

SVA Recount

Christopher Lo

9/19/2022

Introduction

What happens to the number of Surrogate Variables as we increase the number of studies in recount3?

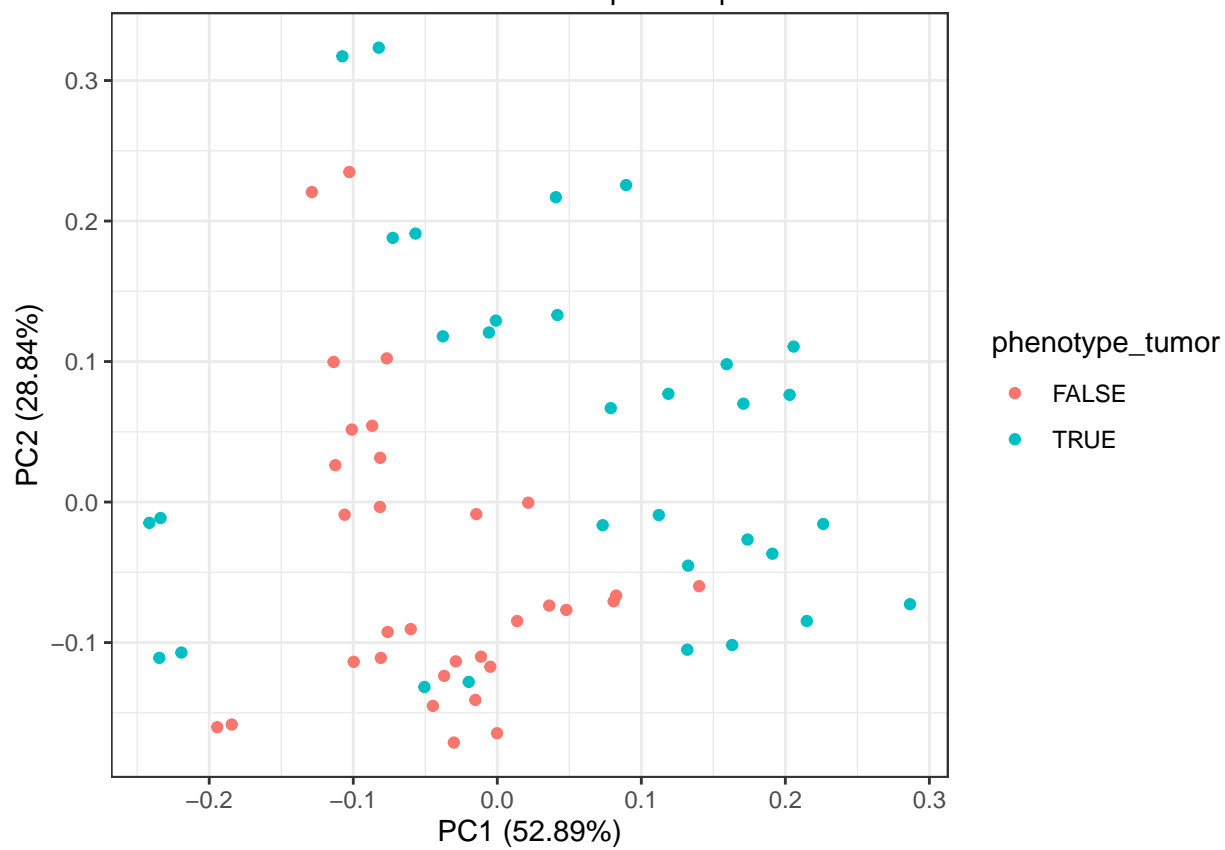
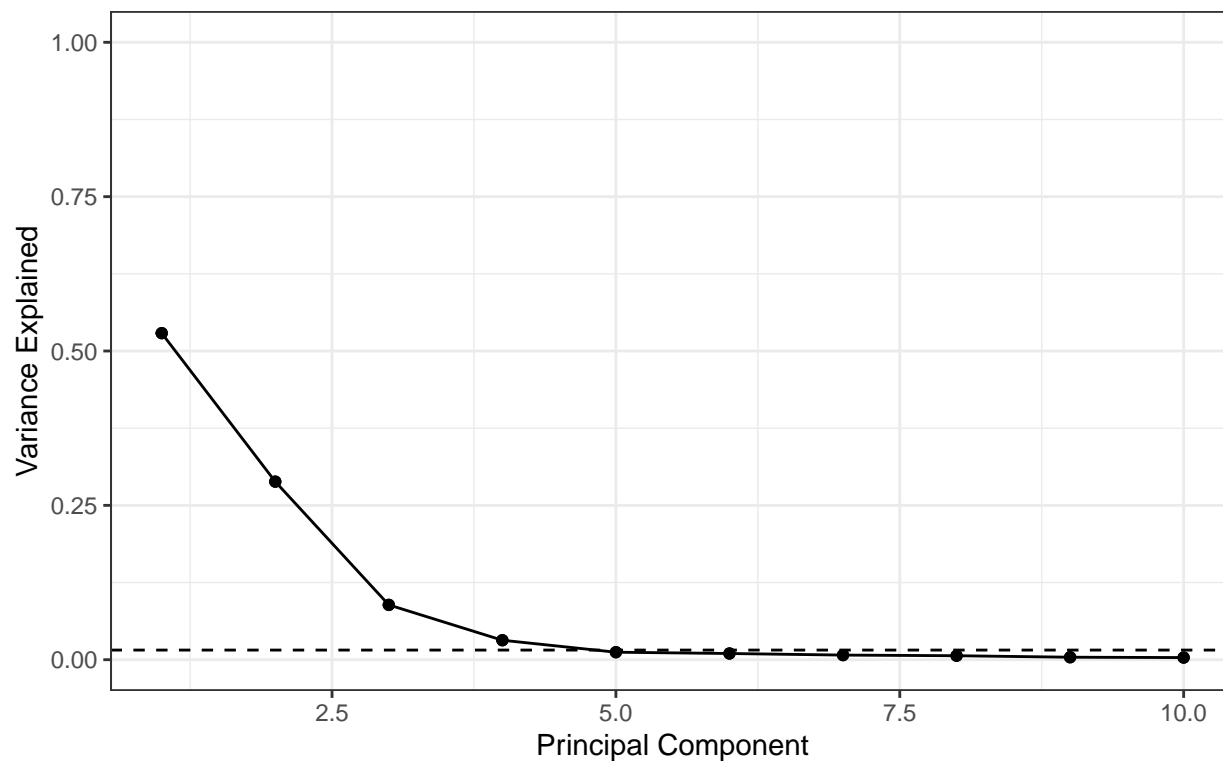
To illustrate this, we examined 4 studies that involved case/control prostate tumor and normal samples, with a minimal of 25 samples per study. For each study, we examined the number of SVs (using method = "be"), and we also aggregated them one by one and examined the number of SVs.

