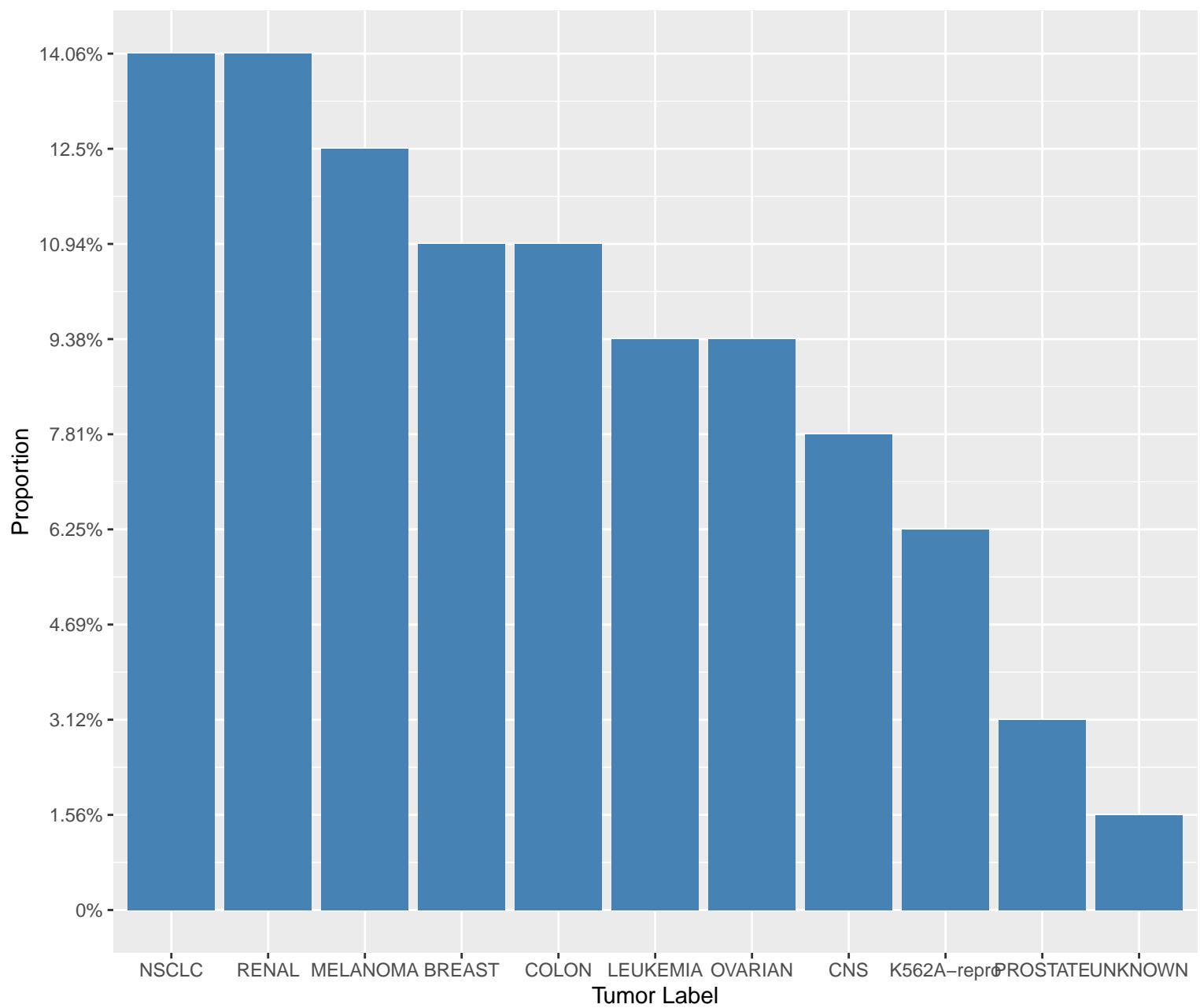
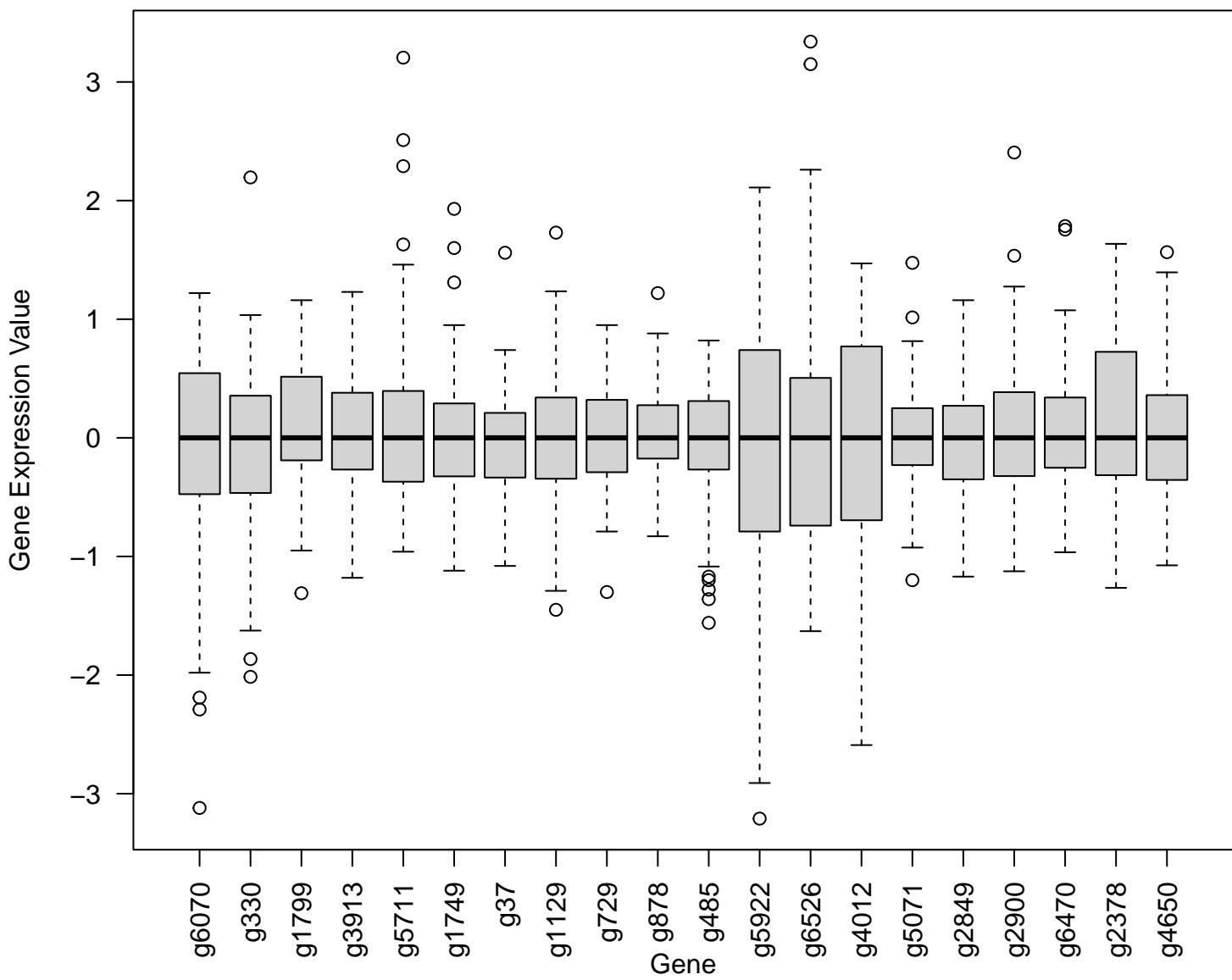


