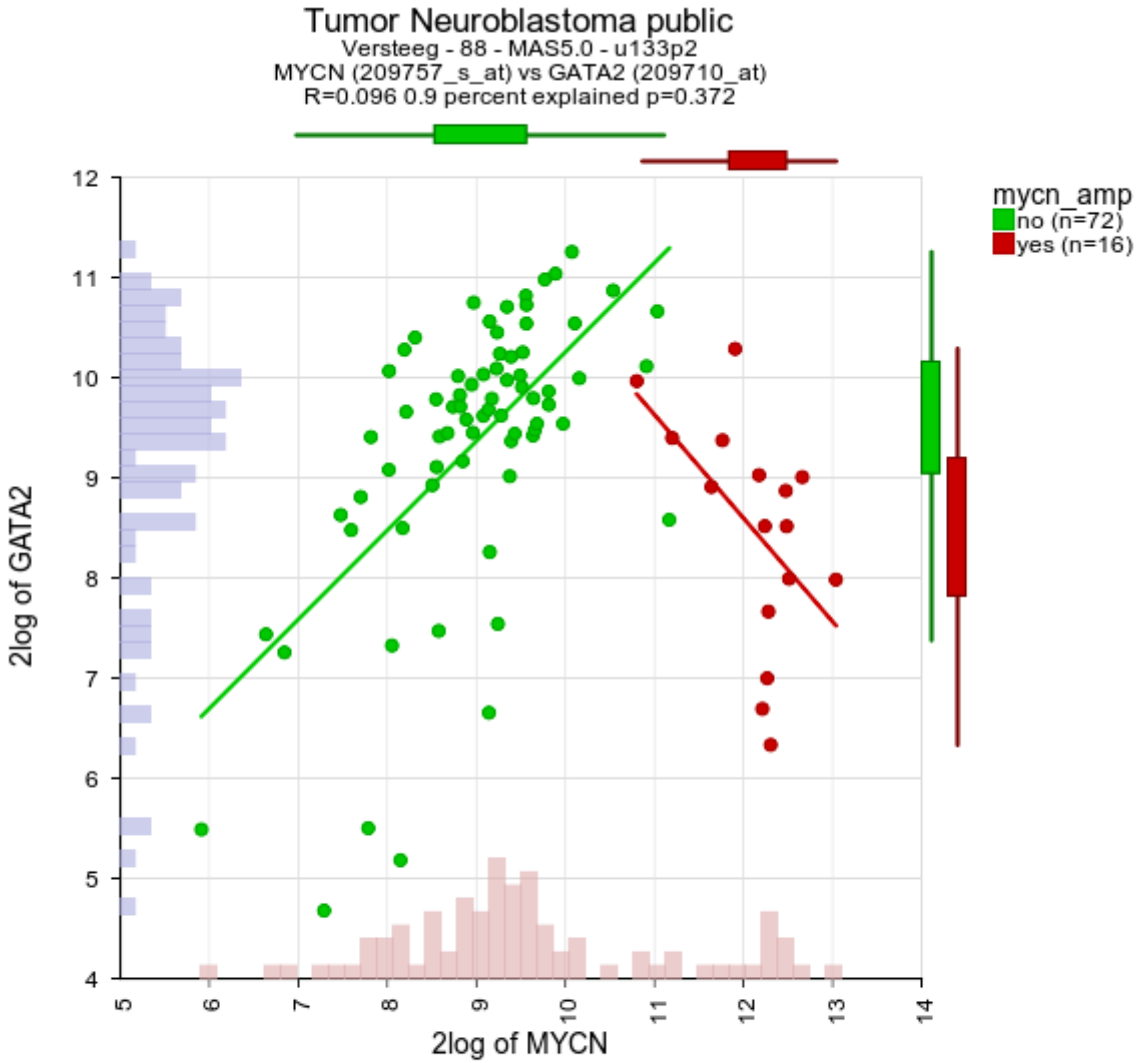
# X-gene-view


R2: Two Gene View for MYCN with MCM2
Online Tutorial
Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2 public
Tumor Neuroblastoma public
Versteeg - 88 - MAS5.0 - u133p2
MYCN (209757_s_at) vs MCM2 (202107_s_at)
R=0.557 31.0 percent explained p=1.75e-08
2log of MYCN
AgeGroup
histology
inss
mycn_amp
r-value=0.557 p-value=
GeneID Hugo
4171 MCM2 minichromosome maintenanc
4613 MYCN v-myc avian myelocytomatos
Tumor Neuroblastoma public
Versteeg - 88 - MAS5.0 - u133p2
MYCN (209757_s_at) vs MCM2 (202107_s_at)
transform_log2 R=0.557 31.0 percent explained p=1.75e-08
2log of MCM2
2log of MYCN
inss
st1 (n=8)
st2 (n=15)
st3 (n=13)
st4 (n=40)
st4s (n=12)
Pubsniffer
MYCN vs MCM2
GeneCards
MCM2
MYCN
KaplanScanner
MCM2 (overall-2011/03)
MYCN (overall-2011/03)
MCM2 (relapsefree-2011/03)
MYCN (relapsefree-2011/03)
Time Series
Check
ProbePlus
209757_s_at
202107_s_at
CliniSnitch
MCM2
MYCN
Adjustable settings
Analysis type: gene vs gene
Gene / Reporter 1: MYCN 209757_s_at advanced
Gene / Reporter 2: MCM2 202107_s_at advanced
Transformation: Log2
Sample Filter
Subset track:
Graphics
Graph type: XY plot
Samples to mark: comma separated sample names
Sample paths:
Color mode: Color by a Track
Color track: inss (5 cat)
Track Display Selection
Select tracks
More Settings
Submit

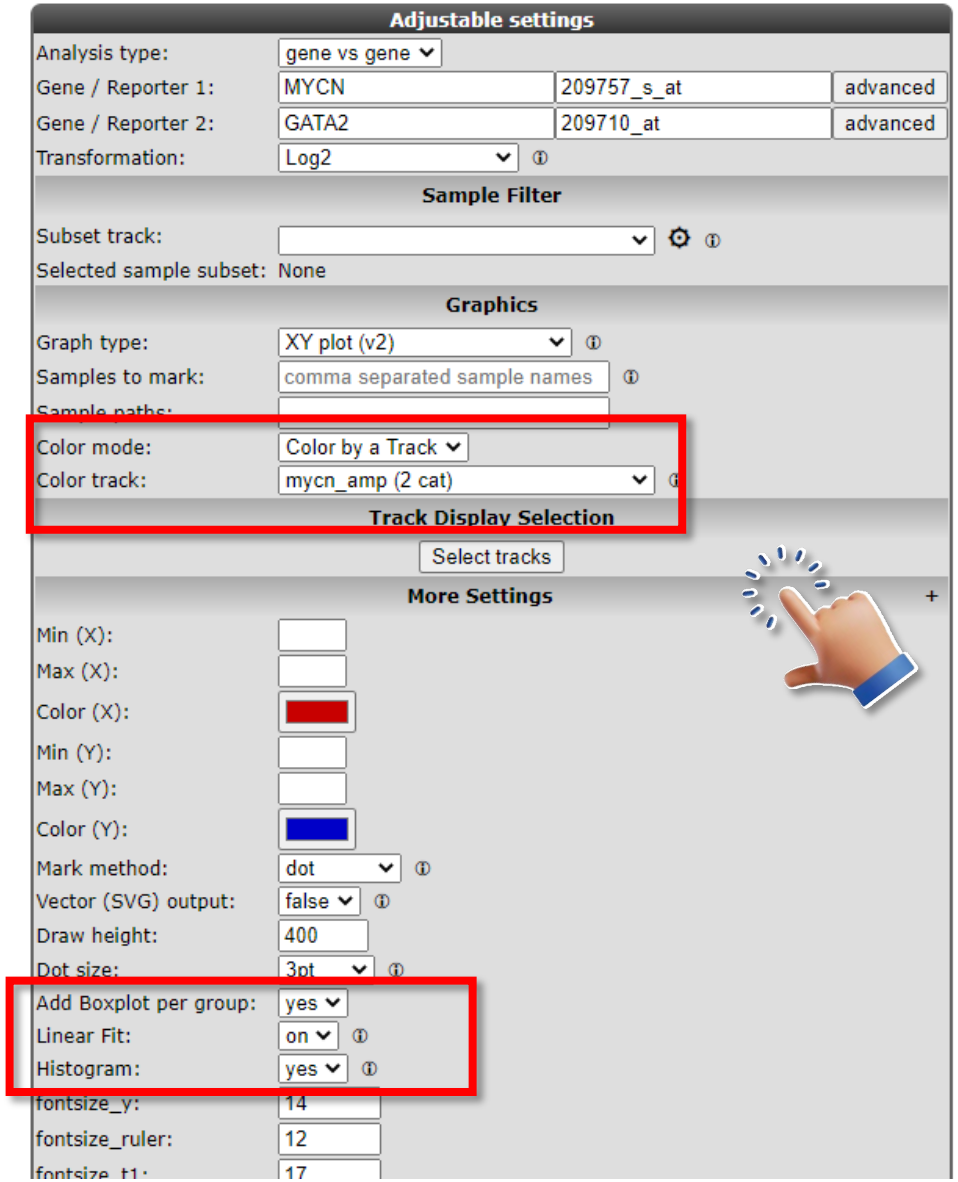
# X-gene-view


Tumor Neuroblastoma public
Versteeg - 88 - MAS5.0 - u133p2
MYCN (209757_s_at) vs GATA2 (209710_at)
transform_log2 R=0.096 0.9 percent explained p=0.372
2log of GATA2
2log of MYCN
12
11
10
9
8
7
6
5
4
5
6
7
8
9
10
11
12
13
14

# X-gene-view



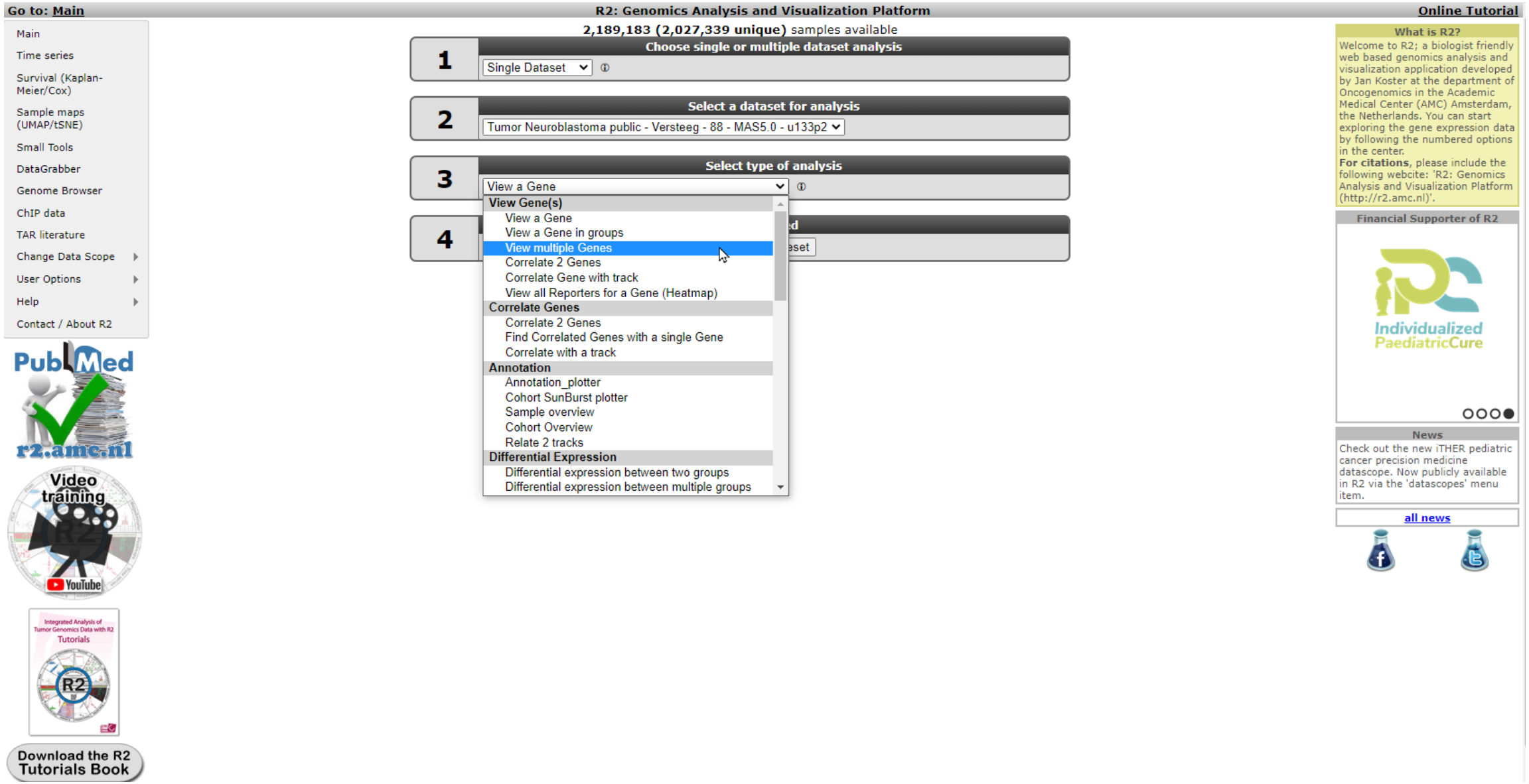

| Trend lines | | |
|---|---|---|
| group | r | rpval |
| no | 0.617 | 8.06e-09 |
| yes | -0.509 | 0.044 |



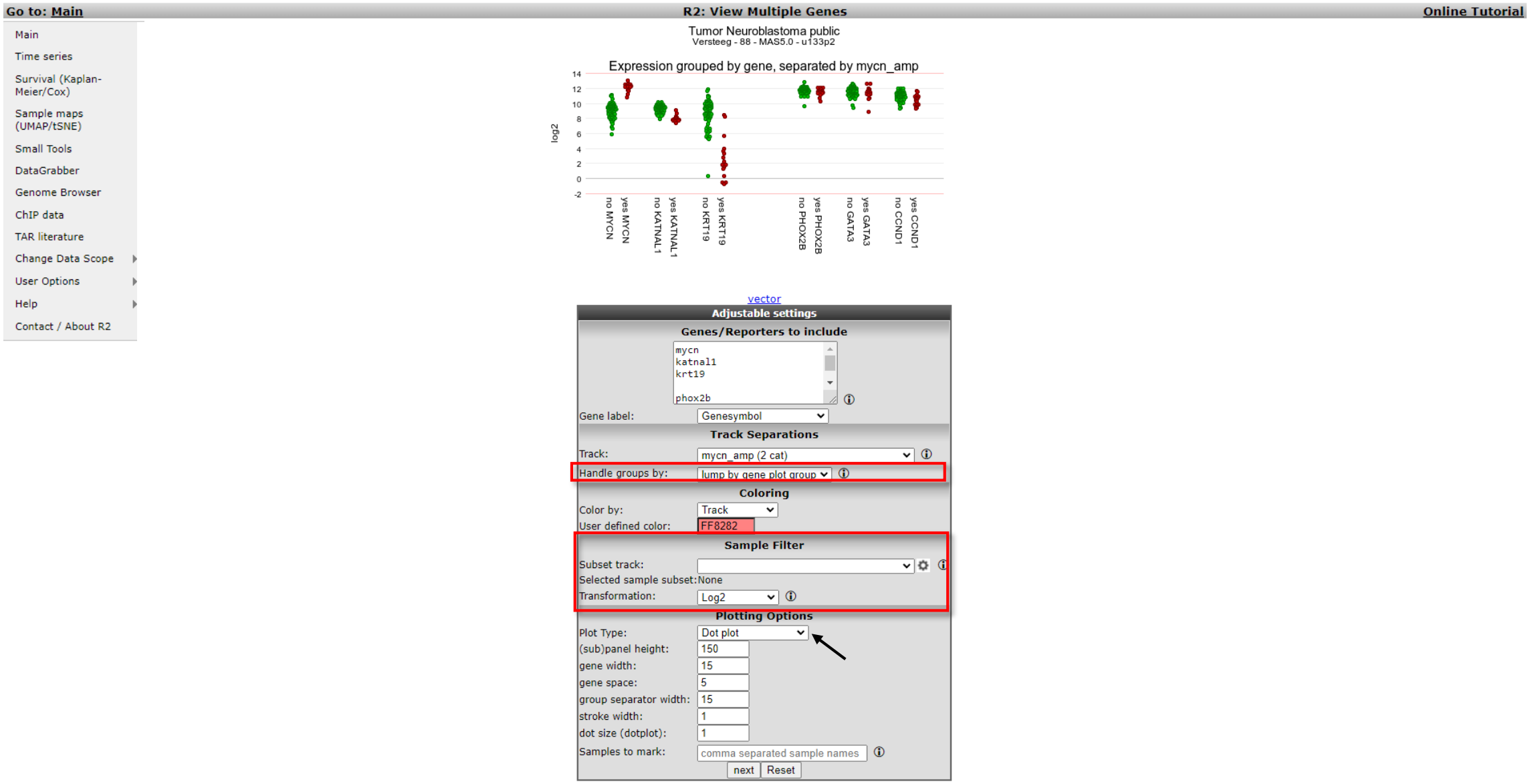

R2-Platform

# View multiple genes

# View multiple genes

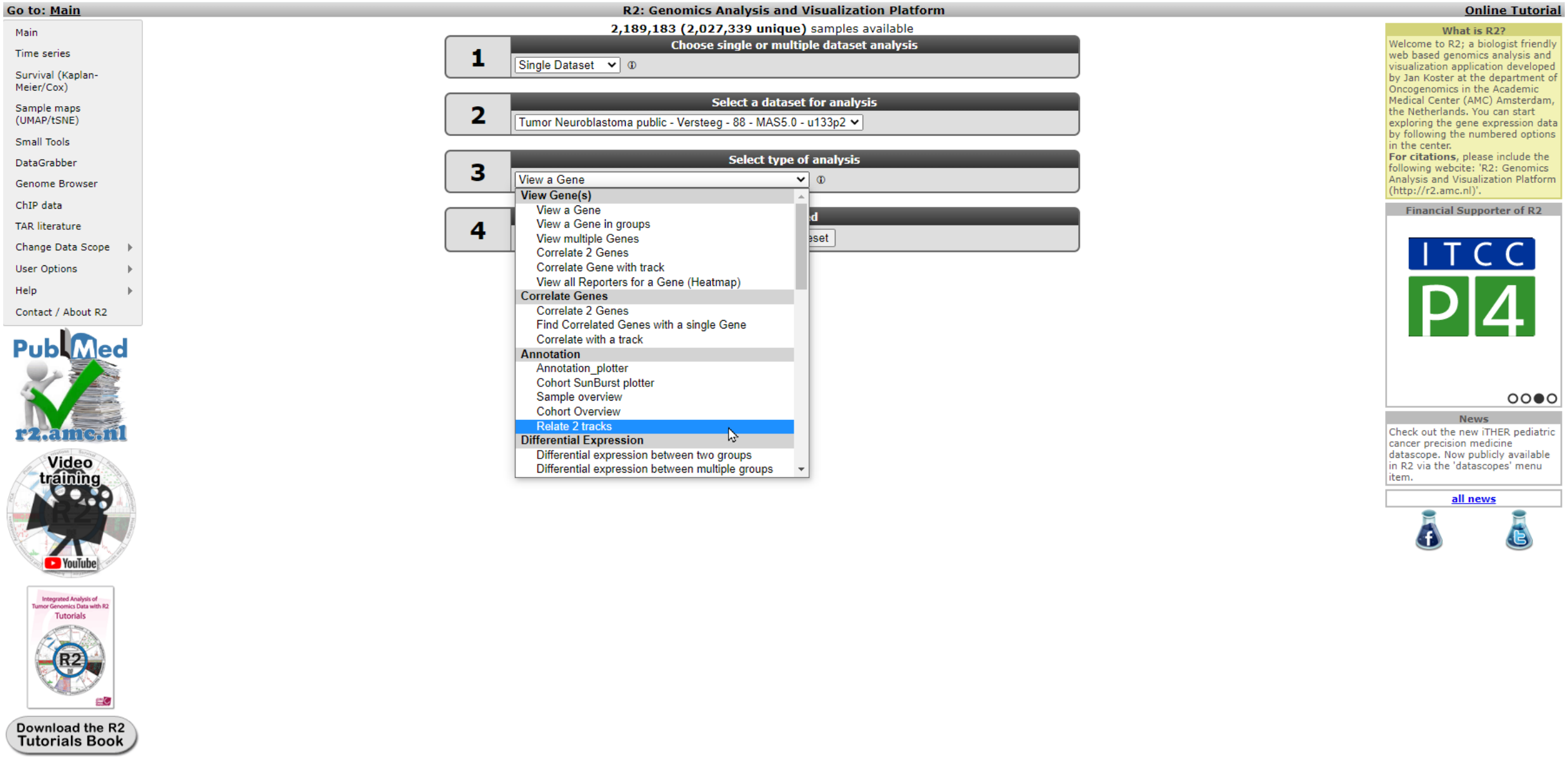


# View multiple genes



# Annotation Views

# Annotation Views



# Relate 2 tracks

**R2: Relate two tracks**

## Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2 public

**Adjustable settings**

| | |
|---|---|
| Analysis type: | track vs track |
| X track: | mycn_amp (2 cat) |
| Y track: | Alive (2 cat) |

**Sample Filter**

| | |
|---|---|
| Subset track: | |
| Selected sample subset: | None |

**Graphics**

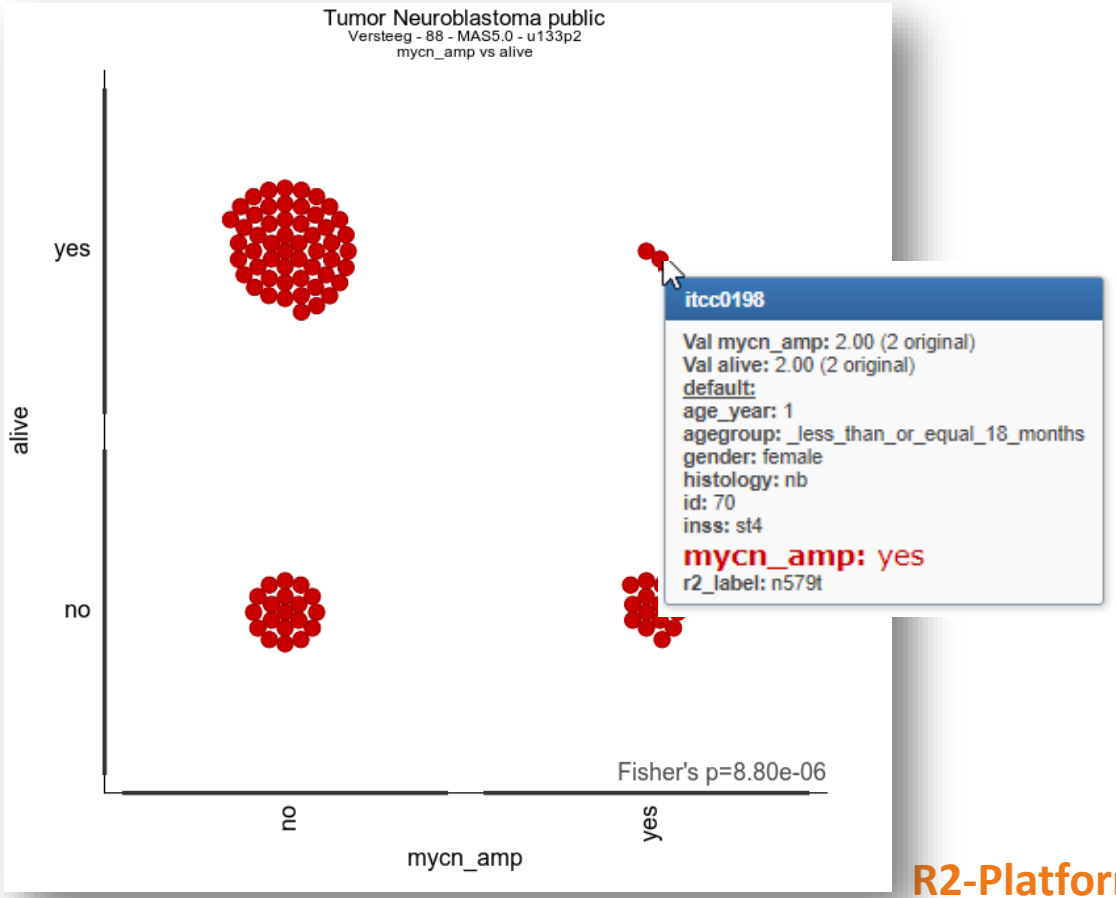
| | |
|---|---|
| Graph type: | XY plot |
| Samples to mark: | comma separated sample names |
| Sample paths: | |
| Group overlap as: | Single samples |
| Color mode: | Default Color |

**Track Display Selection**

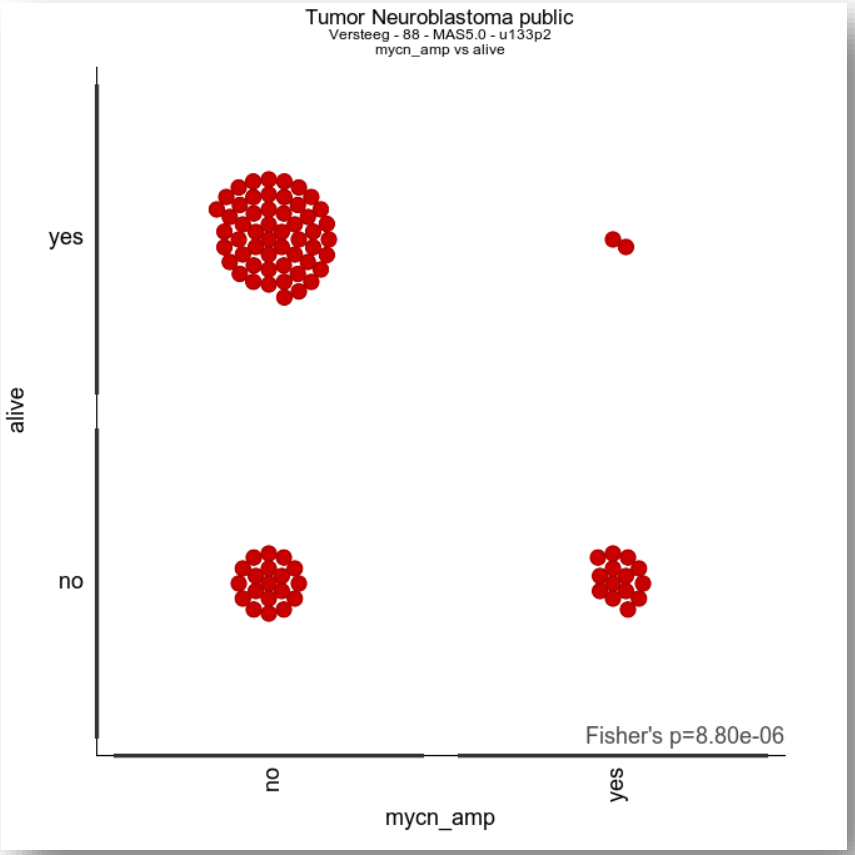
Select tracks

**More Settings** +

Submit



# Relate 2 tracks




**Adjustable settings**

| | |
|---|---|
| Analysis type: | track vs track |
| X track: | mycn_amp (2 cat) |
| Y track: | Alive (2 cat) |

**Sample Filter**

| | |
|---|---|
| Subset track: | |
| Selected sample subset: | None |

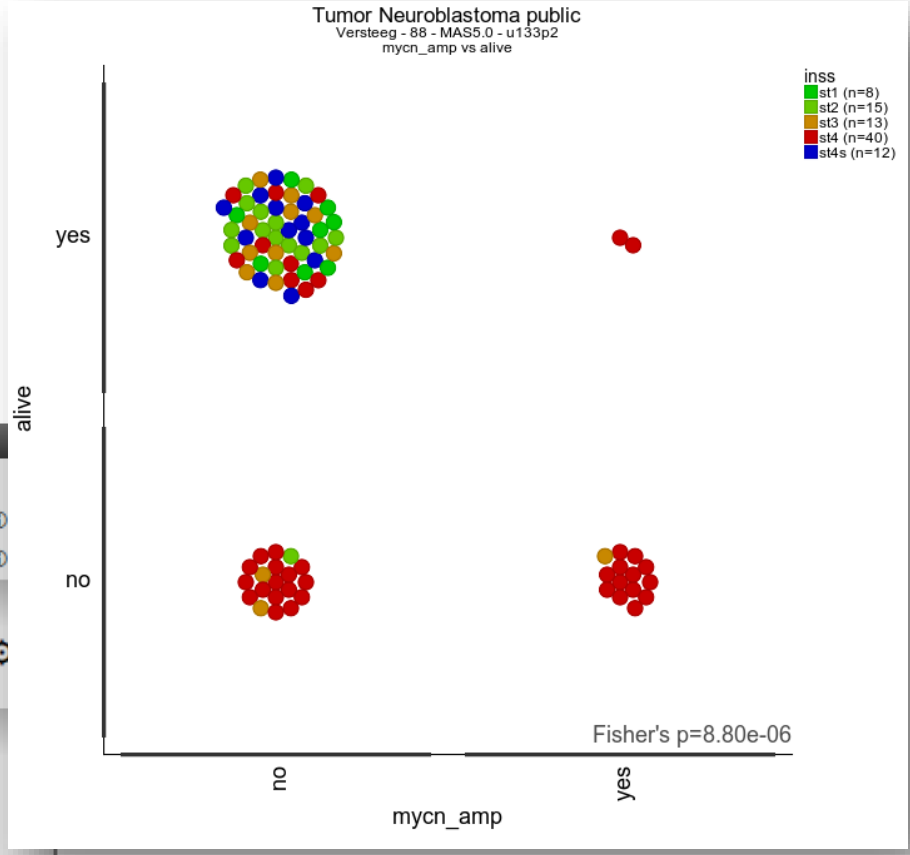
**Graphics**

| | |
|---|---|
| Graph type: | XY plot |
| Samples to mark: | comma separated sample names |
| Sample paths: | |
| Group overlap as: | Single samples |
| Color mode: | Color by a Track |
| Color track: | inss (5 cat) |