1. SRP118614: "Overall design: Matched high-grade (GS=7(4+3)) prostate tumor and adjacent normal specimens from 16 patients (8 AAM and 8 EAM) were subjected to two replicate runs of RNA-sequencing."
2. SRP002628: "Overall design: We sequenced the transcriptome (polyA+) of 20 prostate cancer tumors and 10 matched normal tissues using Illumina GAII platform. Then we used bioinformatic approaches to identify prostate cancer specific aberrations which include gene fusion, alternative splicing, somatic mutation."
3. SRP212704: "Overall design: Strand specific total RNA seq was performed using frozen patient matched prostate cancer tissue in biological duplicates. Purpose: The goal of present study is to compare transcript level changes between normal and tumor of same individuals"
4. SRP027258: "We utilized RNA sequencing to test the hypothesis that SFN modifies the expression of genes that are critical in prostate cancer progression. Normal prostate epithelial cells, and androgen-dependent and androgen-independent prostate cancer cells were treated with 15 μ M SFN and the transcriptome was determined at 6 and 24 hour time points."

Analysis

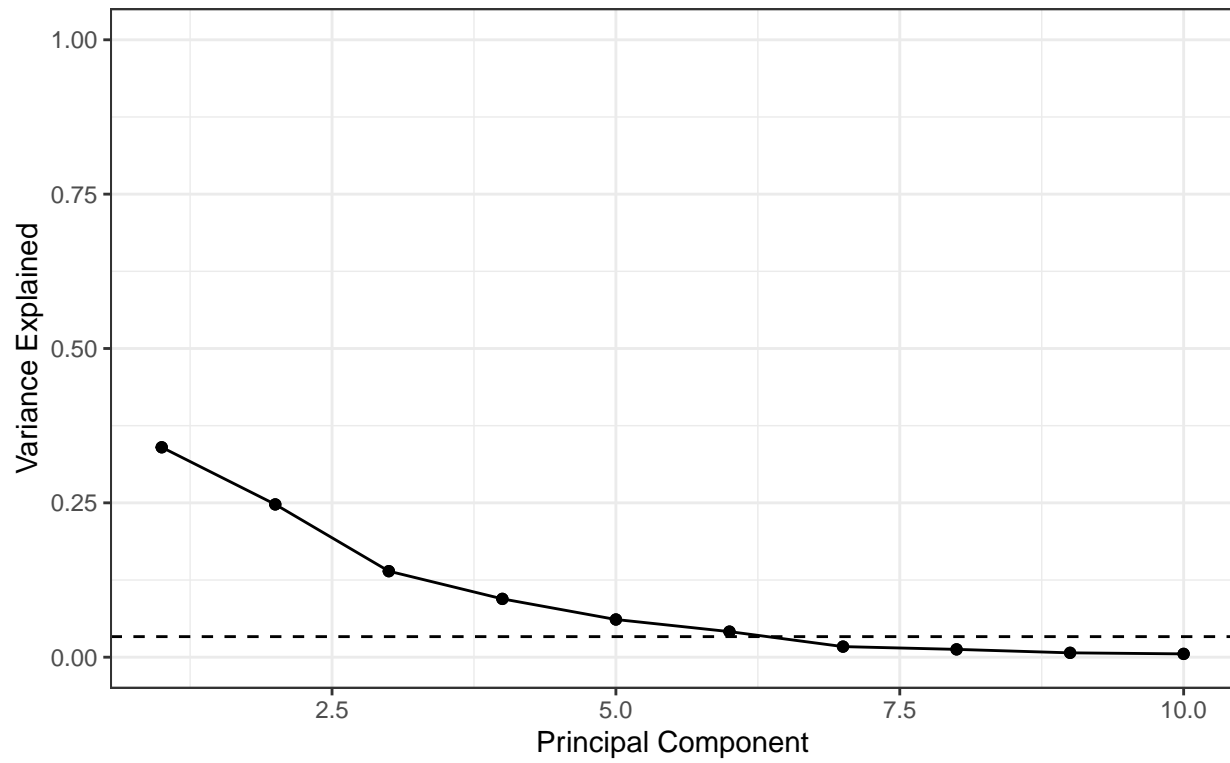
```
## [1] "SizeFactor distribution:"  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
## 0.4696 0.7496  1.0697  1.1054  1.4532  2.0903
```

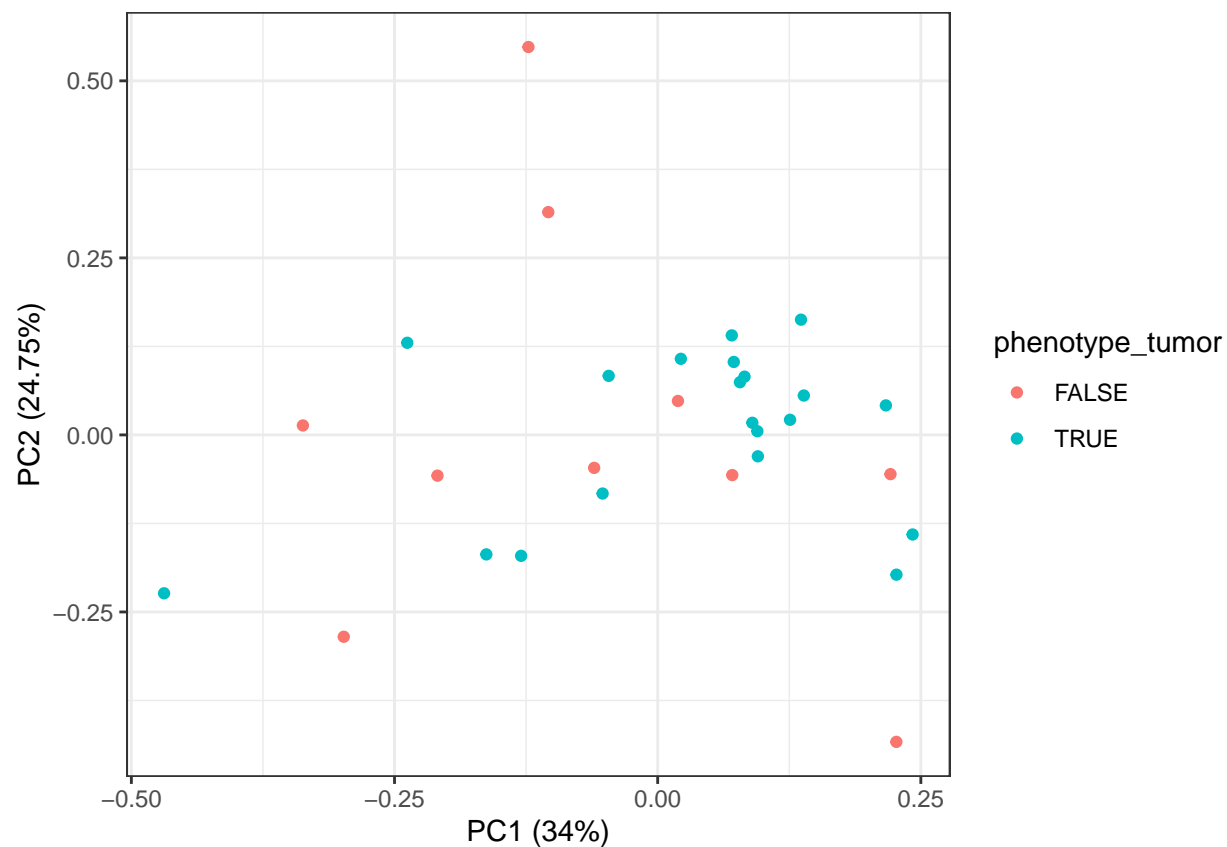
SRP118614
Number of SVs: 3



```
## [1] "SizeFactor distribution:"  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
## 0.2691 0.4948  1.4120  1.2100  1.8021  1.9976
```