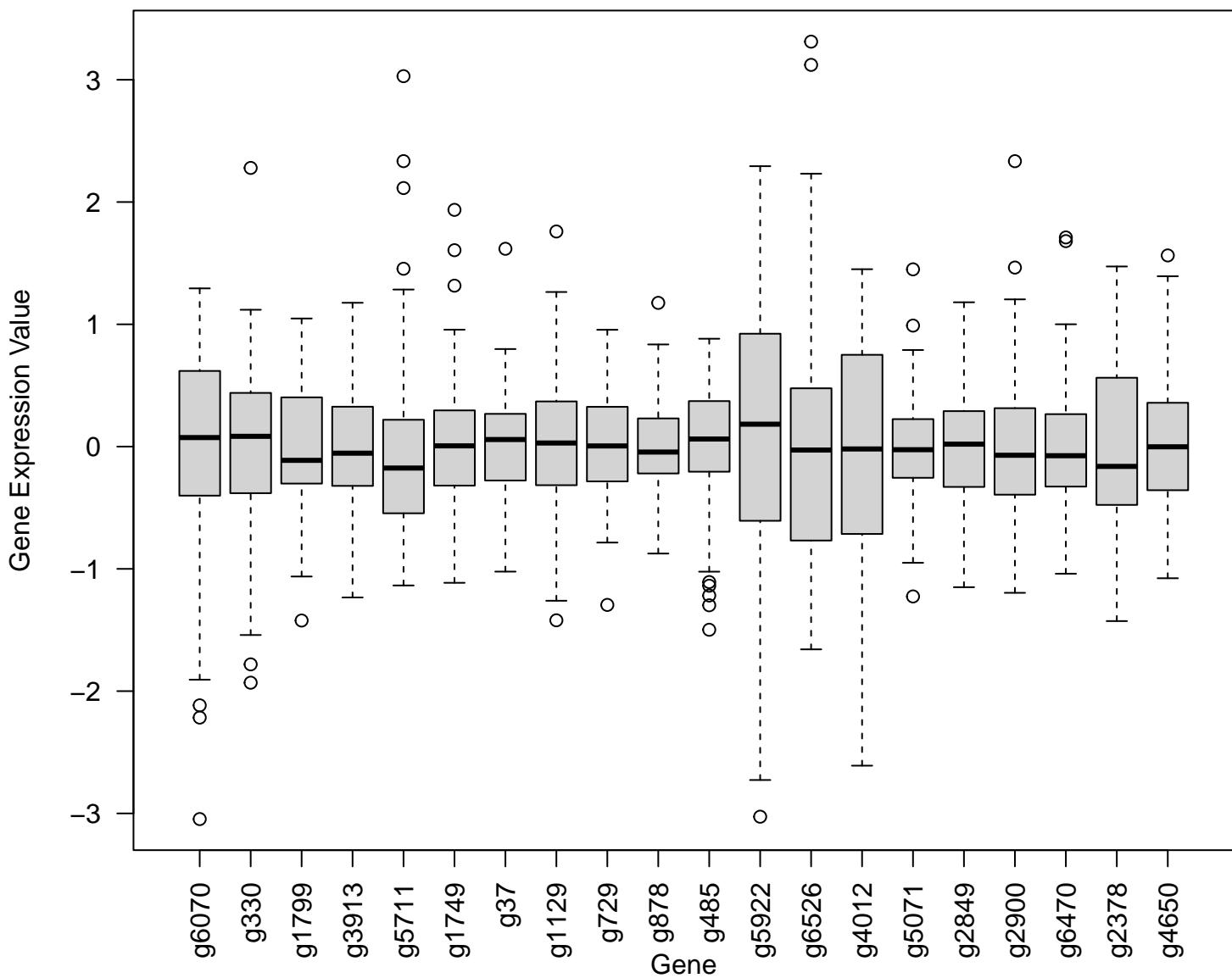
Tumor Label Distribution



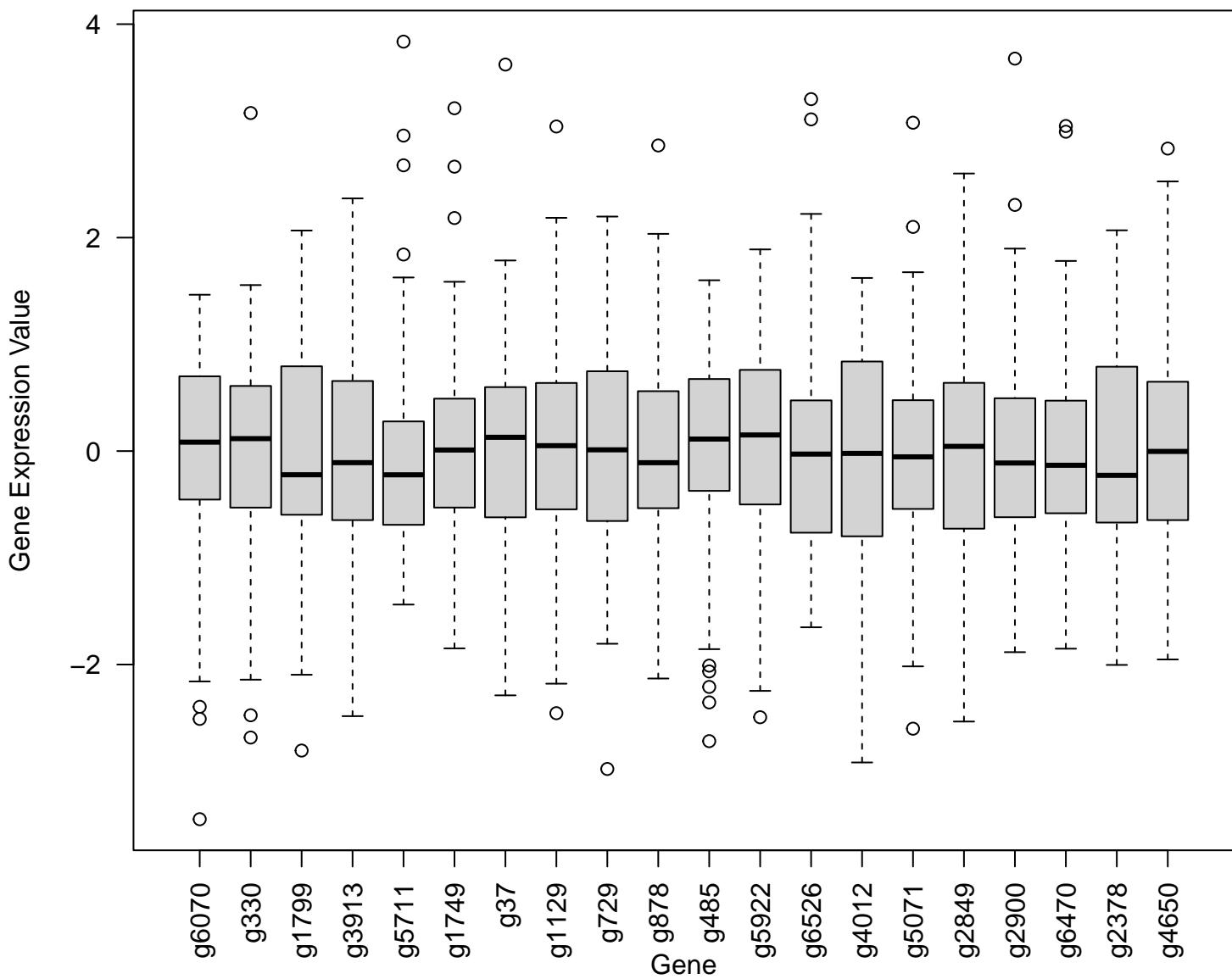
Subset Gene Expression marginalised over patients

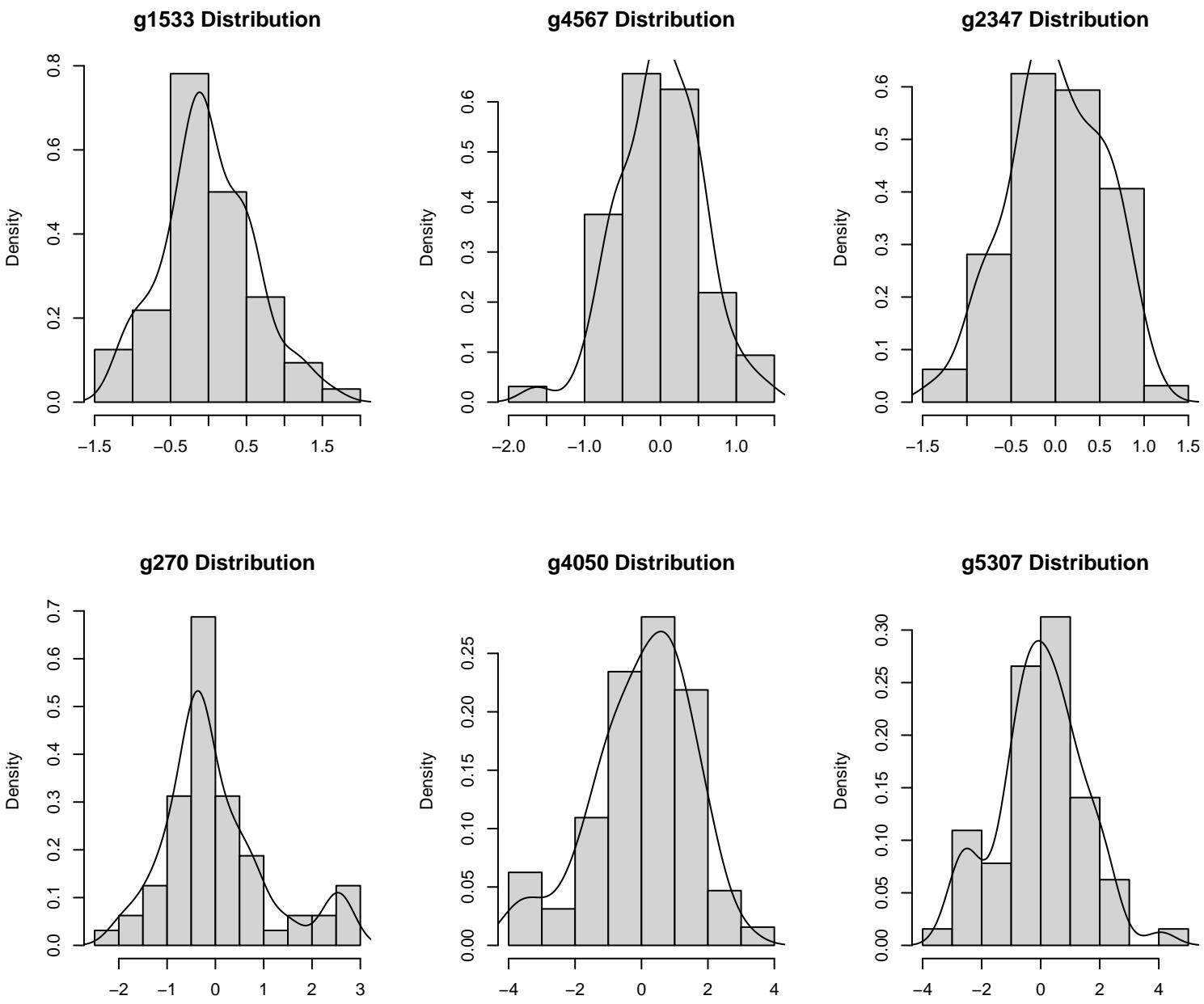


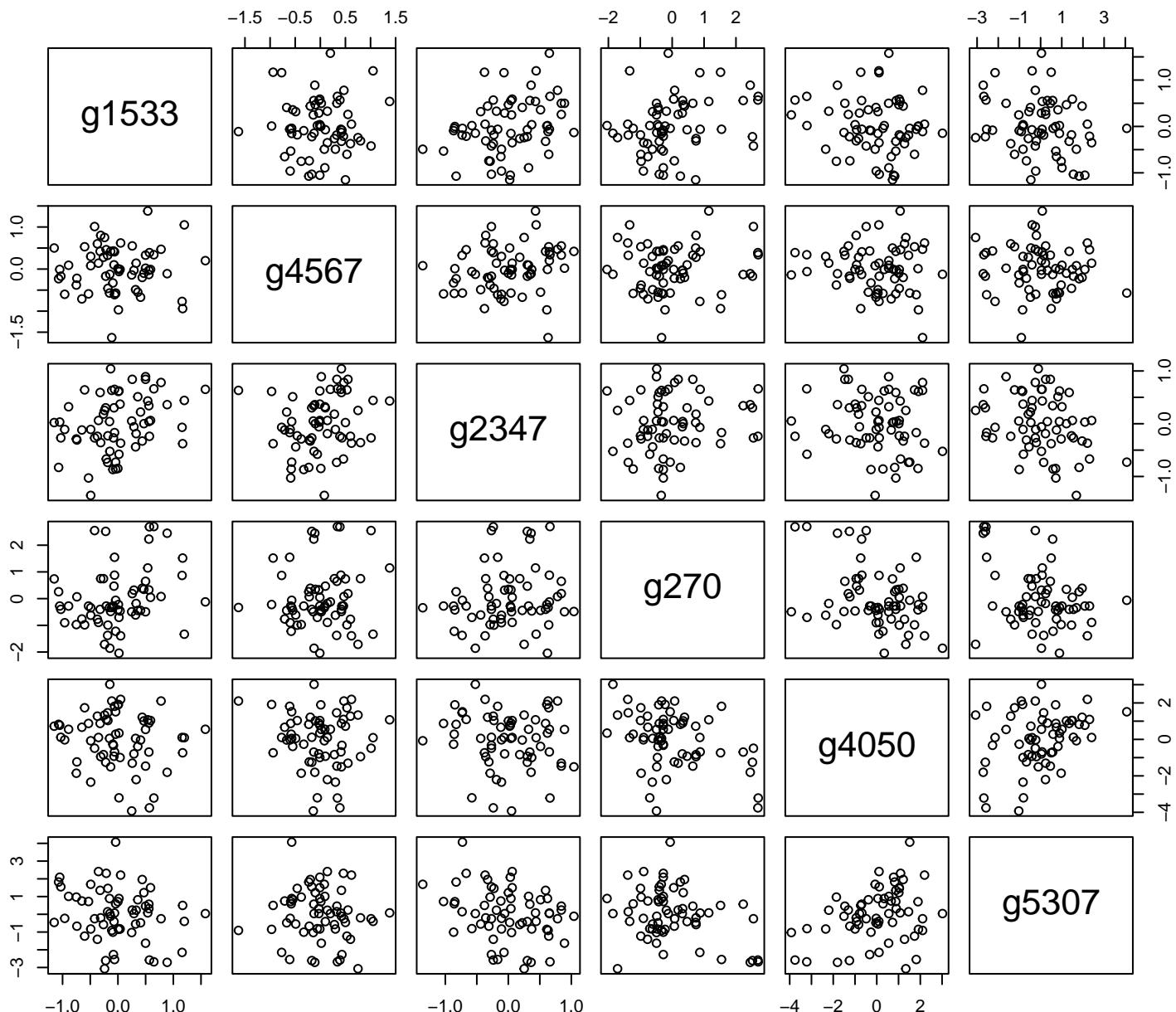
Subset Gene Expression marginalised over patients (centered)

