

RNA sequence data: 308 samples, 20530 genes

Microarray data: 593 samples, 12042 genes

### Groups of interest:

Group 1: survival < 36 months

Group 2: survival > 36 months

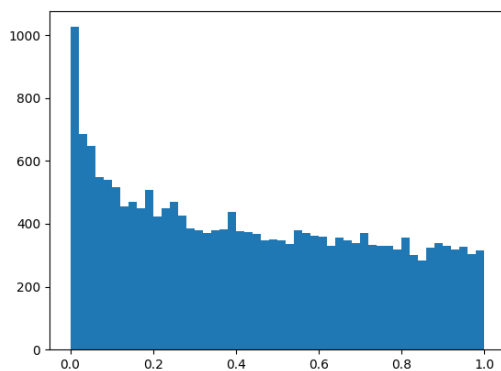
### Statistically significant differentially expressed genes using t-test and Wilcoxon rank-sums test

RNA sequence data

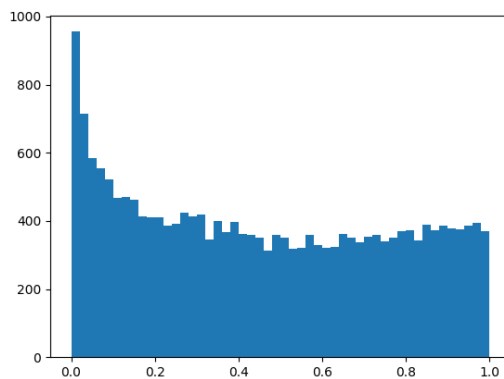
```
SeqData.txt
Total genes selected (t-test, p < 0.05): 2025
Top 10 (t-test)
('CYTH3', 3.884151516771013e-06)
('RGS14', 1.65323597669976e-05)
('TSPAN9', 3.094567611203004e-05)
('FBXO4', 3.244039984591451e-05)
('ATP1A2', 3.502434401748714e-05)
('C12orf5', 5.721209590781841e-05)
('CHUK', 6.058872722509734e-05)
('LOC728606', 6.422763501114352e-05)
('BTRC', 7.285300455202598e-05)
('WAPAL', 8.389441468416214e-05)

Total genes selected (wilcoxon, p < 0.05): 1985
Top 10 (wilcoxon)
('ATP1A2', 7.374516883191402e-06)
('CYTH3', 1.2169375890783899e-05)
('RGS14', 4.247305576581257e-05)
('BTRC', 5.04646345905835e-05)
('LOC728606', 5.656294434655543e-05)
('TEX261', 6.039217830299695e-05)
('SMTNL2', 7.215012531851376e-05)
('ASAP3', 7.533006711227477e-05)
('CAMK2G', 9.532513932786531e-05)
('TSPAN9', 0.00011290979206230389)
```

T-test:



Wilcoxon:

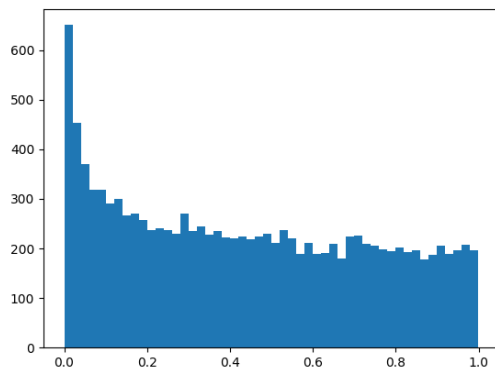


## Microarray data

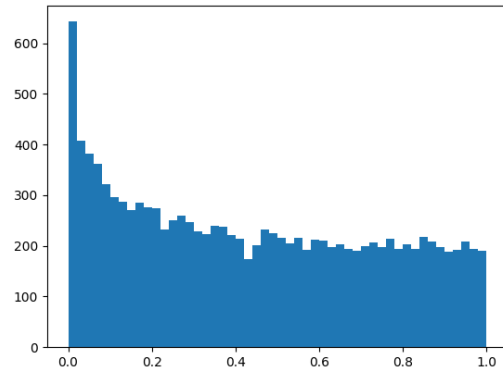
```
ArrayData.txt
Total genes selected (t-test, p < 0.05): 1293
Top 10 (t-test)
('NDRG3', 3.1949594057421897e-06)
('AXL', 3.296830152318338e-06)
('GALNT10', 4.014044131962738e-06)
('NOTCH4', 2.39301174177878e-05)
('RABAC1', 3.0972791623414485e-05)
('OLFML3', 4.9863521898687265e-05)
('SLC37A4', 5.342980011513751e-05)
('PART1', 5.823830536784539e-05)
('LRIG1', 6.301041947880915e-05)
('C10orf56', 9.40988885592525e-05)

Total genes selected (wilcoxon, p < 0.05): 1257
Top 10 (wilcoxon)
('NDRG3', 2.5036917799264073e-07)
('GALNT10', 1.5342589217731108e-05)
('AXL', 2.0333048460213605e-05)
('RABAC1', 3.252723076276983e-05)
('C20orf3', 5.281790018204805e-05)
('ANXA4', 5.2987871344300294e-05)
('CTSA', 5.631749318349664e-05)
('SLC37A4', 7.591666375132883e-05)
('LRIG1', 8.554170012936551e-05)
('AKAP12', 0.00010408266184932014)
```

T-test:



Wilcoxon:



## Multiple comparison correction using Bonferroni correction

```
RNA seq
Selected genes (p < 0.05): 0
```

```
Microarray
Selected genes (p < 0.05): 3
NDRG3 3.1949594057421897e-06
AXL 3.296830152318338e-06
GALNT10 4.014044131962738e-06
```

## False discovery rate of first n significant genes

```
RNA seq
Genes selected: 20
Upper bound of fdr: 0.1635608658327113
Genes selected: 50
Upper bound of fdr: 0.1917984776162816
Genes selected: 100
Upper bound of fdr: 0.244720881525863
Genes selected: 200
Upper bound of fdr: 0.2851315559110384
```

```
Microarray
Genes selected: 20
Upper bound of fdr: 0.12110455187345989
Genes selected: 50
Upper bound of fdr: 0.15422371437862506
Genes selected: 100
Upper bound of fdr: 0.20429195475530018
Genes selected: 200
Upper bound of fdr: 0.25638295262317234
```

## Controlling FDR < 0.05

```
RNA seq
Genes selected: 0
Upper bound of fdr: 0.0797416306393089
```

```
Microarray
12042
2.39301174177878e-05
Genes selected: 3
Upper bound of fdr: 0.01611237314569843
```

**Overlap of significant genes (t-test) in first n significant genes from RNA seq and microarray data**