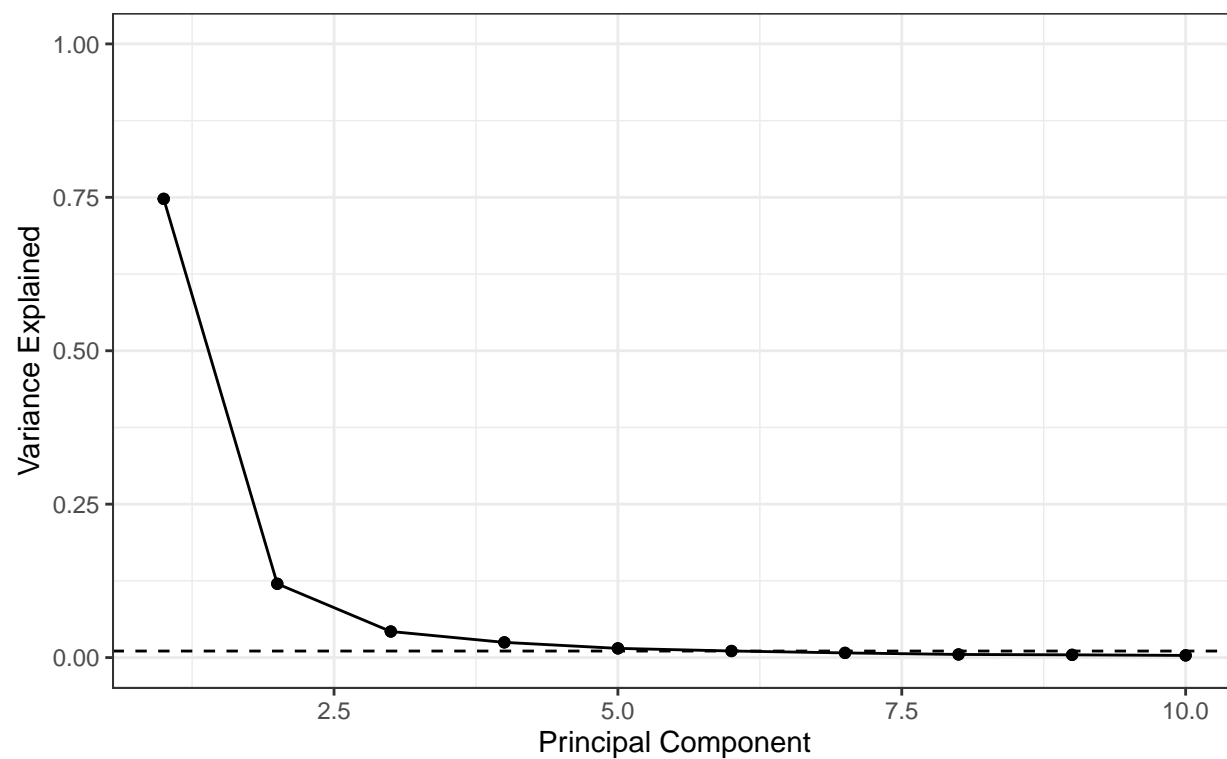
SRP002628
Number of SVs: 4