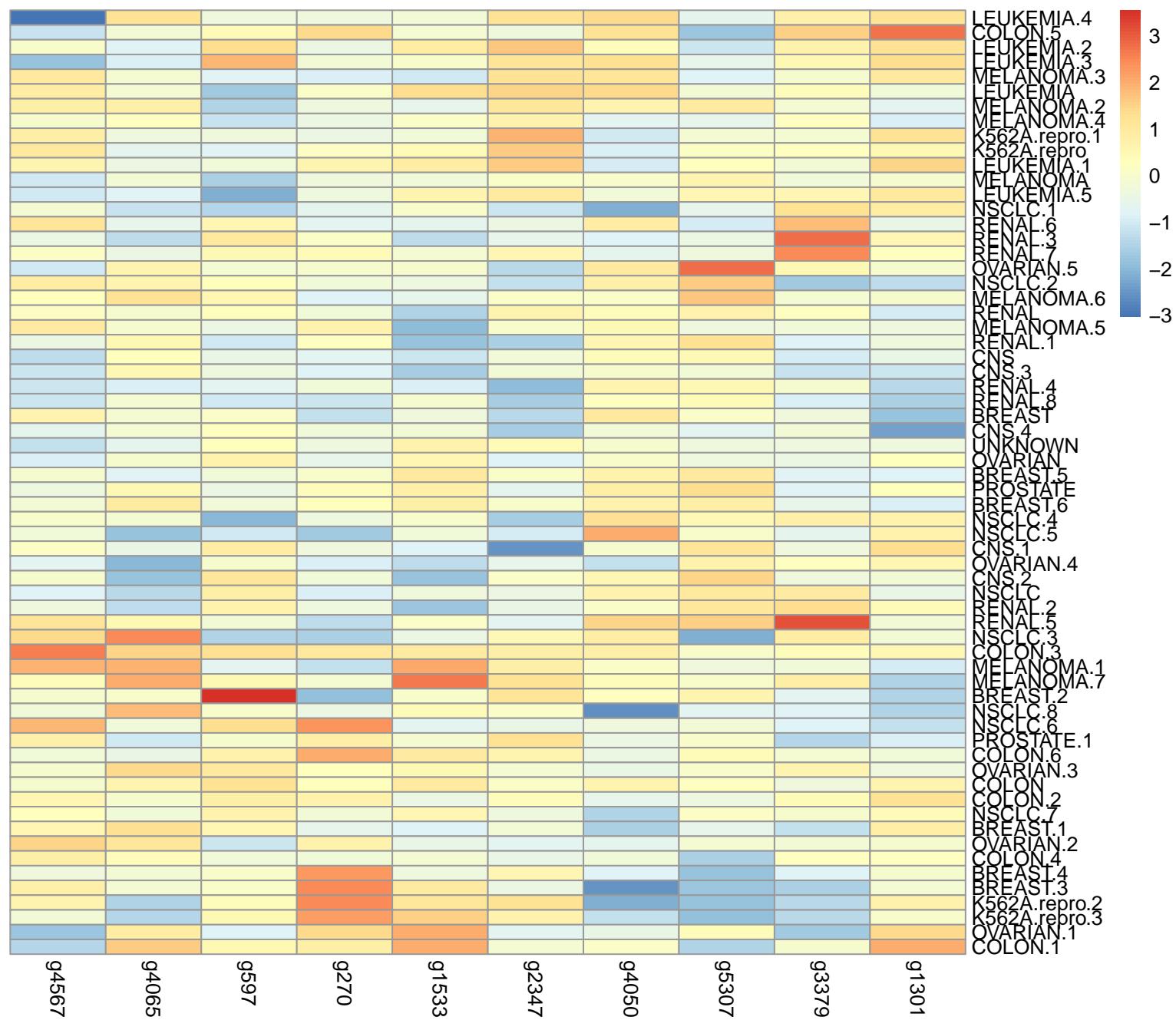


Subset Gene Expression marginalised over patients (centered and scaled)