Select tracks

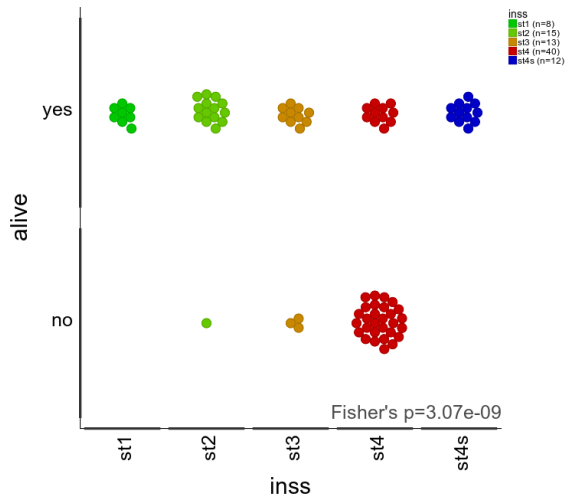
**More Settings** +

Submit

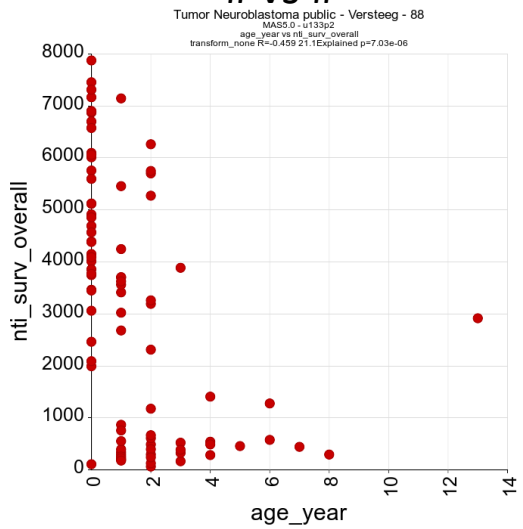


# Relate 2 tracks

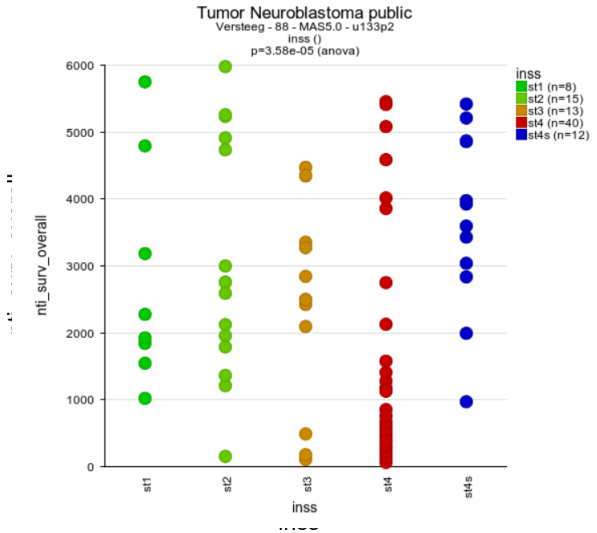
## Cat vs cat

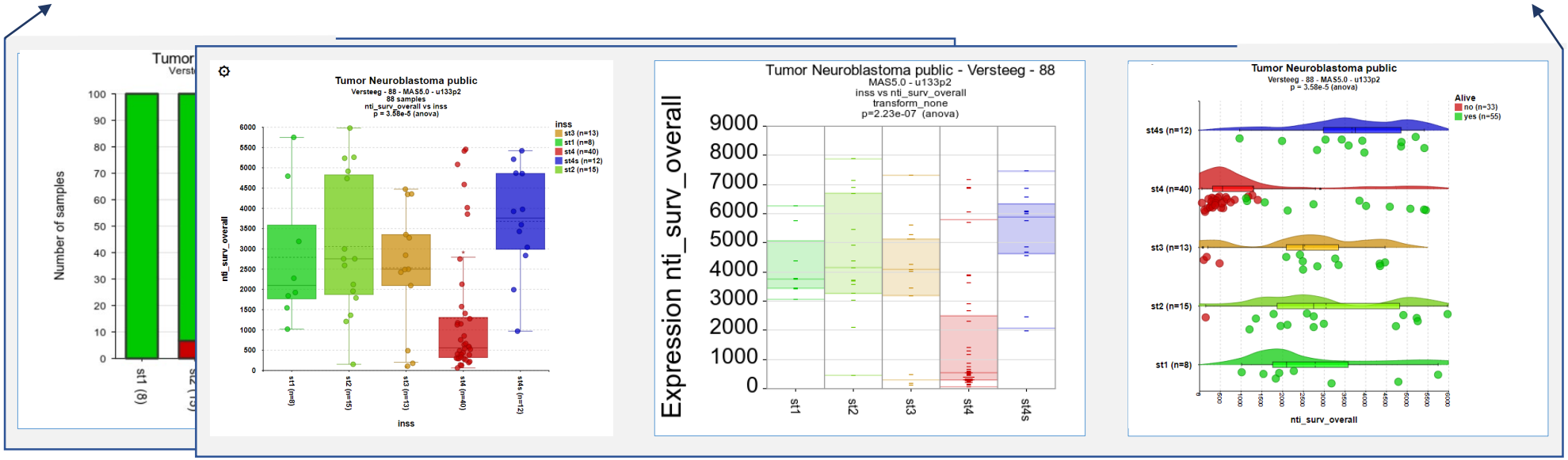



## # vs #

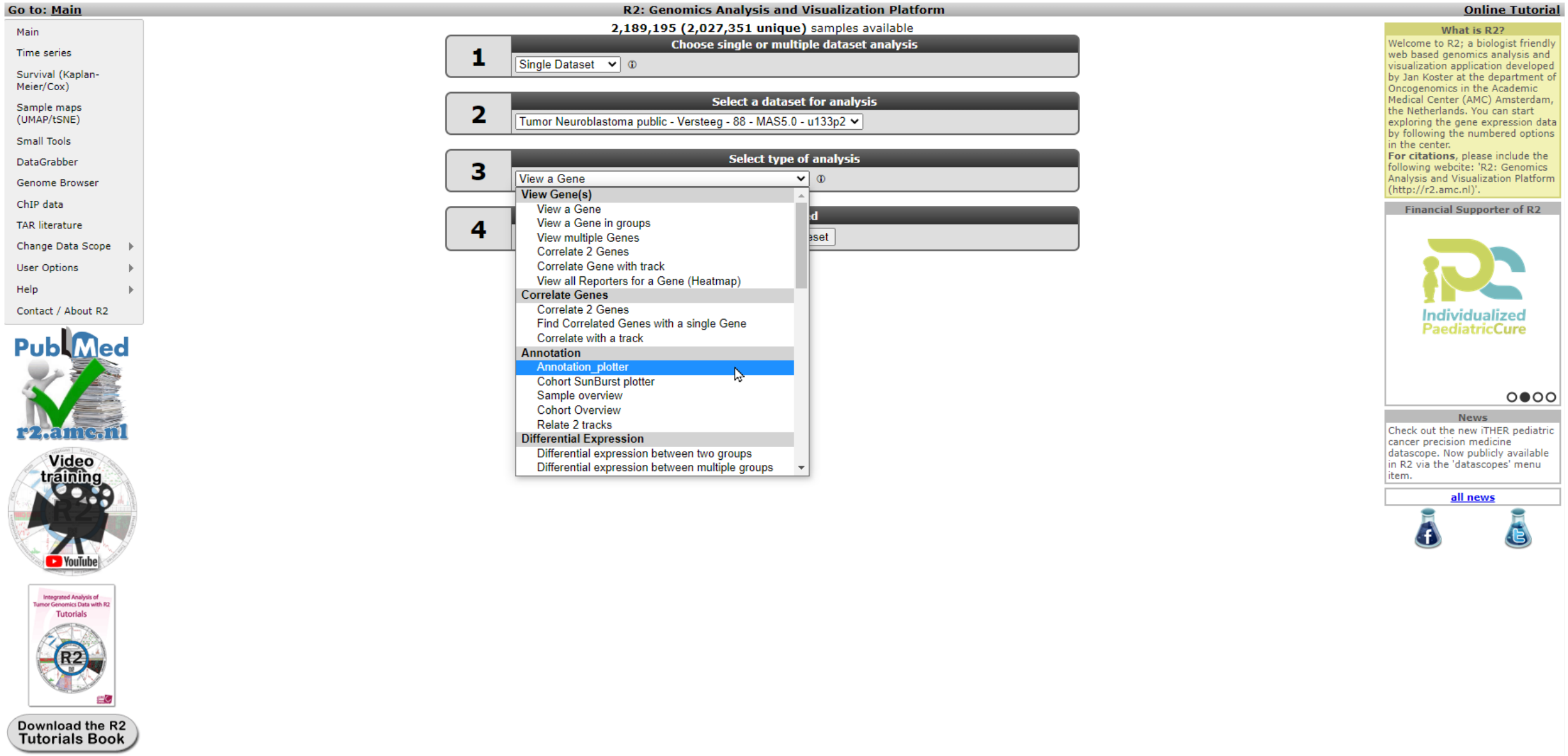



## Cat vs #

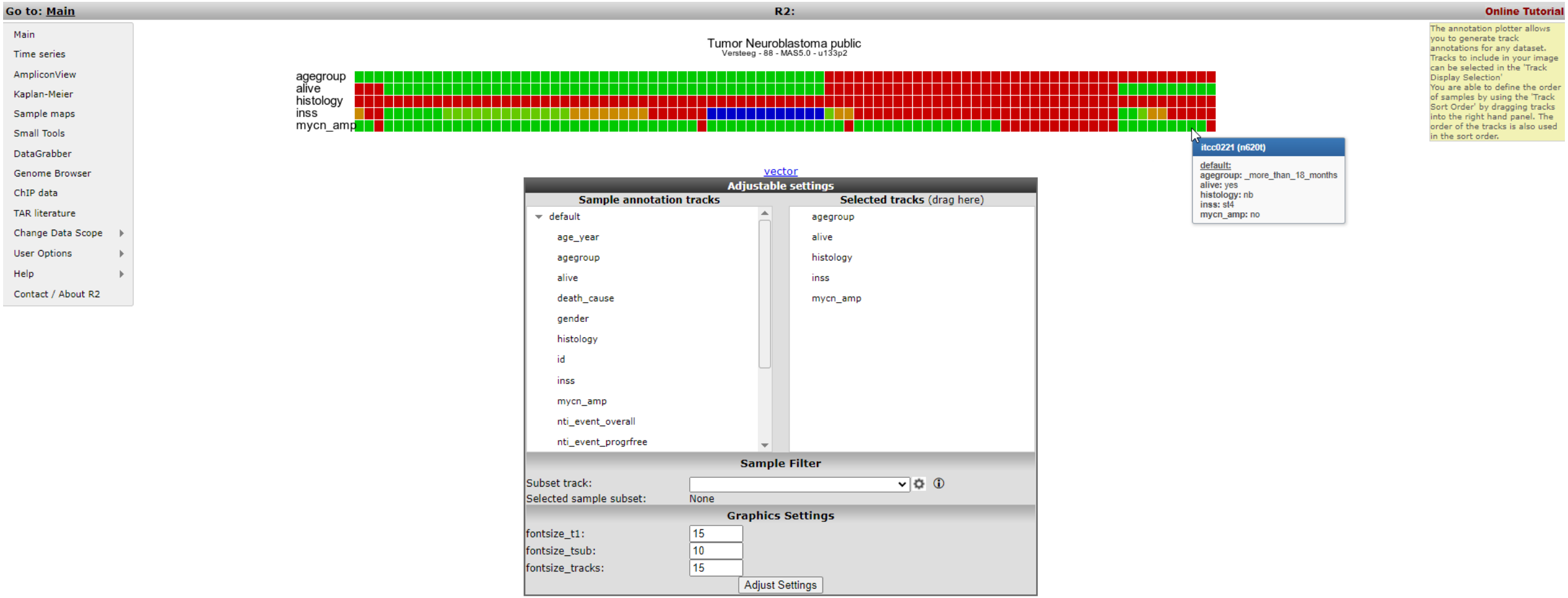


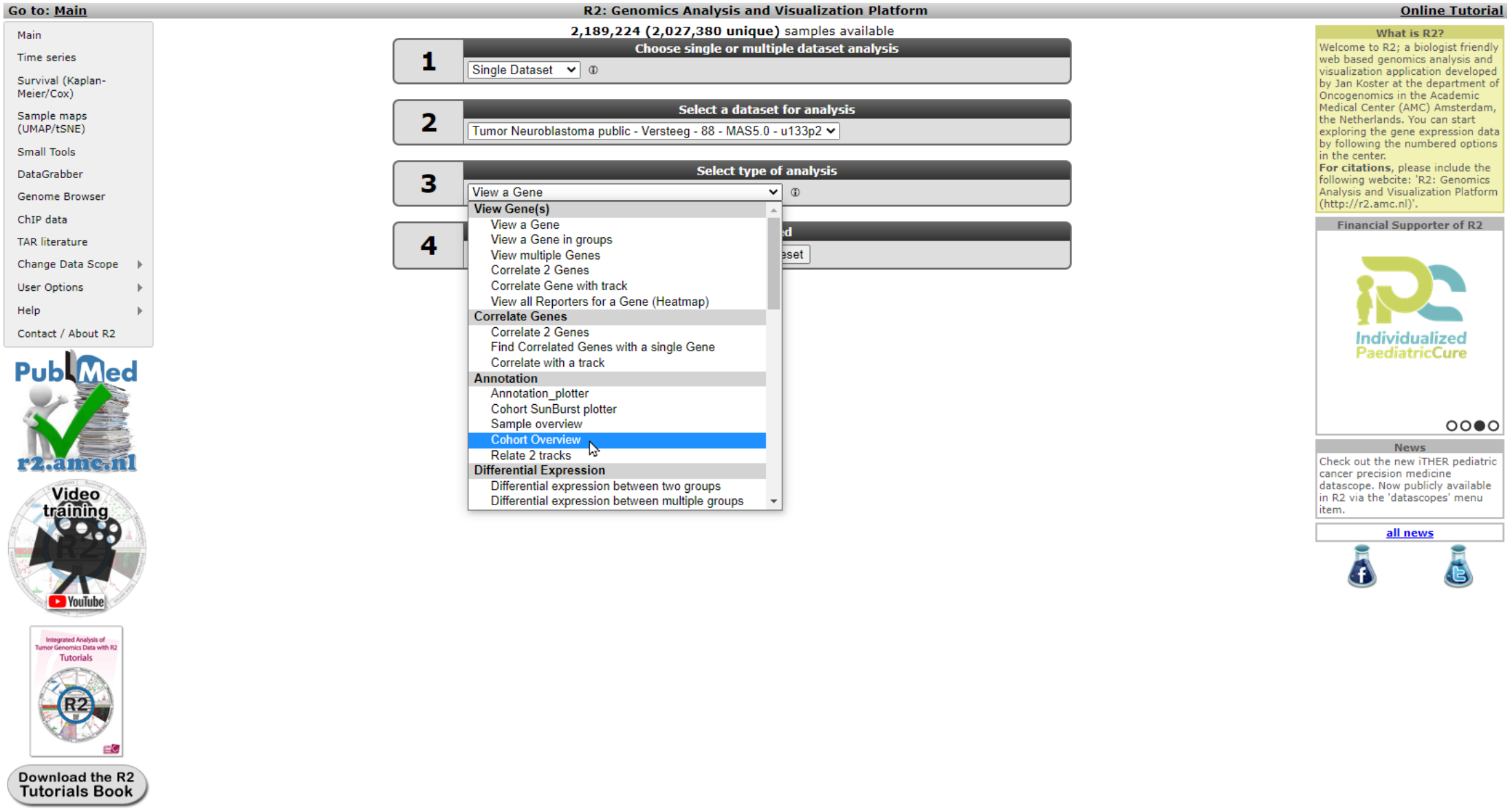
# Annotation Plotter



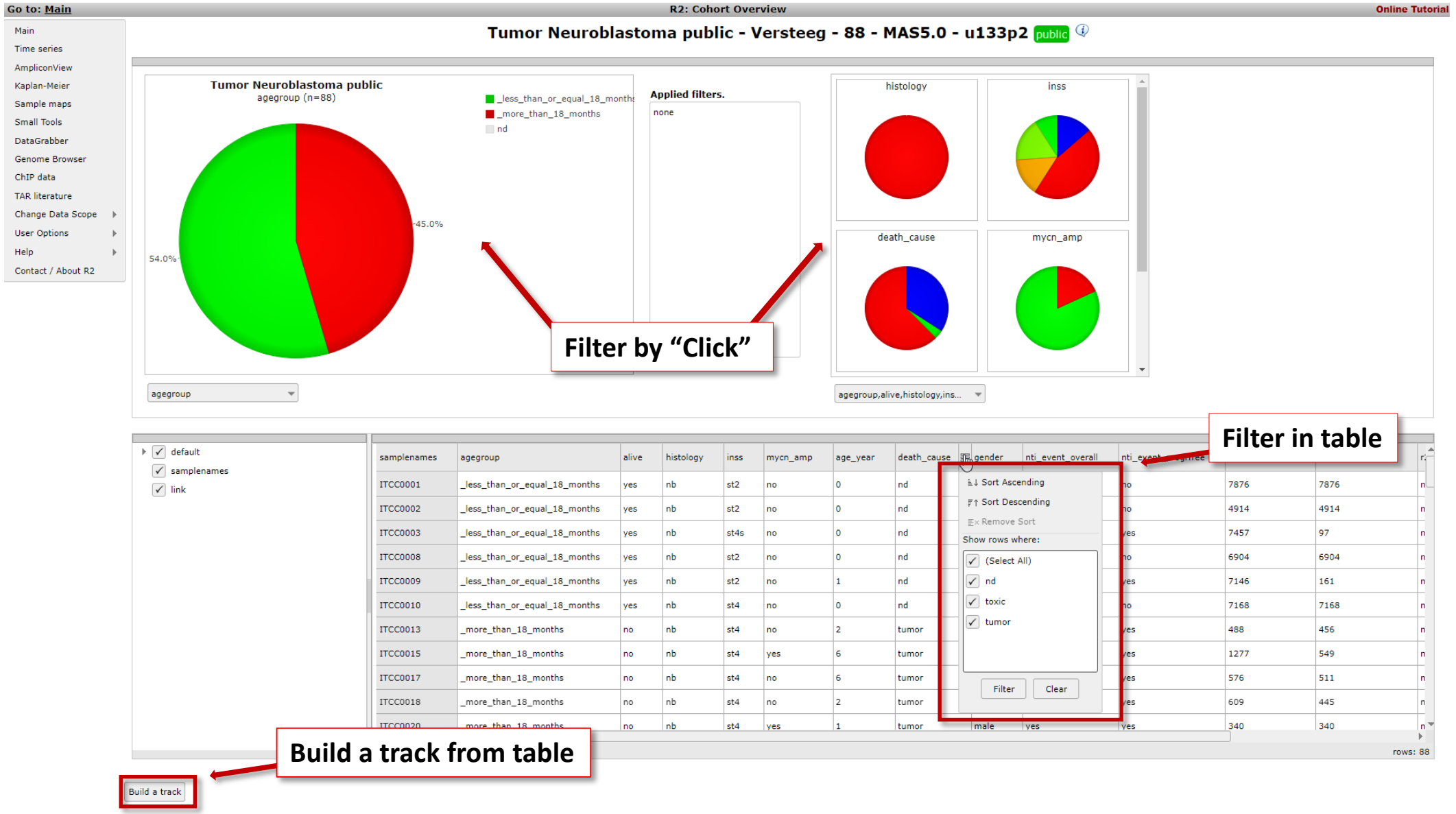
# Annotation Plotter



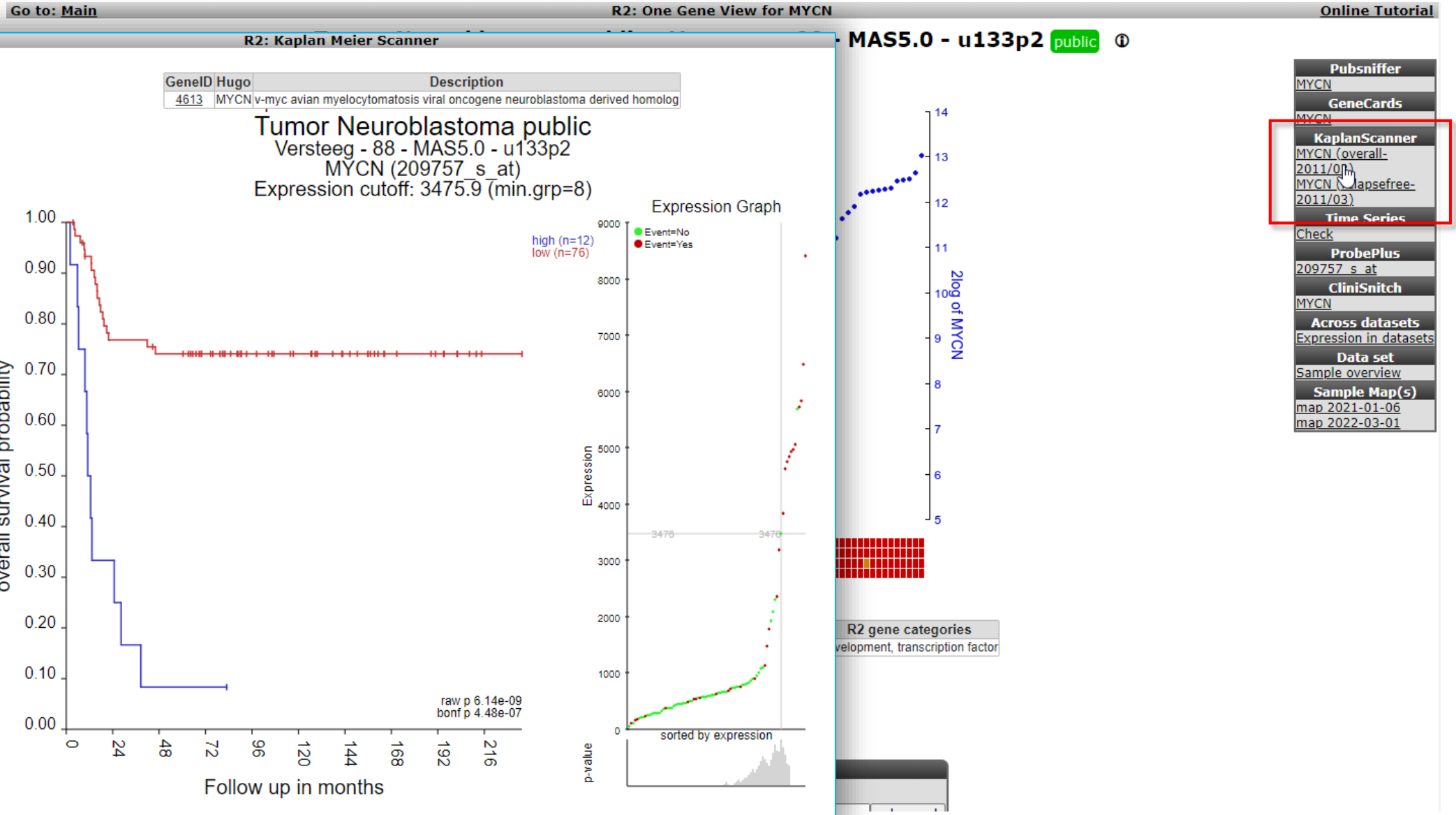
# Cohort Overview



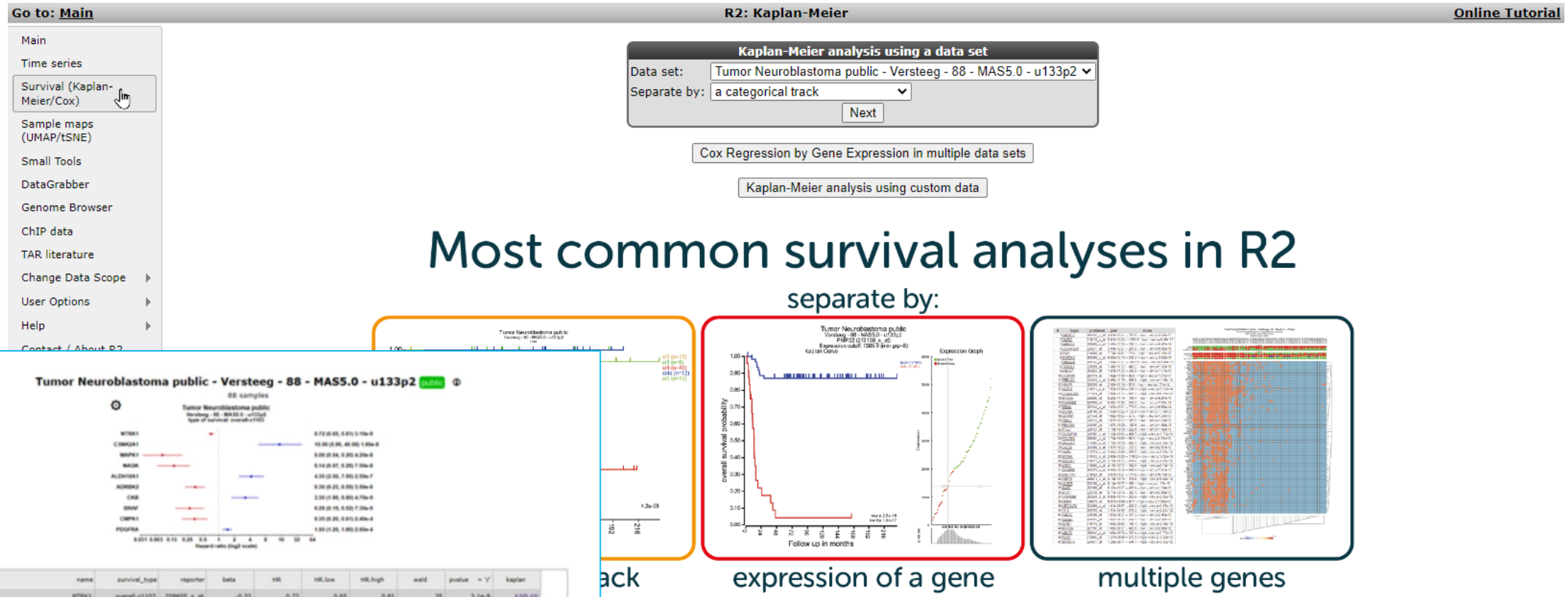
# Cohort Overview



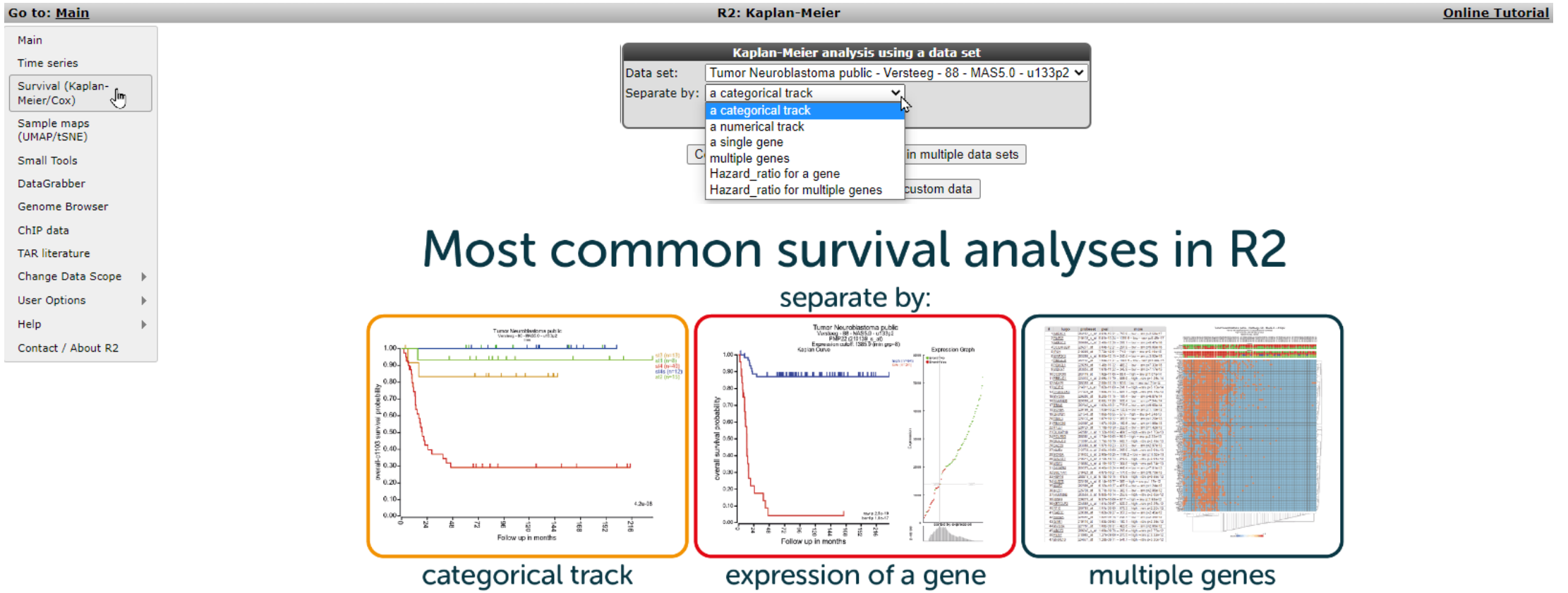
# Recap Result page panel for follow-up analyses



# Survival analysis (Kaplan)



# Survival analysis (Kaplan)



# Kaplan Meier by categorical track

**R2: Kaplan Meier**

**Using dataset Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2**

**Adjustable settings**

**Survival**

Type of Survival: overall-c1103

Separate by: track

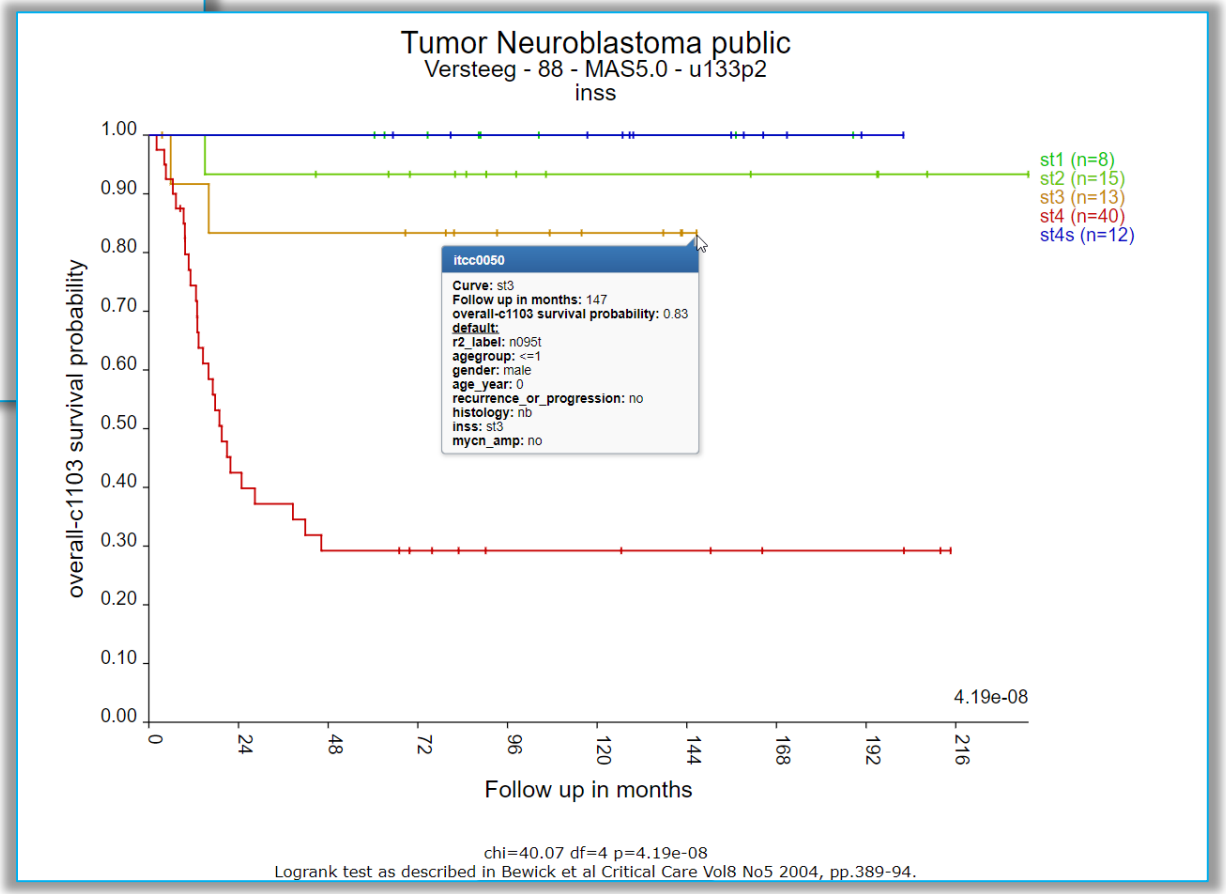
Track: inss (5 cat)

**Sample Filter**

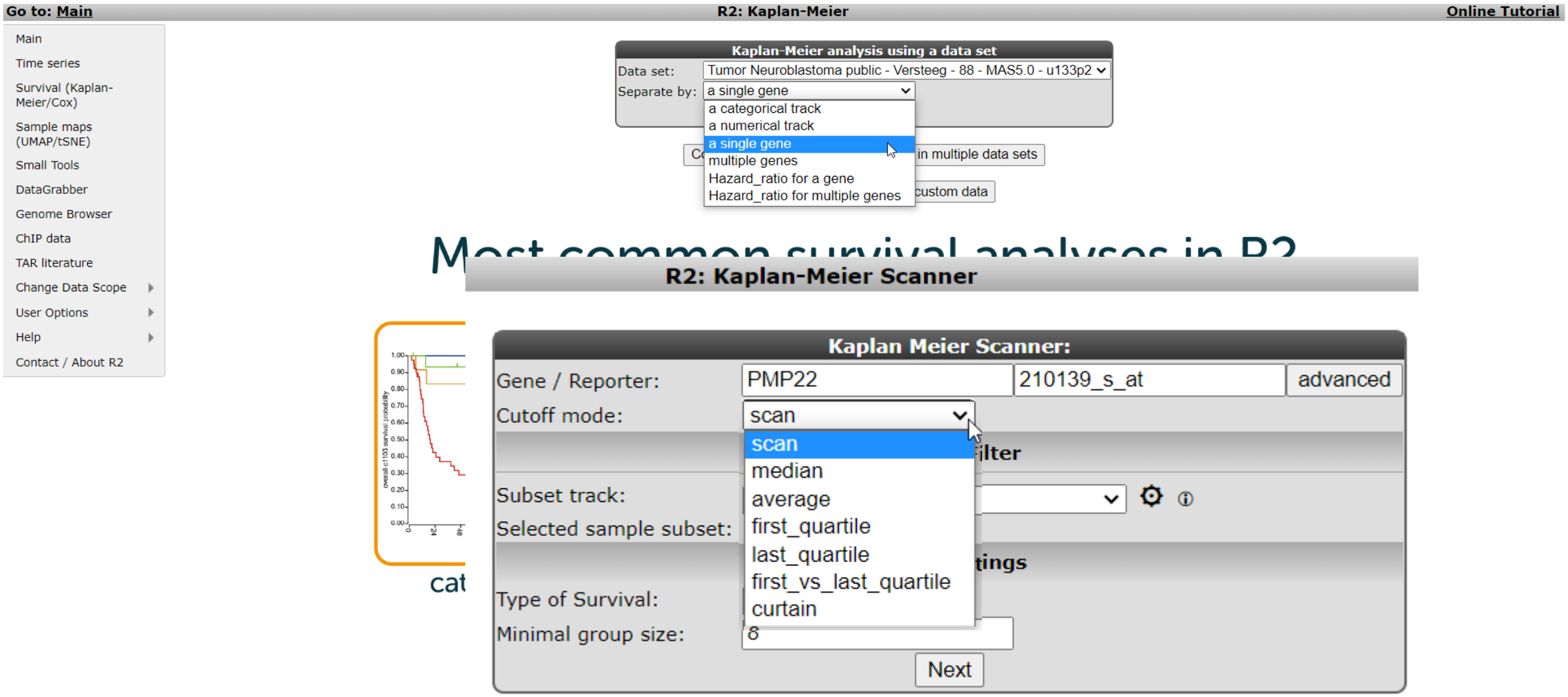
Subset track:

Selected sample subset: None

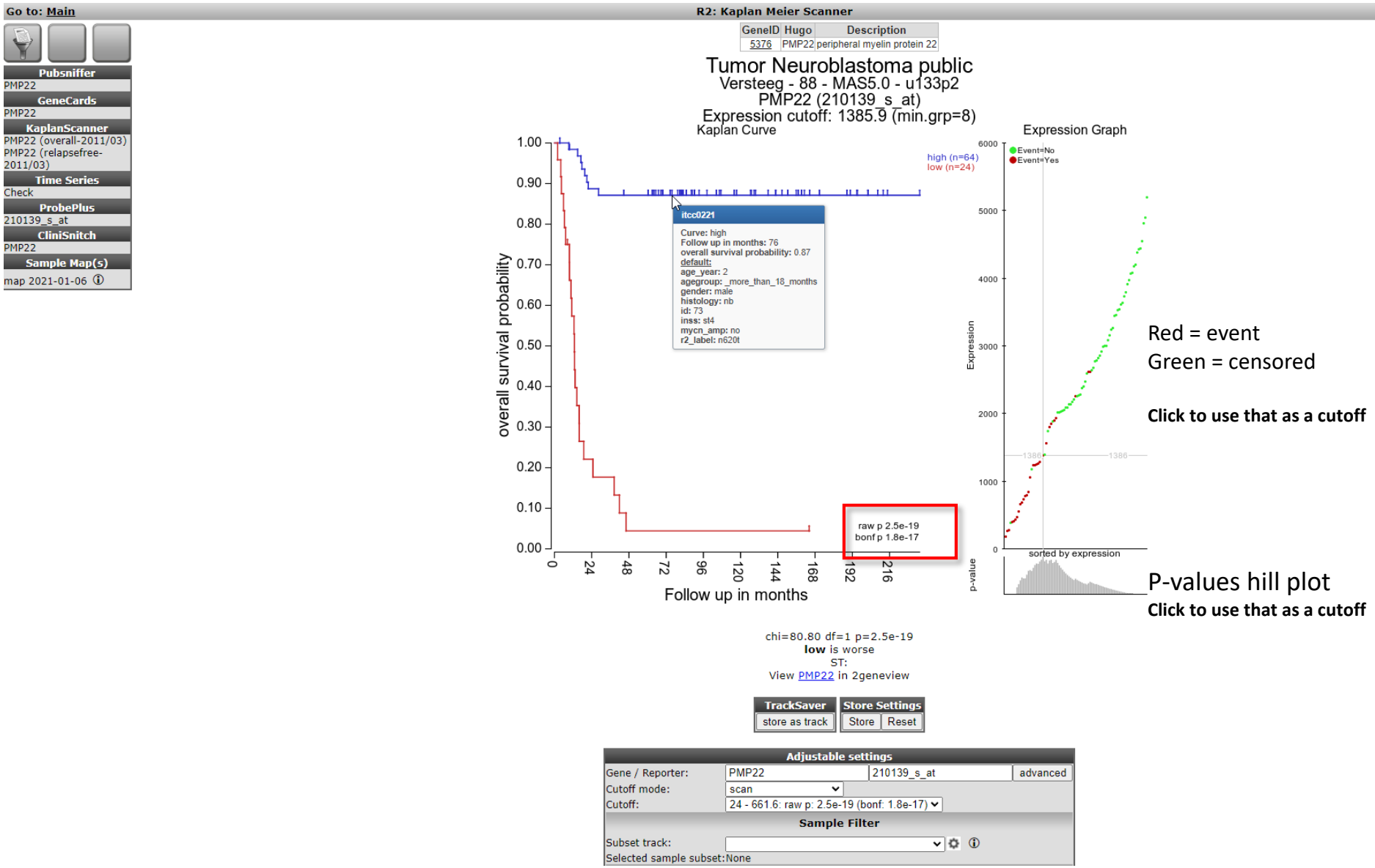
Next Reset



# KaplanScan


Go to: Main
R2: Kaplan-Meier
Online Tutorial
Main
Time series
Survival (Kaplan-Meier/Cox)
Sample maps (UMAP/tSNE)
Small Tools
DataGrabber
Genome Browser
ChIP data
TAR literature
Change Data Scope
User Options
Help
Contact / About R2
Kaplan-Meier analysis using a data set
Data set:
Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2
Separate by:
a single gene
a categorical track
a numerical track
a single gene
multiple genes
Hazard_ratio for a gene
Hazard_ratio for multiple genes
in multiple data sets
custom data
Most common survival analyses in R2
R2: Kaplan-Meier Scanner
Kaplan Meier Scanner:
Gene / Reporter:
PMP22
210139_s_at
advanced
Cutoff mode:
scan
median
average
first_quartile
last_quartile
first_vs_last_quartile
curtain
Subset track:
Selected sample subset:
Type of Survival:
Minimal group size:
6
Next

