**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

# Single Cell Service Data Analysis Report

PO Number:

Customer first name and last name:

Customer company/ institute:

Date of report (DD/MM/YY):

Operator (initials):

## Pre-alignment QC

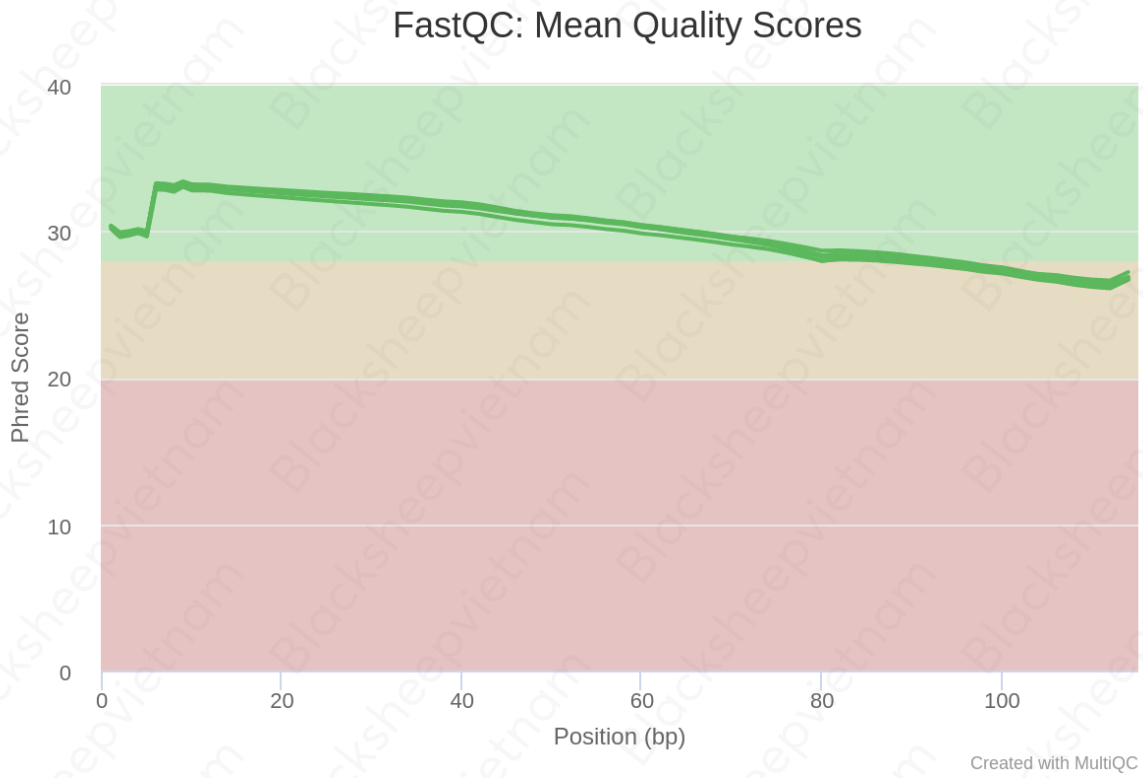


Figure 1. FASTQC - Mean quality score

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

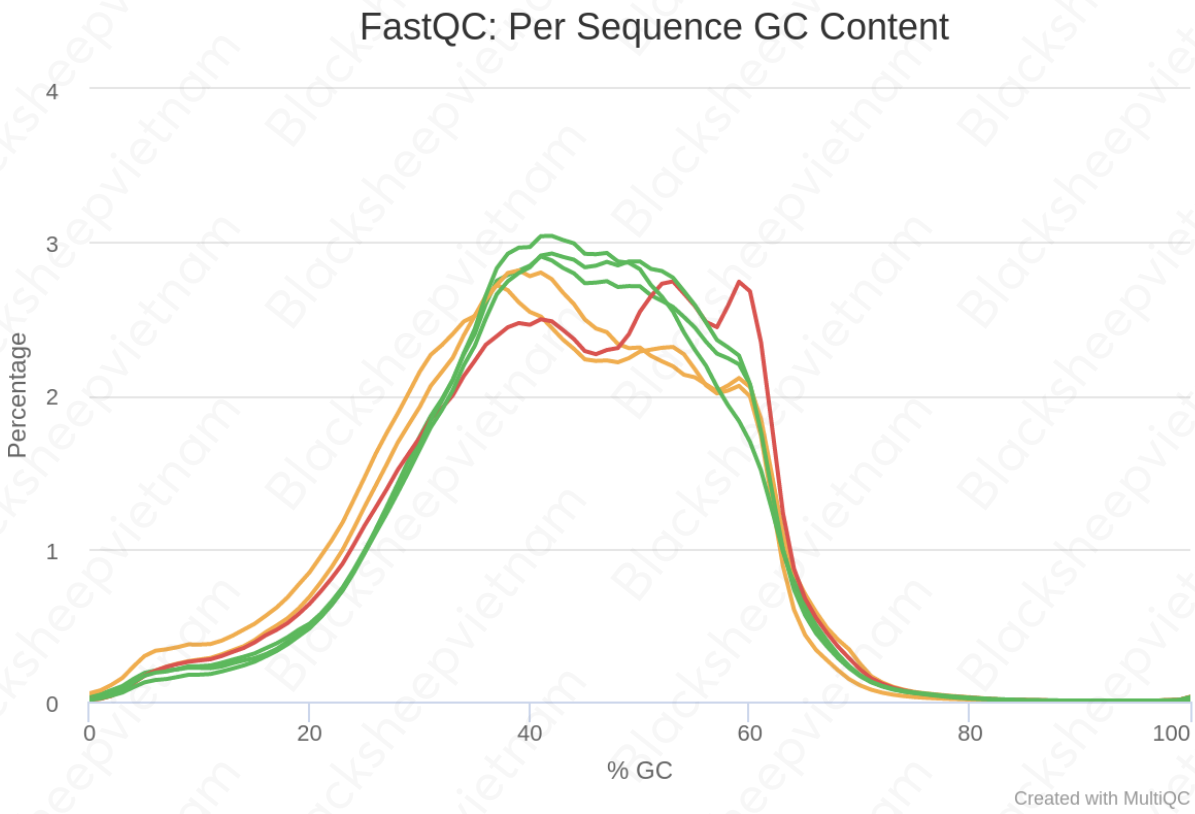


Figure 2. FASTQC - Per sequence GC content

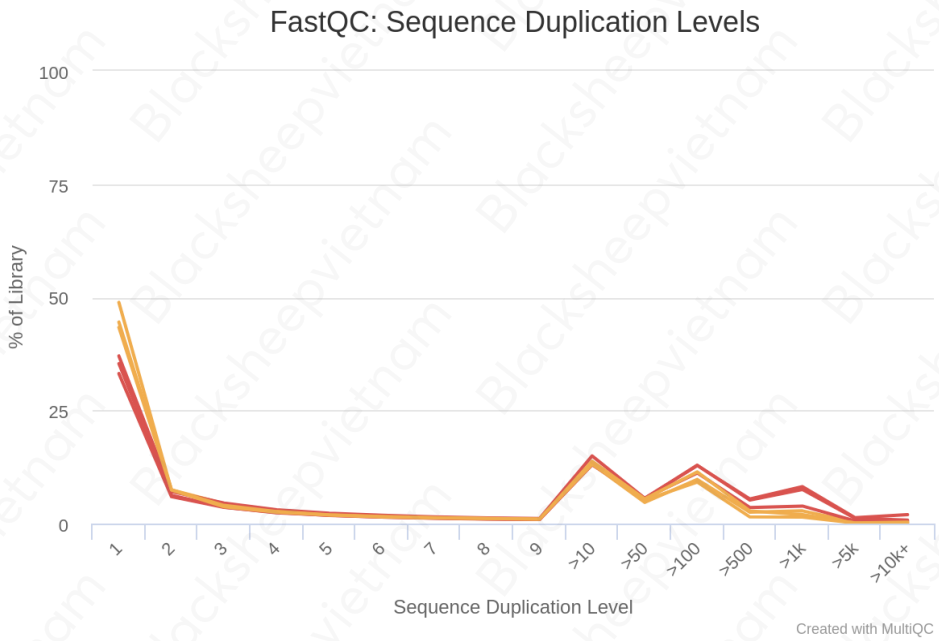


Figure 3. FASTQC - Sequence Duplication levels

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

Table 1. Pre-alignment raw read QC summary

| Sample(s) | Number of reads | %sequences with Phred score >=20 | GC content (%) | % duplication (%) | % trimmed reads (%) |
|---|---|---|---|---|---|
| CG-FK_S1_R2 | 24295502 | 98.65 | 43 | 47.6 | 16.1 |
| LGC-FK_S2_R2 | 29195716 | 98.98 | 40 | 55.0 | 24.2 |
| CG-DW_S3_R2 | 34793649 | 98.92 | 42 | 57.9 | 23.0 |
| LGC-DW_S4_R2 | 43112271 | 99.02 | 43 | 60.3 | 27.9 |
| UB2-control_S5_R2 | 15115431 | 98.99 | 42 | 49 | 14.2 |
| UB2-uberstrainer_S6_R2 | 11872831 | 99.09 | 43 | 43.7 | 17.5 |

Conclusion regarding Phred score: Large proportion (over 98%) of all samples has phred score over 20 meaning most sequences have base-call accuracy over 99%.

GC content apparent contamination:

- According to Figure 2, per sequence GC content of only sample 1, 5 and 6 follow the normal distribution pattern, which indicates potential library contamination or some kind of bias.
- According to Table 1, 40 to 43% GC content of 6 samples fall into the range of normal GC content in human genome, which is about 30 to 60 percent (https://www.nature.com/articles/35057062#Sec31)

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

# Alignment QC

Alignment was performed against [name of species] - [genome or transcriptome]

Reads were mapped with [name of alignment tool, e.g. STAR, Kallisto, Rapmap]