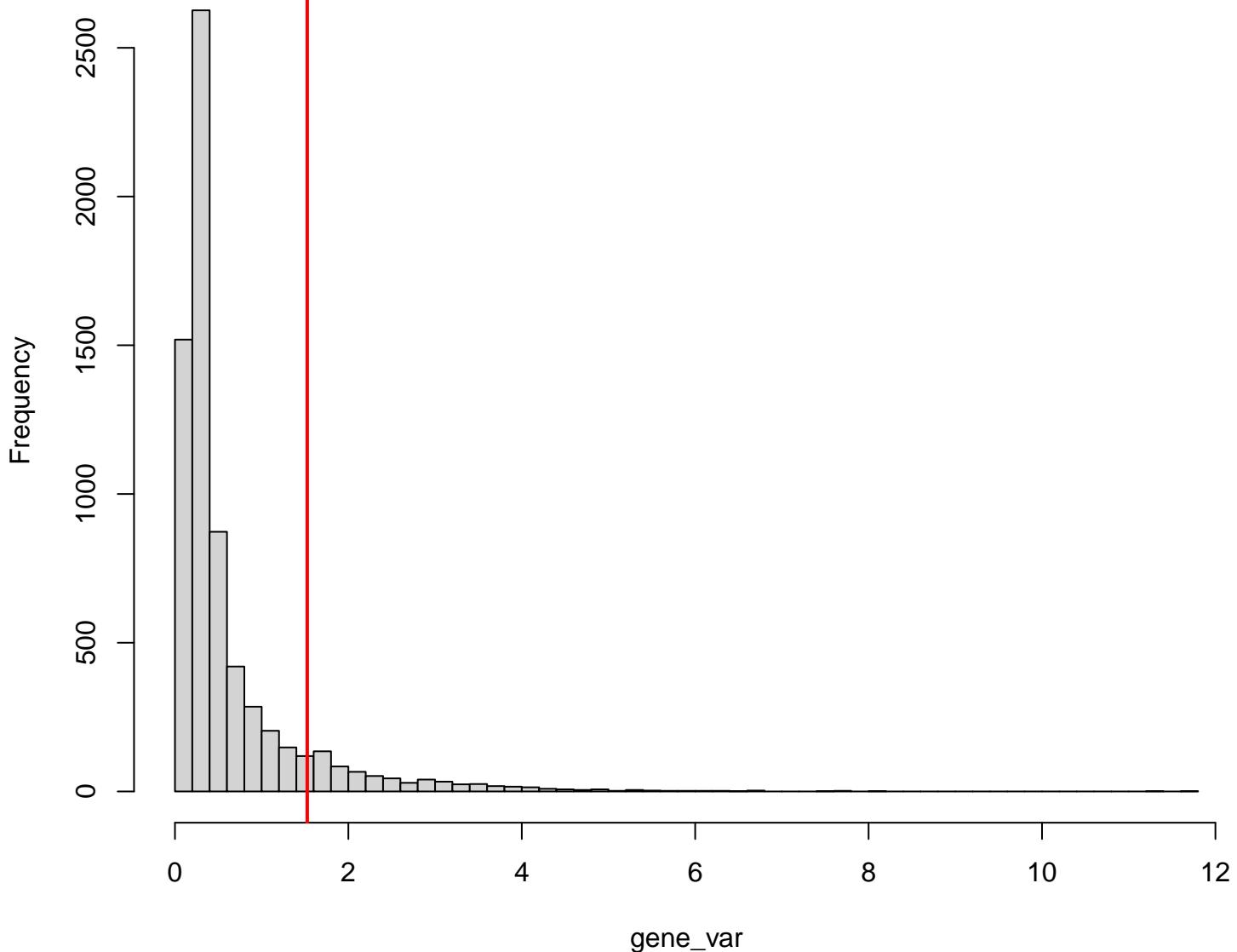




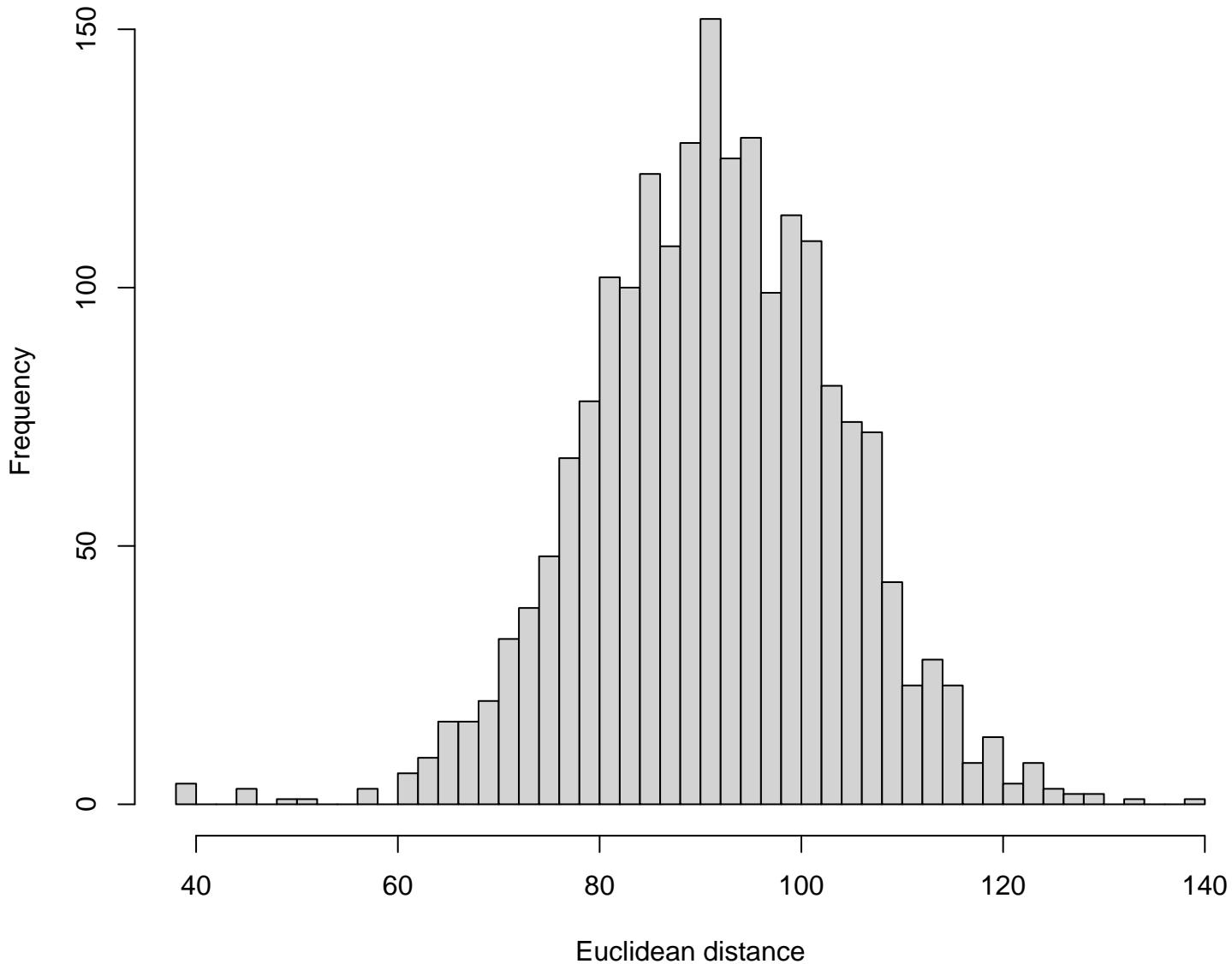




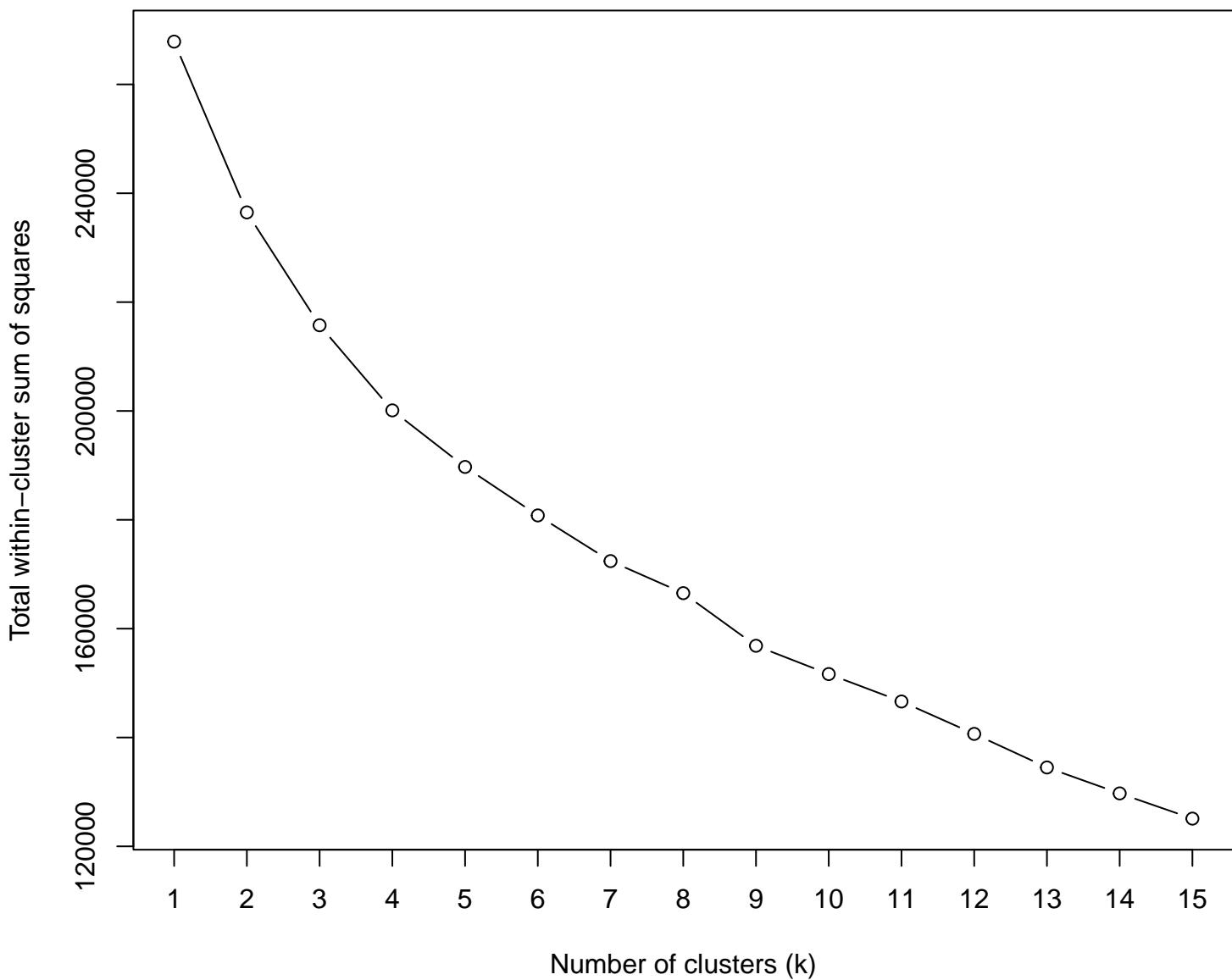
Gene-wise variance distribution



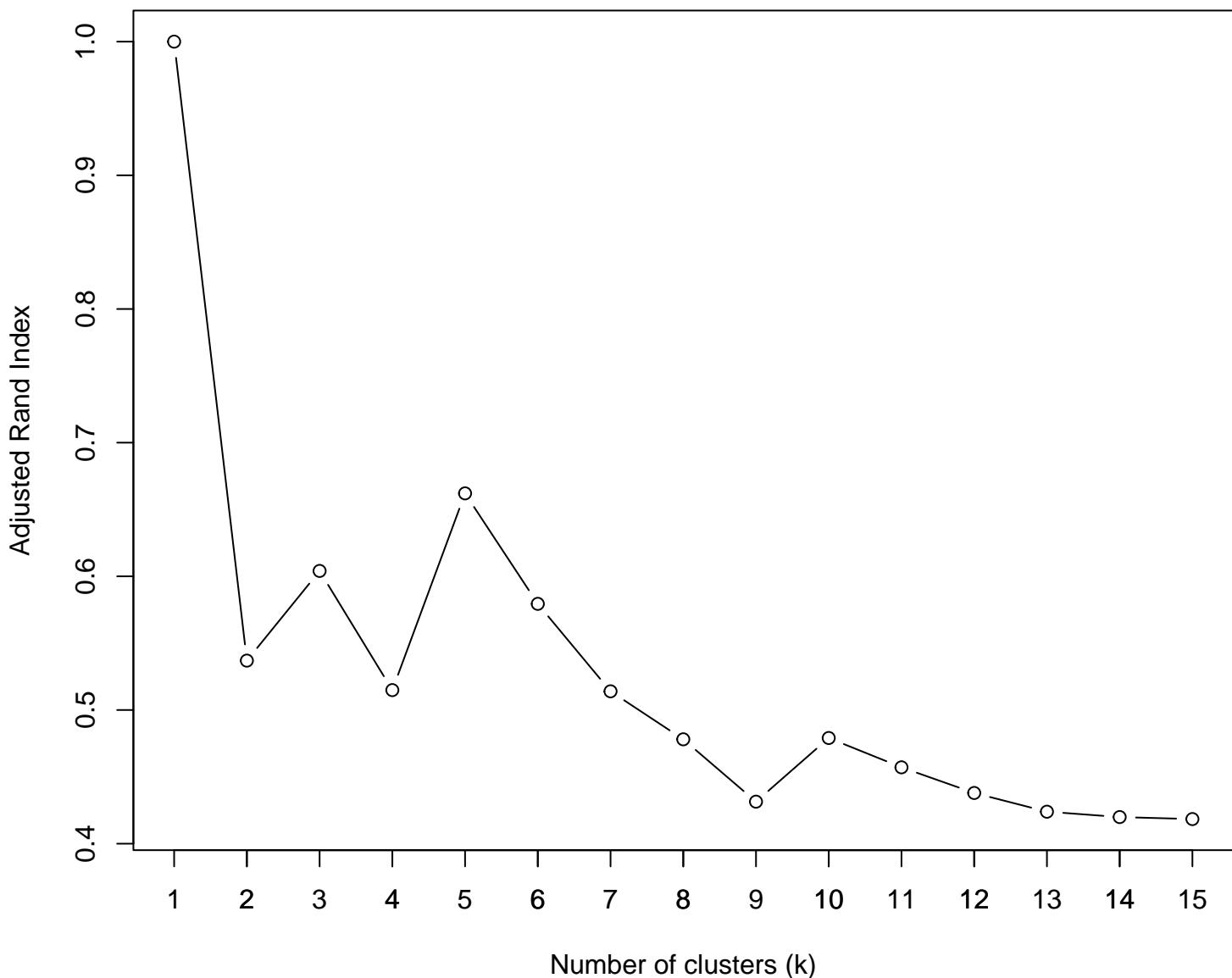
Pairwise distances between samples



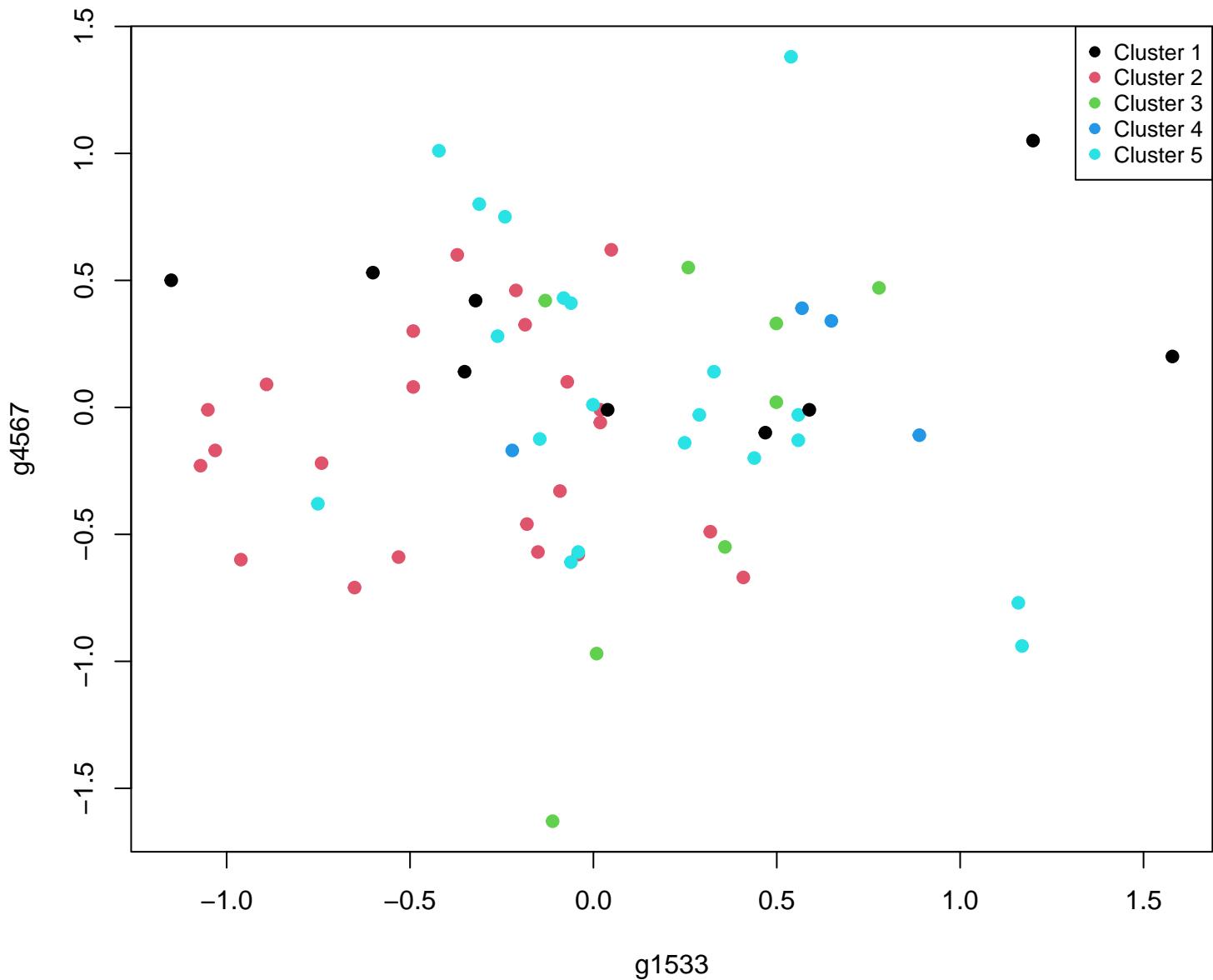