# KaplanScan



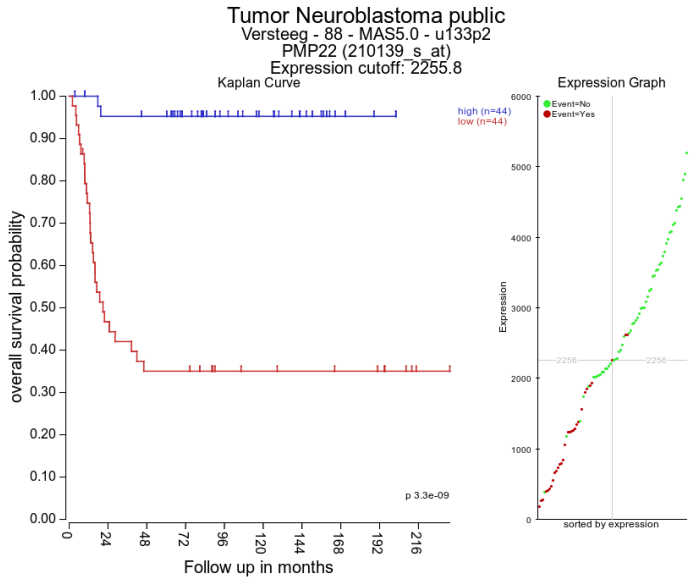

Red = event
Green = censored

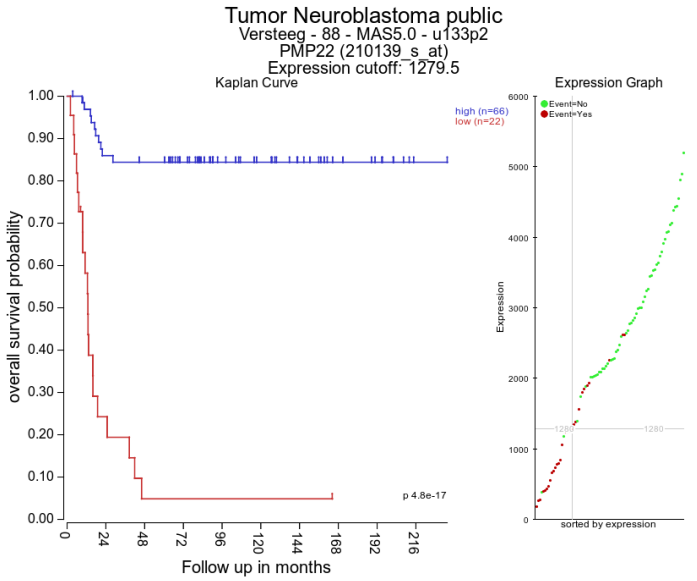
**Click to use that as a cutoff**

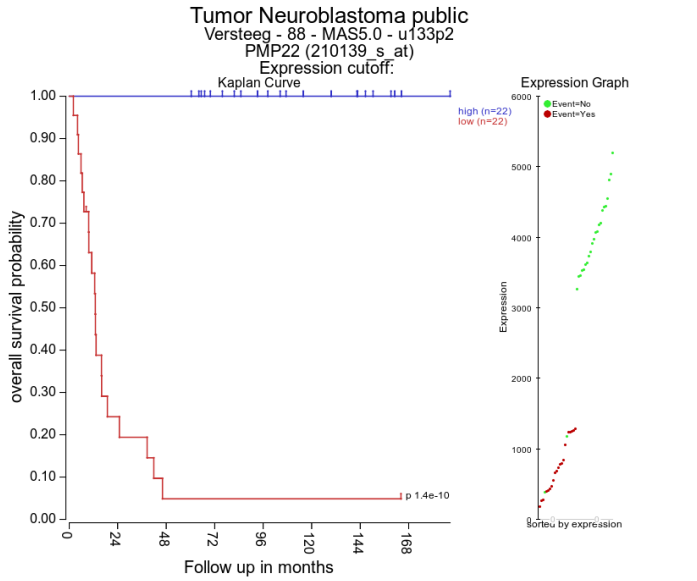
P-values hill plot

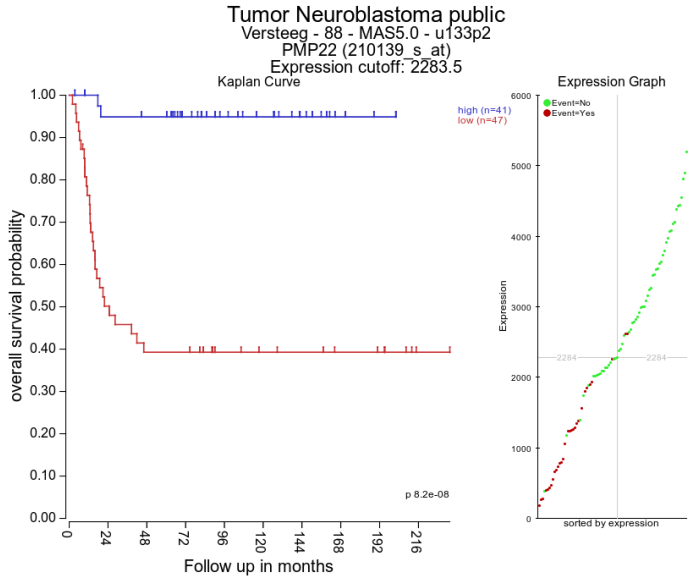
**Click to use that as a cutoff**

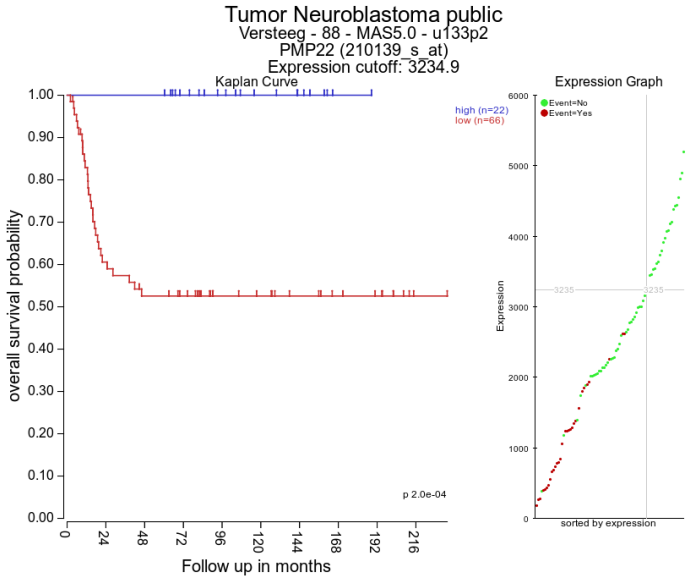
# Other separation methods

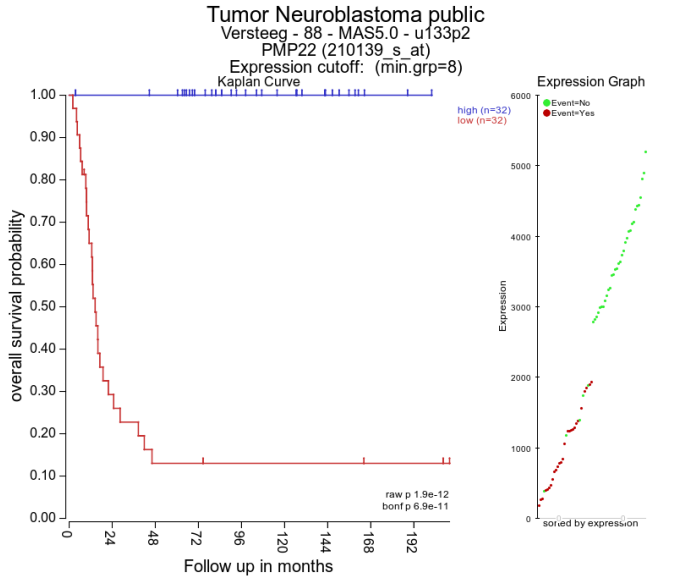


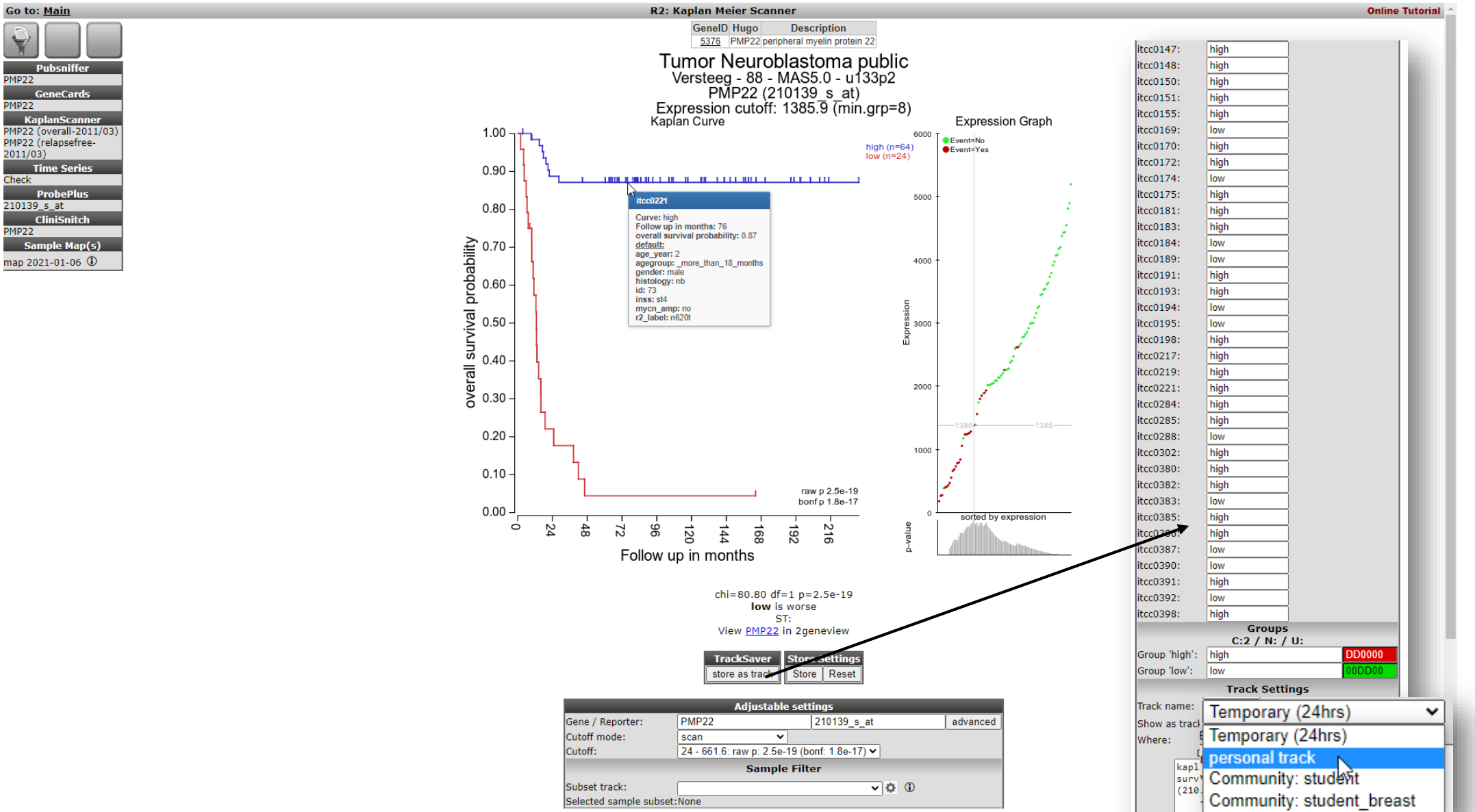




r2-support@amsterdamumc.nl

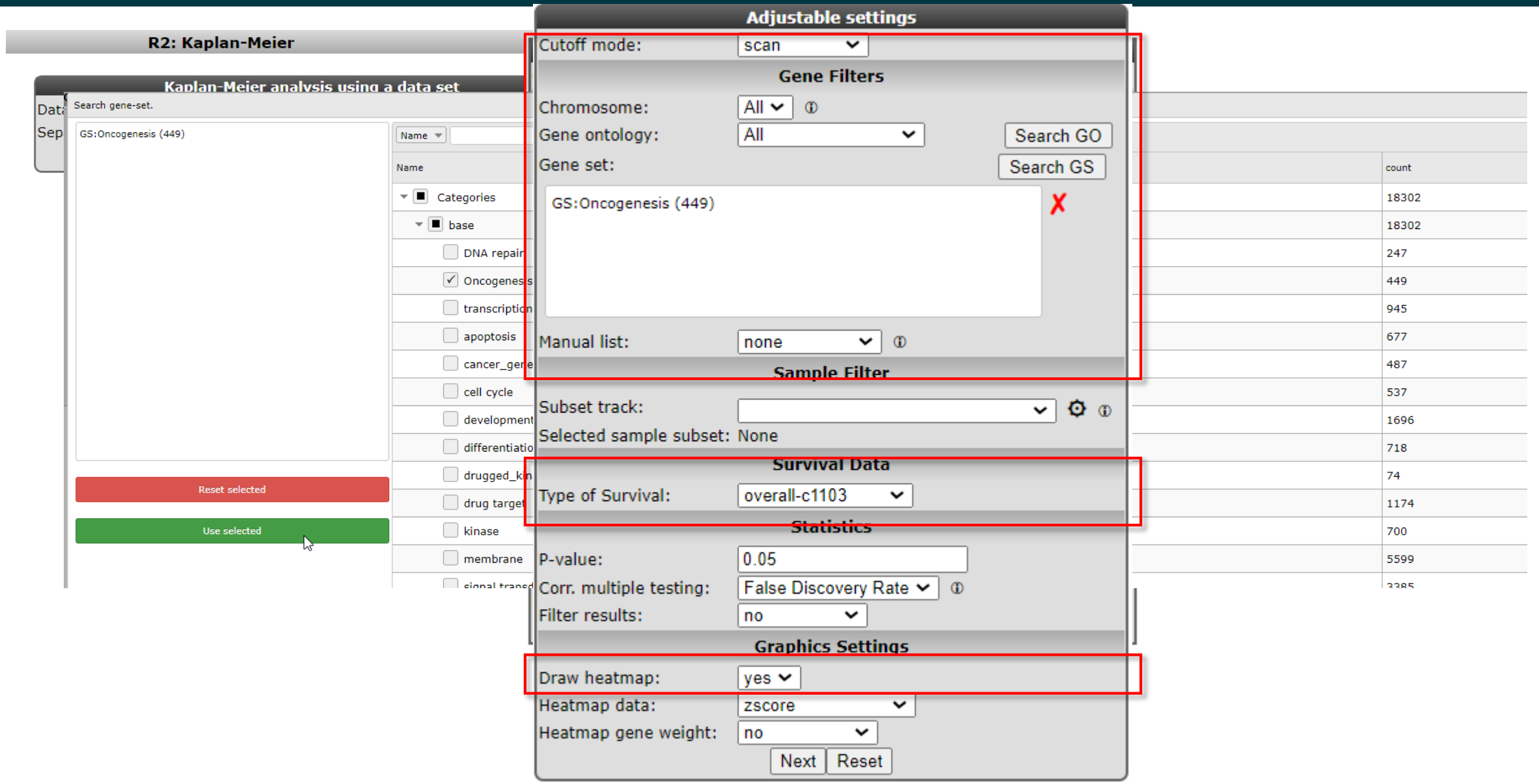
Scan: 1.8 x $10^{-17}$

# KaplanScan




r2-support@amsterdamumc.nl

# KaplanScanner multiple genes



# KaplanScanner multiple genes

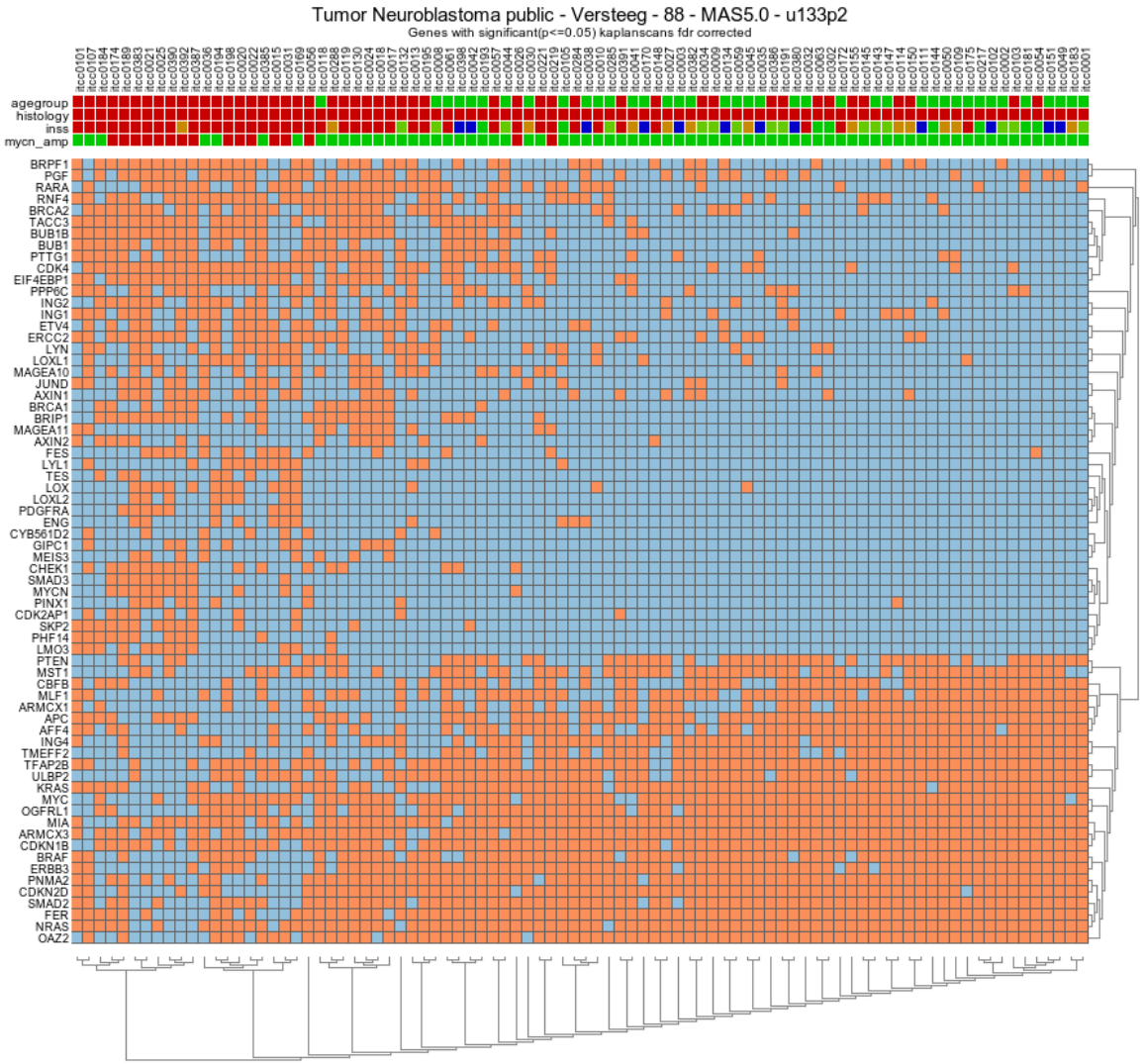
Go to: Main | R2: Kaplan Meier Scanner Pro | Online Tutorial

Kaplan start | Adapt settings

## Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2 public

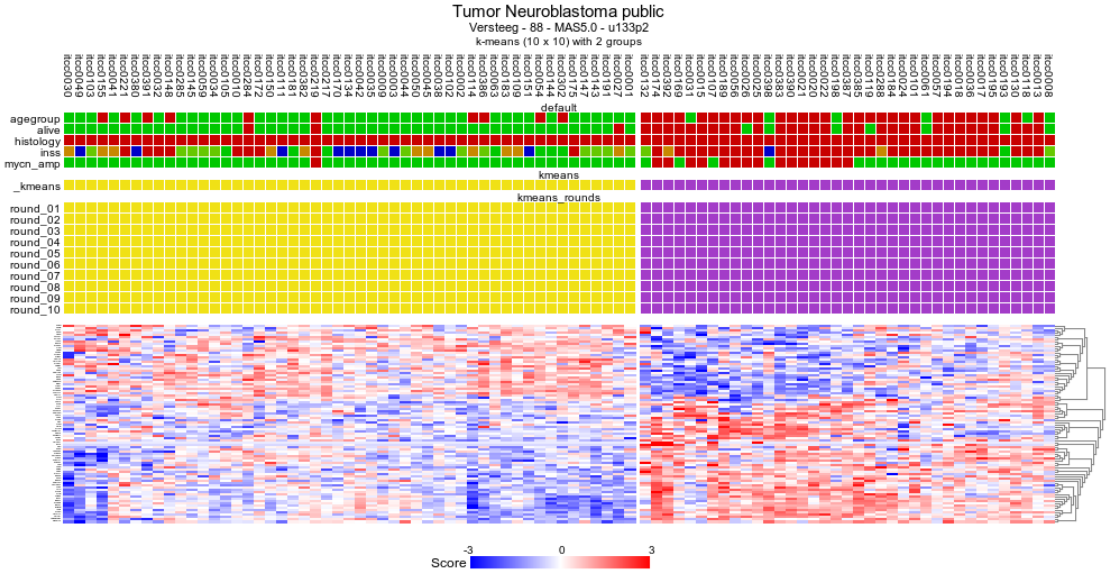
88 samples, gene set: Oncogenesis
Results corrected for multiple gene testing by fdr
68 out of 315 genes found where p<=0.05

| # | hugo | probeset | pval | more |
|---|---|---|---|---|
| 1 | OAZ2 | 201365_at | 2.74e-08 | 14 -- 283.5 -- low -- raw p=8.69e-11 |
| 2 | LMO3 | 204424_s_at | 4.65e-08 | 76 -- 68.1 -- high -- raw p=2.95e-10 |
| 3 | EIF4EBP1 | 221539_at | 5.25e-08 | 56 -- 284.3 -- high -- raw p=5.00e-10 |
| 4 | SKP2 | 203625_x_at | 7.06e-08 | 74 -- 511.2 -- high -- raw p=8.96e-10 |
| 5 | PHF14 | 228095_at | 7.29e-07 | 77 -- 866.7 -- high -- raw p=1.16e-08 |
| 6 | TMEFF2 | 223557_s_at | 4.63e-06 | 37 -- 457.2 -- low -- raw p=8.83e-08 |
| 7 | PINX1 | 223907_s_at | 8.11e-06 | 79 -- 115.3 -- high -- raw p=1.80e-07 |
| 8 | ERBB3 | 226213_at | 8.44e-06 | 19 -- 3.3 -- low -- raw p=2.41e-07 |
| 9 | CDK4 | 202246_s_at | 9.20e-06 | 47 -- 723.3 -- high -- raw p=2.34e-07 |
| 10 | FER | 227579_at | 1.31e-05 | 11 -- 214.6 -- low -- raw p=4.58e-07 |
| 11 | CDKN1B | 209112_at | 1.39e-05 | 9 -- 917.4 -- low -- raw p=4.42e-07 |
| 12 | MYCN | 209757_s_at | 2.03e-05 | 76 -- 3475.9 -- high -- raw p=7.72e-07 |
| 13 | MEIS3 | 228327_x_at | 6.23e-05 | 80 -- 234 -- high -- raw p=2.57e-06 |
| 14 | ULBP2 | 238542_at | 7.27e-05 | 22 -- 22.6 -- low -- raw p=3.23e-06 |
| 15 | SMAD3 | 218284_at | 1.03e-04 | 79 -- 101.6 -- high -- raw p=4.92e-06 |
| 16 | ETV4 | 1554576_a_at | 1.19e-04 | 63 -- 16.7 -- high -- raw p=6.07e-06 |
| 17 | ING4 | 48825_at | 1.60e-04 | 27 -- 149.4 -- low -- raw p=8.66e-06 |
| 18 | BUB1 | 209642_at | 1.70e-04 | 63 -- 133.5 -- high -- raw p=9.73e-06 |
| 19 | APC | 203525_s_at | 1.96e-04 | 26 -- 318.7 -- low -- raw p=1.18e-05 |
| 20 | AXIN2 | 222696_at | 2.02e-04 | 72 -- 158.8 -- high -- raw p=1.28e-05 |
| 21 | OGFRL1 | 226810_at | 2.87e-04 | 18 -- 274.7 -- low -- raw p=1.91e-05 |
| 22 | BUB1B | 203755_at | 3.09e-04 | 54 -- 294 -- high -- raw p=2.16e-05 |
| 23 | TACC3 | 218308_at | 4.81e-04 | 56 -- 133 -- high -- raw p=3.51e-05 |
| 24 | MAGEA10 | 210295_at | 6.26e-04 | 65 -- 54 -- high -- raw p=4.77e-05 |
| 25 | ARMCX3 | 222444_at | 7.39e-04 | 9 -- 374.6 -- low -- raw p=5.87e-05 |
| 26 | CHEK1 | 205394_at | 9.15e-04 | 71 -- 142.3 -- high -- raw p=7.84e-05 |
| 27 | CYB561D2 | 209665_at | 9.38e-04 | 80 -- 101.2 -- high -- raw p=7.74e-05 |
| 28 | MIA | 206560_s_at | 9.55e-04 | 8 -- 14.7 -- low -- raw p=8.49e-05 |
| 29 | PDGFRA | 203131_at | 9.77e-04 | 79 -- 920.6 -- high -- raw p=8.99e-05 |
| 30 | SMAD2 | 203075_at | 1.13e-03 | 13 -- 409.9 -- low -- raw p=1.11e-04 |
| 31 | BRAF | 226391_at | 1.14e-03 | 20 -- 240.7 -- low -- raw p=1.15e-04 |
| 32 | BRIP1 | 235609_at | 1.15e-03 | 64 -- 327.5 -- high -- raw p=1.10e-04 |
| 33 | RNF4 | 212696_s_at | 1.19e-03 | 48 -- 253.3 -- high -- raw p=1.28e-04 |
| 34 | CDKN2D | 210240_s_at | 1.19e-03 | 16 -- 55.2 -- low -- raw p=1.24e-04 |
| 35 | ING1 | 208415_x_at | 1.19e-03 | 60 -- 154.9 -- high -- raw p=1.33e-04 |
| 36 | ERCC2 | 235399_at | 1.55e-03 | 56 -- 41.3 -- high -- raw p=1.78e-04 |
| 37 | CDK2AP1 | 201938_at | 2.09e-03 | 78 -- 2456.8 -- high -- raw p=2.46e-04 |
| 38 | BRPF1 | 204481_at | 2.77e-03 | 49 -- 107.9 -- high -- raw p=3.34e-04 |
| 39 | AXIN1 | 212849_at | 3.34e-03 | 64 -- 56.7 -- high -- raw p=4.13e-04 |
| 40 | BRCA1 | 204531_s_at | 3.54e-03 | 73 -- 382.4 -- high -- raw p=4.49e-04 |
| 41 | JUND | 203752_s_at | 3.79e-03 | 68 -- 1944.3 -- high -- raw p=4.93e-04 |
| 42 | FES | 205418_at | 3.88e-03 | 73 -- 61.9 -- high -- raw p=5.17e-04 |
| 43 | TFAP2B | 214451_at | 4.51e-03 | 16 -- 683.8 -- low -- raw p=6.16e-04 |
| 44 | GIPC1 | 207525_s_at | 9.63e-03 | 78 -- 121.3 -- high -- raw p=1.38e-03 |
| 45 | LYN | 210754_s_at | 9.71e-03 | 62 -- 129.4 -- high -- raw p=1.36e-03 |
| 46 | LOXL1 | 203570_at | 0.010 | 65 -- 129.4 -- high -- raw p=1.56e-03 |



# Add-on analyses on the list of genes

- Correlate the results with dataset
- use genes in k-means with dataset
- Save current selection as TXT file
- Gene set analysis
- Known interactions
- Store result as custom gene set
- Chromosome Map
- Gene Ontology Analysis
- DataAdder




high, low,

| GoPath | R# | # | p_value_ | Goid-Desc | GeneSymbols |
|---|---|---|---|---|---|
| 6978 | 18 | 5 | 3.4e-42 | 1: DNA damage response, signal transduction by p53 class mediator resulting in transcription of p21 class mediator (7.9) | CHEK2, MUC1, BRCA1, BRCA2, TP53 |
| 42772 | 19 | 5 | 5.8e-40 | 1: DNA damage response, signal transduction resulting in transcription (6.8) | CHEK2, MUC1, BRCA1, BRCA2, TP53 |
| 50681 | 43 | 7 | 5.0e-36 | 2: androgen receptor binding (7.8) | DAXX, DDX5, FOXP1, SMARCA4, BRCA1, NCOA1, CCNE1 |
| 30521 | 63 | 8 | 2.9e-32 | 1: androgen receptor signaling pathway (6.10) | DAXX, DDX5, FOXP1, SMARCA4, BRCA1, NCOA1, CCNE1, MED12 |
| 217 | 24 | 5 | 1.3e-31 | 2: DNA secondary structure binding (6.6) | MSH6, MEN1, BLM, WRN, RECQL4 |
| 32392 | 82 | 9 | 2.0e-31 | 1: DNA geometric change (7.7) | DDB2, ERCC2, HMGA1, HNRNPA2B1, BLM, TP53, WRN, BRIP1, RECQL4 |
| 9314 | 396 | 20 | 1.4e-30 | 1: response to radiation (4.4) | CHEK2, CREB1, DDB2, DNMT3A, ERCC2, FANCD2, GNAQ, MSH6, KRAS, MEN1, MYC, NTRK1, PIK3R1, CCND1, BLM, BRCA1, BRAF, BRCA2, TP53, WRN |
| 732 | 26 | 5 | 3.7e-29 | 1: strand displacement (6.9) | BLM, BRCA1, BRCA2, WRN, BRIP1 |
| 10332 | 53 | 7 | 4.8e-29 | 1: response to gamma radiation (6.6) | CHEK2, FANCD2, MEN1, MYC, BRCA2, TP53, WRN |
| 44427 | 784 | 27 | 4.7e-26 | 3: chromosomal part (4.9) | CDK4, CHEK2, CREB1, DAXX, DDX6, DNMT3A, MSH6, HMGA1, HNRNPA2B1, APC, MEN1, MUC1, MYC, MYCN, CCNB1IP1, RARA, BLM, SMARCA4, BRCA1, BRCA2, TCF3, BUB1B, TP53, TPR, WRN, NCOA1, RECQL4 |
| 6259 | 897 | 29 | 7.3e-26 | 1: DNA metabolic process (5.7) | CHEK2, DDB2, DNMT3A, ERCC2, FANCA, FANCD2, FANCF, MSH6, HMGA1, HNRNPA2B1, MEN1, MYC, NFIB, NONO, PDGFRA, CCNB1IP1, BLM, MAP2K4, SET, BRCA1, BRCA2, TCF3, TFRC, TP53, WHSC1, WRN, BRIP1, CCNE1, RECQL4 |
| 5694 | 904 | 29 | 1.3e-25 | 3: chromosome (5.8) | CDK4, CHEK2, CREB1, DAXX, DDX6, DNMT3A, FANCD2, MSH6, HMGA1, HNRNPA2B1, APC, MEN1, MUC1, MYC, MYCN, CCNB1IP1, RARA, BLM, SMARCA4, BRCA1, BRCA2, TCF3, BUB1B, TP53, TPR, WHSC1, WRN, NCOA1, RECQL4 |
| 10212 | 148 | 11 | 1.4e-25 | 1: response to ionizing radiation (5.5) | CHEK2, DNMT3A, FANCD2, MEN1, MYC, CCND1, BLM, BRCA1, BRCA2, TP53, WRN |
| 30518 | 128 | 10 | 1.9e-24 | 1: intracellular steroid hormone receptor signaling pathway (5.9) | DAXX, DDX5, FOXP1, SMARCA4, BRCA1, RUNX1, NCOA1, CBFB, CCNE1, MED12 |
| 9411 | 130 | 10 | 4.9e-24 | 1: response to UV (6.6) | DDB2, ERCC2, MSH6, MEN1, MYC, PIK3R1, CCND1, BRCA2, TP53, WRN |
| 3684 | 65 | 7 | 1.7e-23 | 2: damaged DNA binding (6.6) | DDB2, MSH6, BLM, BRCA1, TP53, WRN, RECQL4 |
| 9416 | 261 | 14 | 9.3e-23 | 1: response to light stimulus (5.5) | CREB1, DDB2, ERCC2, GNAQ, MSH6, KRAS, MEN1, MYC, PIK3R1, CCND1, BRAF, BRCA2, TP53, WRN |
| 4003 | 34 | 5 | 3.1e-22 | 2: ATP-dependent DNA helicase activity (5.11) | ERCC2, BLM, WRN, BRIP1, RECQL4 |
| 35690 | 306 | 15 | 5.3e-22 | 1: cellular response to drug (5.5) | CDK4, CHEK2, DAXX, DNMT3A, GNAQ, MYB, MYC, NTRK1, PTEN, BLM, BRCA1, BRAF, TFRC, TP53, NCOA1 |
| 32870 | 619 | 22 | 5.5e-22 | 1: cellular response to hormone stimulus (5.6) | CDK4, NCOA2, CHEK2, CREB1, DAXX, DDX5, FOXP1, GNAQ, APC, MEN1, MSN, PIK3R1, PTEN, RARA, SMARCA4, BRCA1, BRIP1, RUNX1, NCOA1, CBFB, CCNE1, MED12 |
| 70035 | 91 | 8 | 6.5e-22 | 2: purine NTP-dependent helicase activity (9.9) | DDX5, DDX6, DDX10, ERCC2, BLM, WRN, BRIP1, RECQL4 |
| 45935 | 1497 | 36 | 1.1e-21 | 1: positive regulation of nucleobase-containing compound metabolic process (5.7) | NCOA2, CHEK2, CREB1, DDX5, ERCC2, ETV4, CAMTA1, MSH6, HMGA1, HNRNPA2B1, LYL1, MEN1, MUC1, MYB, MYC, MYCN, NFIB, PAX5, PDGFRA, PIK3R1, RARA, BLM, MAP2K4, SMARCA4, BRCA1, BRCA2, TCF3, TFRC, TP53, WHSC1, WRN, RUNX1, NCOA1, CBFB, CCNE1, MED12 |
| 10604 | 2649 | 50 | 1.2e-21 | 1: positive regulation of macromolecule metabolic process (4.6) | CDK4, NCOA2, CHEK2, CREB1, DAXX, DDX5, EPS15, ERCC2, ETV4, CAMTA1, MSH6, HMGA1, HNRNPA2B1, APC, IL6ST, KRAS, LYL1, MEN1, MSN, MUC1, MYB, MYC, MYCN, NFIB, NTRK1, PAX5, PDGFRA, PIK3R1, FEV, PTEN, RARA, CCND1, BLM, MAP2K4, SMARCA4, BRCA1, BRAF, BRCA2, TCF3, TFRC, TP53, WHSC1, WRN, ASPSCR1, RUNX1, NCOA1, CBFB, FUBP1, CCNE1, MED12 |
| 35257 | 144 | 10 | 1.7e-21 | 2: nuclear hormone receptor binding (5.6) | NCOA2, DAXX, DDX5, FOXP1, HMGA1, SMARCA4, BRCA1, NCOA1, CCNE1, MED12 |
| 9725 | 857 | 26 | 2.9e-21 | 1: response to hormone (4.5) | CDK4, NCOA2, CHEK2, CREB1, DAXX, DDX5, DNMT3A, FOXP1, GNAQ, APC, KRAS, MEN1, MSN, PIK3R1, PTEN, RARA, CCND1, SMARCA4, BRCA1, BRAF, BRIP1, RUNX1, NCOA1, CBFB, CCNE1, MED12 |
| 43401 | 175 | 11 | 3.0e-21 | 1: steroid hormone mediated signaling pathway (5.8) | DAXX, DDX5, FOXP1, RARA, SMARCA4, BRCA1, RUNX1, NCOA1, CBFB, CCNE1, MED12 |
| 32508 | 72 | 7 | 4.9e-21 | 1: DNA duplex unwinding (8.8) | DDB2, ERCC2, HMGA1, BLM, WRN, BRIP1, RECQL4 |
| 6974 | 754 | 24 | 7.3e-21 | 1: cellular response to DNA damage stimulus (5.5) | CHEK2, DDB2, DDX5, ERCC2, FANCA, FANCD2, FANCF, MSH6, HMGA1, APC, MEN1, MUC1, MYC, NONO, PIK3R1, CCND1, BLM, BRCA1, BRCA2, TP53, WHSC1, WRN, BRIP1, RECQL4 |
| 9893 | 2818 | 51 | 1.0e-20 | 1: positive regulation of metabolic process (3.5) | CDK4, NCOA2, CHEK2, CREB1, DAXX, DDX5, EPS15, ERCC2, ETV4, CAMTA1, MSH6, HMGA1, HNRNPA2B1, APC, IL6ST, KRAS, LYL1, MEN1, MSN, MUC1, MYB, MYC, MYCN, NFIB, NTRK1, PAX5, PDGFRA, PIK3R1, FEV, PTEN, RARA, CCND1, BLM, MAP2K4, SMARCA4, BRCA1, BRAF, BRCA2, TCF3, TFRC, TP53, WHSC1, WRN, ASPSCR1, SLC45A3, RUNX1, NCOA1, CBFB, FUBP1, CCNE1, MED12 |
| 43565 | 875 | 26 | 1.1e-20 | 2: sequence-specific DNA binding (6.6) | NCOA2, CREB1, ETV4, CAMTA1, FOXP1, HMGA1, HNRNPA2B1, HOXD11, LYL1, MUC1, MYB, MYC, MYCN, NFIB, NONO, PAX5, FEV, RARA, BLM, SMARCA4, TCF3, TP53, WHSC1, WRN, RUNX1, MED12 |
| 51276 | 1077 | 29 | 2.9e-20 | 1: chromosome organization (5.5) | DAXX, DDB2, DNMT3A, ERCC2, FANCD2, MSH6, HMGA1, HNRNPA2B1, APC, MEN1, MUC1, MYB, MYC, PAX5, CCNB1IP1, BLM, SET, SMARCA4, BRCA1, BRCA2, BUB1B, TP53, TPR, WHSC1, WRN, BRIP1, NCOA1, CCNE1, RECQL4 |