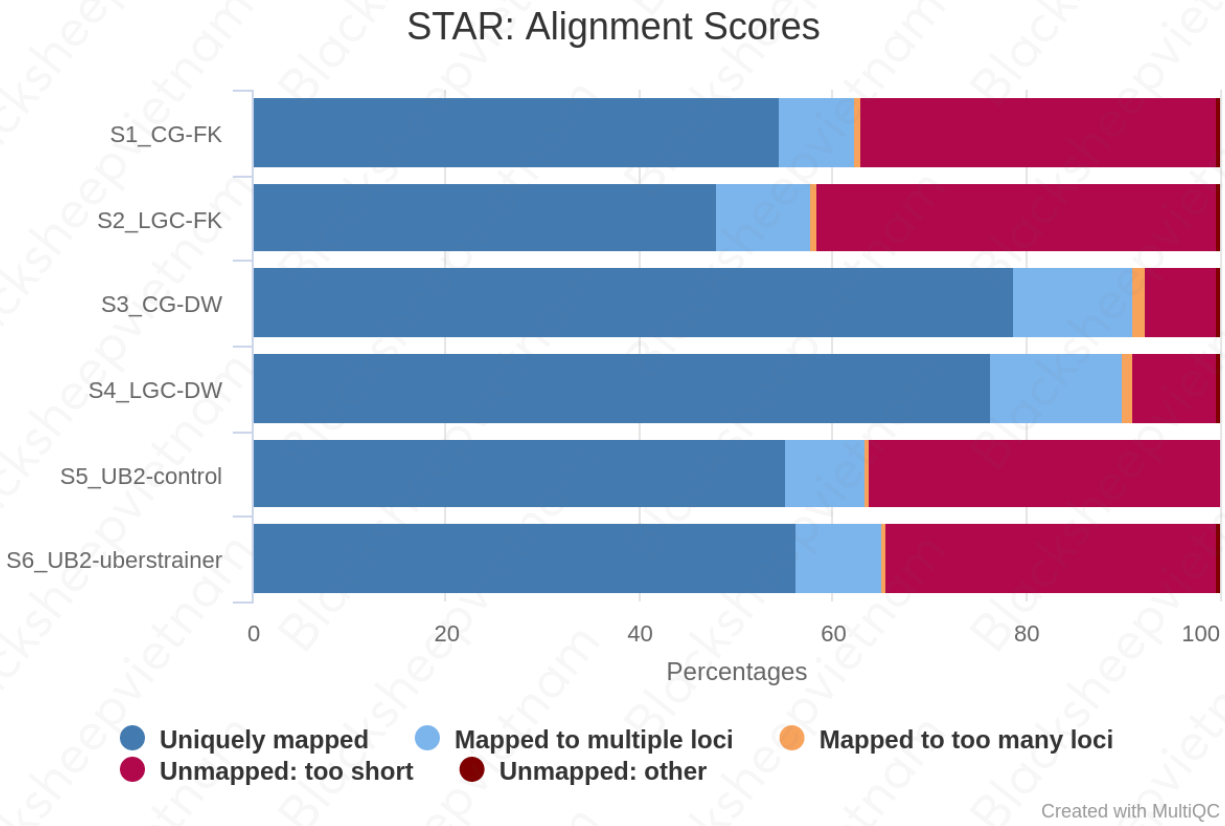


Figure 4. Alignment summary. Percentage of uniquely mapped reads: CG-FK_S1: 54.5%; LGC-FK_S2: 48.1%; CG-DK_S3: 78.8%; LGC-DW_S4: 76.5%; UB2-control_S5: 55.3%; UB2-uberstrainer_S6: 56.4%.

Percentage of uniquely mapped reads: LGC-FK_S2 sample is the only sample that has %uniquely mapped reads under 50% (specifically, 48.1%) [number, ideally >50] %