Elbow plot for k-means clustering



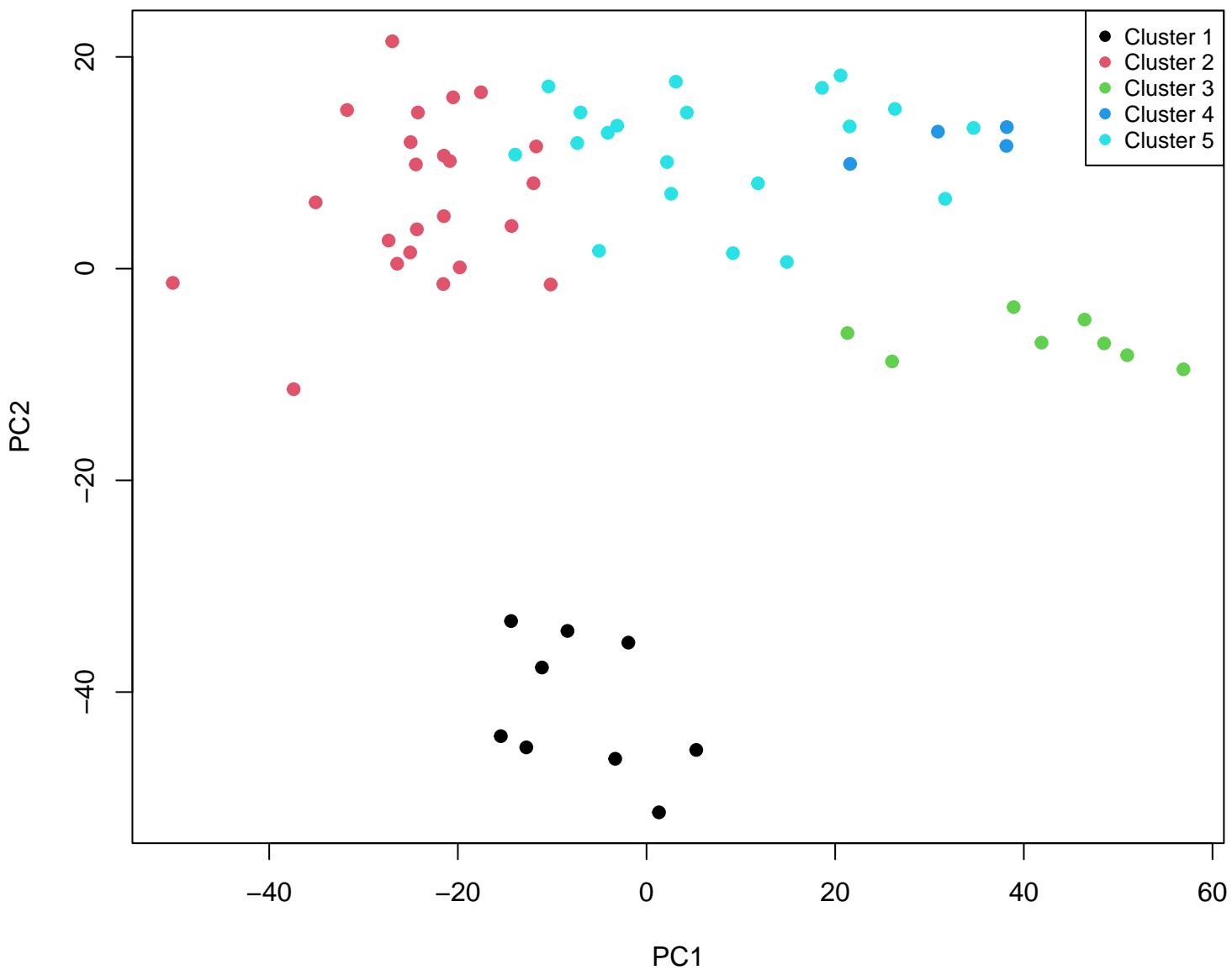
ARI vs k clusters



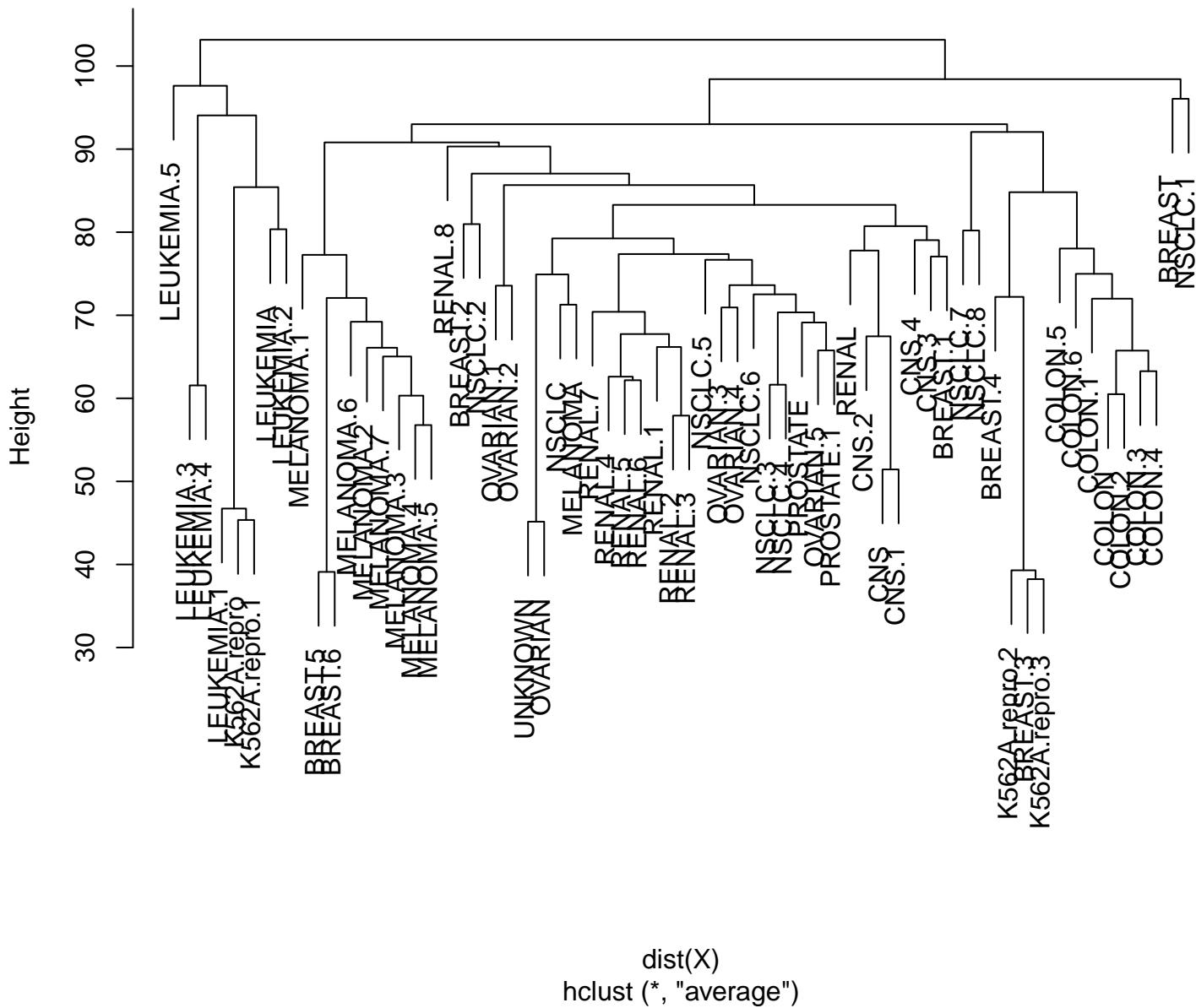
k-means clusters with 5 clusters (using 2 randomly sampled genes)



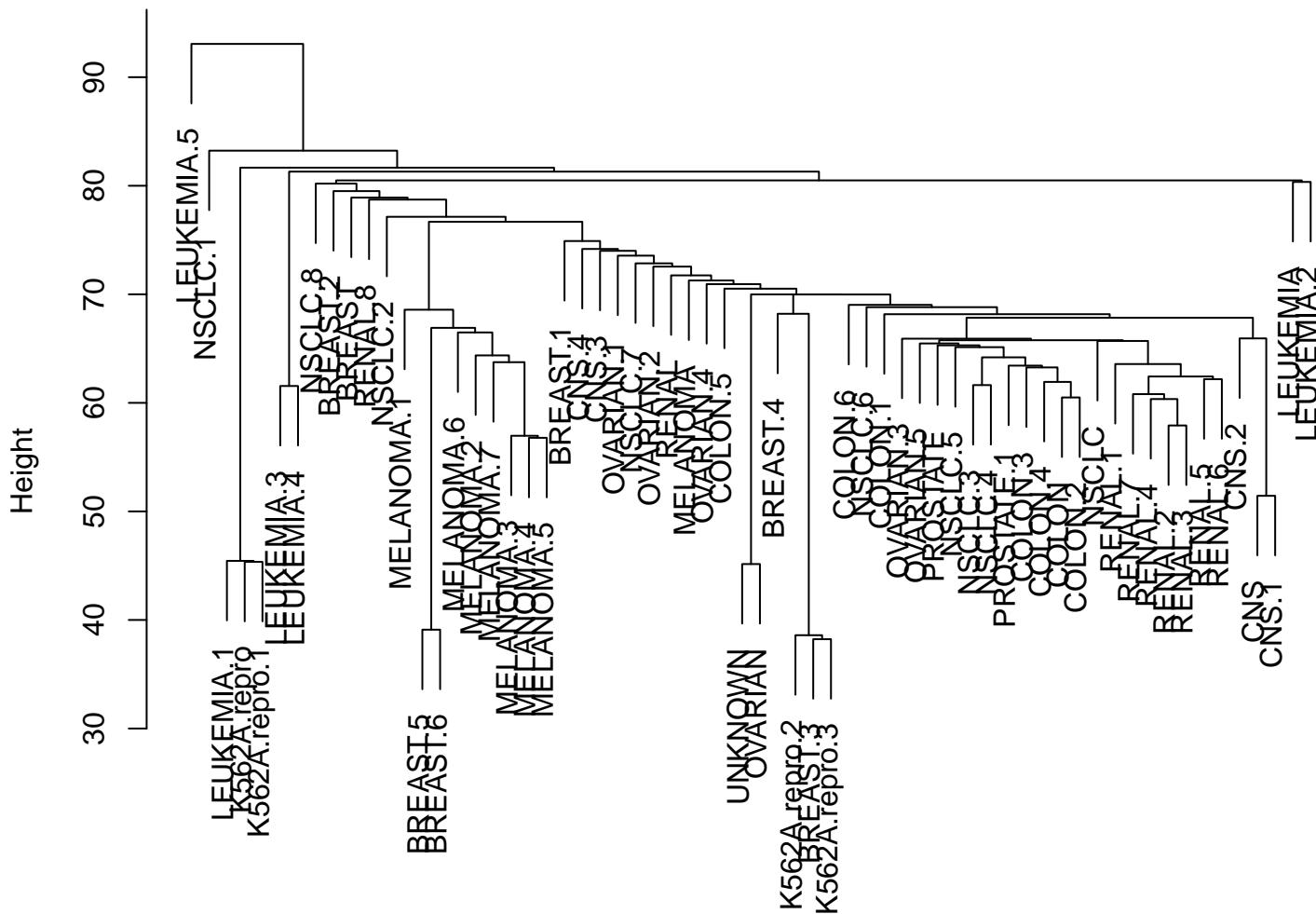
K-means clusters (first 2 PCs, raw clustering)