# KaplanScanner Custom Data

**R2: Kaplan-Meier**

**Kaplan-Meier analysis using a data set**

Data set: Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2

Separate by: a categorical track

Next

Kaplan-Meier analysis using custom data

**R2: KaplanScan on UserDefined Data**

**KaplanScan:**

Paste Survival information samplename; survtime(days); event(1/0/YES/NO); expression_value (tab or ; separated)

```
#Copy your data below the #H: line.
#H:samplename;survival_time;event;Gene_x
GSM1542334      7876    no      11.03011533
GSM1542335      4914    no      11.67083326
GSM1542336      7457    no      11.22562922
GSM1542337      6904    no      10.44429021
GSM1542338      7146    no      11.08308037
GSM1542339      7168    no      10.98299357
GSM1542340      488     yes     10.28262513
GSM1542341      1277    yes     10.30309576
GSM1542342      576     yes     10.03974158
GSM1542343      609     yes     11.35524096
GSM1542344      340     yes     10.27484403
GSM1542345      296     yes     9.412993173
GSM1542346      394     yes     8.68474862
GSM1542347      1408    yes     9.510170751
GSM1542348      196     yes     8.737416366
GSM1542349      2916    no      11.14140469
GSM1542350      108     no      11.44315076
GSM1542351      4013    no      12.34396294
GSM1542352      866     yes     10.85447883
GSM1542353      6872    no      10.8752886
GSM1542354      5458    no      11.21170582
GSM1542355      6013    no      11.59091447
GSM1542356      165     yes     9.62516088
GSM1542357      5759    no      11.81750335
GSM1542358      5118    no      11.43728393
GSM1542359      4693    no      11.54530237
GSM1542360      4153    no      10.97806689
```

Cutoff mode: scan

Next Reset

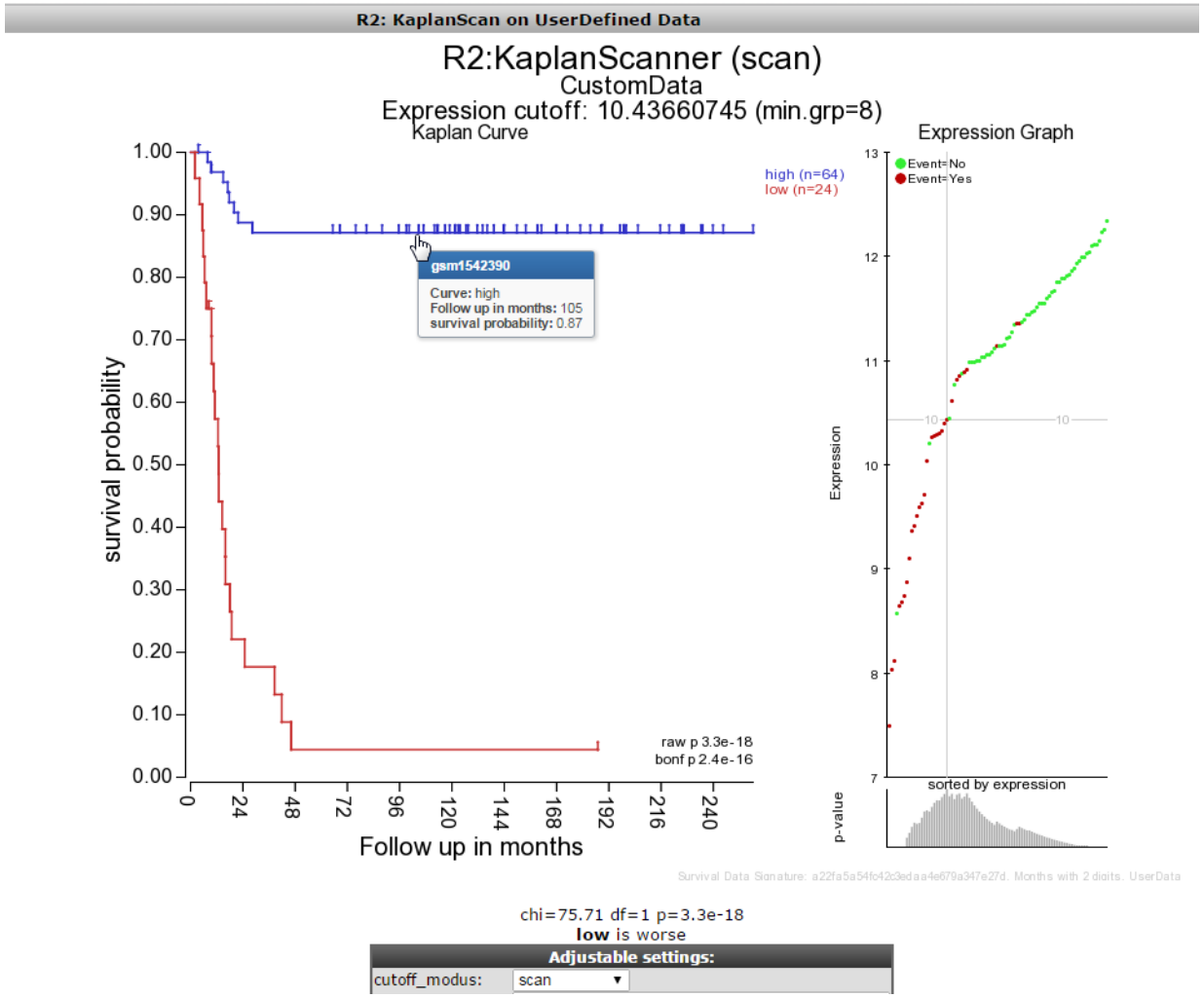
Survival time
In days

Event
yes/no/0/1

Numerical values
e.g. Expression Gene X

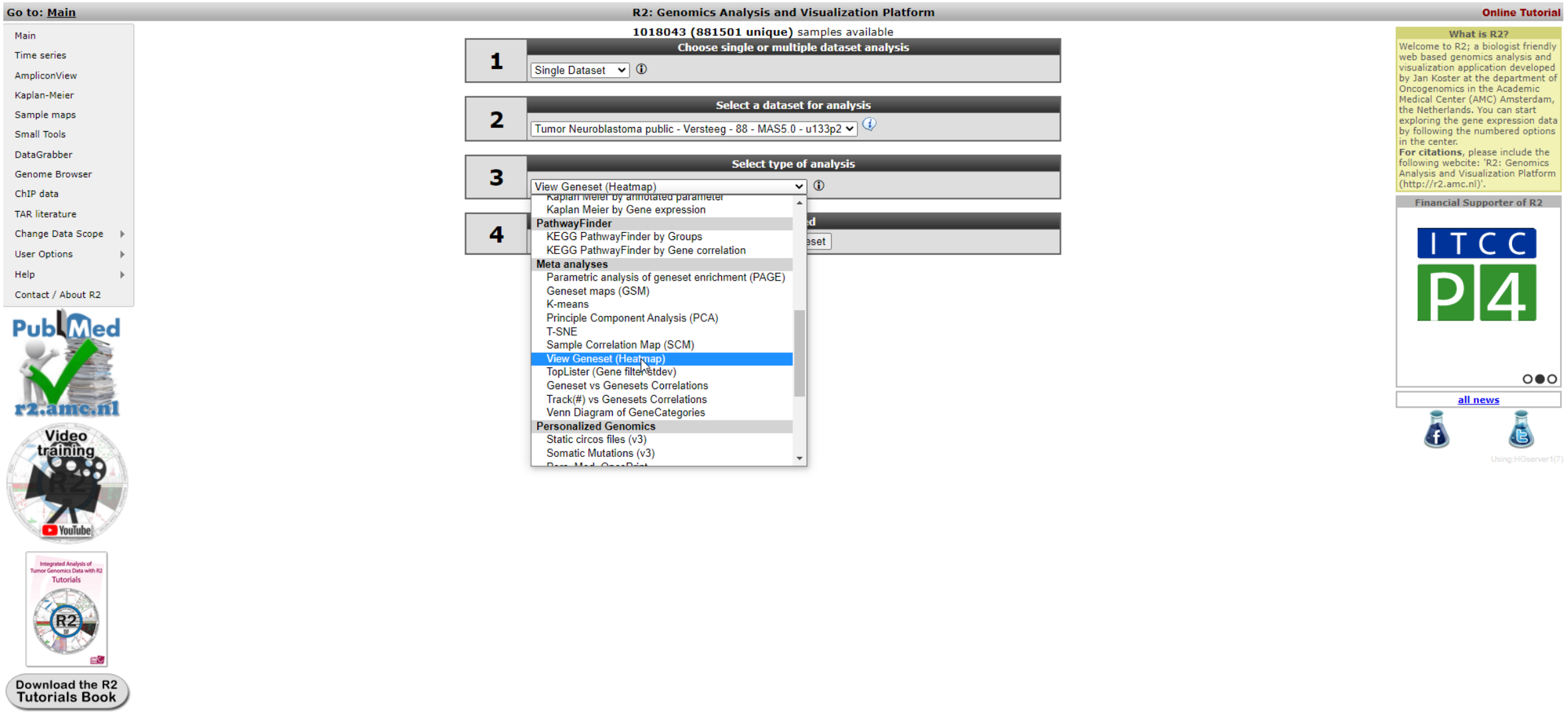
r2-support@amsterdamumc.nl

# KaplanScanner Custom Data

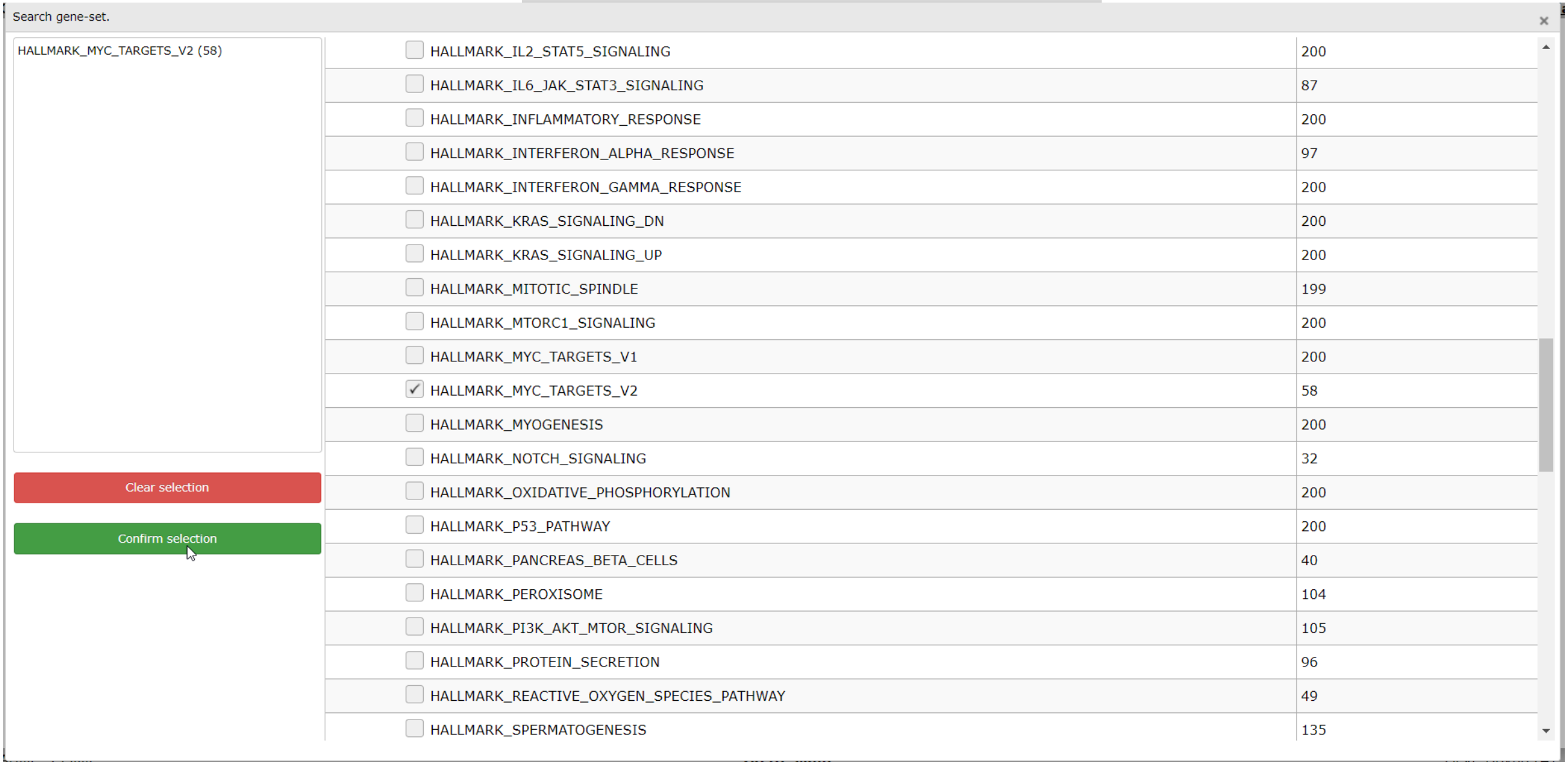


# View a GeneSet (Heatmap)

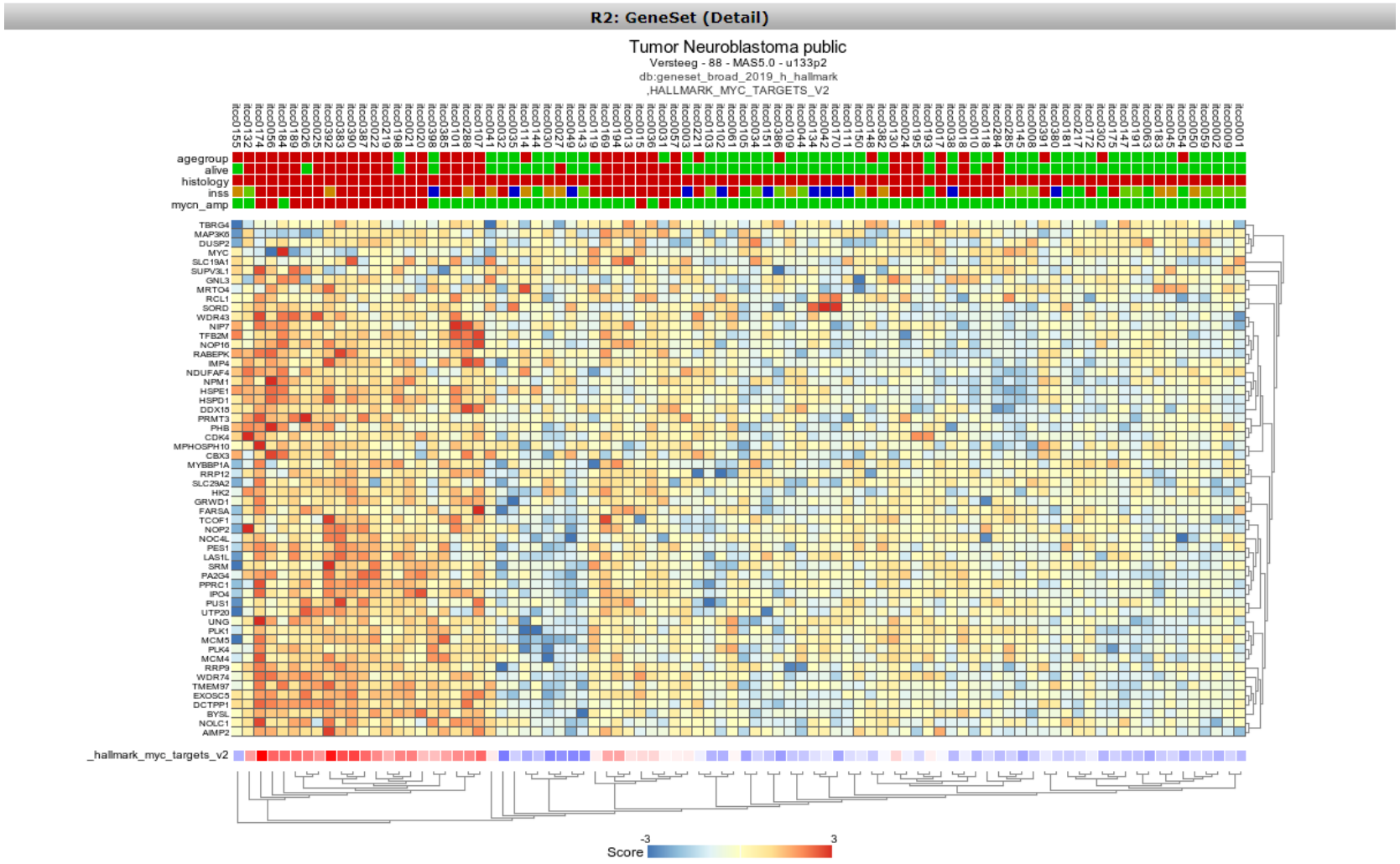
# View a GeneSet



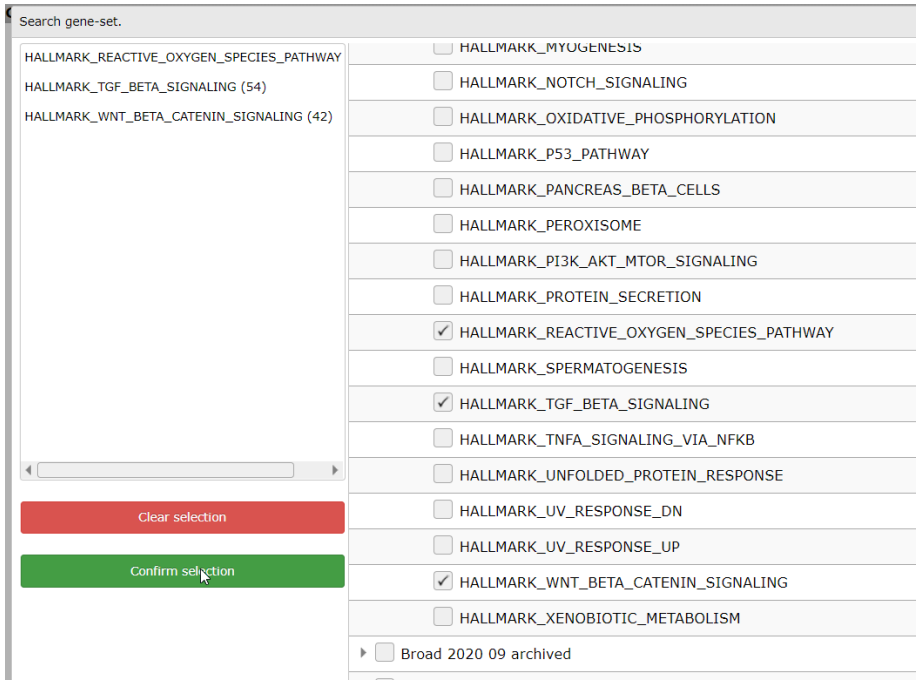
# cluster

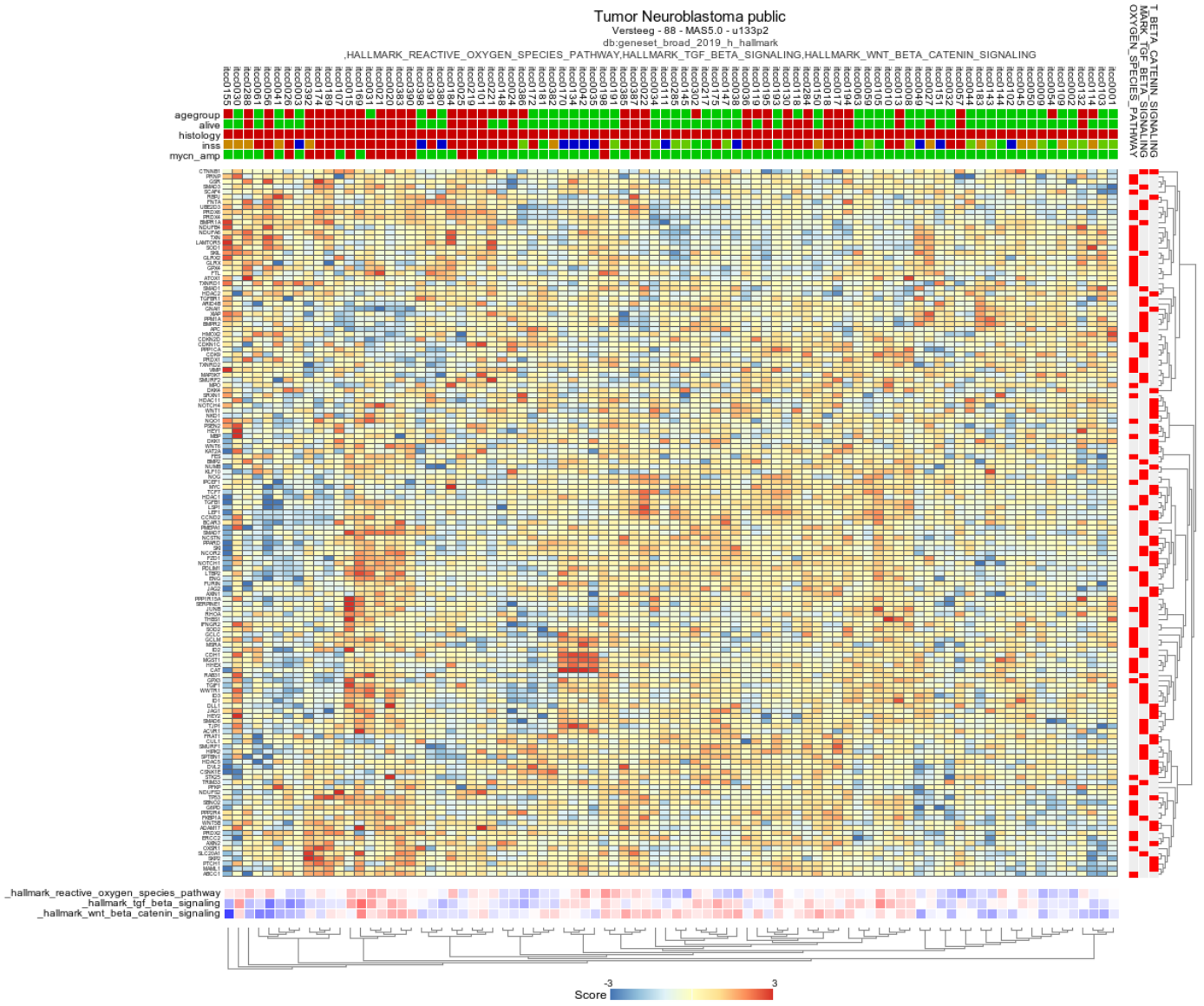

Search gene-set.
HALLMARK_MYC_TARGETS_V2 (58)
Clear selection
Confirm selection
HALLMARK_IL2_STAT5_SIGNALING 200
HALLMARK_IL6_JAK_STAT3_SIGNALING 87
HALLMARK_INFLAMMATORY_RESPONSE 200
HALLMARK_INTERFERON_ALPHA_RESPONSE 97
HALLMARK_INTERFERON_GAMMA_RESPONSE 200
HALLMARK_KRAS_SIGNALING_DN 200
HALLMARK_KRAS_SIGNALING_UP 200
HALLMARK_MITOTIC_SPINDLE 199
HALLMARK_MTORC1_SIGNALING 200
HALLMARK_MYC_TARGETS_V1 200
HALLMARK_MYC_TARGETS_V2 58
HALLMARK_MYOGENESIS 200
HALLMARK_NOTCH_SIGNALING 32
HALLMARK_OXIDATIVE_PHOSPHORYLATION 200
HALLMARK_P53_PATHWAY 200
HALLMARK_PANCREAS_BETA_CELLS 40
HALLMARK_PEROXISOME 104
HALLMARK_PI3K_AKT_MTOR_SIGNALING 105
HALLMARK_PROTEIN_SECRETION 96
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY 49
HALLMARK_SPERMATOGENESIS 135

# cluster

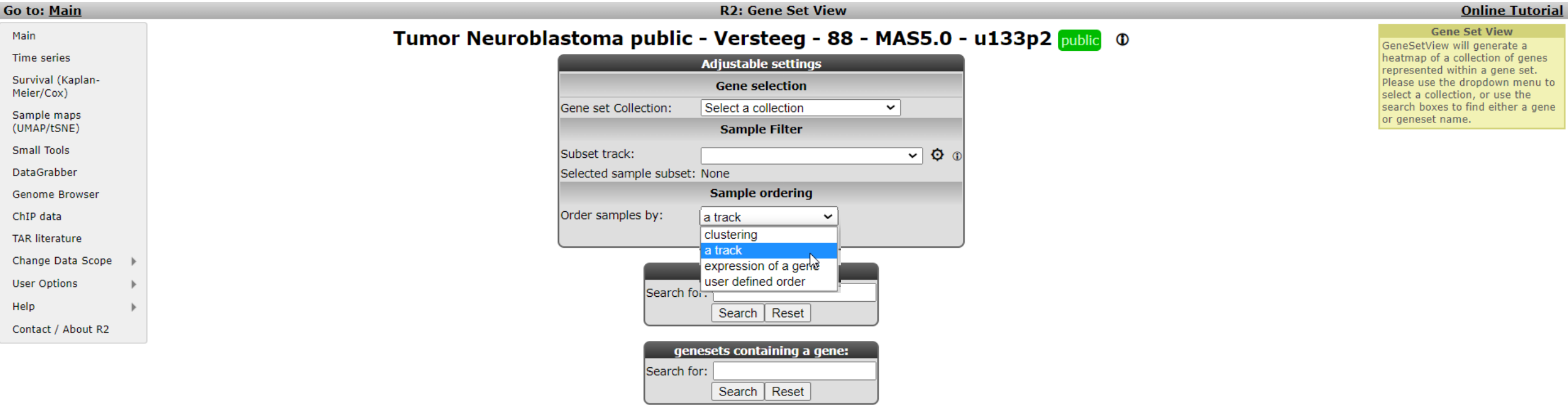

R2: GeneSet (Detail)
Tumor Neuroblastoma public
Versteeg - 88 - MAS5.0 - u133p2
db:geneset_broad_2019_h_hallmark
,HALLMARK_MYC_TARGETS_V2
agegroup
alive
histology
inss
mycn_amp
_hallmark_myc_targets_v2
-3
3
Score
distance:e; linkage:a;
Sort Order Listing
DetailView hyperlink settings
Transformation: Log2 z-score
Gene set values