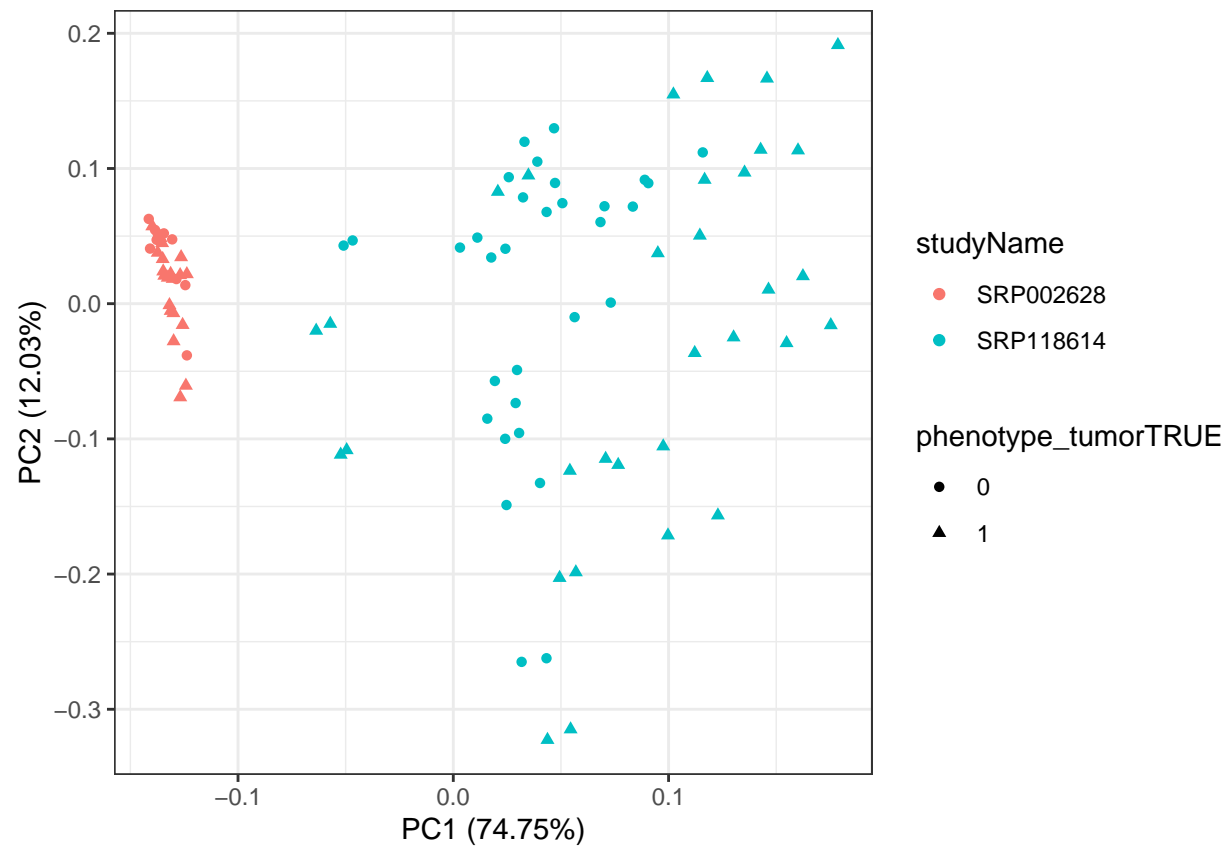




SRP118614 + SRP002628

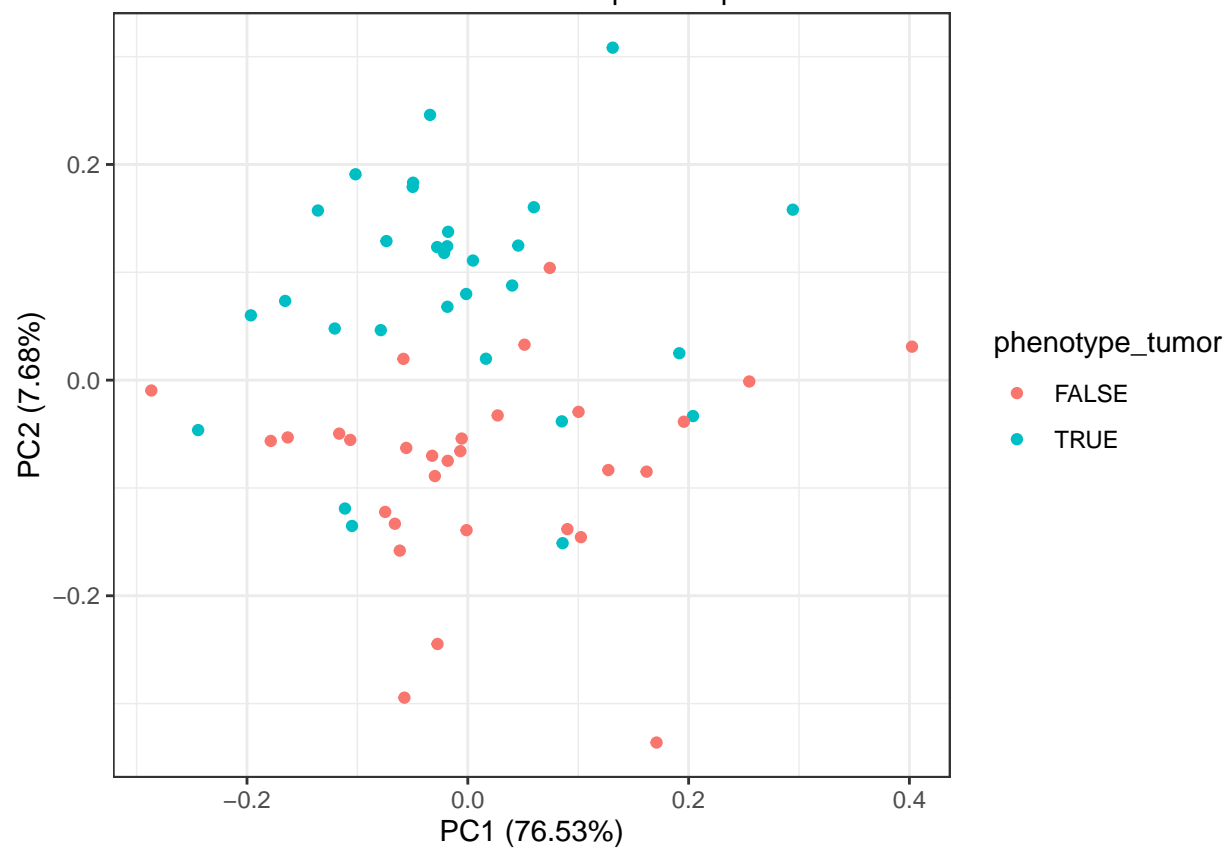