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

# Post-alignment QC

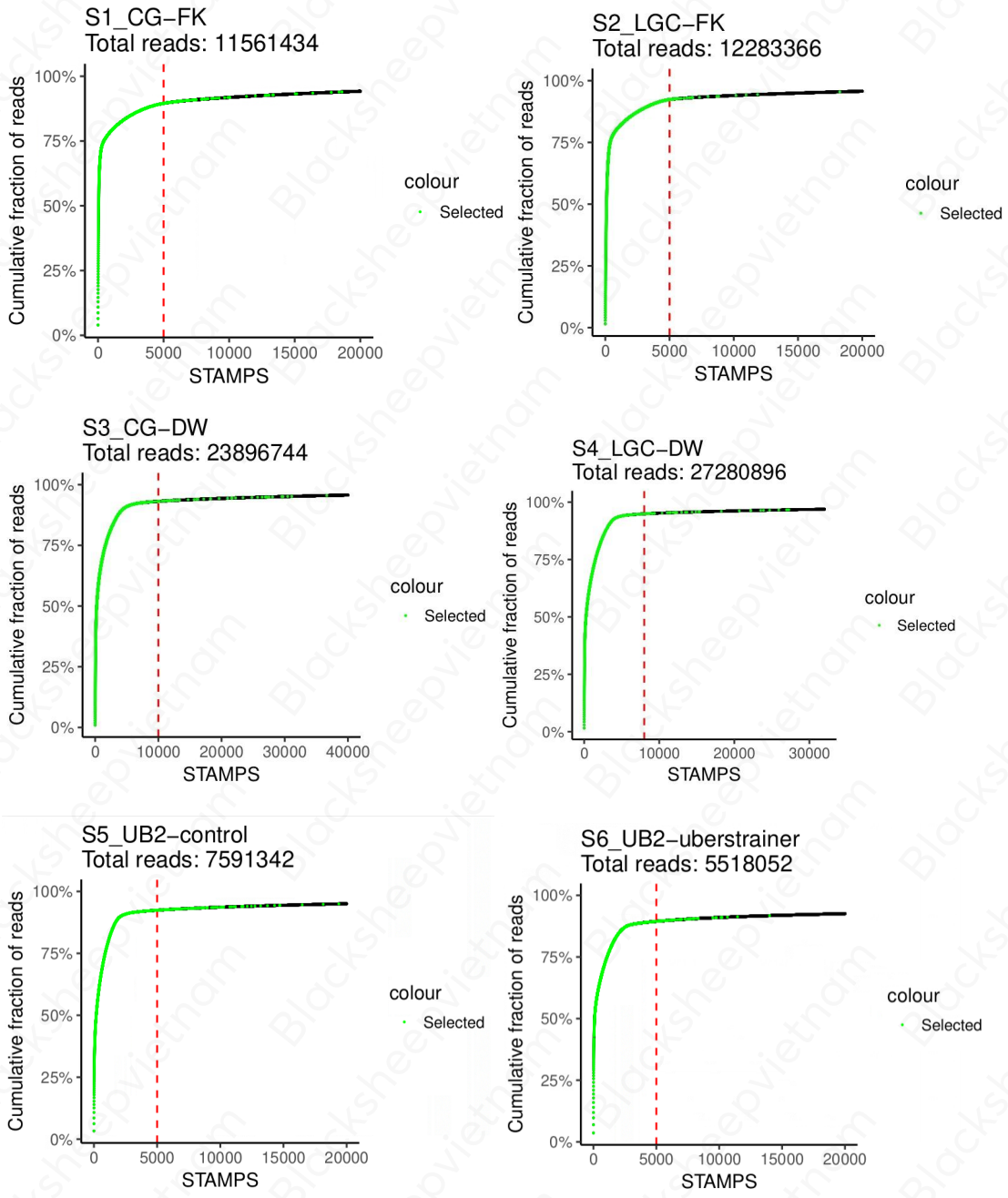


Figure 5. Knee-plot, from which the number of captured single-cell transcriptomes can be deducted. Green dots represent chosen STAMPs

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

Number of captured single-cell transcriptomes based on knee-plot:

Table 2. Post-alignment QC summary

| Sample(s) | Number of STAMPs* | Number of reads/ cell | Median number of UMIs/ cell | Median number of genes | Number of UMI/ genes |
|---|---|---|---|---|---|
| CG-FK_S1 | 5000 | 287 | 244.18 | 34 | 1.14 |
| LGC-FK_S2 | 5000 | 307.5 | 275.47 | 41 | 1.14 |
| CG-DW_S3 | 10000 | 835 | 496.56 | 147 | 1.22 |
| LGC-DW_S4 | 8000 | 1494.5 | 537.02 | 194 | 1.22 |
| UB2-control_S5 | 5000 | 643 | 578.31 | 187 | 1.29 |
| UB2-uberstrainer_S6 | 5000 | 473 | 642.74 | 196 | 1.27 |

*STAMP = Single-cell transcriptome attached to micro-particle. i.e. single-cell transcriptome captured on an mRNA capture bead.

Expressions vary vastly between genes and cells.

