

# **Análise da montagem**

Oswaldo Reis Junior

# Métricas

```
#TrinityStats.pl trinity_out_dir.Trinity.fasta
#####
## Counts of transcripts, etc.
#####
Total trinity 'genes': 52637
Total trinity transcripts: 65967
Percent GC: 41.00

#####
Stats based on ALL transcript contigs:
#####

Contig N10: 4962
Contig N20: 3766
Contig N30: 3072
Contig N40: 2547
Contig N50: 2100
```

# Métricas

**Median contig length: 863**

**Average contig: 1317.27**

**Total assembled bases: 86896671**

#####

**## Stats based on ONLY LONGEST ISOFORM per 'GENE':**

#####

**Contig N10: 4269**

**Contig N20: 3231**

**Contig N30: 2596**

**Contig N40: 2121**

**Contig N50: 1770**

**Median contig length: 660**

**Average contig: 1085.56**

**Total assembled bases: 57140728**

# Full-length Transcript

#hit_pct_cov_bin	count_in_bin	>bin_below
100	14397	14397
90	1235	15632
80	1012	16644
70	1162	17806
60	1062	18868
50	875	19743
40	752	20495
30	649	21144
20	418	21562
10	76	21638

# Transrate



## USAGE:

transrate <options>

## OPTIONS:

-a, --assembly=<s>	Assembly file(s) in FASTA format, comma-separated
-l, --left=<s>	Left reads file in FASTQ format
-r, --right=<s>	Right reads file in FASTQ format
-e, --reference=<s>	Reference proteome or transcriptome file in FASTA format
-t, --threads=<i>	Number of threads to use (default: 8)
-m, --merge-assemblies=<s>	Merge best contigs from multiple assemblies into file
-o, --outfile=<s>	Prefix filename to use for CSV output (default: transrate)
-g, --loglevel=<s>	Log level. One of [error, info, warn, debug] (default: info)
-i, --install-deps=<s>	Install any missing dependencies. One of [all, read, ref]
-x, --examples	Show some example commands with explanations
-v, --version	Print version and exit
-h, --help	Show this message

# Transrate

fragments	41,525,052
fragments mapped	40,384,985
p fragments mapped	0.97
good mappings	39,072,068
p good mapping	0.94
bad mappings	1,312,917
potential bridges	33,973
bases uncovered	10,828,407
p bases uncovered	0.12
contigs uncovbase	36,839
p contigs uncovbase	0.56
contigs uncovered	7,406
p contigs uncovered	0.11
contigs lowcovered	32,381
p contigs lowcovered	0.49
contigs segmented	7,072
p contigs segmented	0.11

n seqs	65,967
smallest	224
largest	16,592
n bases	86,896,671
mean len	1,317.27
n under 200	0
n over 1k	30,021
n over 10k	36
n with orf	33,246
mean orf percent	56.92
n90	558
n70	1,385
n50	2,100
n30	3,072
n10	4,962
gc	0.41
gc skew	0.01
at skew	0
cpg ratio	1.6
bases n	0
proportion n	0
linguistic complexity	0.22

TRANSRATE ASSEMBLY SCORE	364
TRANSRATE OPTIMAL SCORE	602
TRANSRATE OPTIMAL CUTOFF	0.4111
good contigs	48330
p good contigs	0.73

# Transdecoder

<https://transdecoder.github.io/>

# Trinotate

<https://trinotate.github.io/>



# Gene Ontology Enrichment

[http://trinityrnaseq.github.io/analysis/run\\_GOseq.html](http://trinityrnaseq.github.io/analysis/run_GOseq.html)