# Cluster multiple gene sets

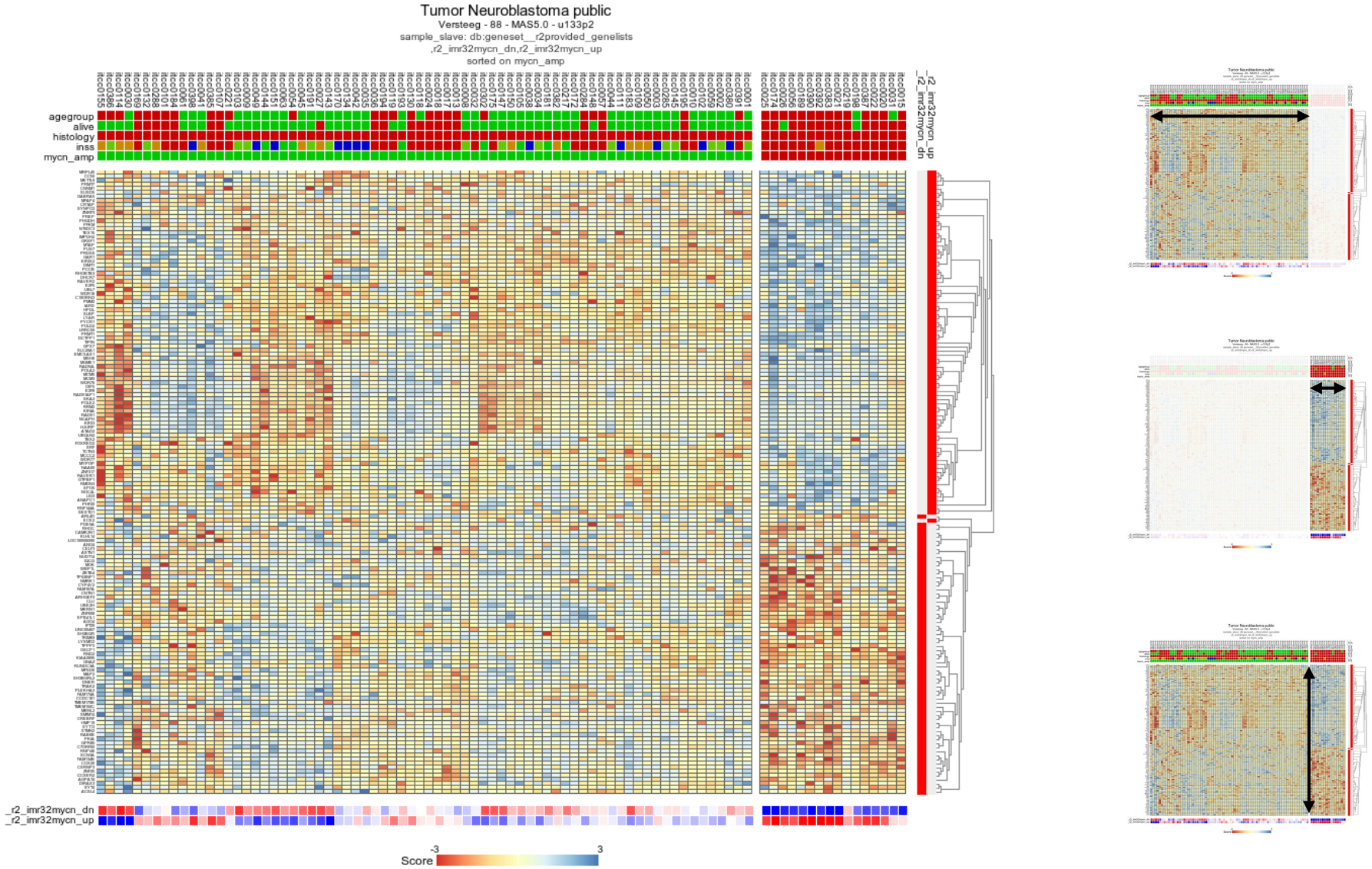


# Sub-cluster by track



# Sub-cluster by track

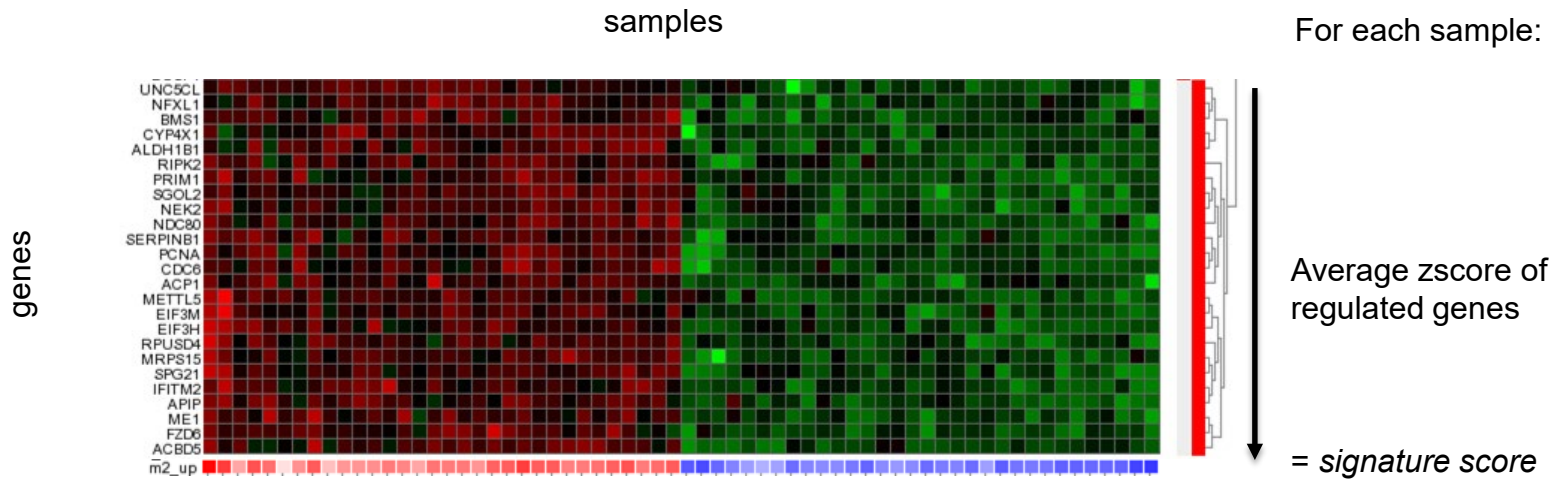


# Metagenes

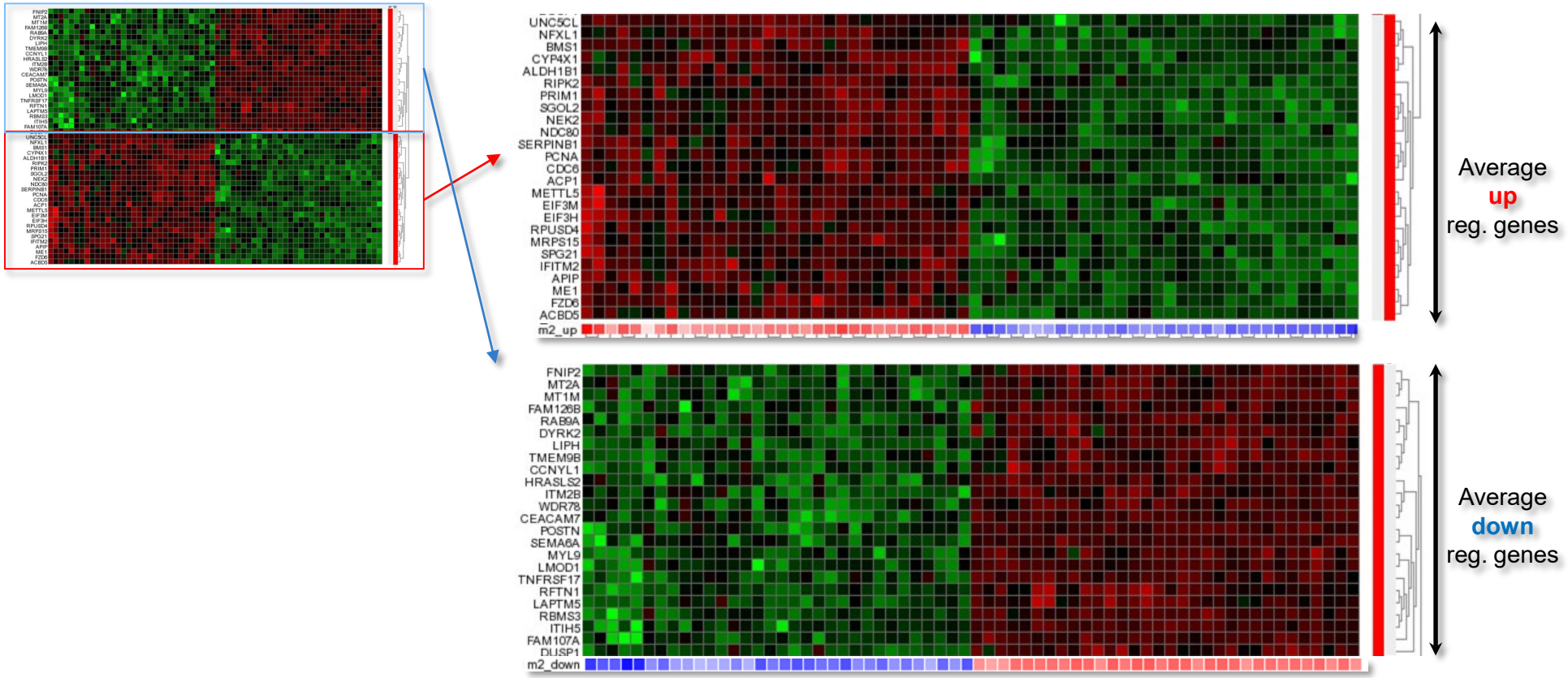
- Convert any list of genes (signature) into a single value, a *signature score*
  - a “summary expression value” of a group of genes for each sample of a dataset
  - Source can be any group of genes (e.g. genecategory)
  - Can be used as if it is one gene (=metagene) to represent a pathway activity for example

# Metagene

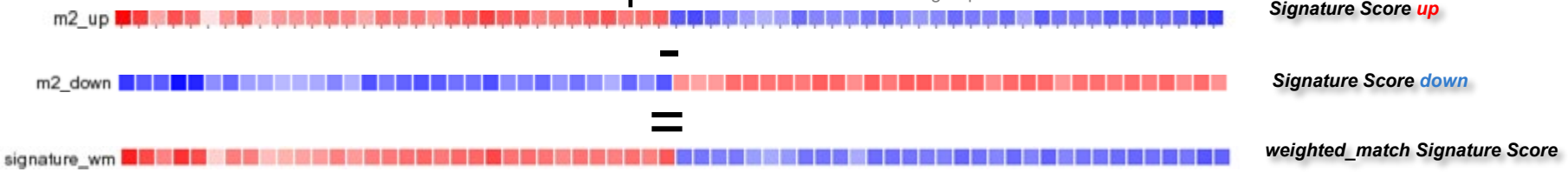
- How does it work?
  - When all genes in your signature have a common direction



# Signature scores are generated in the ‘view a geneset’ module

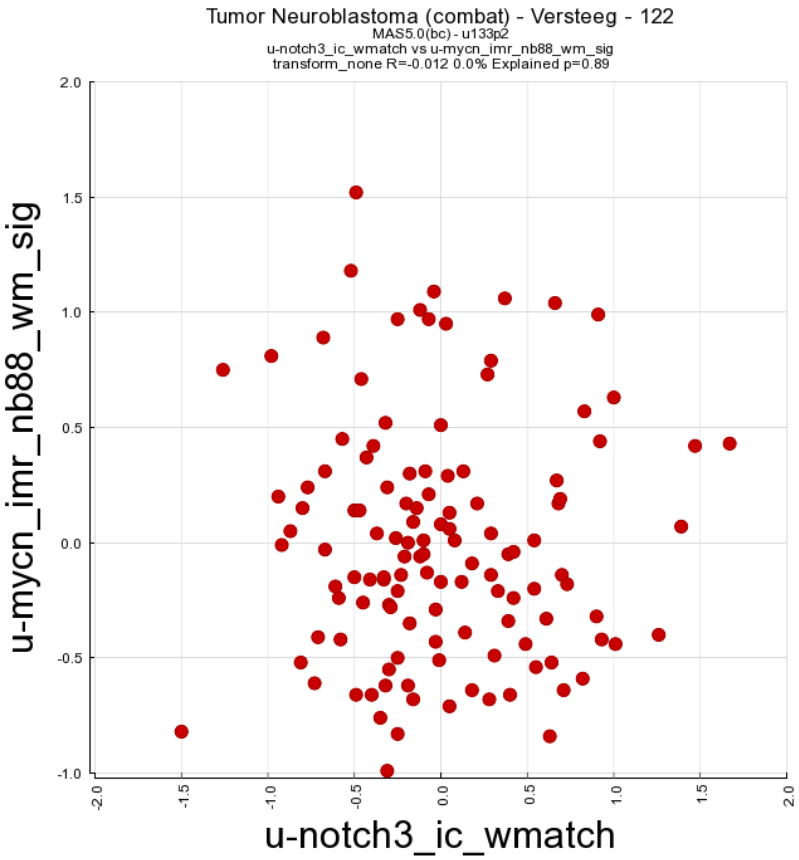


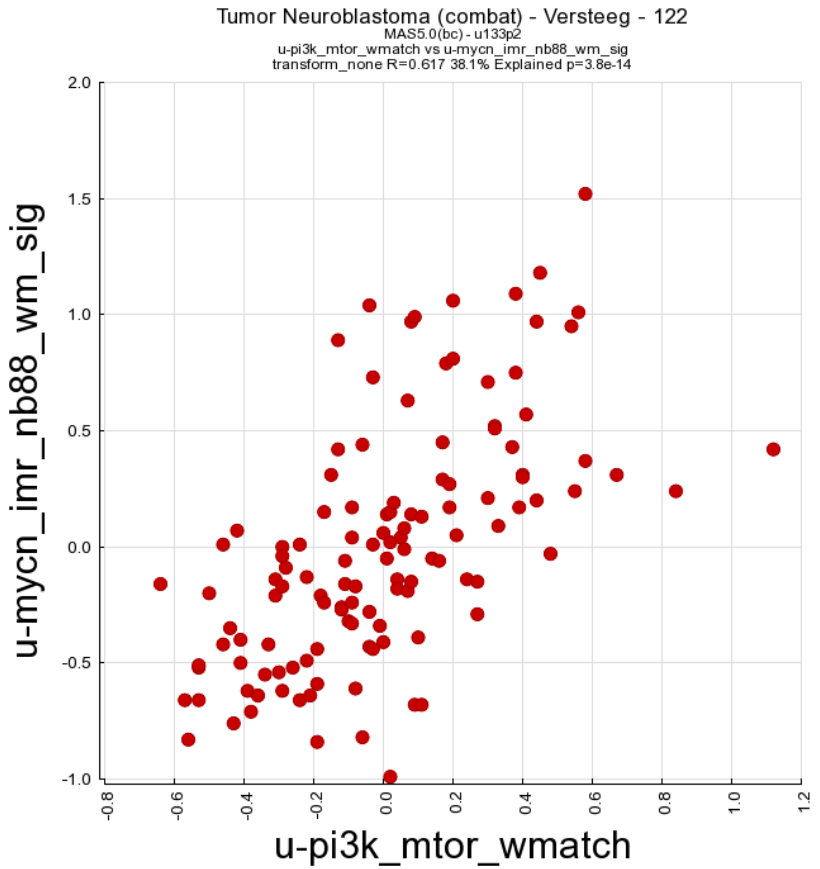

## Up - Down corrected for group size



# Store Signature Scores as Tracks in R2

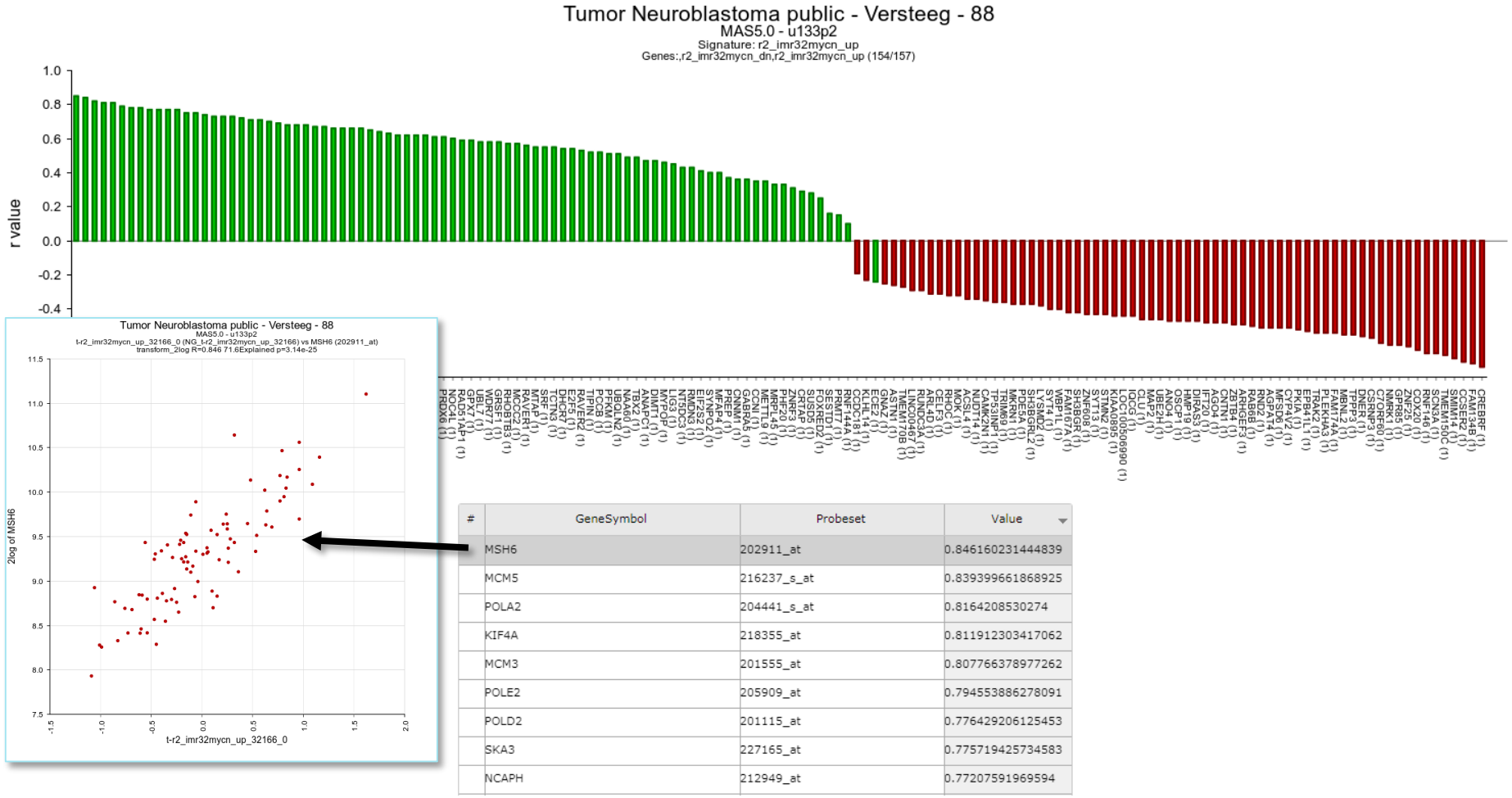
# Compare 2 signatures (via relate 2 tracks)



# Find best matching signature gene

# Signature(s) vs signature genes



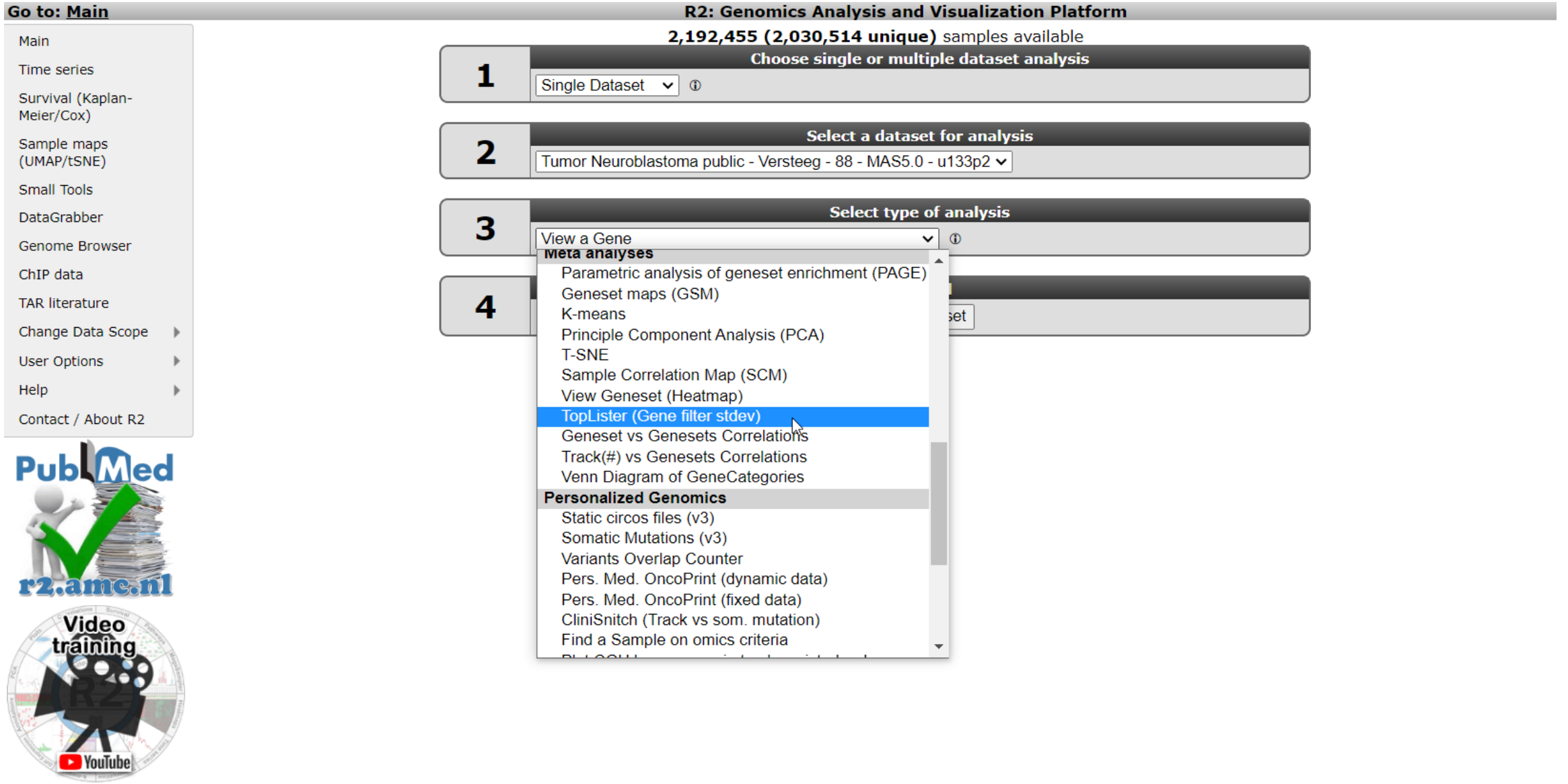

| # | GeneSymbol | Probeset | Value |
|---|---|---|---|
| | MSH6 | 202911_at | 0.846160231444839 |
| | MCM5 | 216237_s_at | 0.839399661868925 |
| | POLA2 | 204441_s_at | 0.8164208530274 |
| | KIF4A | 218355_at | 0.811912303417062 |
| | MCM3 | 201555_at | 0.807766378977262 |
| | POLE2 | 205909_at | 0.794553886278091 |
| | POLD2 | 201115_at | 0.776429206125453 |
| | SKA3 | 227165_at | 0.775719425734583 |
| | NCAPH | 212949_at | 0.77207591969594 |

# Cohort subgroups

# Cohort subgroups: Exploratory dataset analyses

- Finding / visualizing subgroups within a dataset often performed on cohorts
- Different ways to do this in R2
  - Hierarchical clustering
  - K-means clustering
  - Principle component analysis
  - t-SNE/ UMAP maps

# Toplister


Go to: Main
R2: Genomics Analysis and Visualization Platform
Main
Time series
Survival (Kaplan-Meier/Cox)
Sample maps (UMAP/tSNE)
Small Tools
DataGrabber
Genome Browser
ChIP data
TAR literature
Change Data Scope
User Options
Help
Contact / About R2
2,192,455 (2,030,514 unique) samples available
1
Choose single or multiple dataset analysis
Single Dataset
2
Select a dataset for analysis
Tumor Neuroblastoma public - Versteeg - 88 - MAS5.0 - u133p2
3
Select type of analysis
View a Gene
Meta analyses
Parametric analysis of geneset enrichment (PAGE)
Geneset maps (GSM)
K-means
Principle Component Analysis (PCA)
T-SNE
Sample Correlation Map (SCM)
View Geneset (Heatmap)
TopLister (Gene filter stdev)
Geneset vs Genesets Correlations
Track(#) vs Genesets Correlations
Venn Diagram of GeneCategories
Personalized Genomics
Static circos files (v3)
Somatic Mutations (v3)
Variants Overlap Counter
Pers. Med. OncoPrint (dynamic data)
Pers. Med. OncoPrint (fixed data)
CliniSnitch (Track vs som. mutation)
Find a Sample on omics criteria
4
PubMed
r2.amc.nl
Video training
YouTube

# Toplister

**Adjustable settings**

| | |
|---|---|
| Which set: | Standard Deviation (SD) |
| Modus: | normal |
| How many genes | 100 |
| Floor value: | 0 |
| Transformation: | Log2 |

**Sample Filter**

| | |
|---|---|
| Subset track: | |
| Selected sample subset: | None |

**Gene Filters**

| | |
|---|---|
| HugoOnce mode: | yes |
| Min. # Present calls: | 1 |
| Minimal maximum value: | |
| Minimal range size: | 0 |
| Chromosome: | All |
| Gene ontology: | All |
| Gene set: | |
| Manual list: | none |

Search GO | Search GS | Next | Reset

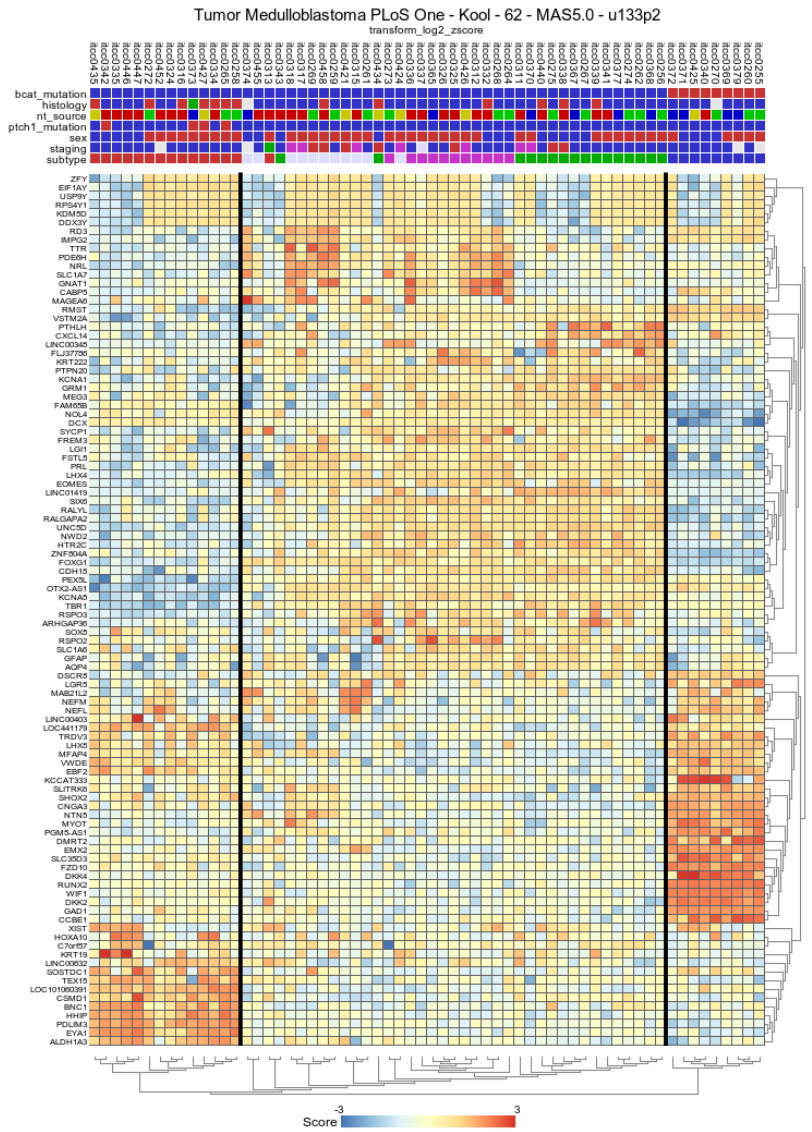
Go to: Main | R2: TopLister | Online Tutorial

**Tumor Medulloblastoma PLoS One - Kool - 62 - MAS5.0 - u133p2** public

Top 100 standard_deviation normal
, transform_log2, present>=1

Save current selection as TXT file | Heatmap(zscore) | Gene Ontology Analysis | Store result as custom gene set

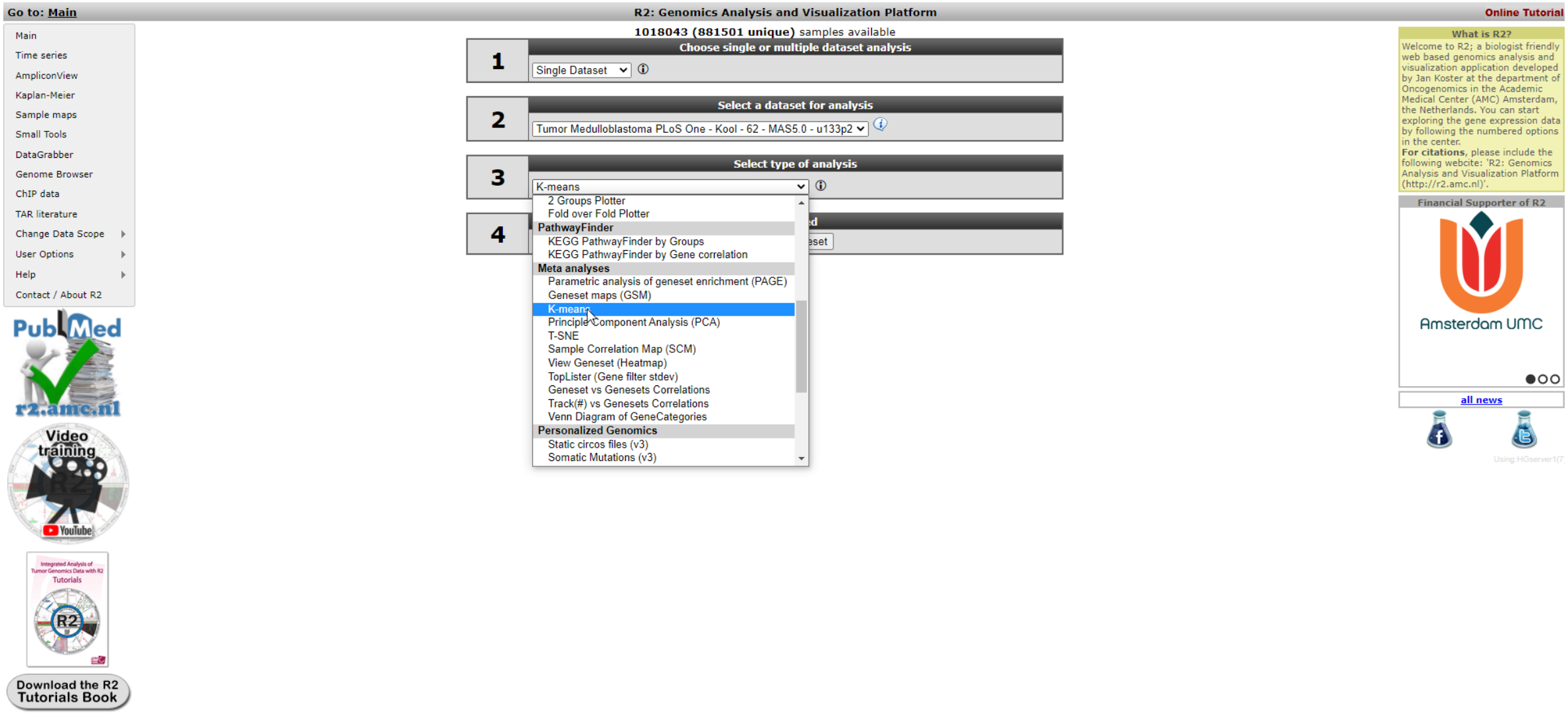
| View | Rank | Gene | Reporter | Value |
|---|---|---|---|---|
| | 1 | LINC01419 | 1559213_at | 5.2946... |
| | 2 | XIST | 224588_at | 4.3756... |
| | 3 | FOXG1 | 206018_at | 4.3274... |
| | 4 | RPS4Y1 | 201909_at | 4.1058... |
| | 5 | RALGAPA2 | 234314_at | 4.0906... |
| | 6 | RALYL | 213967_at | 3.9871... |
| | 7 | SHOX2 | 210135_s_at | 3.8976... |
| | 8 | LHX8 | 1569469_a_at | 3.8847... |
| | 9 | GAD1 | 205278_at | 3.7782... |
| | 10 | DDX3Y | 205000_at | 3.7024... |
| | 11 | WIF1 | 204712_at | 3.6922... |
| | 12 | HTR2C | 207307_at | 3.6730... |
| | 13 | EMX2 | 221950_at | 3.6010... |
| | 14 | EOMES | 231776_at | 3.5940... |
| | 15 | DKK2 | 219908_at | 3.5496... |
| | 16 | RMST | 229782_at | 3.5128... |
| | 17 | MFAP4 | 212713_at | 3.4545... |
| | 18 | EIF1AY | 204409_s_at | 3.4509... |
| | 19 | PTPN20 | 215172_at | 3.4277... |
| | 20 | USP9Y | 228492_at | 3.4212... |
| | 21 | FSTL5 | 232010_at | 3.4148... |
| | 22 | CNGA3 | 207261_at | 3.4119... |
| | 23 | IMPG2 | 241856_at | 3.3991... |
| | 24 | TRDV3 | 216191_s_at | 3.3744... |
| | 25 | KCNA5 | 206762_at | 3.3582... |
| | 26 | PDE6H | 206841_at | 3.3328... |
| | 27 | SLC1A6 | 1554593_s_at | 3.3287... |
| | 28 | NEFM | 205113_at | 3.3048... |
| | 29 | LINC00348 | 233170_at | 3.2960... |
| | 30 | LGR5 | 213880_at | 3.2935... |
| | 31 | C7ORF57 | 1557636_a_at | 3.2496... |



# K-means clustering

- You tell the algorithm to subdivide a dataset in a predefined number of groups
- Algorithm has a random start, which may impact the results
  - Keep best solution out of 10 draws
  - Repeat the procedure to assess stability (10X10)

# K-means


Go to: Main
R2: Genomics Analysis and Visualization Platform
Online Tutorial
Main
Time series
AmpliconView
Kaplan-Meier
Sample maps
Small Tools
DataGrabber
Genome Browser
ChIP data
TAR literature
Change Data Scope
User Options
Help
Contact / About R2
1018043 (881501 unique) samples available
1
Choose single or multiple dataset analysis
Single Dataset
2
Select a dataset for analysis
Tumor Medulloblastoma PLoS One - Kool - 62 - MAS5.0 - u133p2
3
Select type of analysis
K-means
2 Groups Plotter
Fold over Fold Plotter
PathwayFinder
KEGG PathwayFinder by Groups
KEGG PathwayFinder by Gene correlation
Meta analyses
Parametric analysis of geneset enrichment (PAGE)
Geneset maps (GSM)
K-means
Principle Component Analysis (PCA)
T-SNE
Sample Correlation Map (SCM)
View Geneset (Heatmap)
TopLister (Gene filter stdev)
Geneset vs Genesets Correlations
Track(#) vs Genesets Correlations
Venn Diagram of GeneCategories
Personalized Genomics
Static circos files (v3)
Somatic Mutations (v3)
4
What is R2?
Welcome to R2; a biologist friendly web based genomics analysis and visualization application developed by Jan Koster at the department of Oncogenomics in the Academic Medical Center (AMC) Amsterdam, the Netherlands. You can start exploring the gene expression data by following the numbered options in the center.
For citations, please include the following webcite: 'R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl)'.
Financial Supporter of R2
Amsterdam UMC
all news
PubMed
r2.amc.nl
Video training
YouTube
Integrated Analysis of Tumor Genomics Data with R2
Tutorials
Download the R2 Tutorials Book

# K-means

**R2: K-means**

## Tumor Medulloblastoma PLoS One - Kool - 62 - MAS5.0 - u133p2 public

**Adjustable settings**

| | |
|---|---|
| Floor value: | 16 |
| Transformation: | Log2 z-score |

**Sample Filter**

| | |
|---|---|
| Subset track: | |
| Selected sample subset: | None |

**Gene Filters**

| | |
|---|---|
| Min. # Present calls: | 1 |
| Minimal maximum value: | 64 |
| Minimal range size: | 48 |
| # highest SD genes: | 1500 |
| Chromosome: | All |
| Gene ontology: | All — Search GO |
| Gene set: | Search GS |
| Manual list: | none |

**Clustering**

| | |
|---|---|
| Number of groups: | 3 |
| Number of passes: | 10 |
| Number of rounds: | 10 |