Average

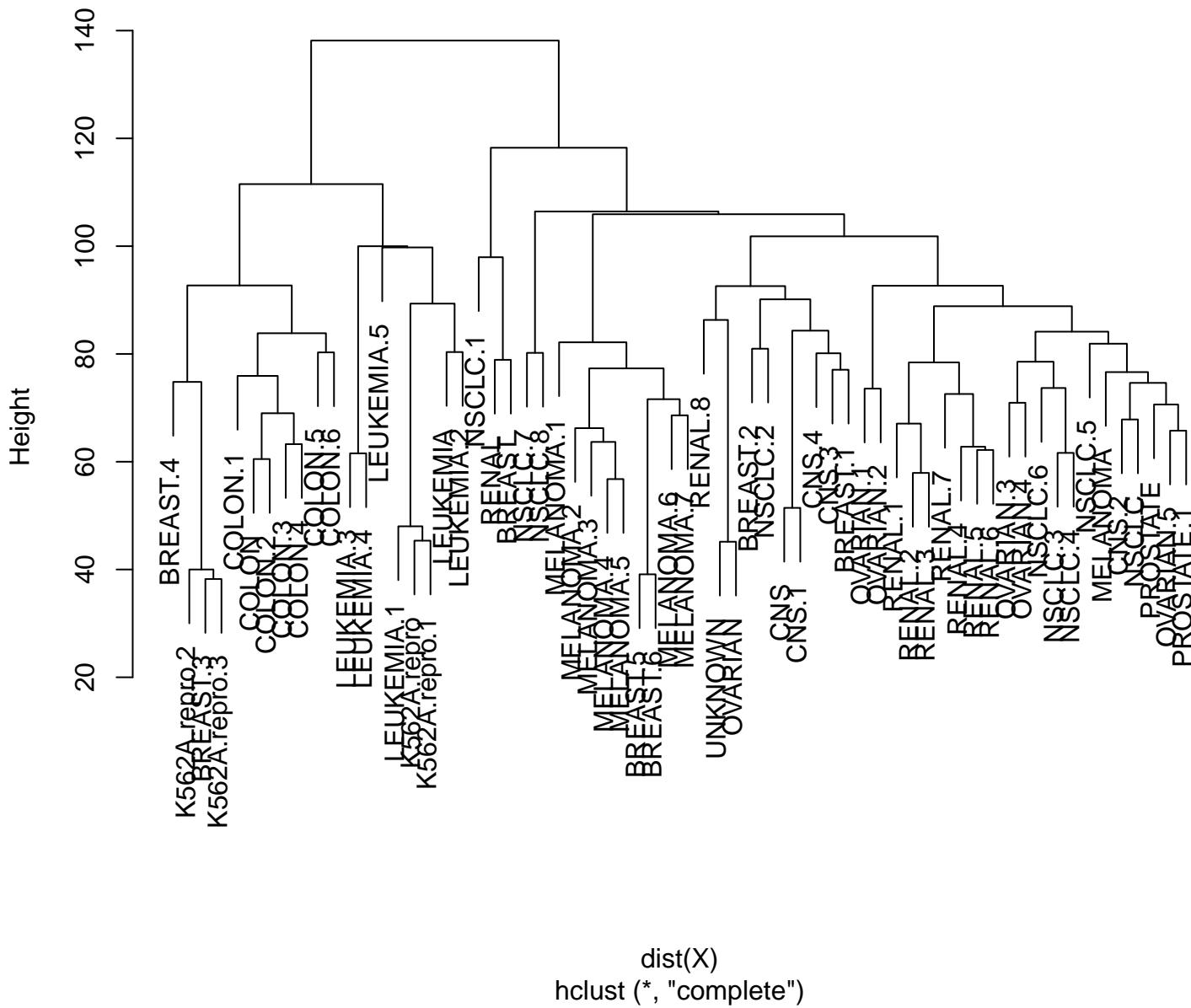


Single

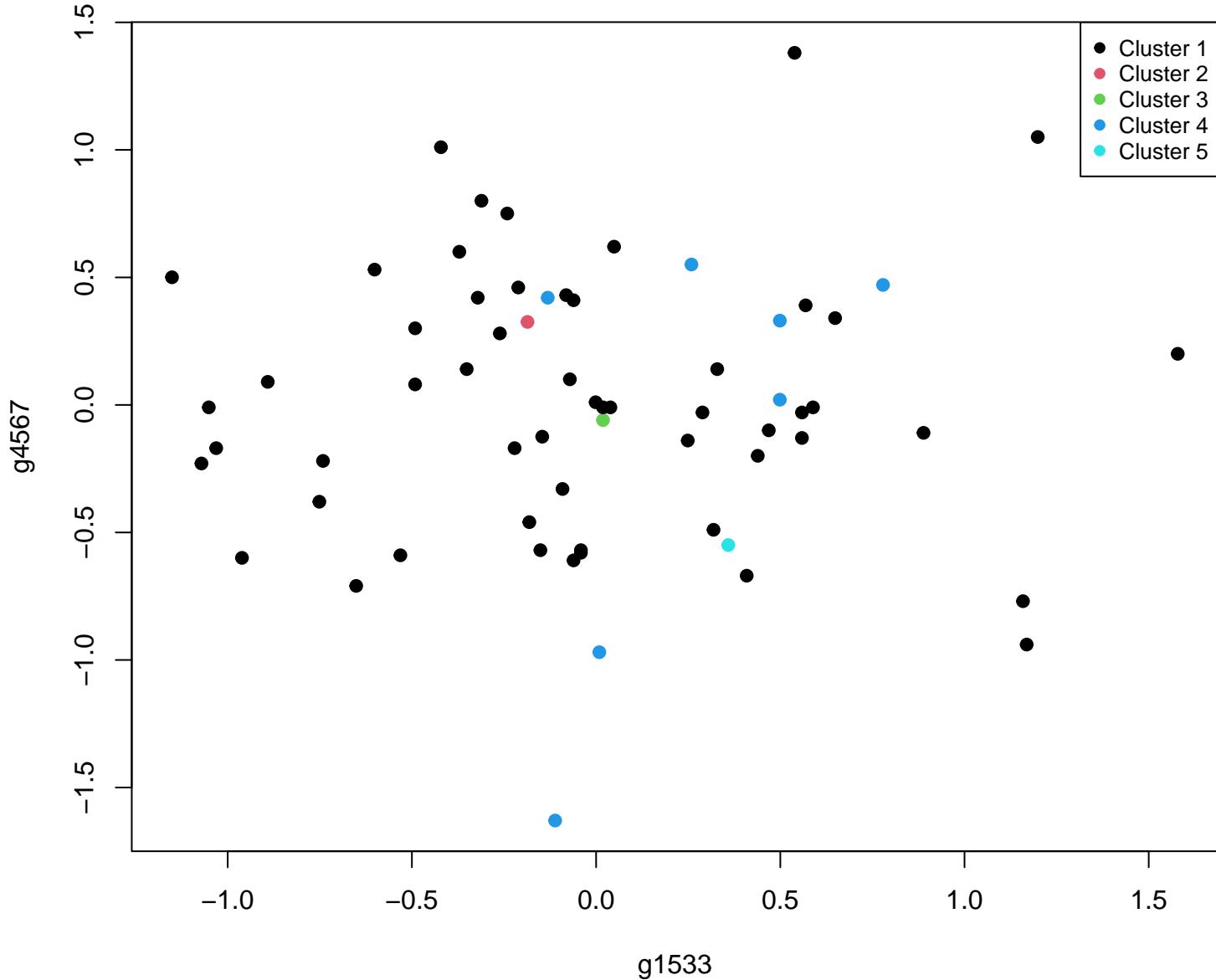


$\text{dist}(X)$
hclust (*, "single")

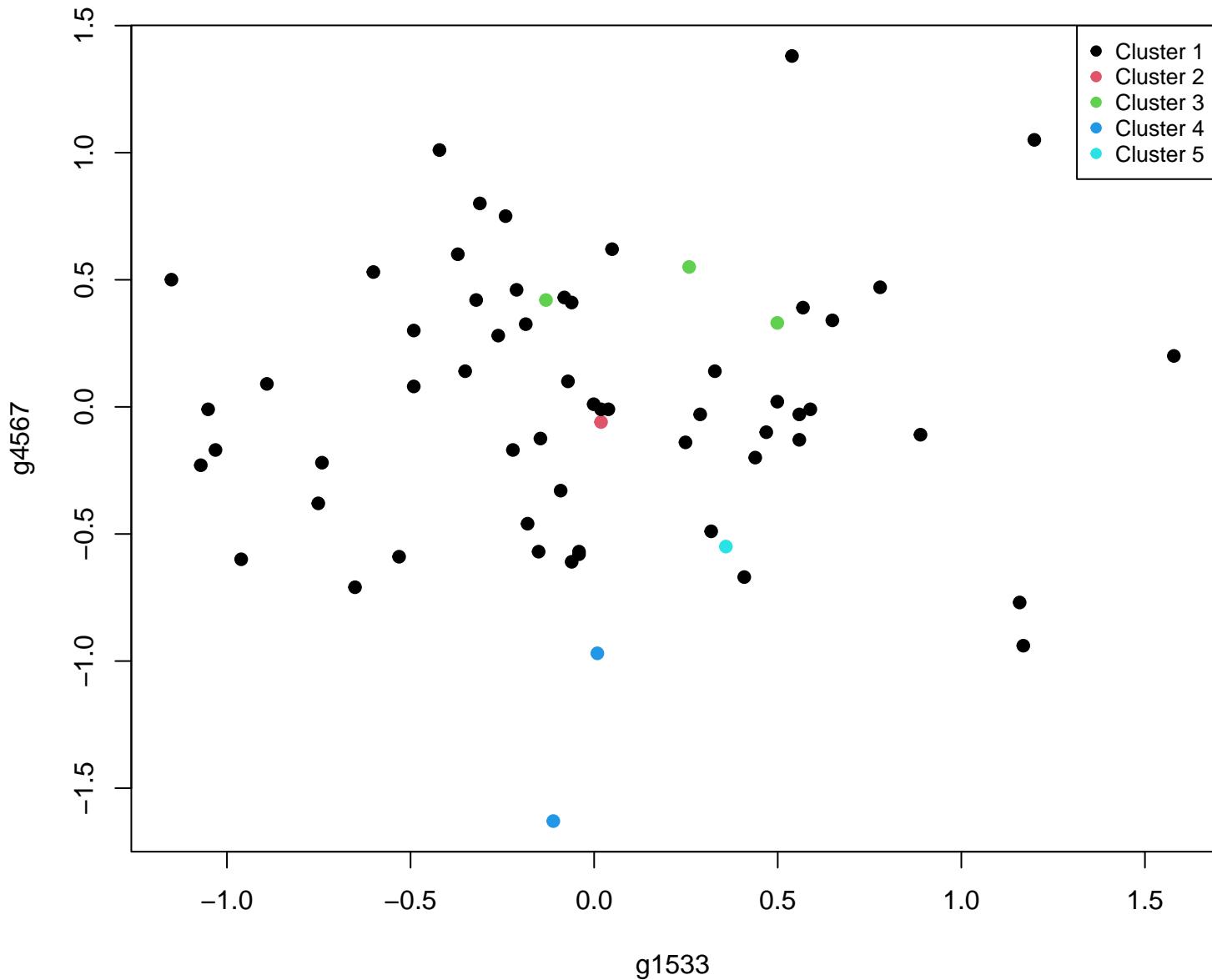
Complete



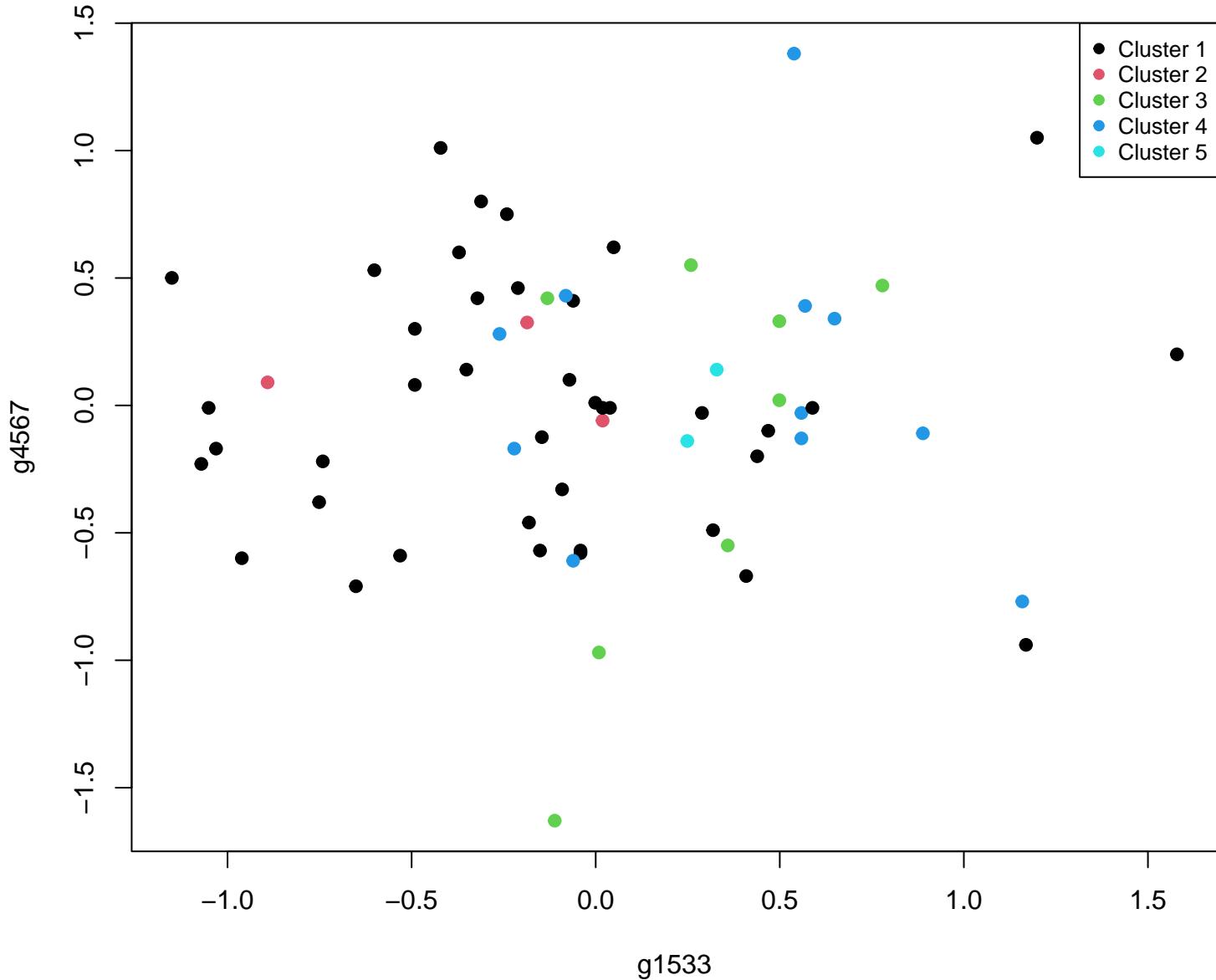
Average linkage clusters with 5 cluster cutoff (using 2 randomly sampled genes)



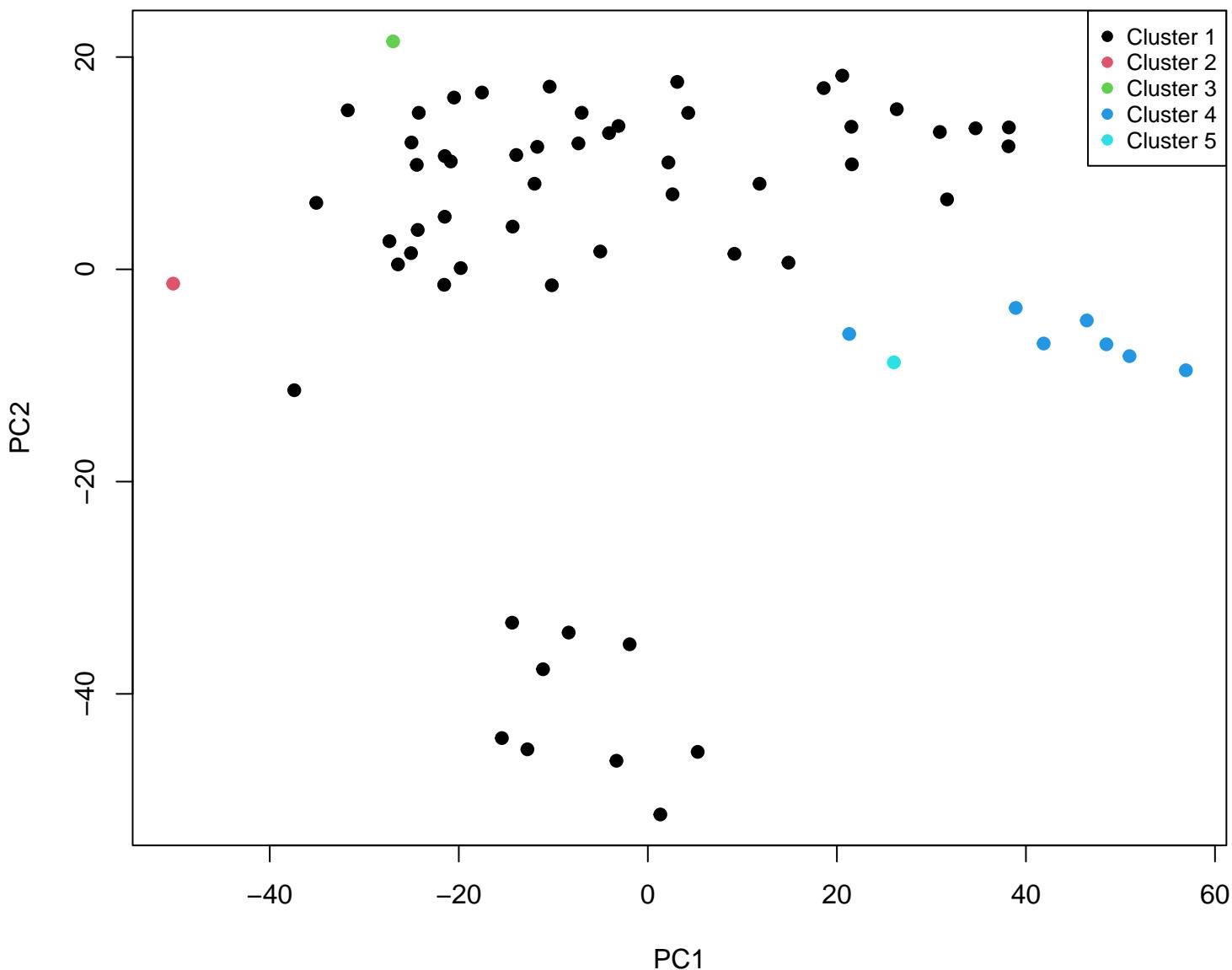
Single linkage clusters with 5 cluster cutoff (using 2 randomly sampled genes)