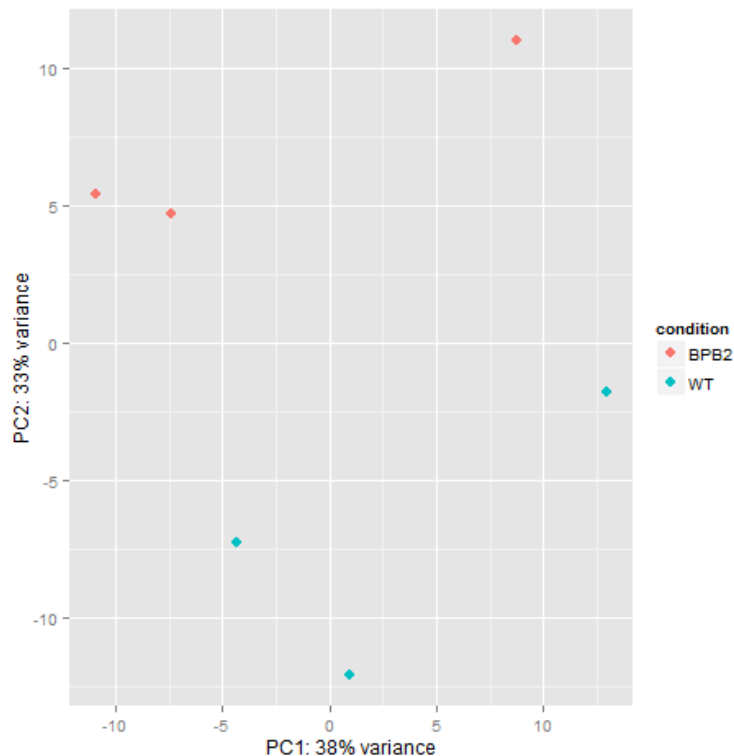
# Abundance Estimation

[http://trinityrnaseq.github.io/analysis/abundance\\_estimation.html](http://trinityrnaseq.github.io/analysis/abundance_estimation.html)

# Differential Expression Analysis

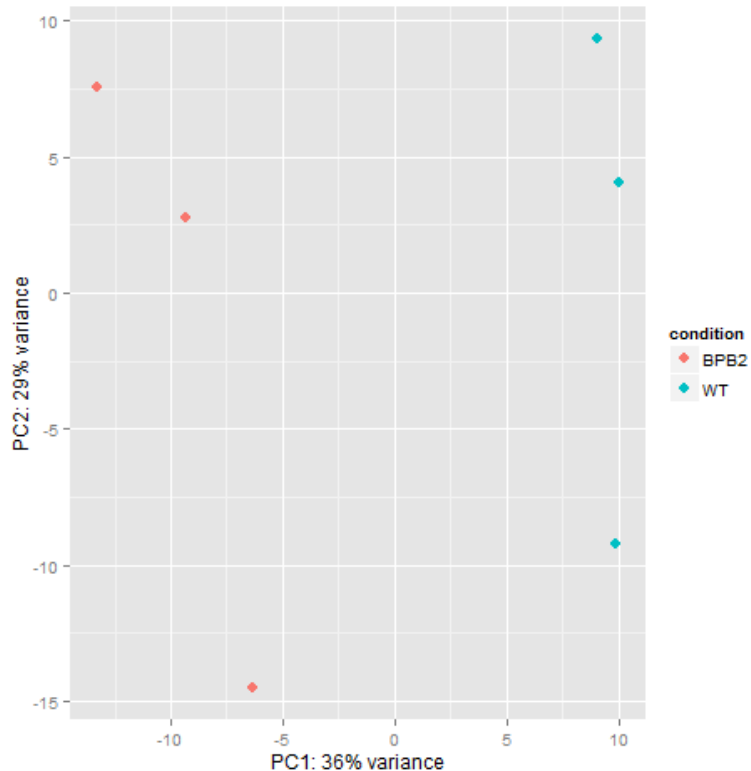
[http://trinityrnaseq.github.io/analysis/diff\\_expression\\_analysis.html](http://trinityrnaseq.github.io/analysis/diff_expression_analysis.html)

# Original



out of 31419 with nonzero total read count  
adjusted p-value < 0.05  
LFC > 0 (up) : 627, 2%  
LFC < 0 (down) : 571, 1.8%  
outliers [1] : 447, 1.4%  
low counts [2] : 9142, 29%  
(mean count < 35.5)

# Trinity



out of 64489 with nonzero total  
read count

adjusted p-value < 0.1

LFC > 0 (up) : 1017, 1.6%

LFC < 0 (down) : 778, 1.2%

outliers [1] : 758, 1.2%

low counts [2] : 41347, 64%

(mean count < 65.8)