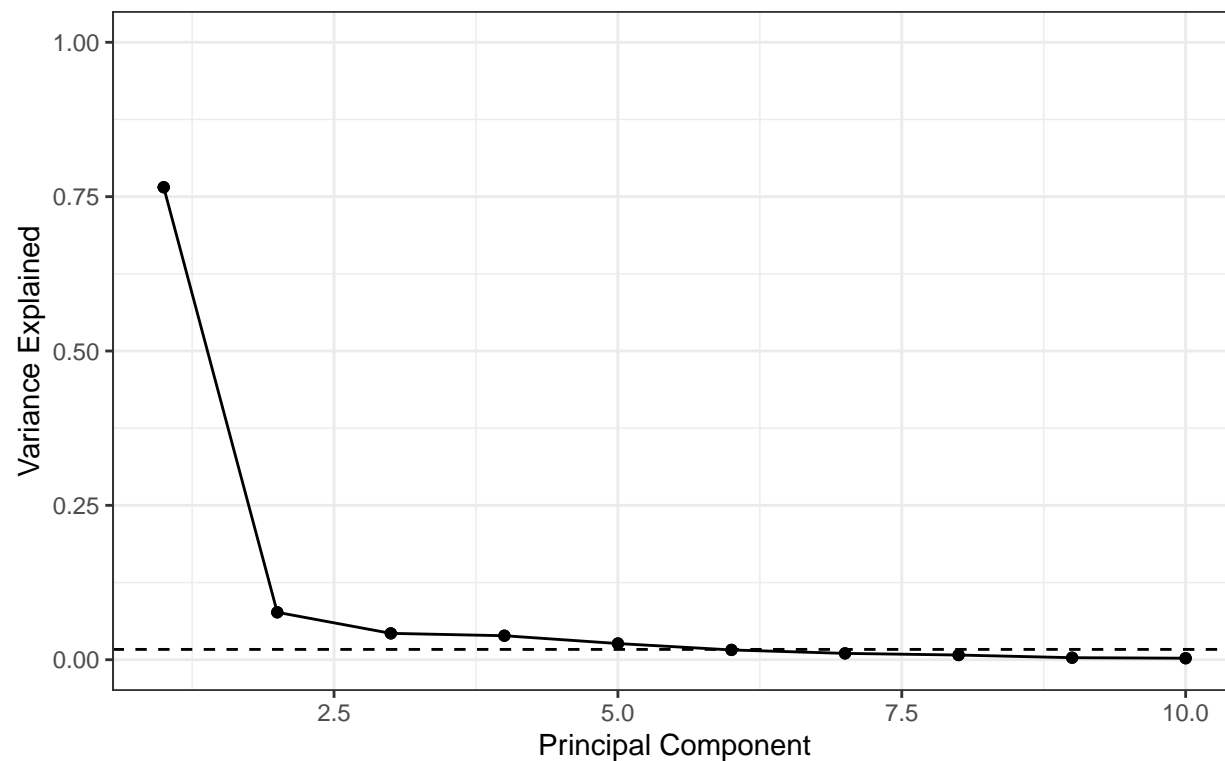
Number of SVs: 2