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

# Data visualization

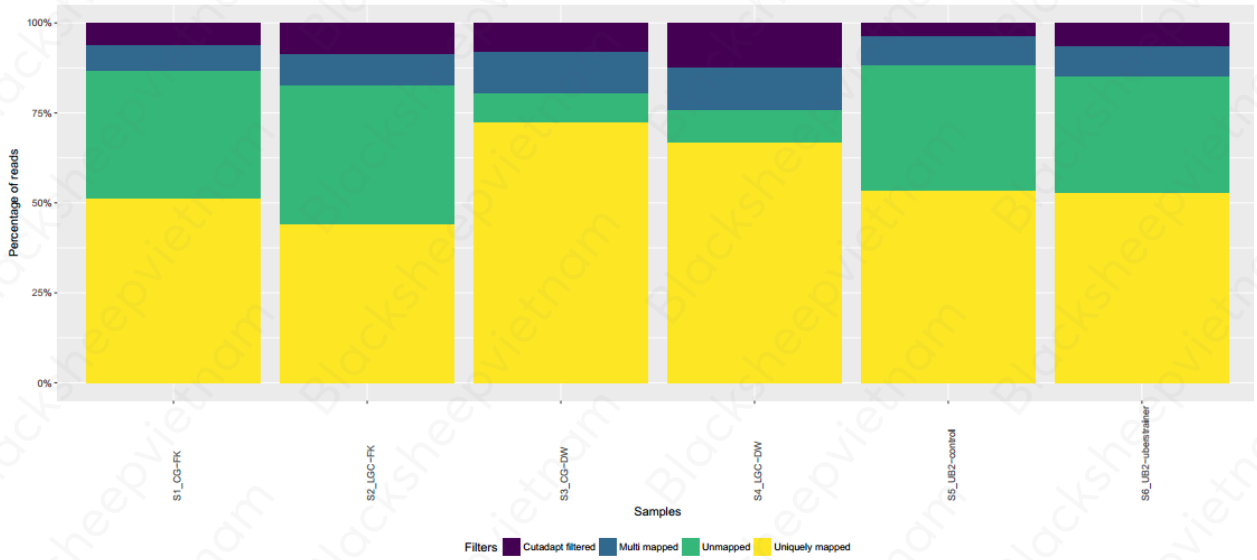


Figure 6. Yield graph

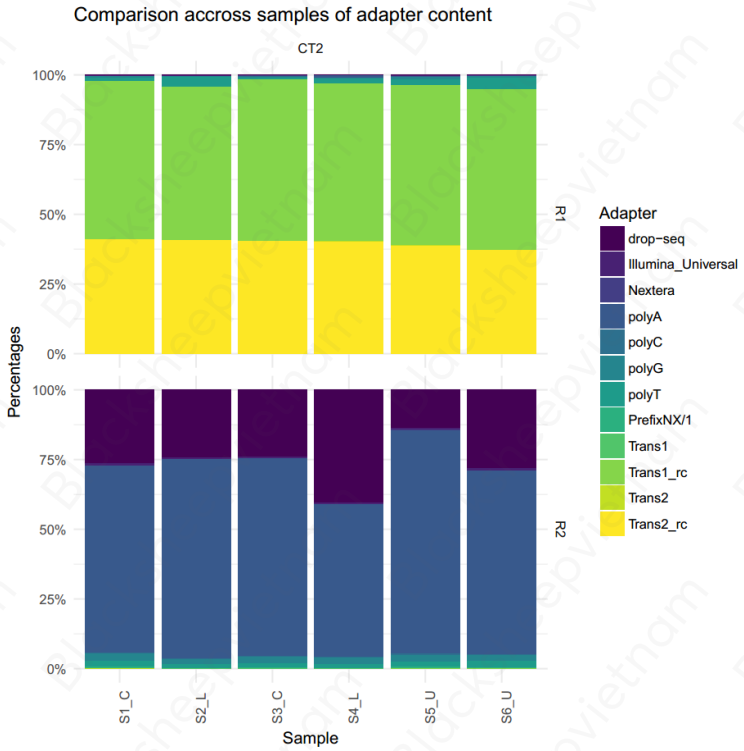


Figure 7. Adapter content

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