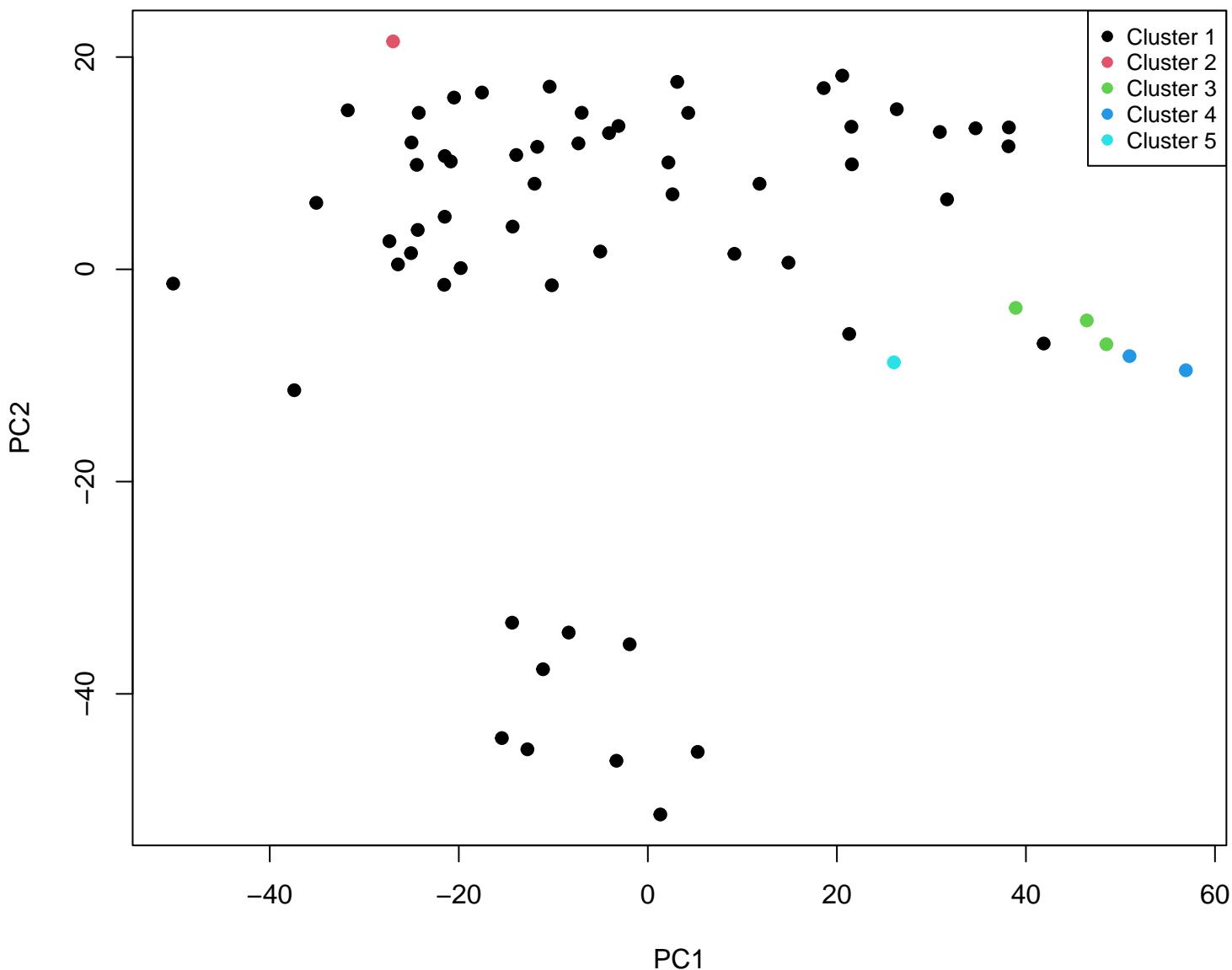
Complete linkage clusters with 5 cluster cutoff (using 2 randomly sampled genes)



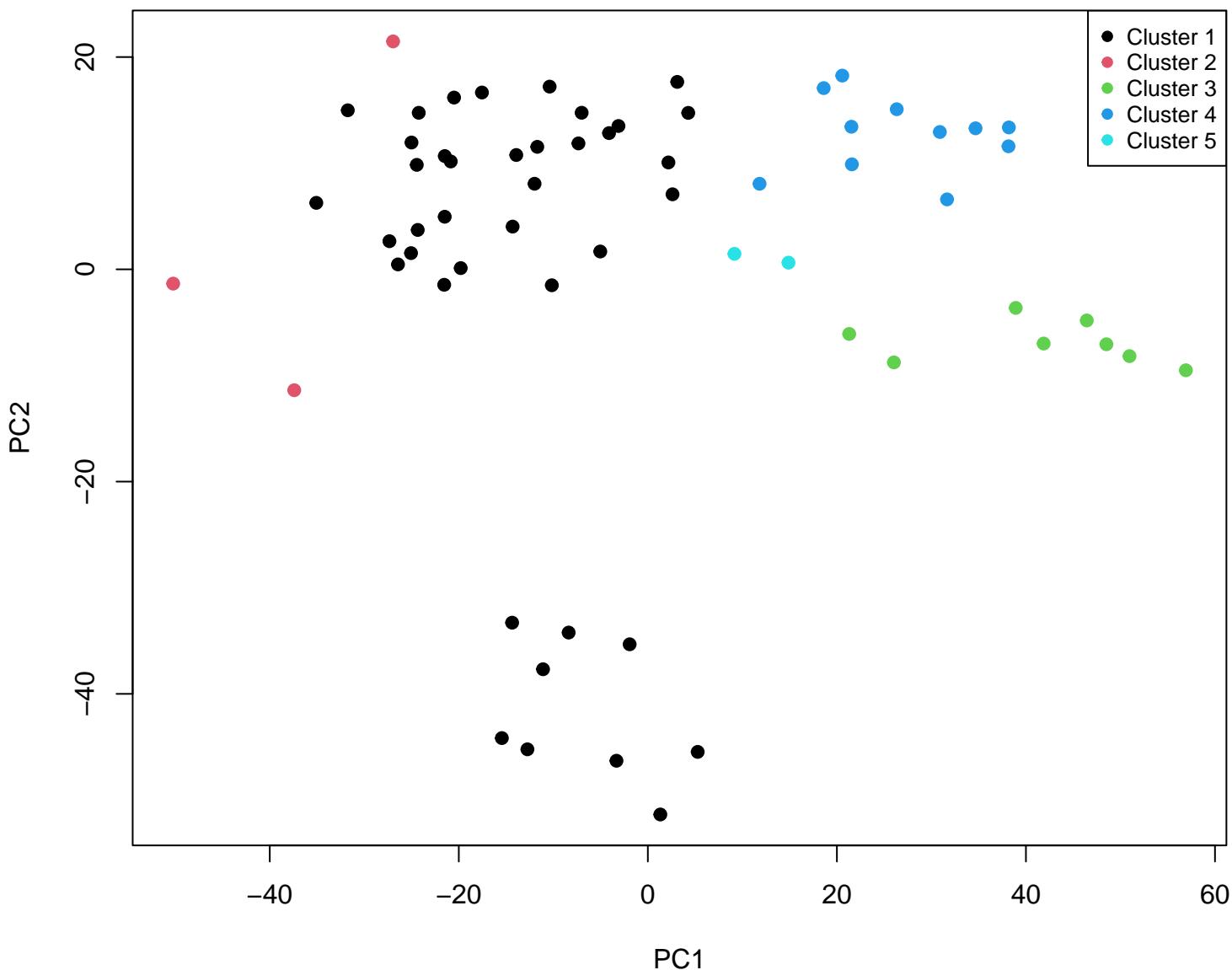
Average linkage clusters (first 2 PCs, 5 clusters cut-off)



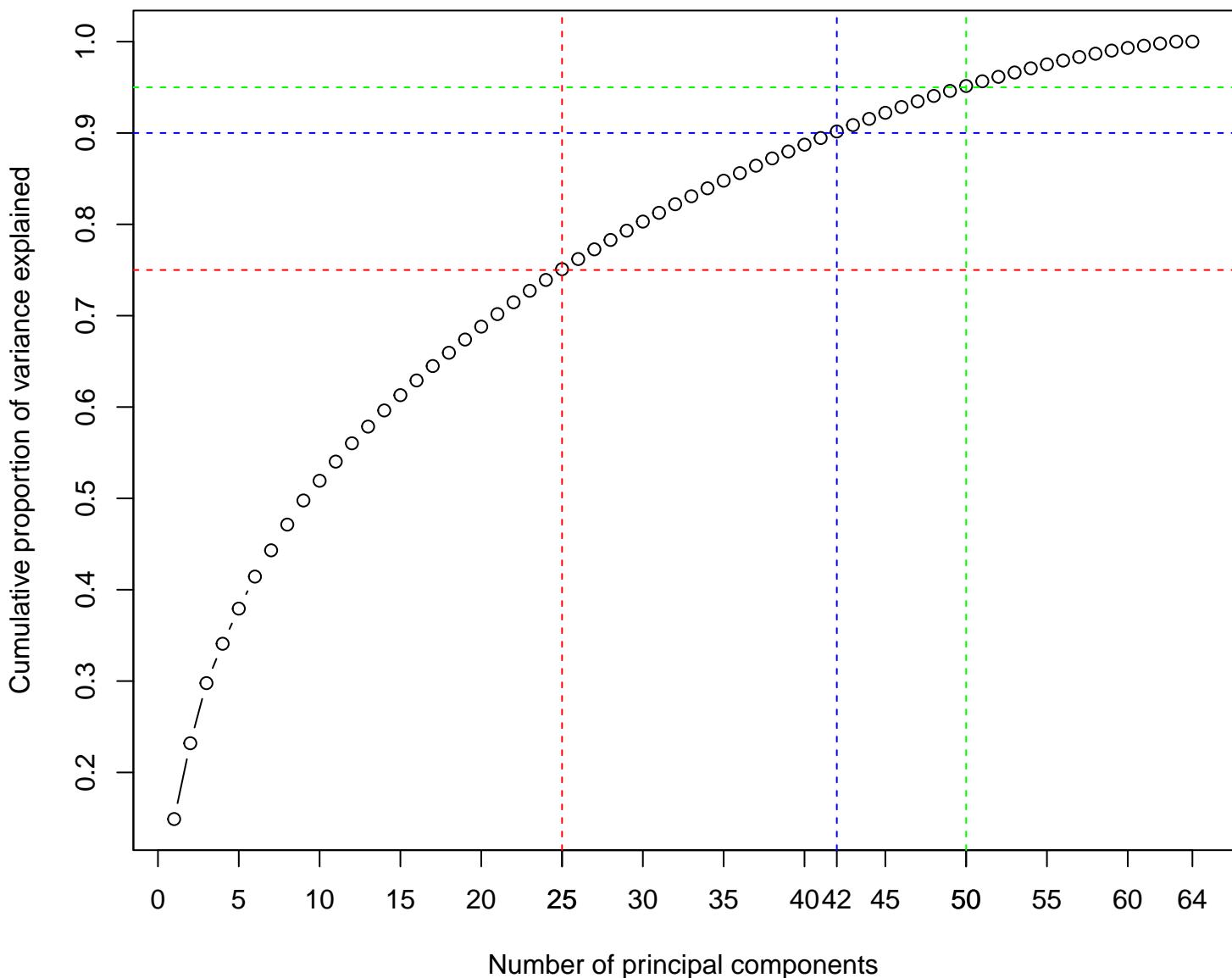
Single linkage clusters (first 2 PCs, 5 clusters cut-off)