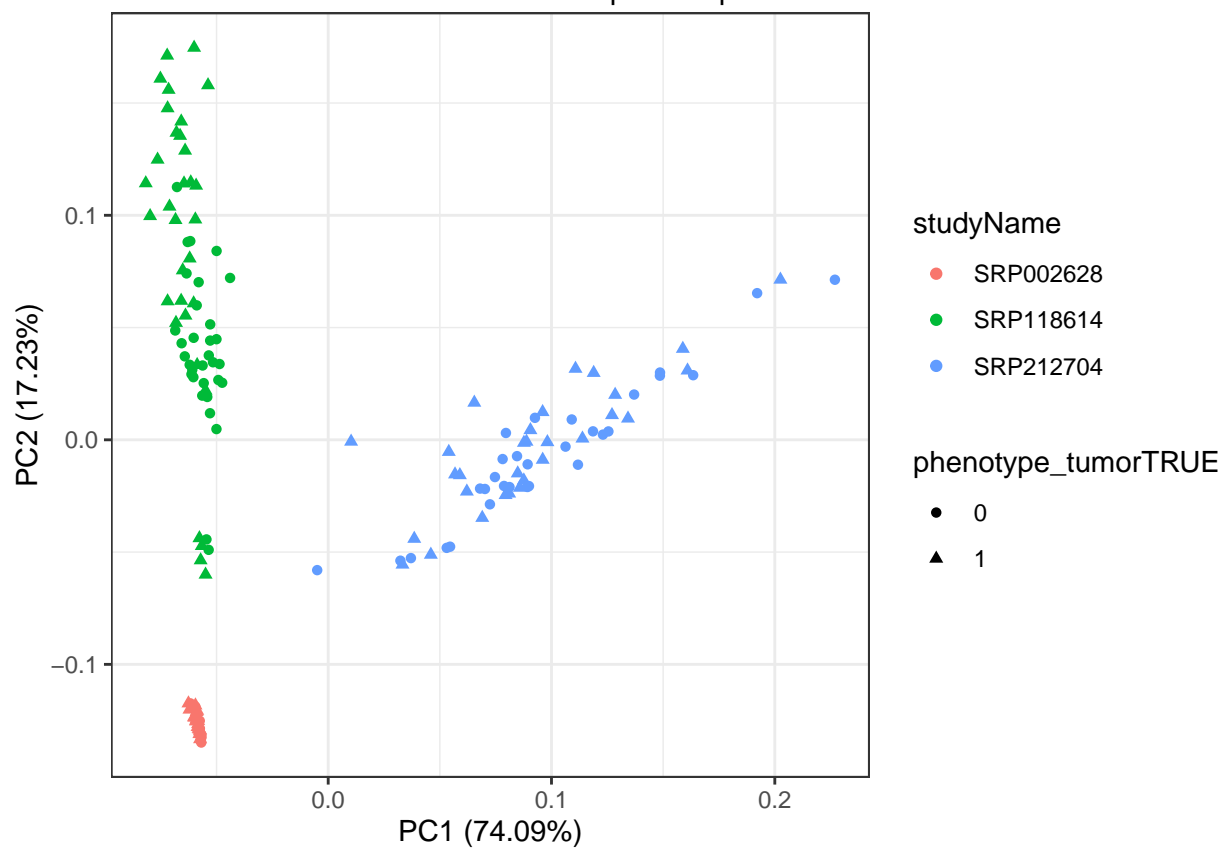
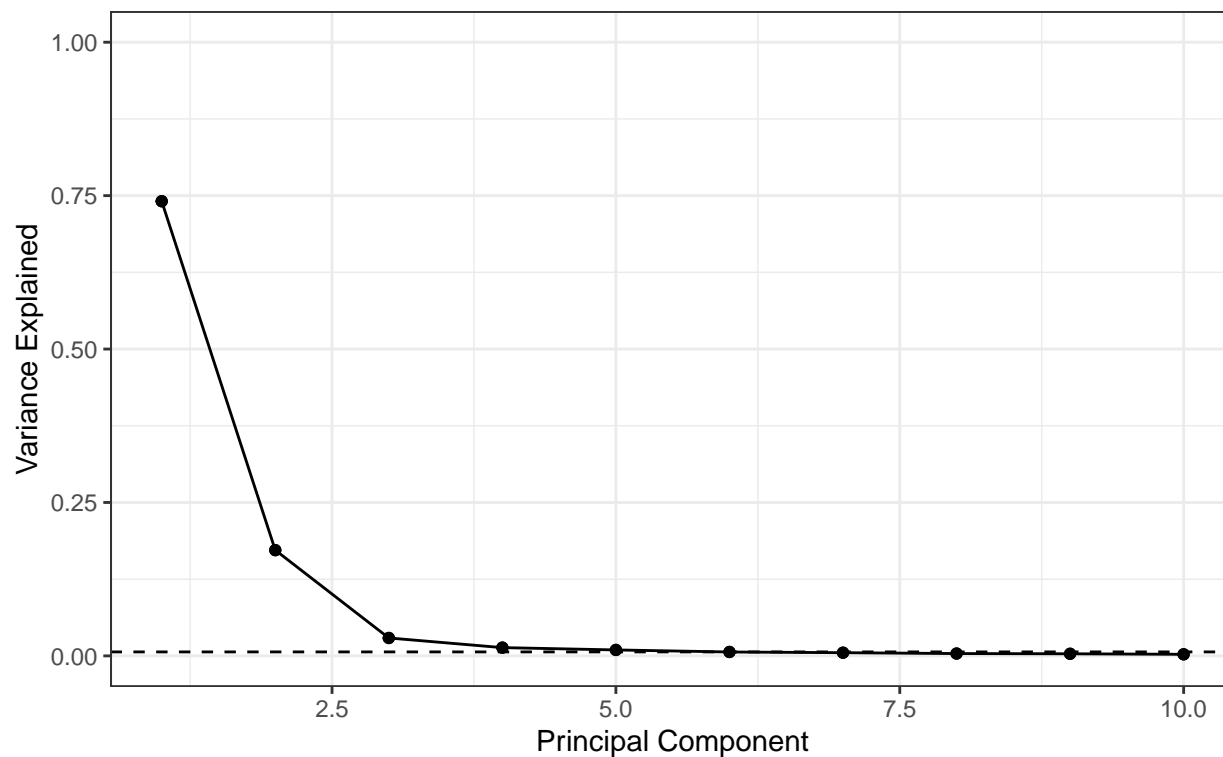
```
## [1] "SizeFactor distribution:"
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.5444 0.8273  1.0219  1.0575  1.2505  1.8502
```

SRP212704
Number of SVs: 1



SRP118614 + SRP002628 + SRP212704

Number of SVs: 1



```
## [1] "SizeFactor distribution:"  
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.  
##  0.7349  0.9231  1.0293  1.0309  1.1130  1.5355
```

SRP027258
Number of SVs: 2