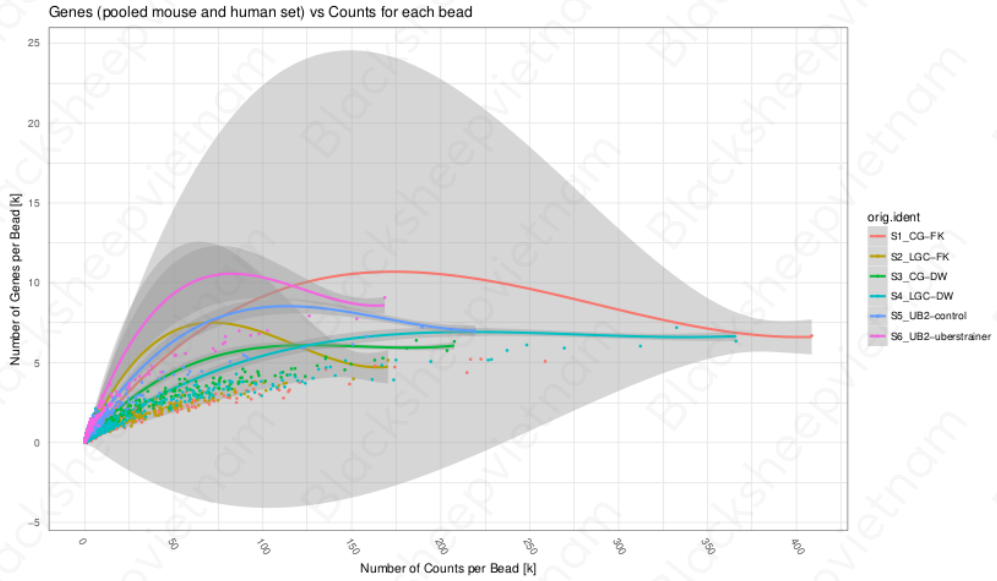


Figure 8. Read count vs genes

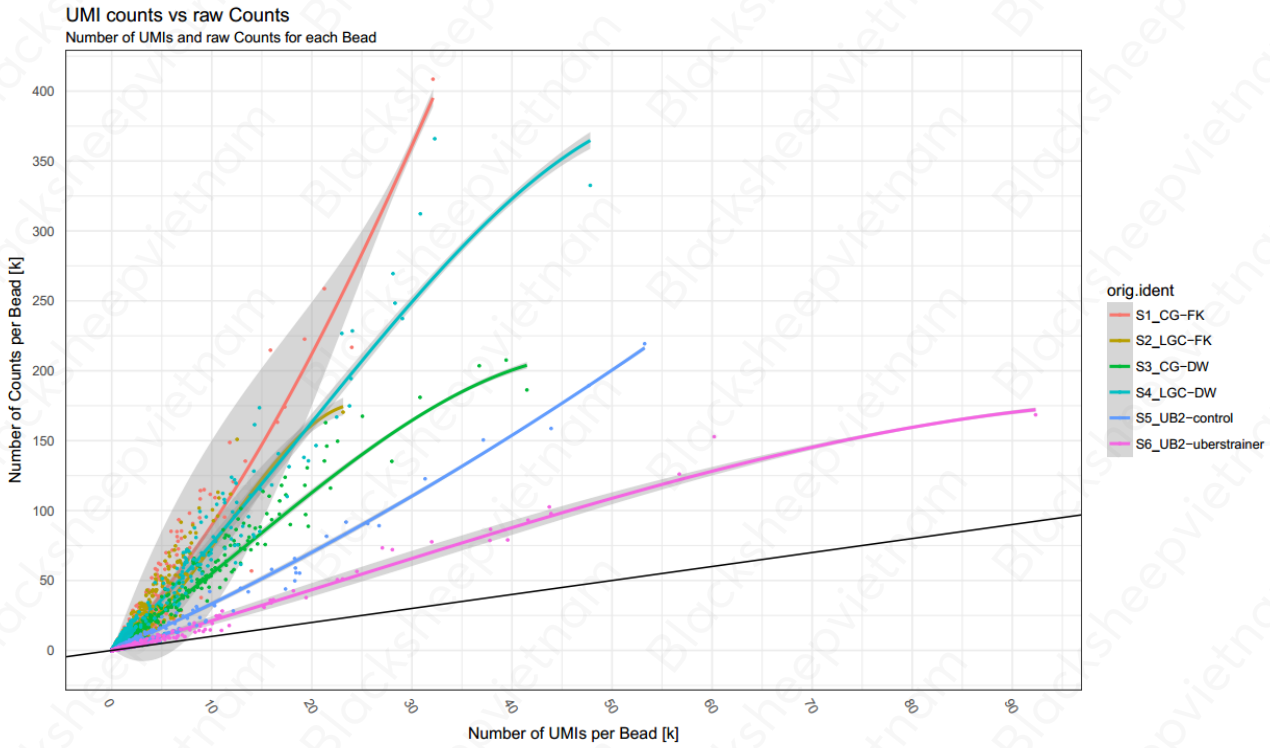


Figure 9. UMI vs read count

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

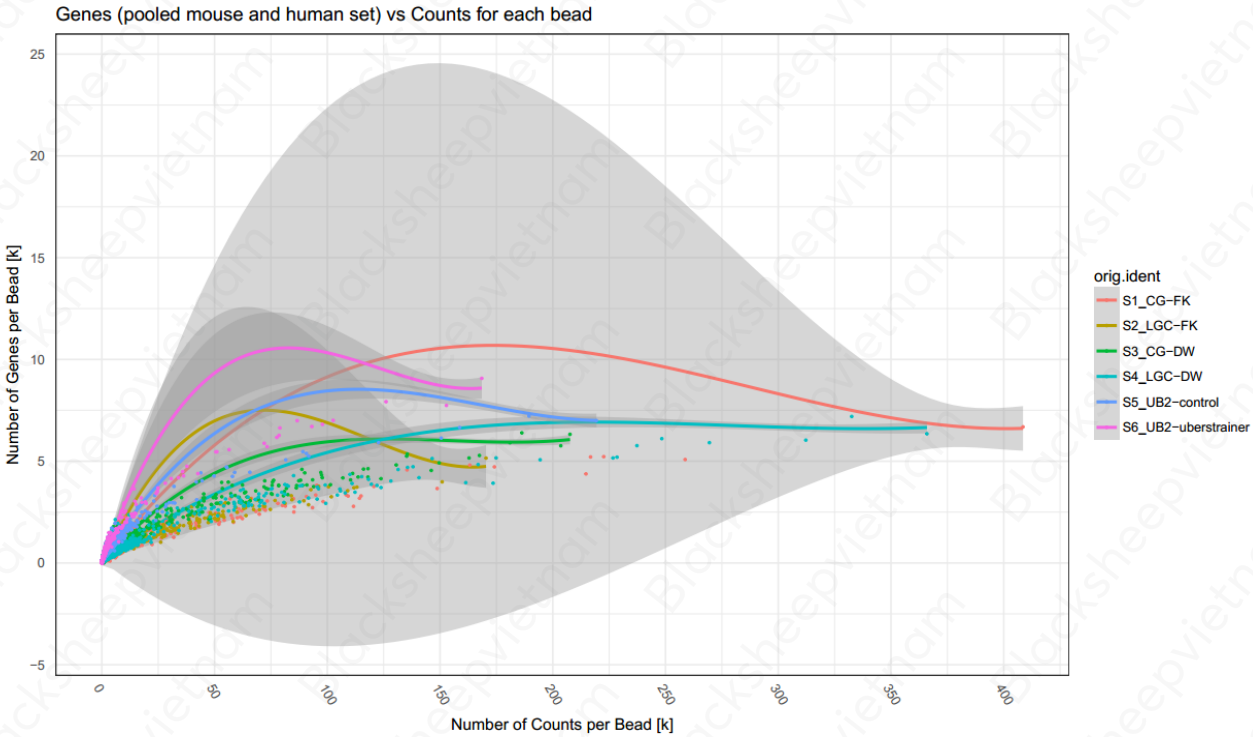


Figure 10. UMI vs gene count