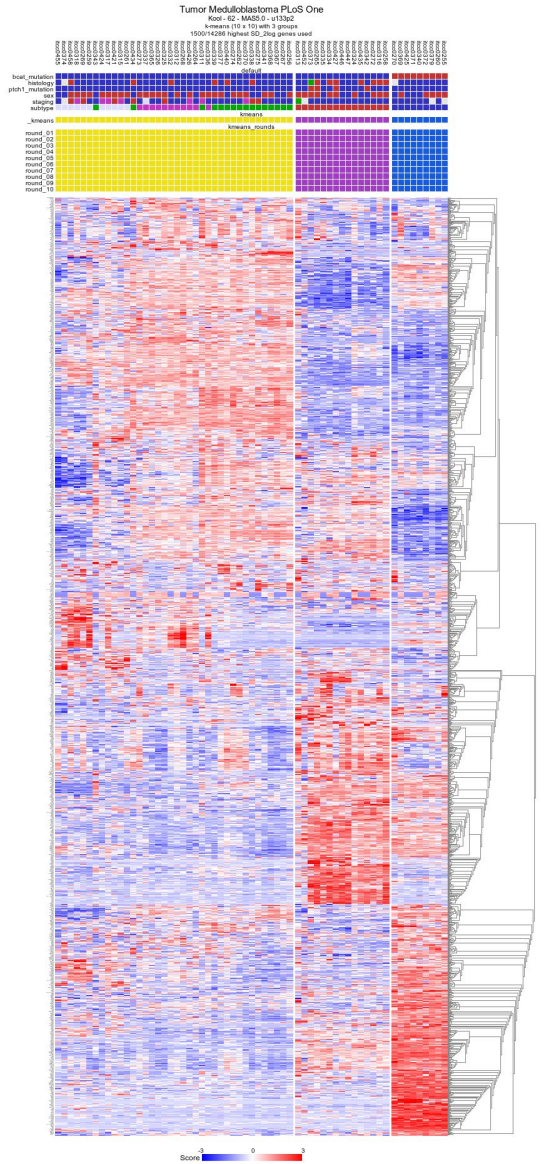
**Graphics**

| | |
|---|---|
| Draw heatmap: | yes |
| Color scheme: | blue-white-red |
| Label track: | samplenames (62 cat) |

**Heatmap Options**

| | |
|---|---|
| Cell width: | auto |
| Cell height: | auto |

next Reset




**kmeans group vs annotation tracks**

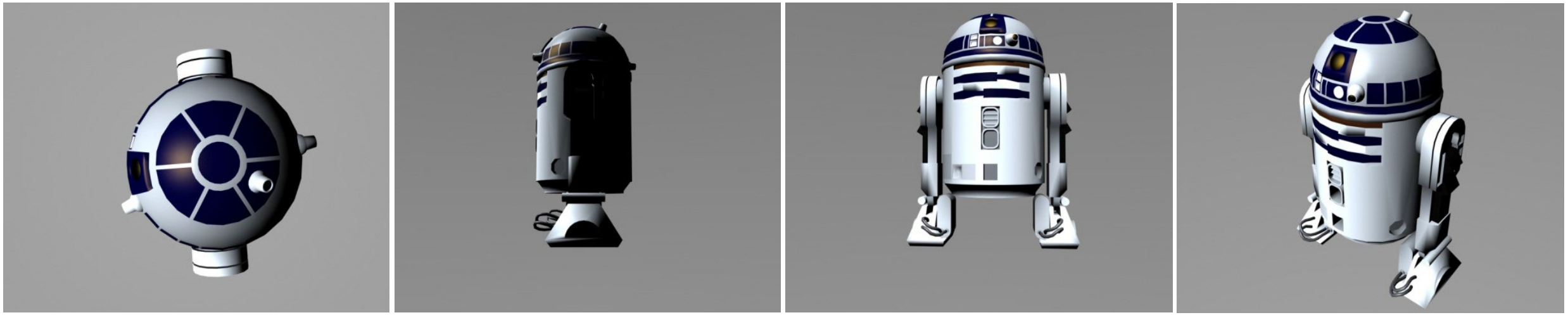
Chi-square tests with Yates correction (if 2x2) and Fisher's Exact tests (#groups<30)
NA or ND values are skipped and test is only performed on 'cat' variables

**Source track: _kmeans**

| track | chi_pvalue | chisquare | df | fisher's exact |
|---|---|---|---|---|
| round_01 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_03 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_04 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_05 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_07 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_08 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_09 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_10 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_02 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| round_06 | 7.47e-26 | 124.000 | 4 | 7.89e-24 |
| subtype | 4.94e-23 | 124.000 | 8 | 7.89e-24 |
| bcat_mutation | 3.44e-14 | 62.000 | 2 | 4.93e-11 |
| ptch1_mutation | 1.23e-03 | 13.398 | 2 | 4.93e-03 |
| histology | 0.016 | 12.141 | 4 | 8.92e-03 |
| staging | 0.080 | 14.057 | 8 | 0.063 |
| nt_source | 0.212 | 8.378 | 6 | 0.195 |
| samplenames | 0.433 | 124.000 | 122 | too many groups(186) |
| sex | 0.763 | 0.540 | 2 | 0.747 |

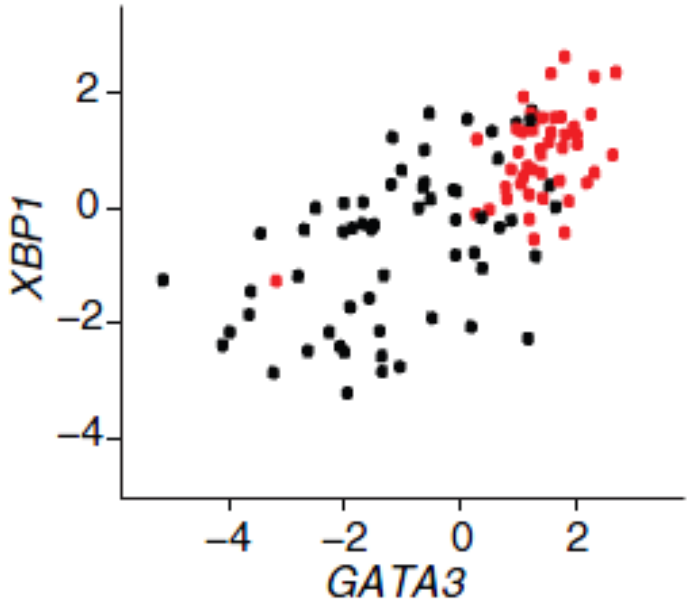
# Principal Components Analysis (PCA)

- Describe samples by a very limited number of variables
- Samples that are similar ‘cluster’ closer together
- Mathematical procedure orthogonal projection (Mathematical procedure) for data reduction
- Nice way of visualizing subgroups

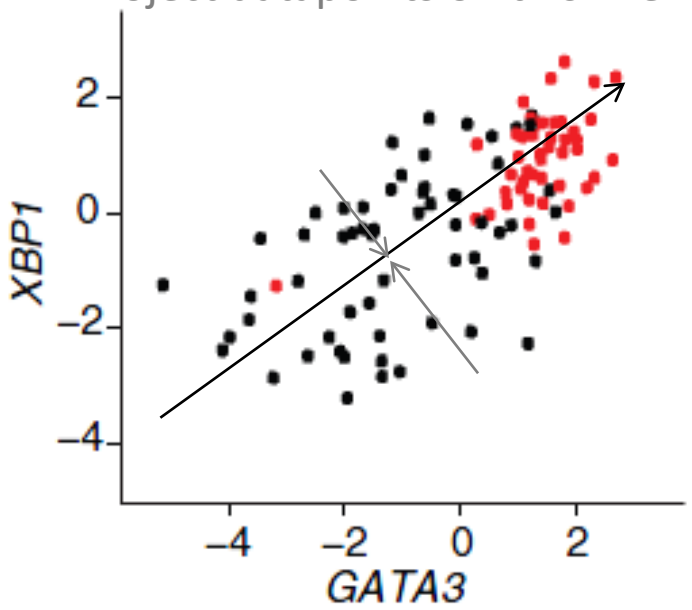


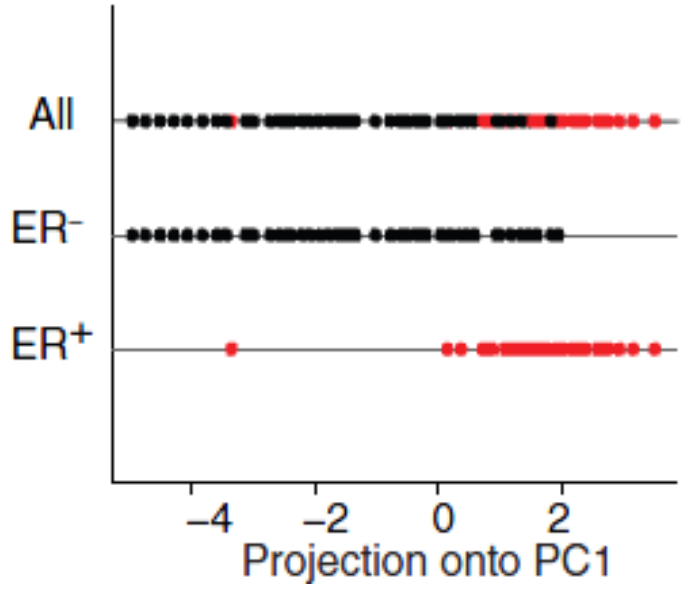
# Principal Components Analysis (PCA)

- Describe samples by a very limited number of variables
- Samples that are similar ‘cluster’ closer together
- Mathematical procedure orthogonal projection (Mathematical procedure) for data reduction
- Nice way of visualizing subgroups

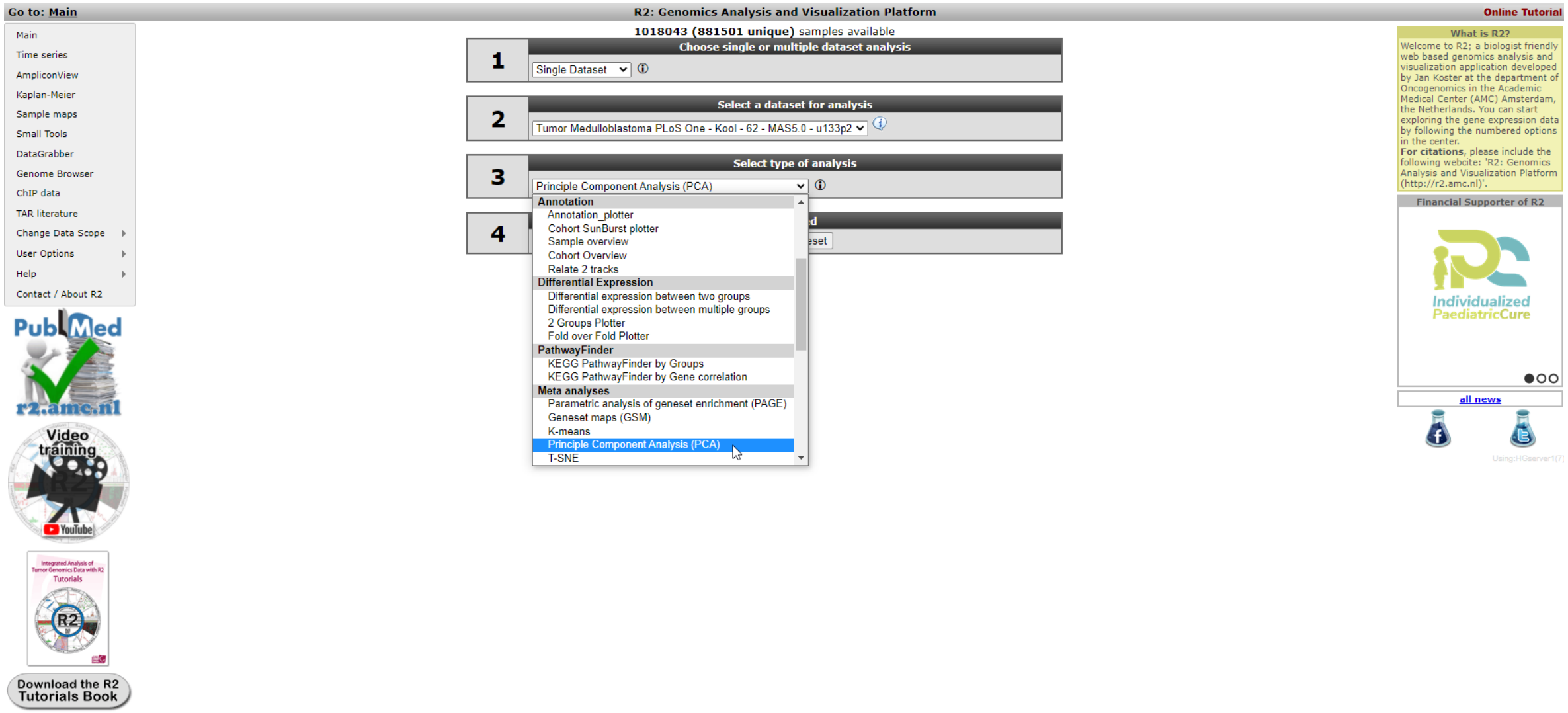



Draw the longest possible line
Through the data = max. variance
Project datapoints on this line

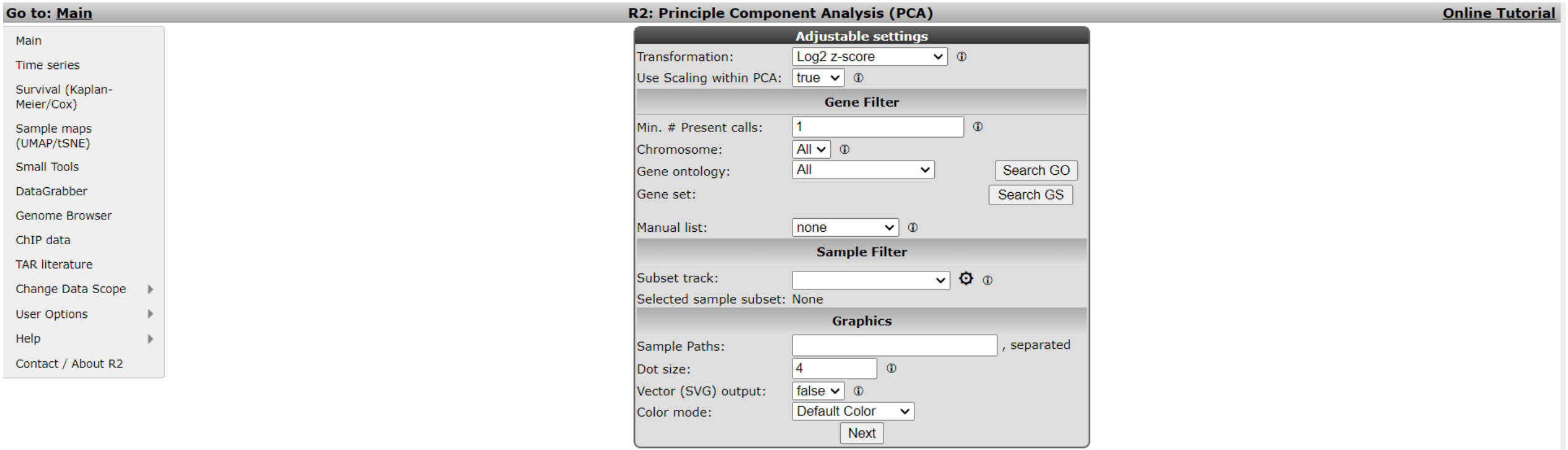


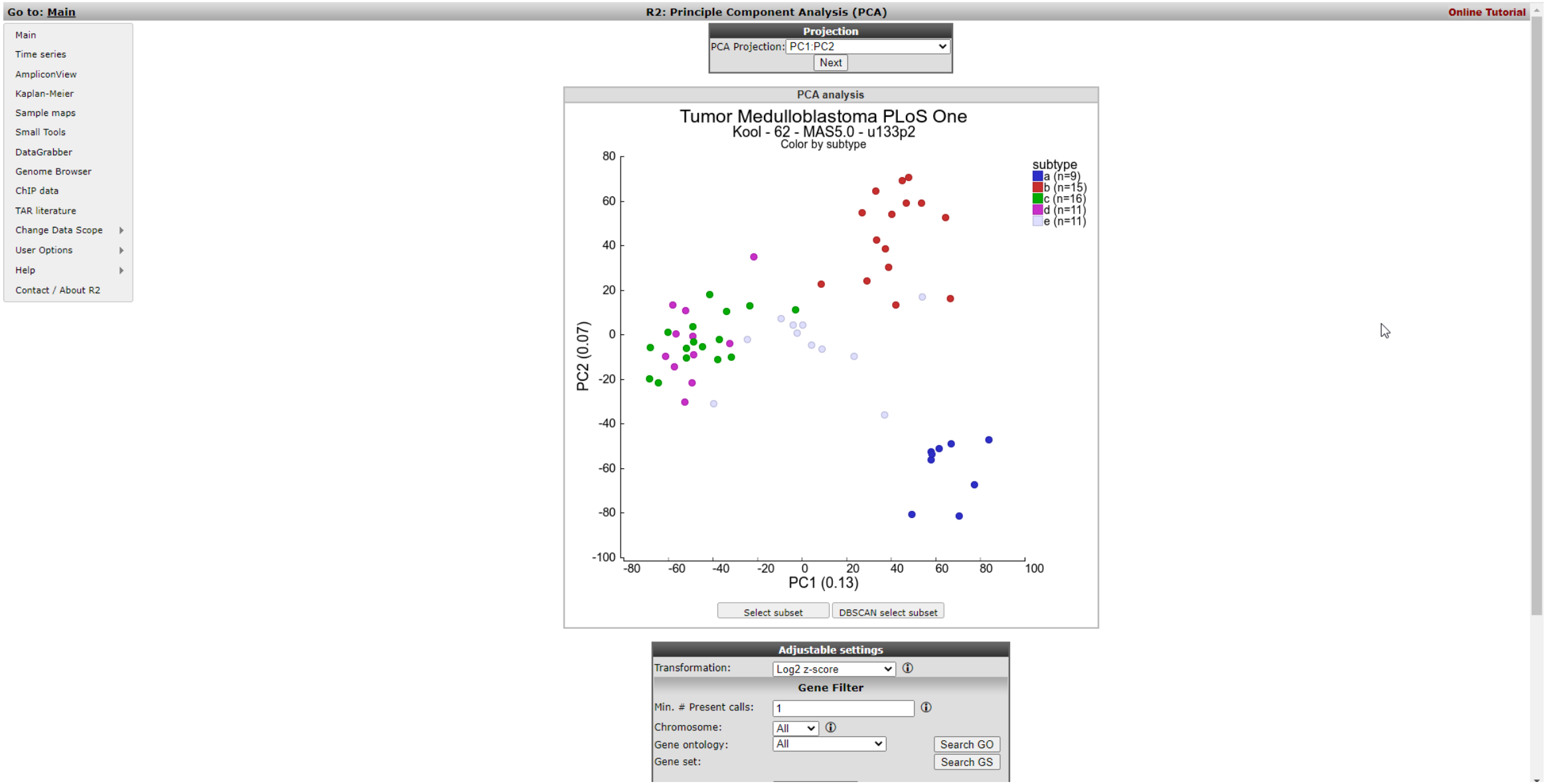
# Principal Components Analysis (PCA)



# Principal Components Analysis (PCA)



# Principal Components Analysis (PCA)



# Principal Components Analysis (PCA)

Go to: Main

R2: Principle Component Analysis (PCA)

Online Tutorial

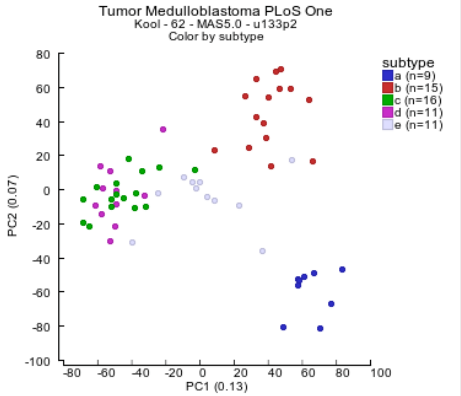
- Main
- Time series
- AmpliconView
- Kaplan-Meier
- Sample maps
- Small Tools
- DataGrabber
- Genome Browser
- ChIP data
- TAR literature
- Change Data Scope
- User Options
- Help
- Contact / About R2

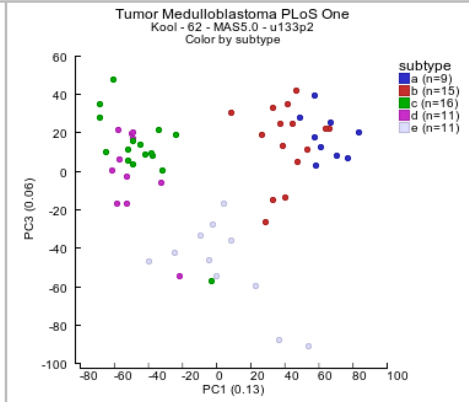
**Projection**

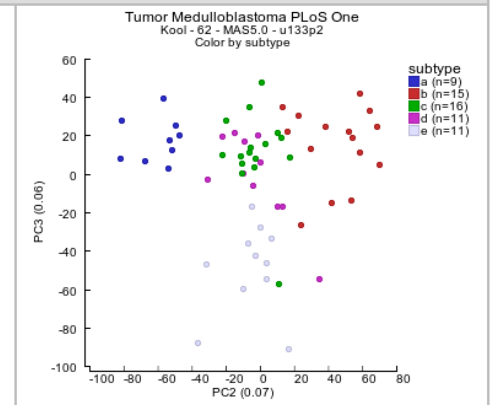
PCA Projection: PC1:PC2::PC1:PC3::PC2:PC3

Next

**PCA analysis**




**Adjustable settings**

Transformation: Log2 z-score

**Gene Filter**

Min. # Present calls: 1

Chromosome: All

Gene ontology: All — Search GO

Gene set: — Search GS

Manual list: none

**Sample Filter**

Subset track:

Selected sample subset: None

**Graphics**

Sample Paths: , separated

Dot size: 4

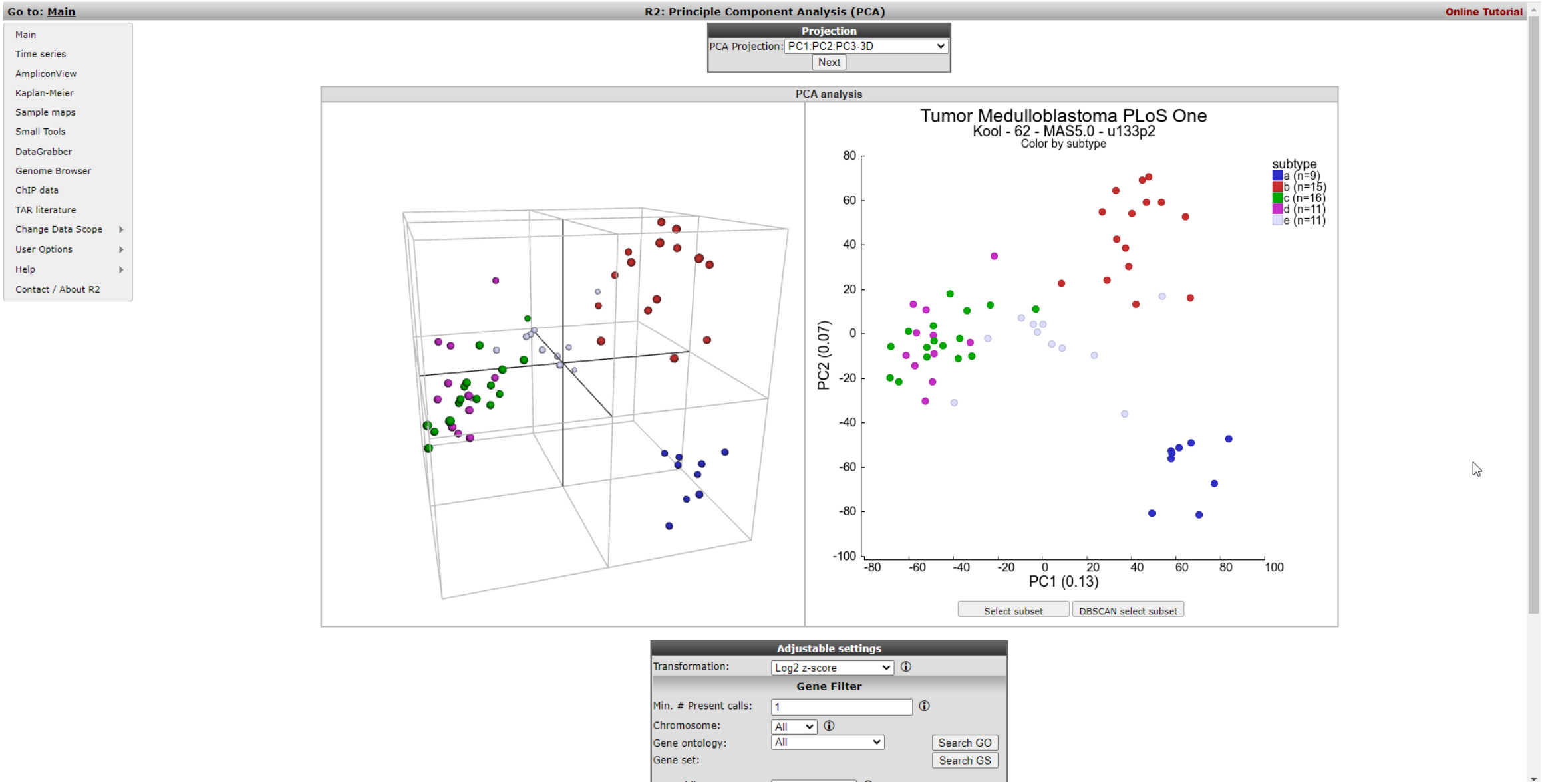
Vector (SVG) output: false

Color mode: Color by Track

Color track: subtype (5 cat)

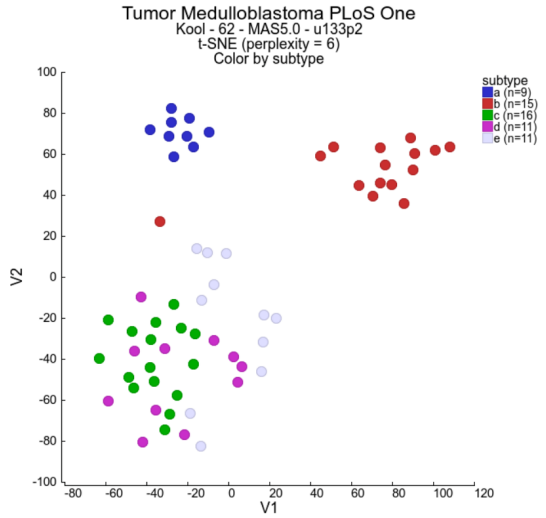
Next

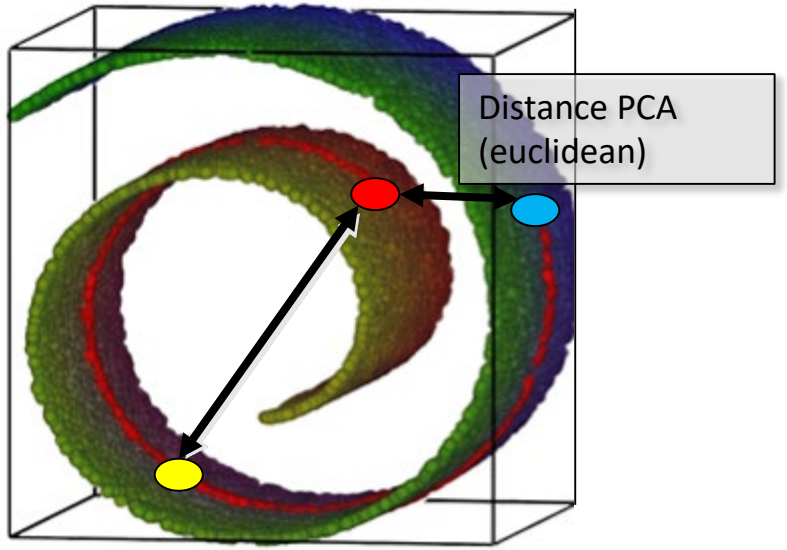
# Principal Components Analysis (PCA)

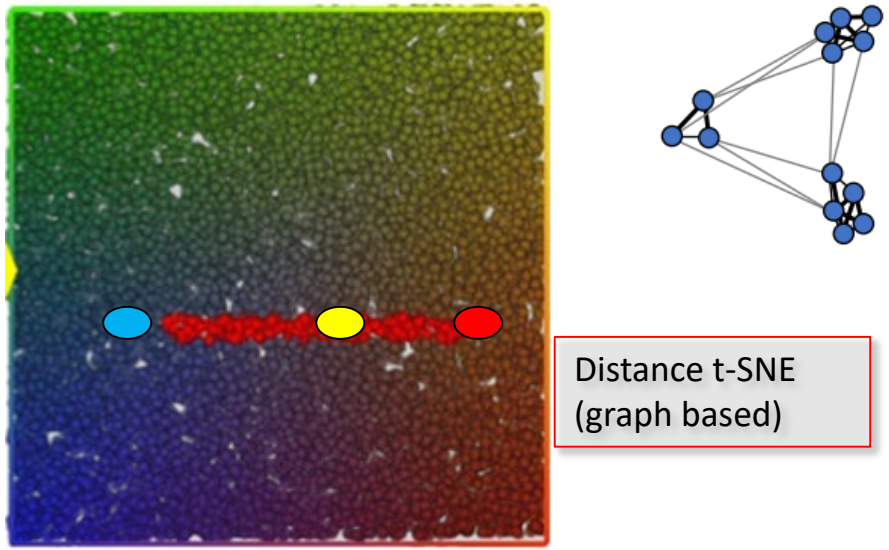


# tSNE/UMAP Analysis

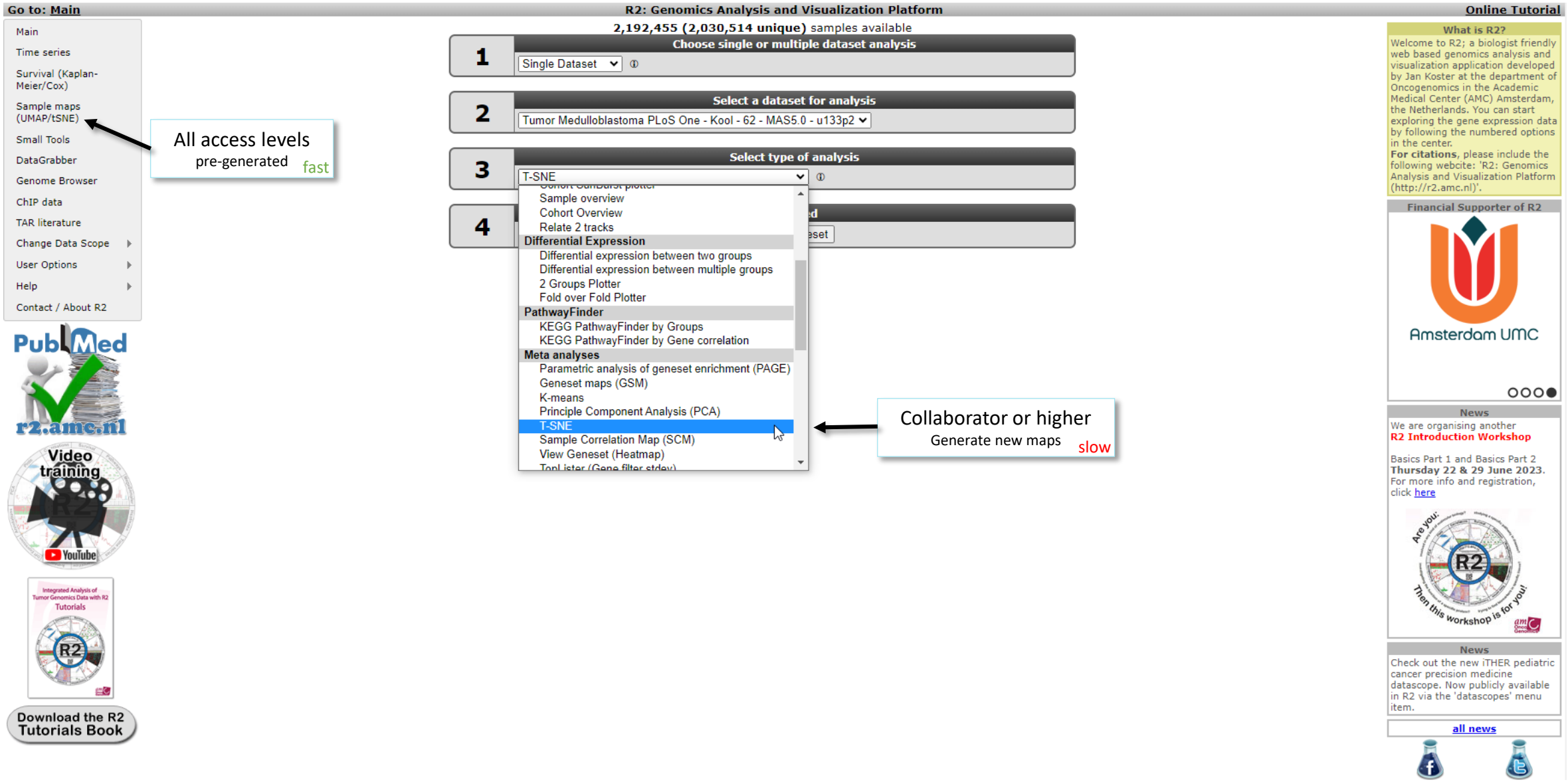
- t-Distributed Stochastic Neighbor Embedding (t-SNE)
  - a technique for dimensionality reduction that is particularly well suited for the visualization of high-dimensional datasets.
  - Has an almost magical ability to create 2-dimensional ‘maps’ of data with thousands of dimensions
  - Non-linear
- Graph based dimensionality reduction: connected points in graph
- UMAP