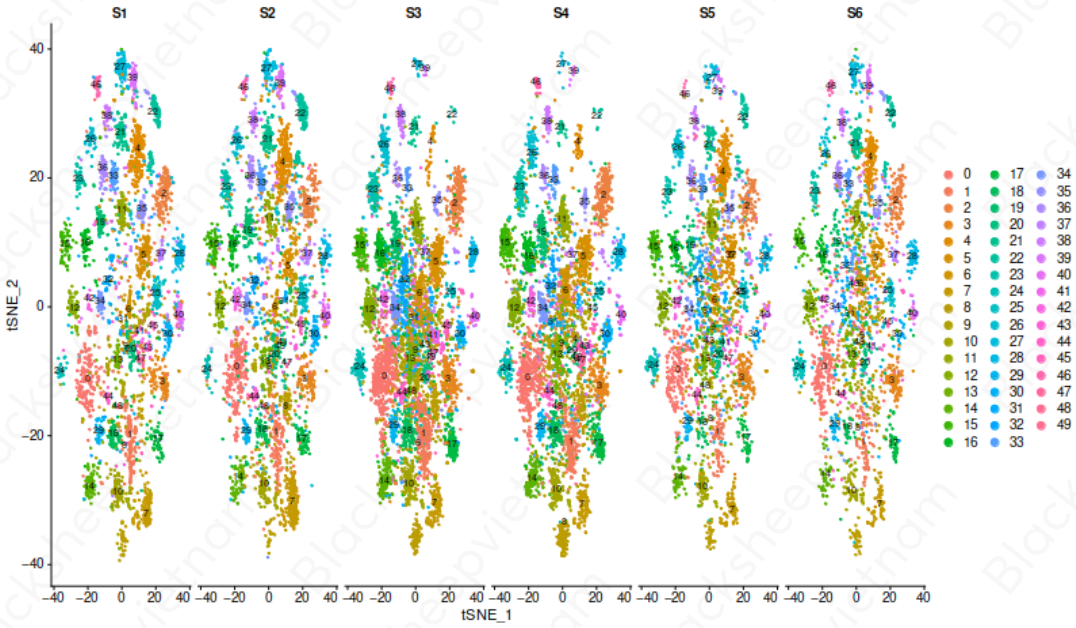


Figure 11. Clustering by t-SNE, split by Sample

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

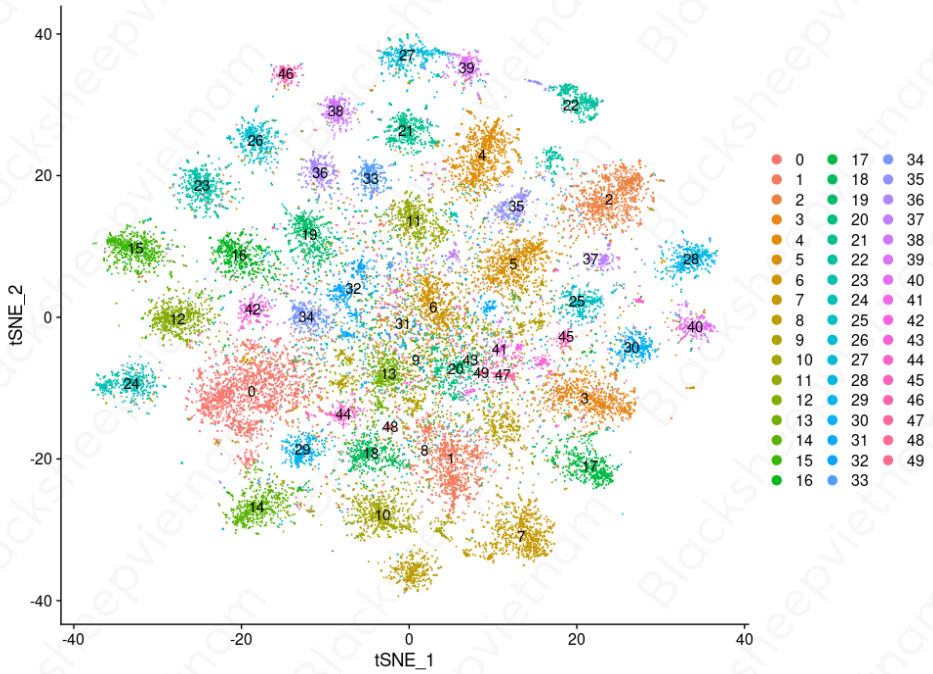


Figure 12. Clustering by t-SNE, all 6 samples

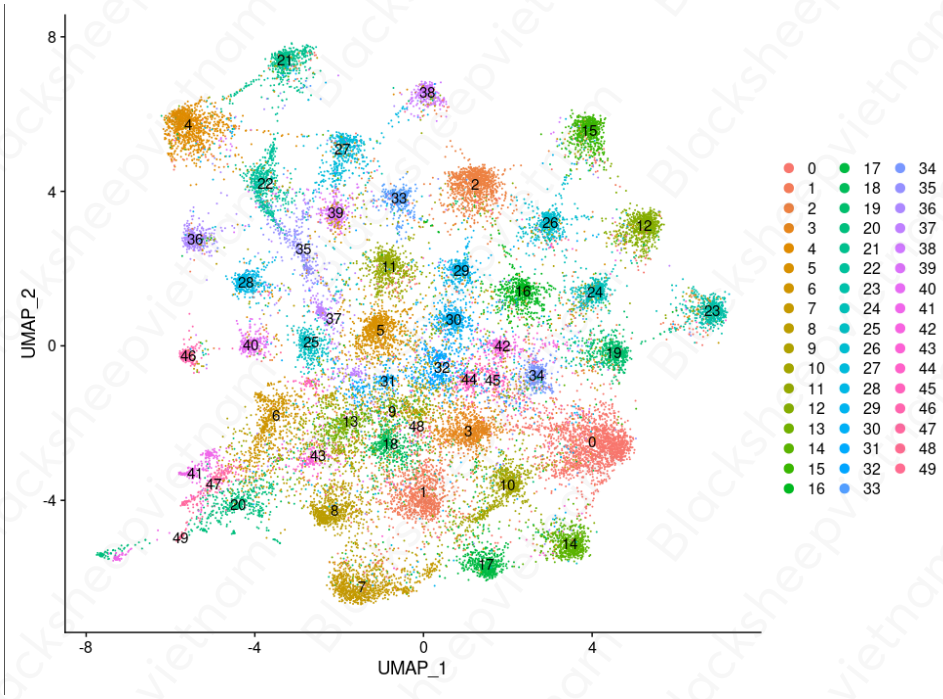


Figure 13. Clustering by UMAP, all 6-samples

**BLACKSHEEP CONSULTING PTE LTD**
30 Cecil Street, #19-08 Prudential Tower
Singapore 04972
M: admin@blacksheeppower.com
W: blacksheeppower.com

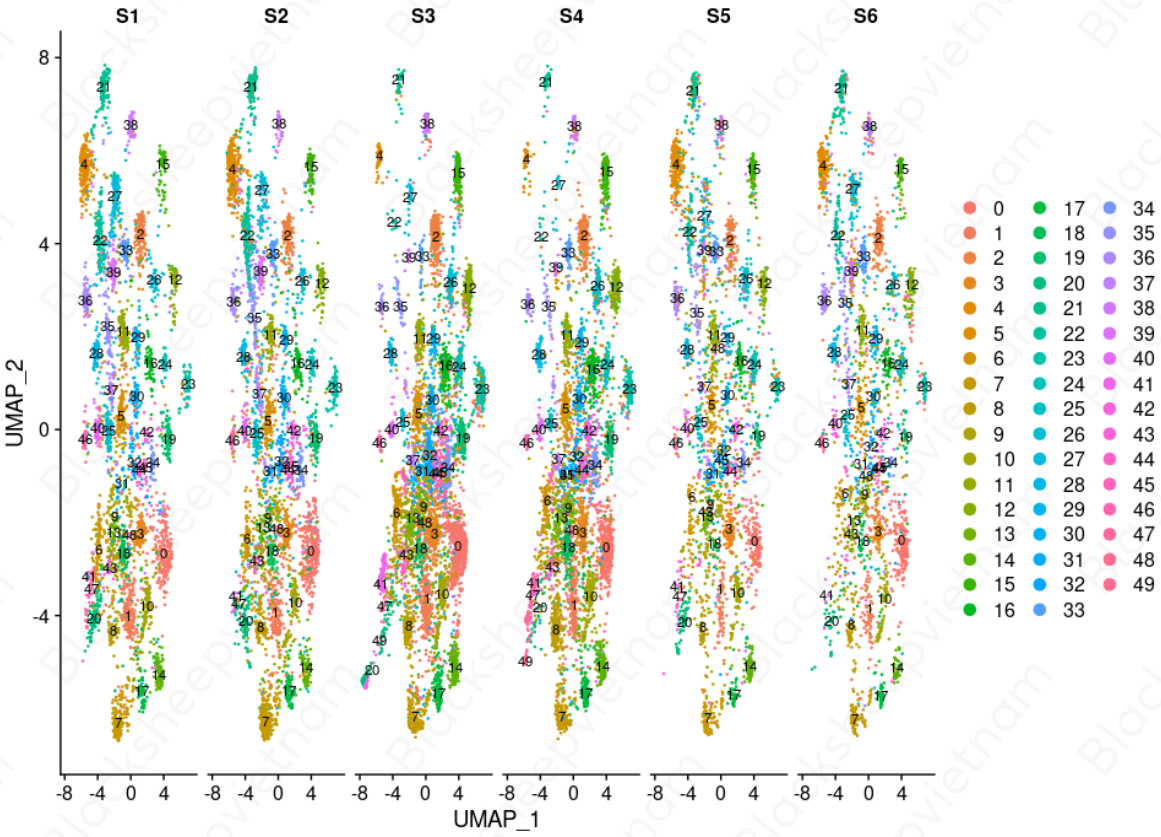


Figure 14. CLustering by UMAP, split by samples