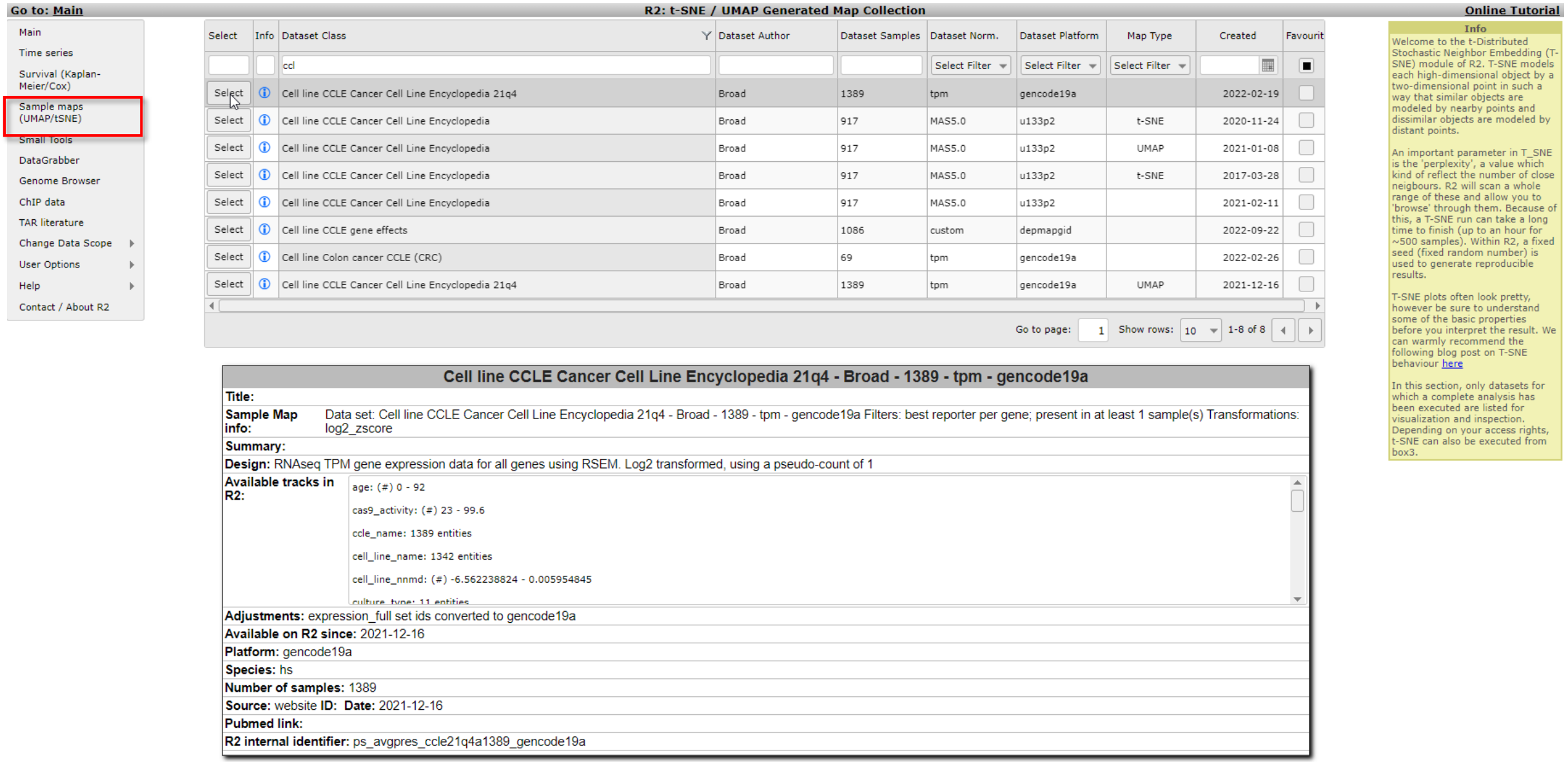



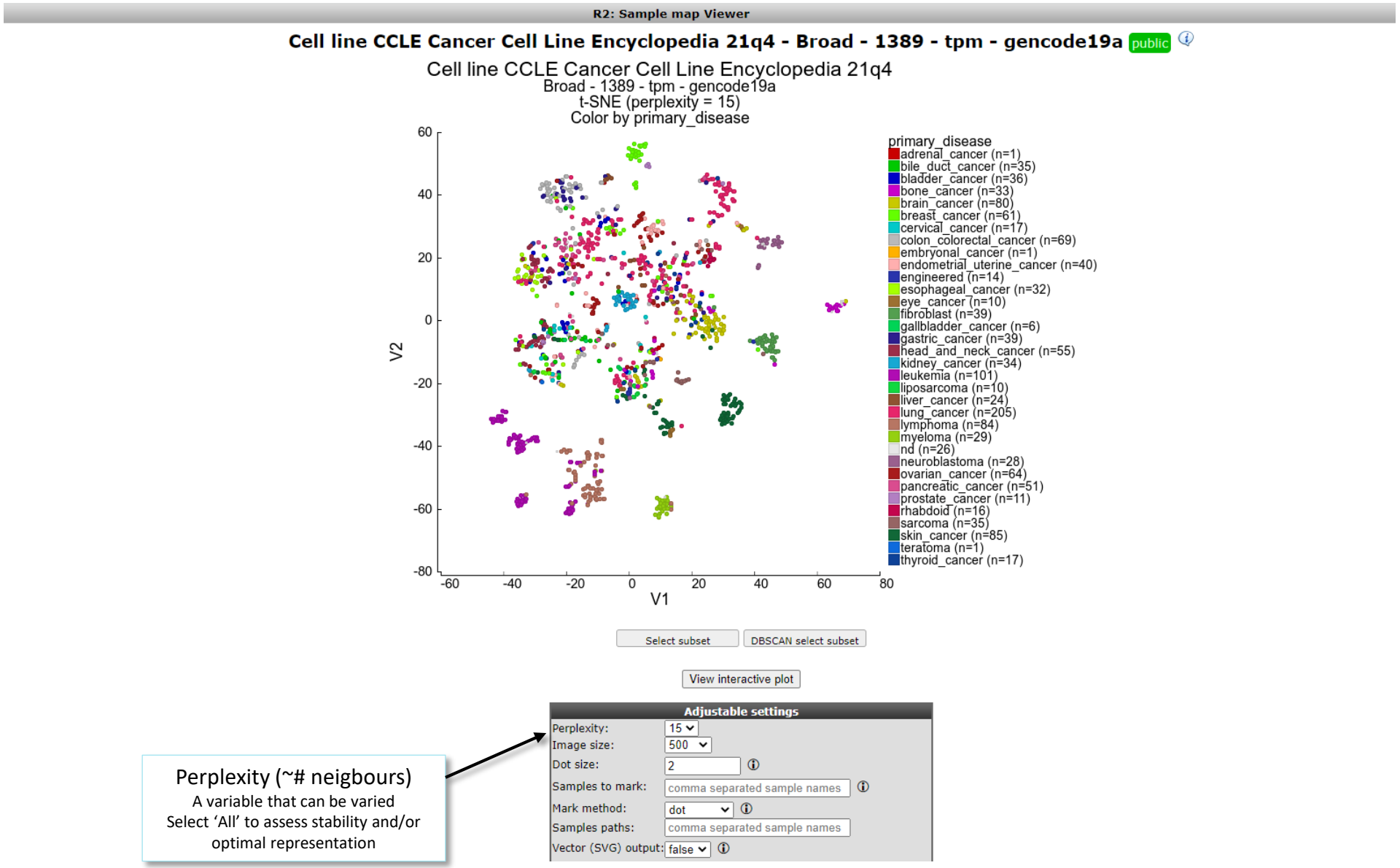
# tSNE/UMAP Analysis



# tSNE/UMAP Analysis: select pregenerated sample maps



# tSNE Analysis result




Perplexity (~# neigbours)

A variable that can be varied

Select 'All' to assess stability and/or optimal representation

# tSNE Analysis: perplexity all

# tSNE Analysis: color by gene expression

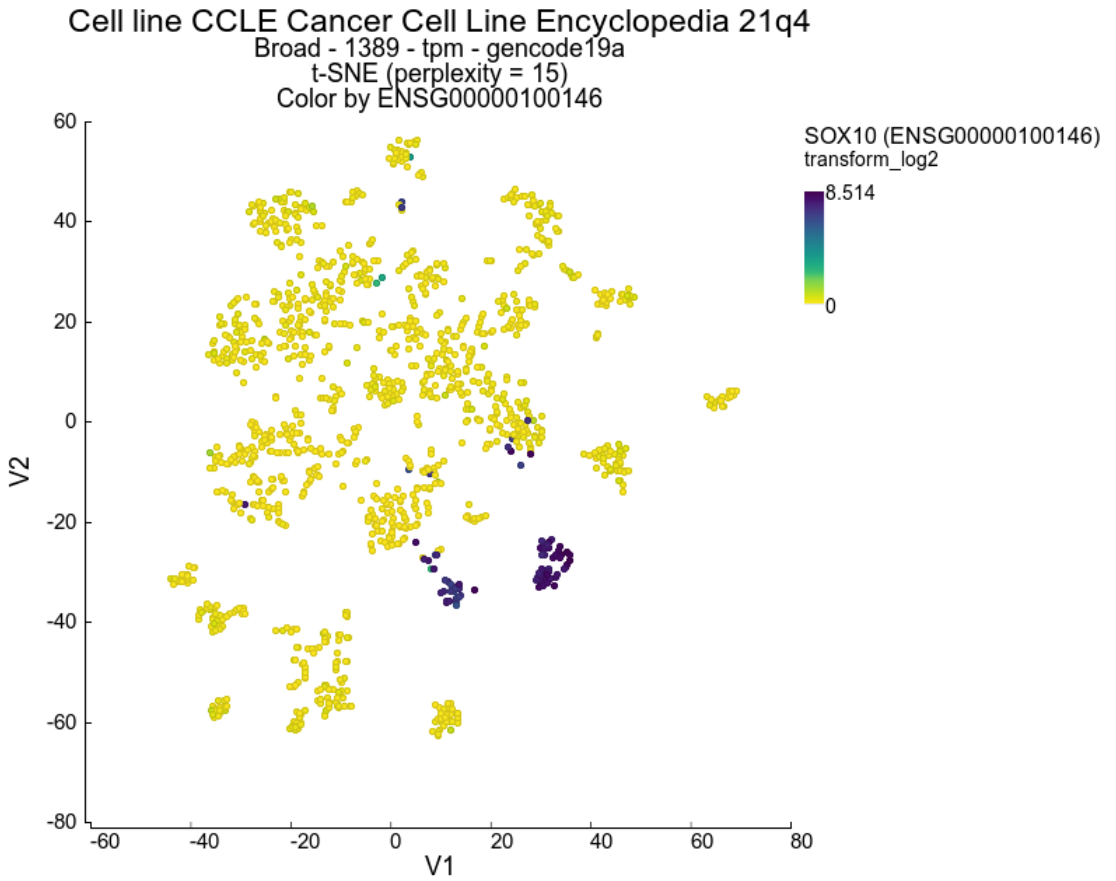
**Adjustable settings**

| | |
|---|---|
| Perplexity: | 15 |
| Image size: | 500 |
| Dot size: | 2 |
| Samples to mark: | comma separated sample names |
| Mark method: | dot |
| Samples paths: | comma separated sample names |
| Vector (SVG) output: | false |
| Enable hovering: | yes |
| x: | min: max: |
| y: | min: max: |
| Color mode: | Color by Gene |
| Color source: | Cell line CCLE Cancer Cell Line Encyclopedia 21q4 - Broad - 1389 - tpm - gencode19a |
| Gene / Reporter: | SOX10 ENSG00000100146 advanced |
| Transformation: | Log2 |
| Color scheme: | viridis |

Submit

**R2: Sample map Viewer**

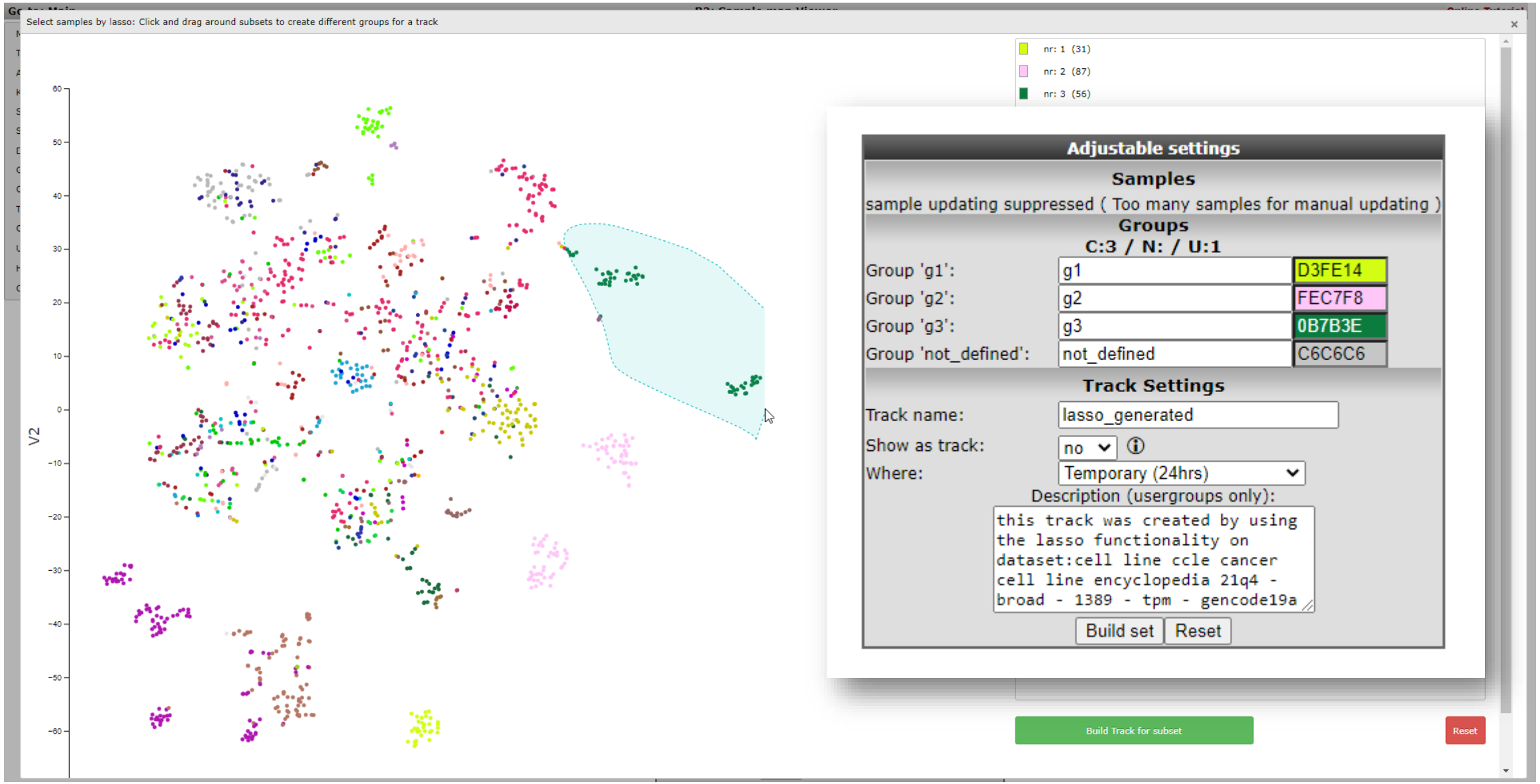
**Cell line CCLE Cancer Cell Line Encyclopedia 21q4 - Broad - 1389 - tpm - gencode19a** public



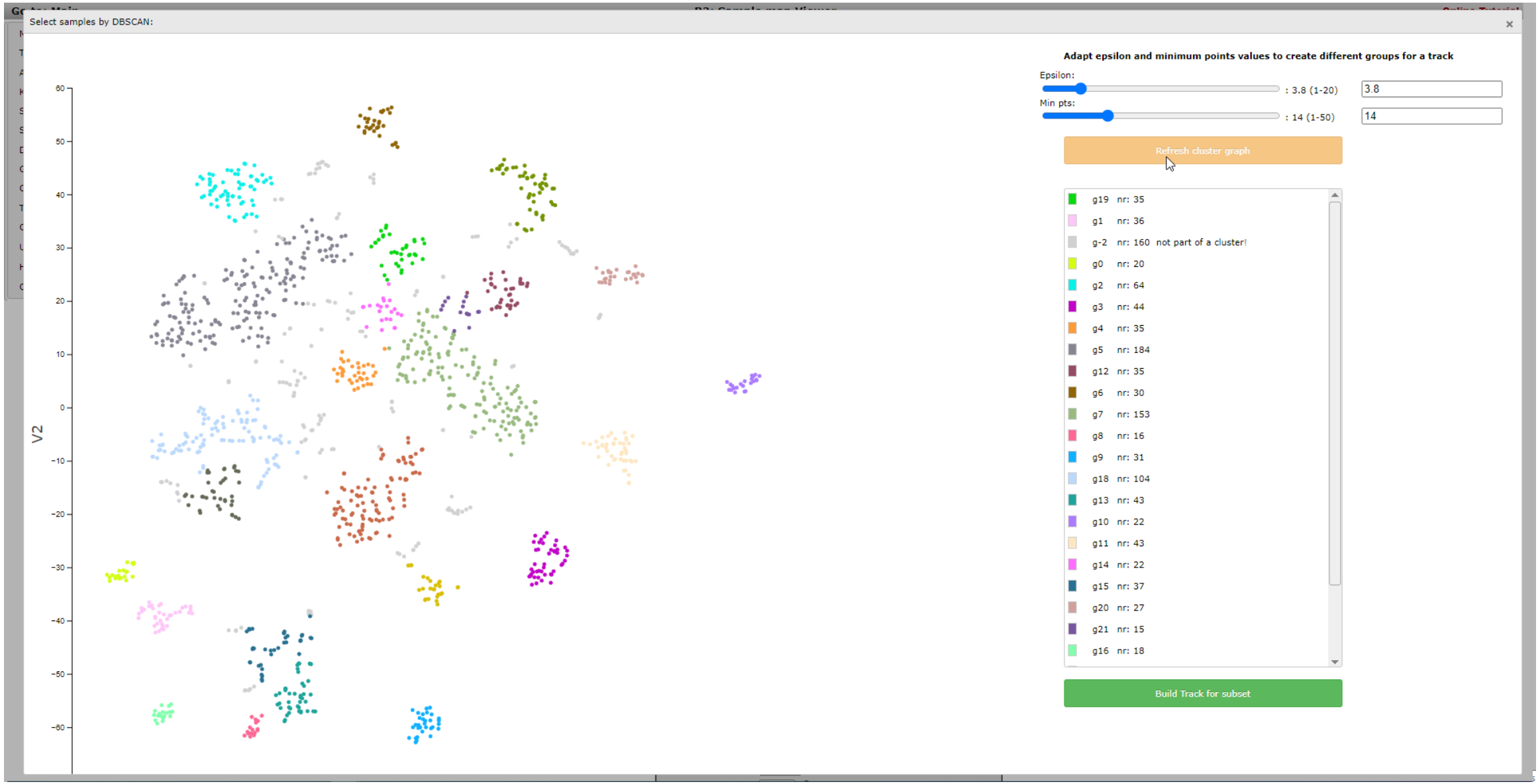

Select subset | DBSCAN select subset

View interactive plot

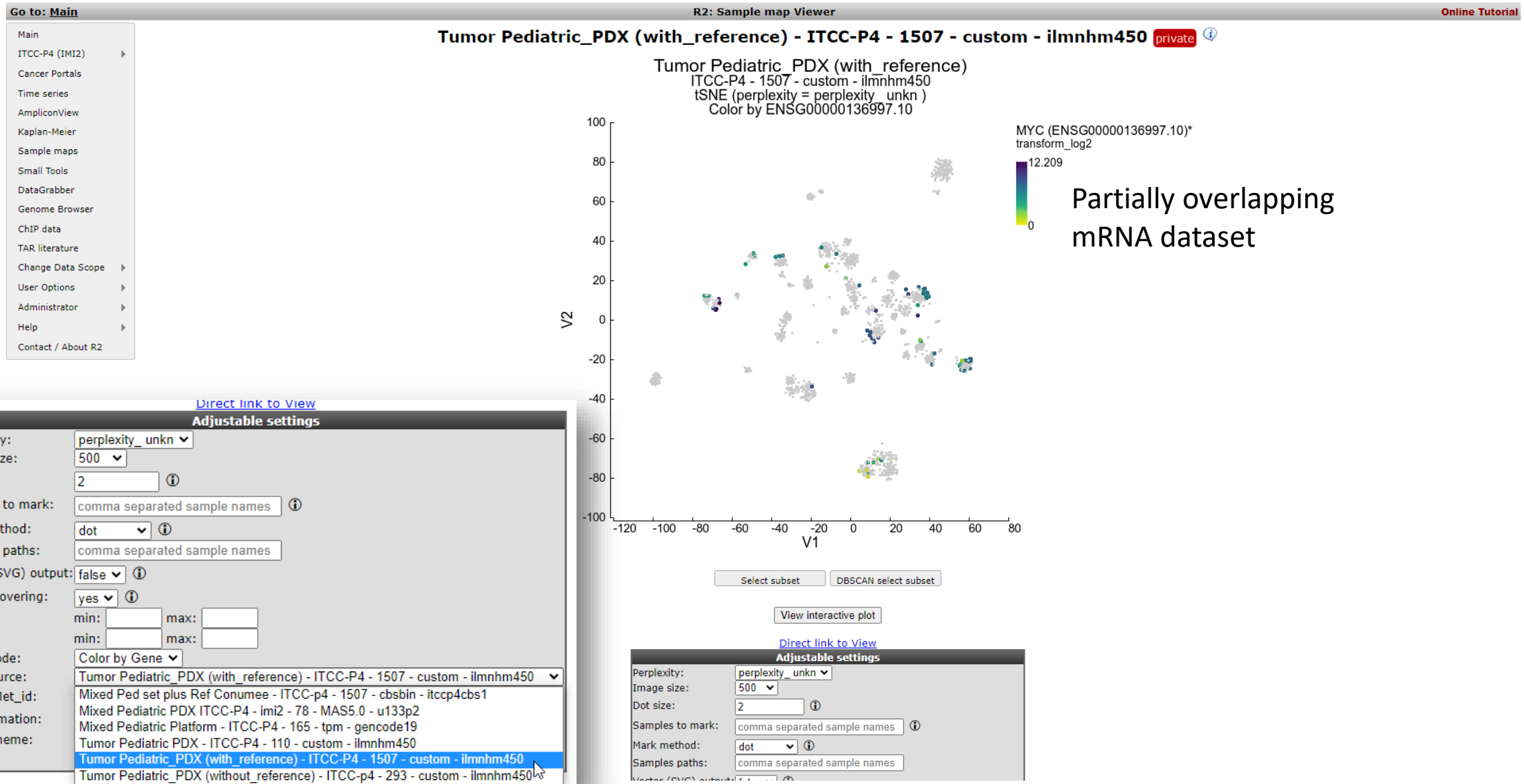
# tSNE Analysis: lasso



# tSNE Analysis



# In multi omics sets ‘color by gene’

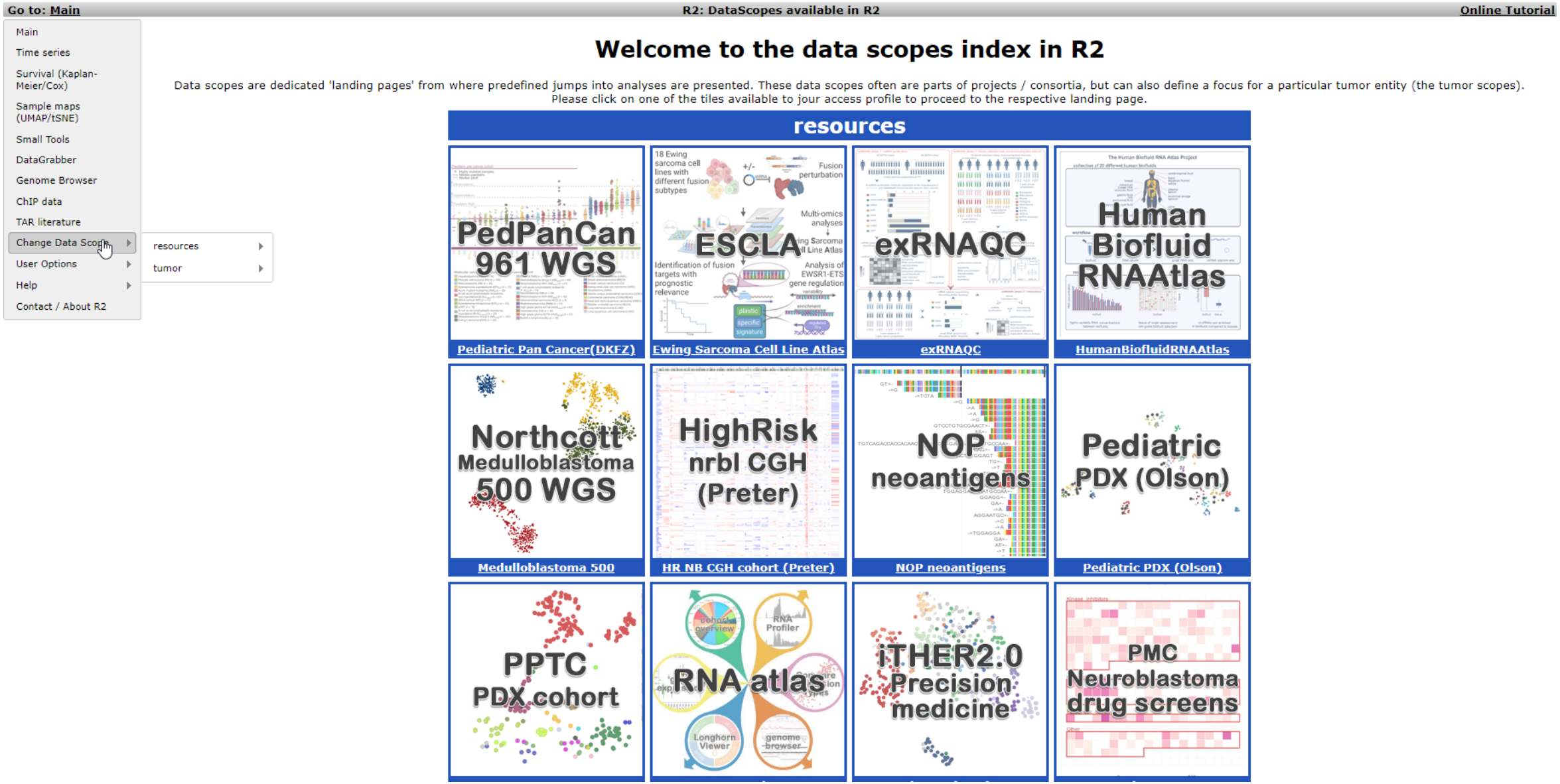



Partially overlapping mRNA dataset

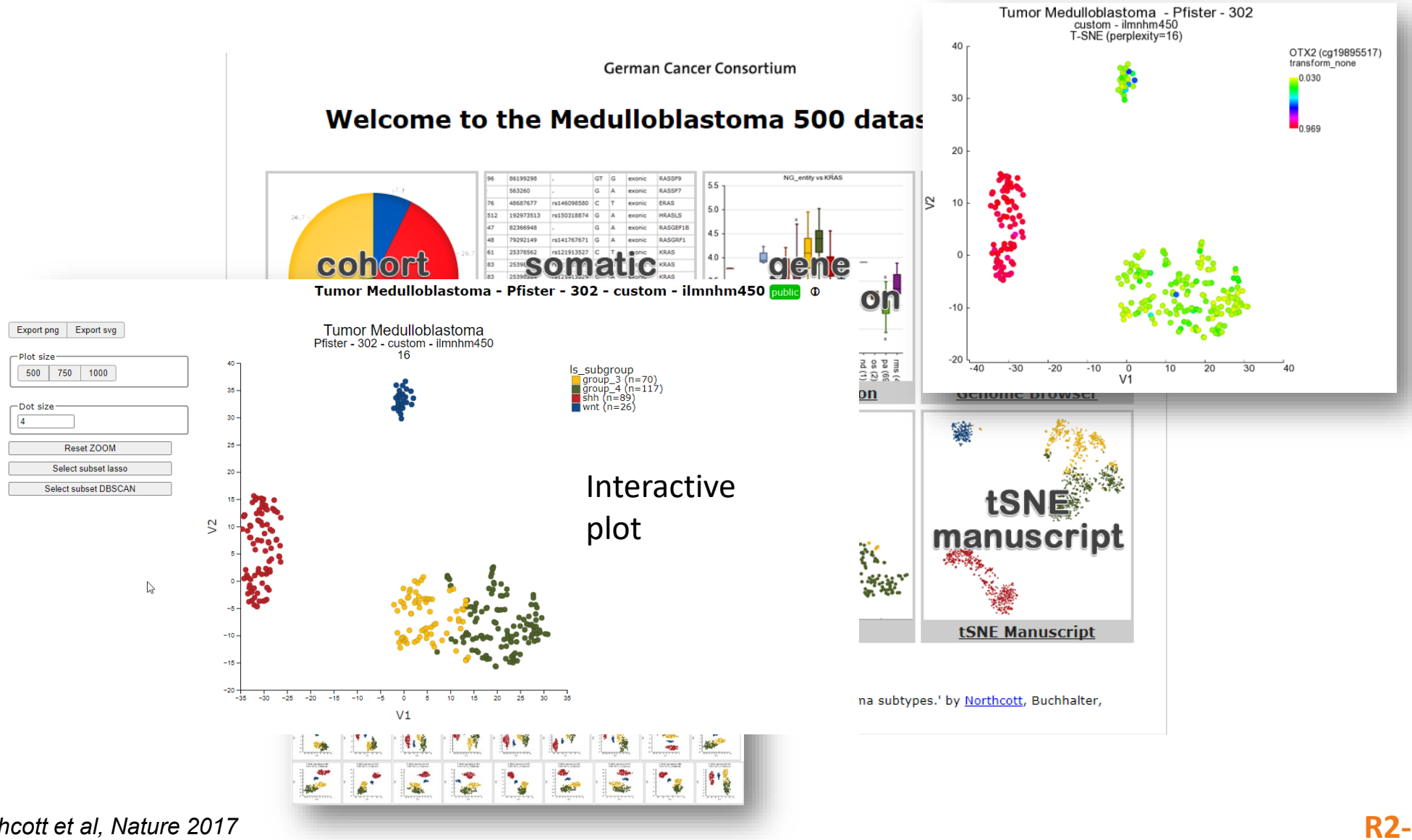
r2-support@amsterdamumc.nl

# DataScopes

# DataScopes (Landing pages with subsection of datasets)

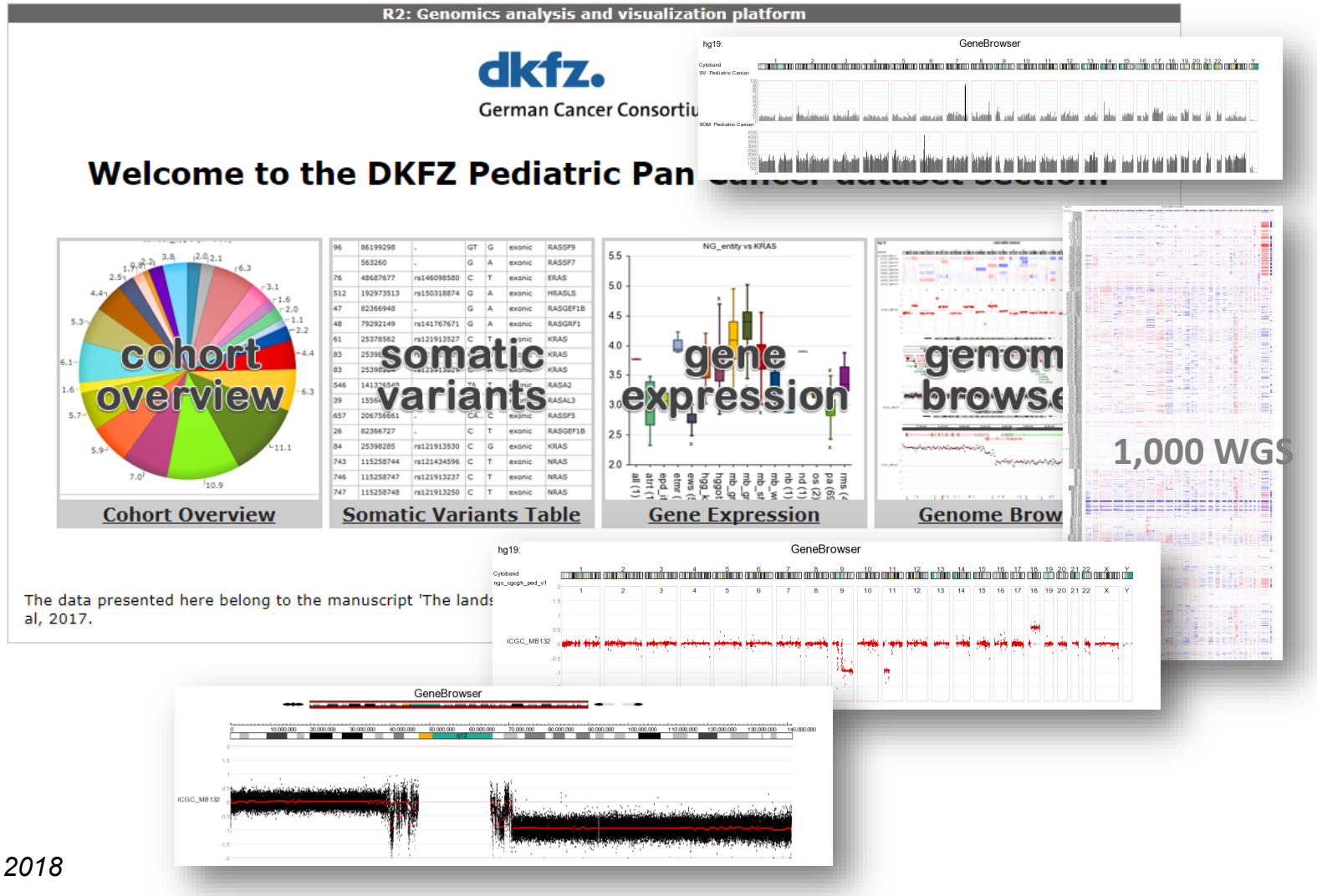


# WGS of 500 medulloblastoma tumors



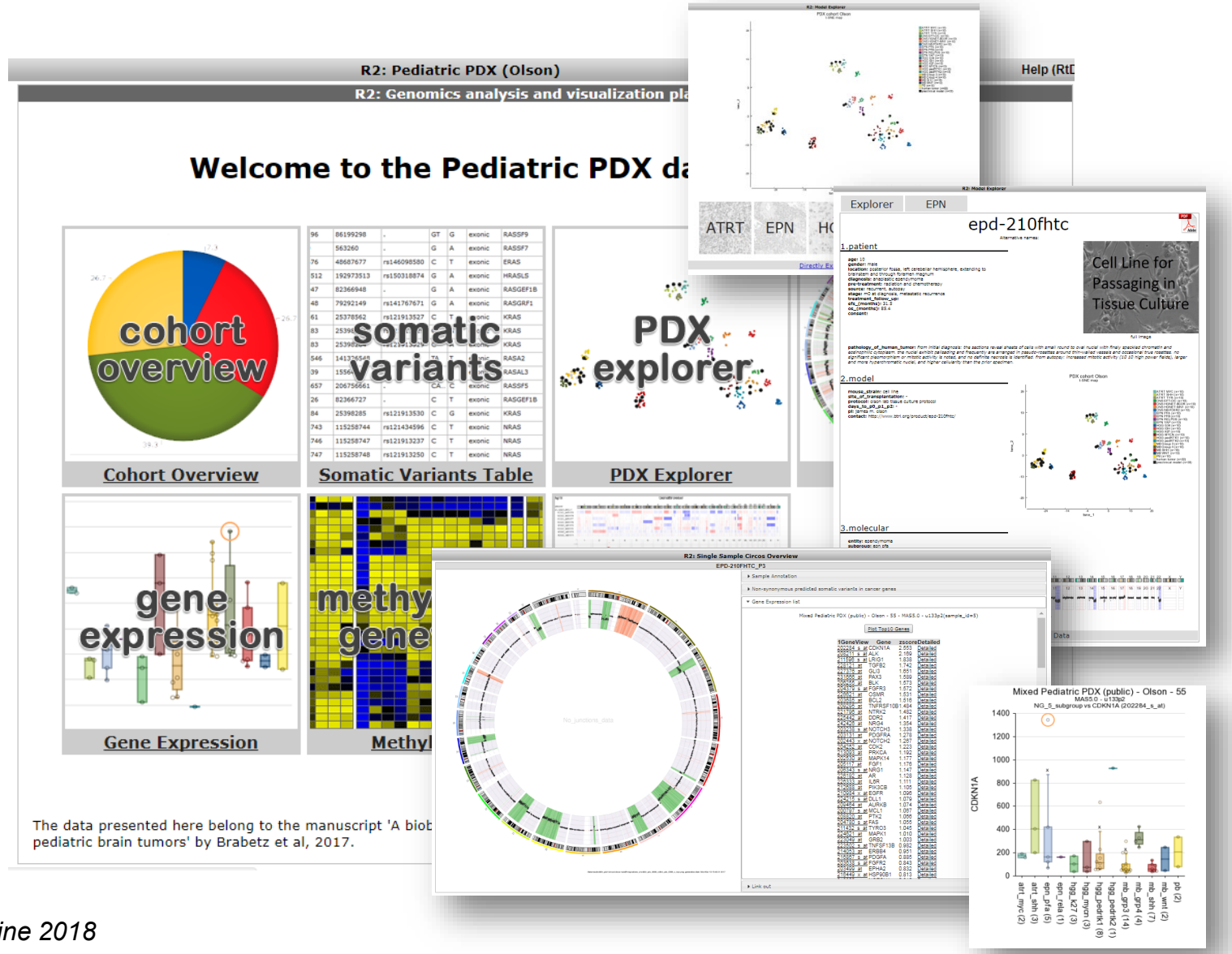

Northcott Nature 2017

# WGS of 1000 pediatric patients



# 556 Neuroblastoma CGH samples

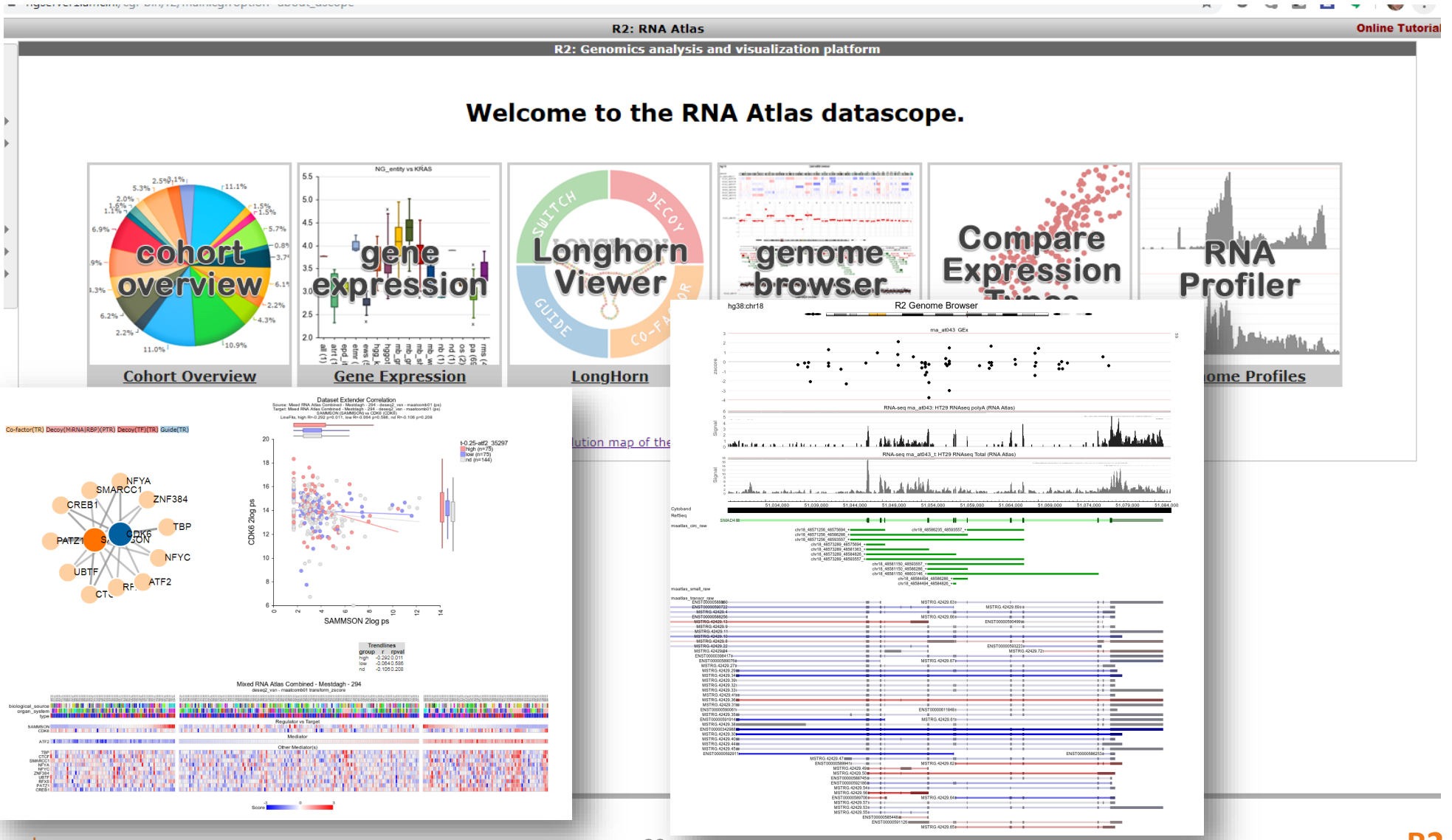
# PDX explorer linked to genomic data



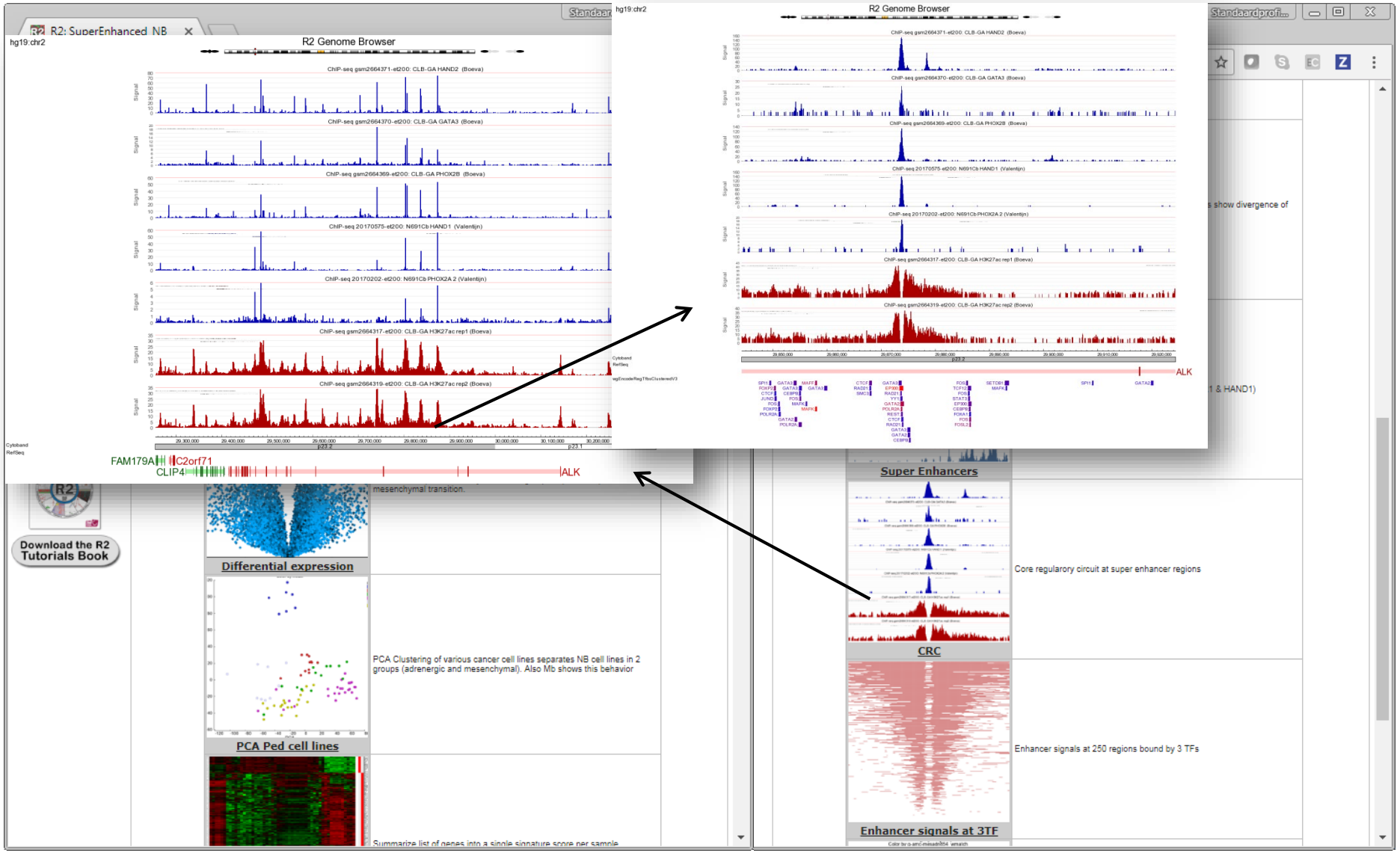

 *Brabetz et al, Nat. Medicine 2018*

**R2-Platform**

# RNA Atlas (300 polyA / RiboDepl / Small)



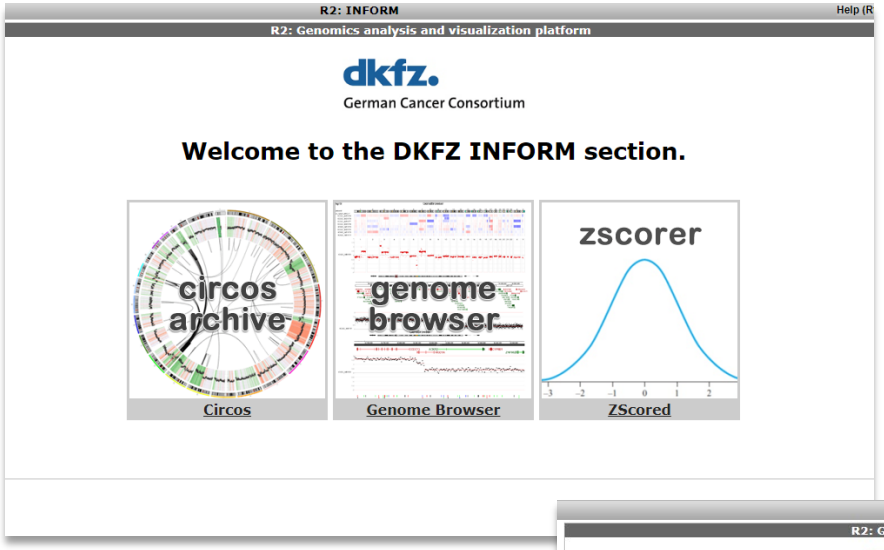
# Links to interactive analyses (like Super Enhancers)

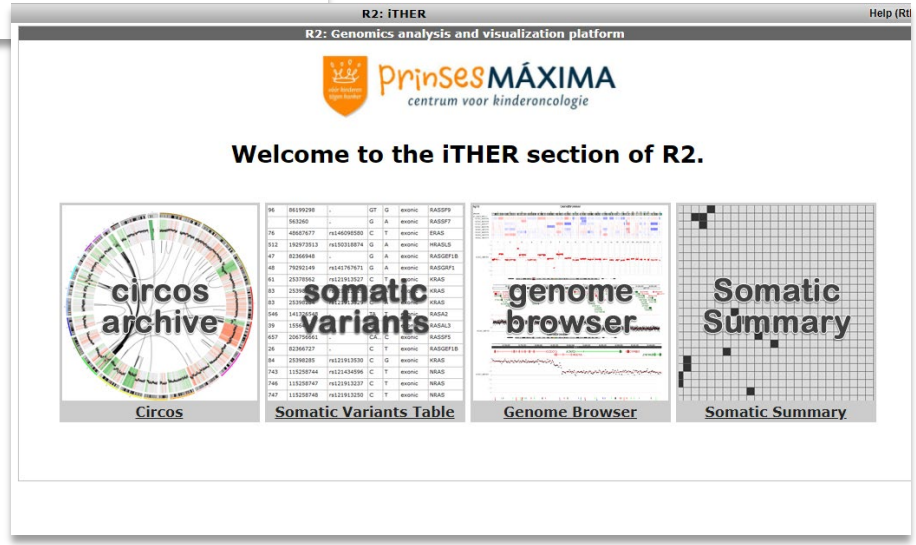





*Van Groningen, Koster & Valentijn et al, Nat. Genet, 2017*

# Personalized Medicine programs



# Interested?

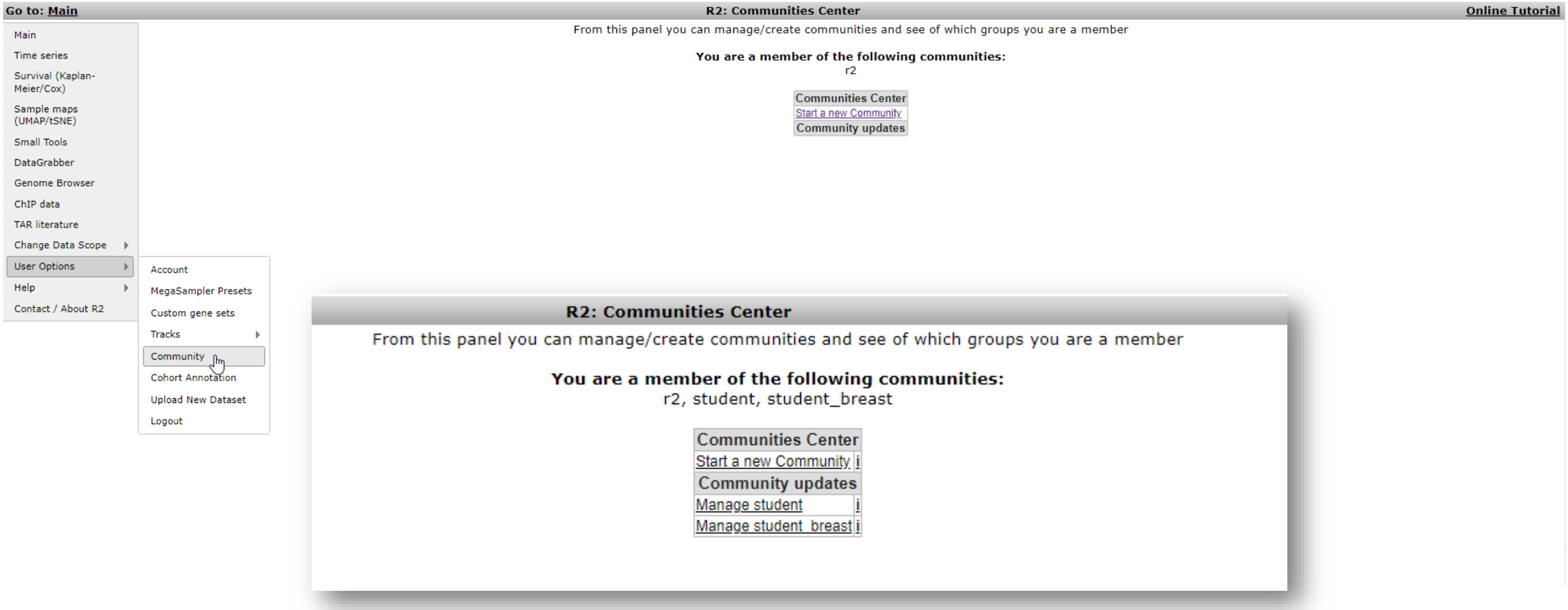
- If you are interested in having a datascope foryour group / consortium?
  - Get in touch with us via r2-support@amsterdamumc.nl

# R2 communities

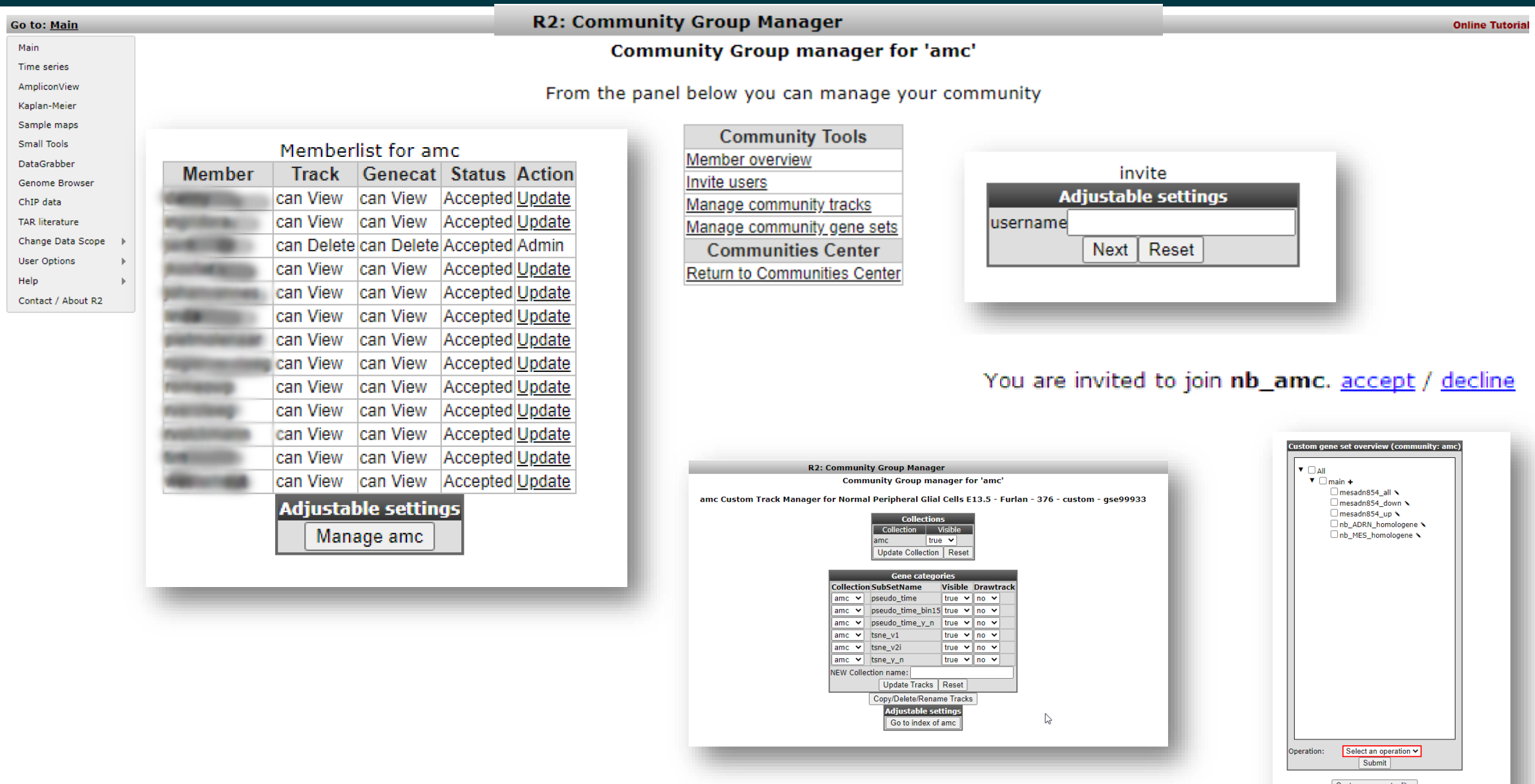
- Users working together on 1 project may want to share
  - MegaSampler presets
  - Tracks
  - Gene Sets

- R2 Community feature
  - Any user can start user groups (as many as you like)
  - Invite other users
  - Share
    - Tracks
    - MegaSampler presets
    - Gene Categories
    - GenomeBrowser profiles

# R2 communities



# R2 communities




r2-support@amsterdamumc.nl

# Recap R2 Support

Go to: Main

- Main
- Time series
- Survival (Kaplan-Meier/Cox)
- Sample maps (UMAP/tSNE)
- Small Tools
- DataGrabber
- Genome Browser
- ChIP data
- TAR literature
- Change Data Scope
- User Options
- Help
- Contact / About R2

Online Documentation (RTD)

Tutorials Book (PDF)

Training Courses

PubMed

r2.amc.nl

Video training

YouTube

Download the R2 Tutorials Book

AUMC: CEMM

r2-training-courses.readthedocs.io/en/latest/

R2 training courses

latest

Search docs

STUDENTS COURSE

1. Investigating Intra-tumor Heterogeneity in Neuroblastoma

GRADUATE COURSE

1. Investigating structural variants
2. Investigating Intra-tumor Heterogeneity

ADDITIONAL COURSE MATERIALS

1. Differential gene expression in micro-array colon cancer data
2. Finding causes in Neuroblastoma genomics data
3. Investigating structural variants
4. BMS38: Computer Practicals



Docs » R2 Training Courses: 2023-02-02

Edit on GitHub

## R2 Training Courses: 2023-02-02

This contains a collection of training courses for R2; a biologist friendly, web based genomics analysis and visualization application developed by Jan Koster at the department of Oncogenomics in the Academic Medical Center (AMC) Amsterdam, the Netherlands. For citations, please include the following webcite: 'R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl http://r2platform.com)'



Table of Contents

### Students Course

- 1. Investigating Intra-tumor Heterogeneity in Neuroblastoma
  - 1.1. Introduction
  - 1.2. Tumors and origins: a first impression of your data
  - 1.3. Urgency of research: patient material
  - 1.4. Which genes make a difference? Creating signatures
  - 1.5. Identifying groups: using signatures to classify other datasets
  - 1.6. Using scores for further characterization
  - 1.7. Finding causes: homing in on transcription factors
  - 1.8. Proving causes: manipulating cell lines
  - 1.9. Creating hypotheses: relating to chromatin modification data
  - 1.10. Suggesting therapy
  - 1.11. Final remarks / future directions

### Graduate Course

- 1. Investigating structural variants
  - 1.1. Introduction
  - 1.2. Exploring the dataset
  - 1.3. Pie Charts
  - 1.4. Somatic mutations in neuroblastoma
  - 1.5. Further use of WGS data; structural variants
  - 1.6. Chromothripsis
  - 1.7. Locations of structural variants, hotspots?

Online Tutorial

What is R2?

Welcome to R2; a biologist friendly web based genomics analysis and visualization application developed by Jan Koster at the department of Oncogenomics in the Academic Medical Center (AMC) Amsterdam, the Netherlands. You can start exploring the gene expression data by following the numbered options in the center.
**For citations**, please include the following webcite: 'R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl)'.

Financial Supporter of R2

Cancer Center Amsterdam

News

Check out the new iTHER pediatric cancer precision medicine datascope. Now publicly available in R2 via the 'datascopes' menu item.

all news

Subscribe

encca (2)

r2-support@amsterdamumc.nl

R2-Platform

# Fun things to do:

**Computer login:**
**func_user**
**Welkom01**

## R2 MSCA-DN Training Courses

- Help => Training Courses => PRESSURE
  - USE hgserver2!
- Read carefully and follow the research line step by step
- Feel free to toy around
  - Who knows

## Do your own research

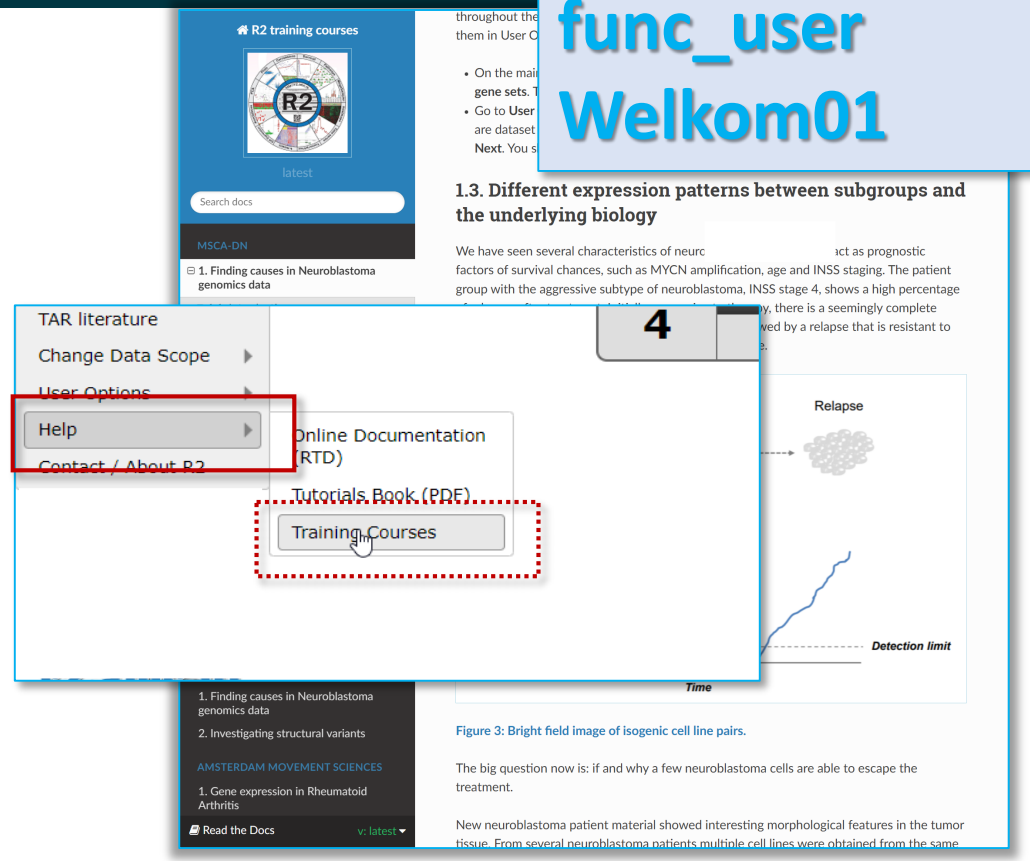
- Find a dataset of your interest in R2 dataset grid
  - **Tissue/Tumor: “Esoph” => 21 public sets!**
  - **Author: Bijlsma | Zalm (access rights?)**
- Think about a good biological research question
- Try to find the analyses that can help you gain insight

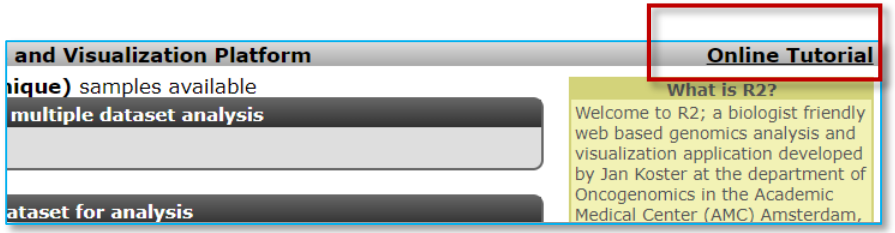
## Tutorial

- Go to a chapter of interest and follow the examples
- Toy around with a similar pipeline on a different dataset
- Perform a follow up analysis from the tutorial

## Datascopes

- Look for datascopes that have similar type of data as yours and look at the kind of analyses that were done for inspiration

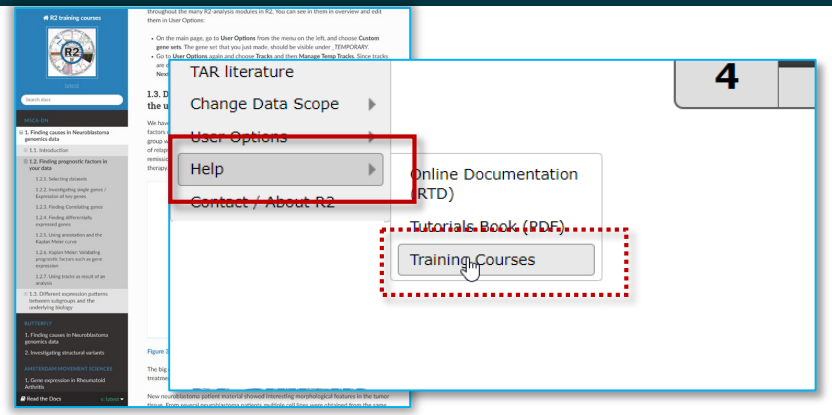


# What's next?

## R2 MSCA-DN Training Courses

- Help => Training Courses => PRESSURE

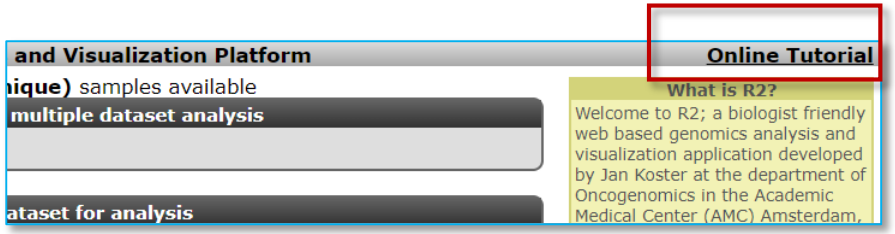



## Do your own research with available datasets in the grid

- Find a dataset of your interest in R2 dataset grid
  - **Tissue/Tumor: "Esoph" => 21 public sets!**
  - **Author: Bijlsma | Zalm (access rights?)**
- Think about a good biological research question
- Try to find the analyses that can help you gain insight

## Tutorial

- Go to a chapter of interest and follow the examples
- Toy around with a similar pipeline on a different dataset
- Perform a follow up analysis from the tutorial




## Next:

- Look into other courses in the course material
- Datascopes
  - Look for datascopes that have similar type of data as yours and look at the kind of analyses that were done for inspiration
- **Find publicly available datasets and email us to get them uploaded**
- **Create you own data and email us to upload**

**r2-support@amsterdamumc.nl**

Amsterdam UMC
University Medical Centers

# http://r2platform.com

**R2 Platform Team**

Romeo Willinge Prins

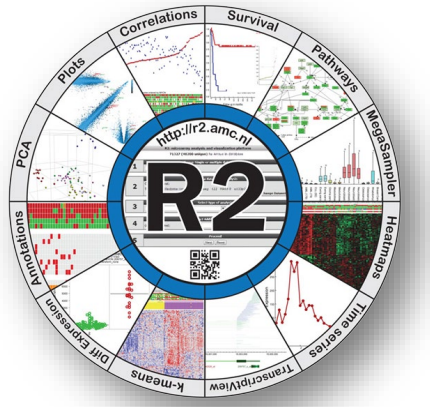
Danny Zwijnenburg

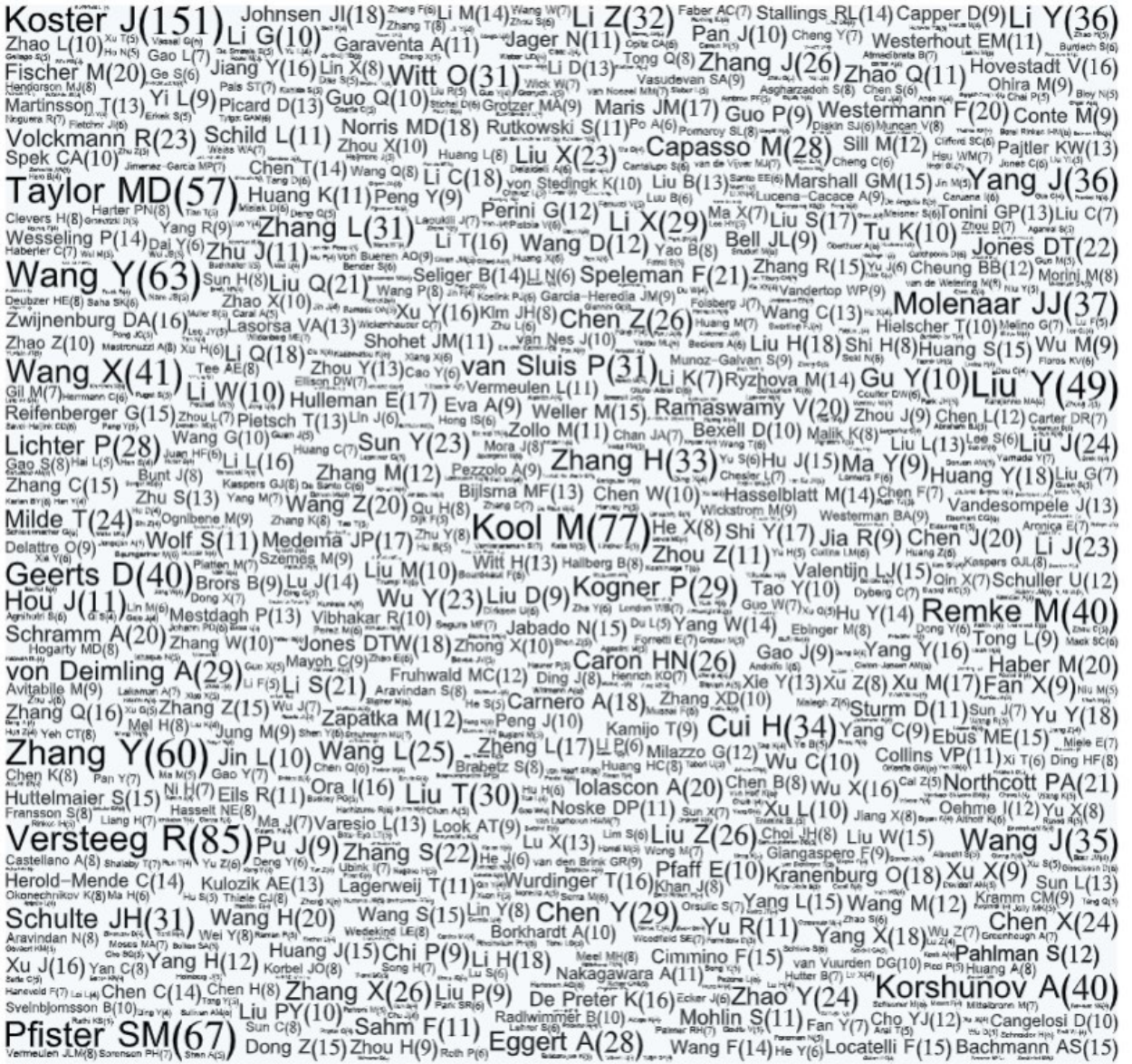
Richard Volckmann

Christian Griffioen

Lieke Hoyng

**Jan Koster**




http://twitter.com/r2_platform
http://facebook.com/r2platform

r2-support@amsterdamumc.nl

Thanks to all authors who used the R2 platform in their research

**R2-Platform**