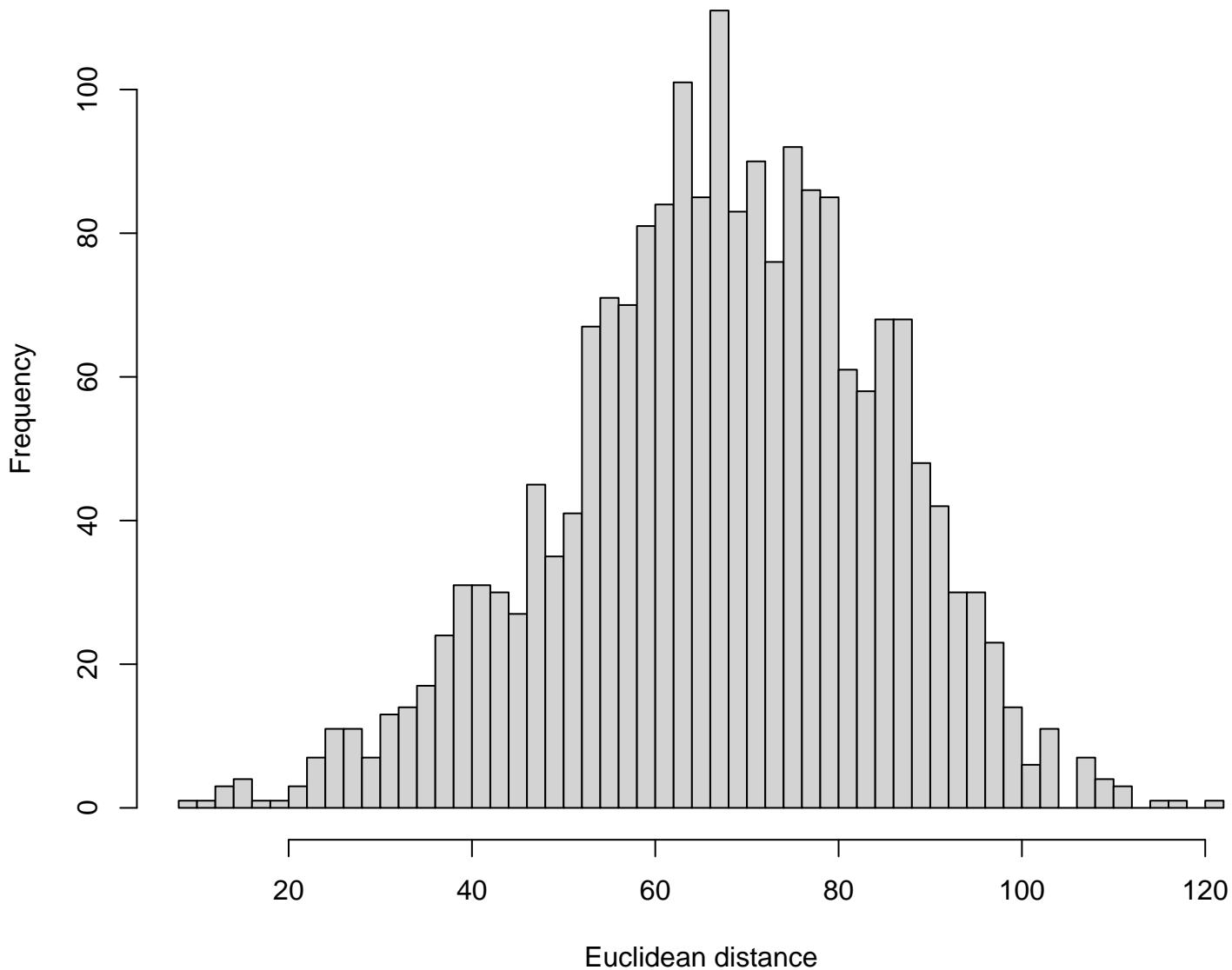
Complete linkage clusters (first 2 PCs, 5 clusters cut-off)



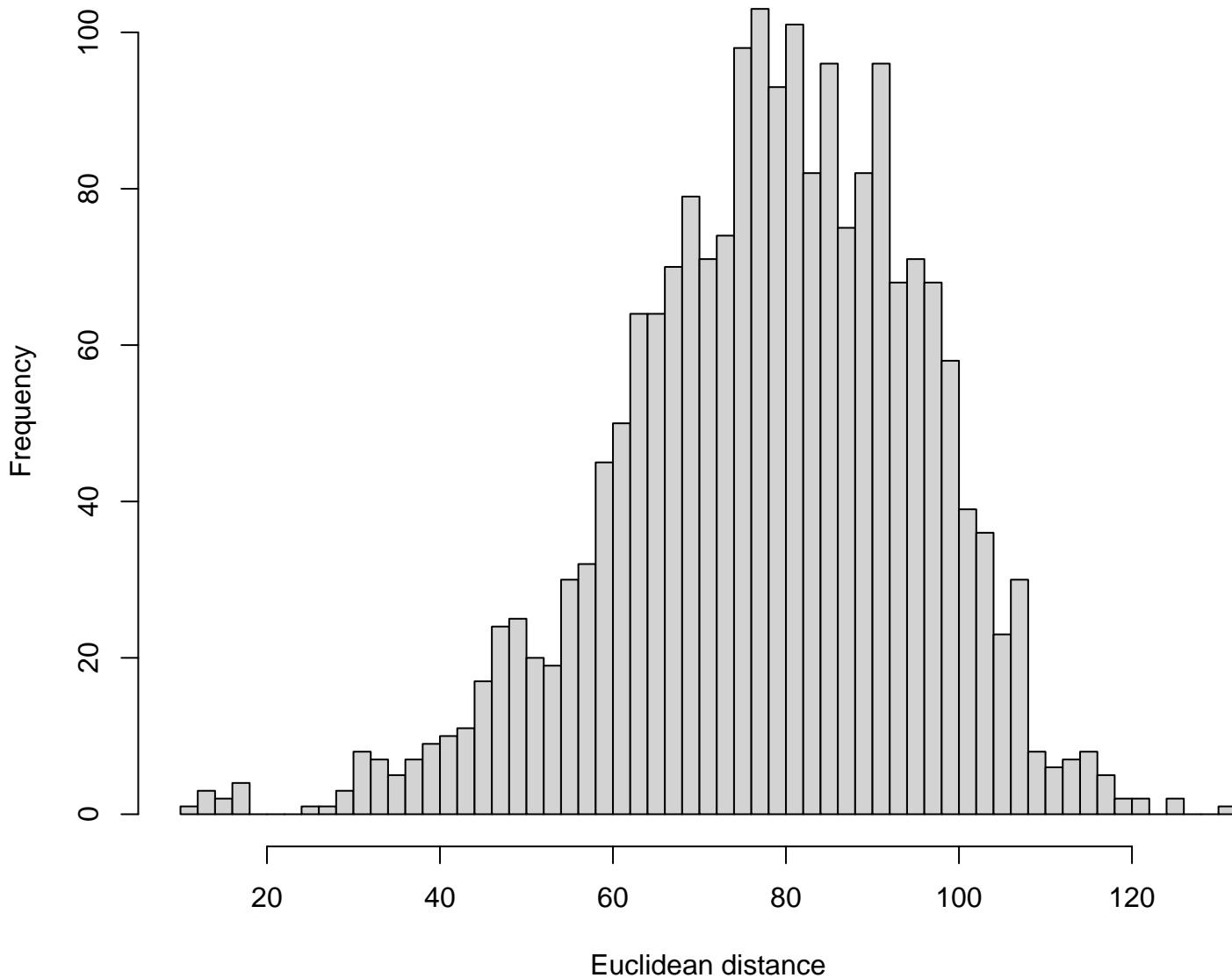
Cumulative variance explained by PCA



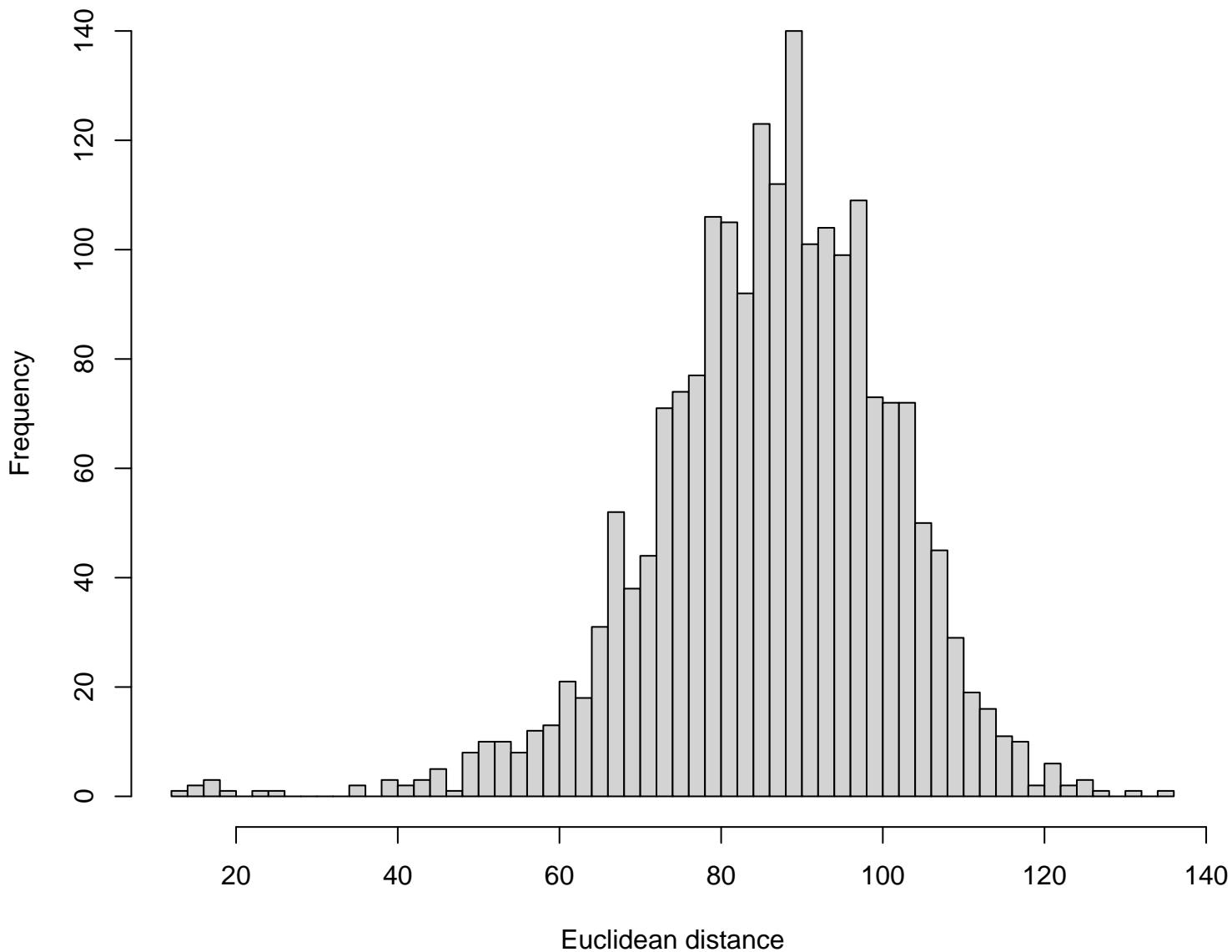
Pairwise distances between samples on reduced data Z_sp



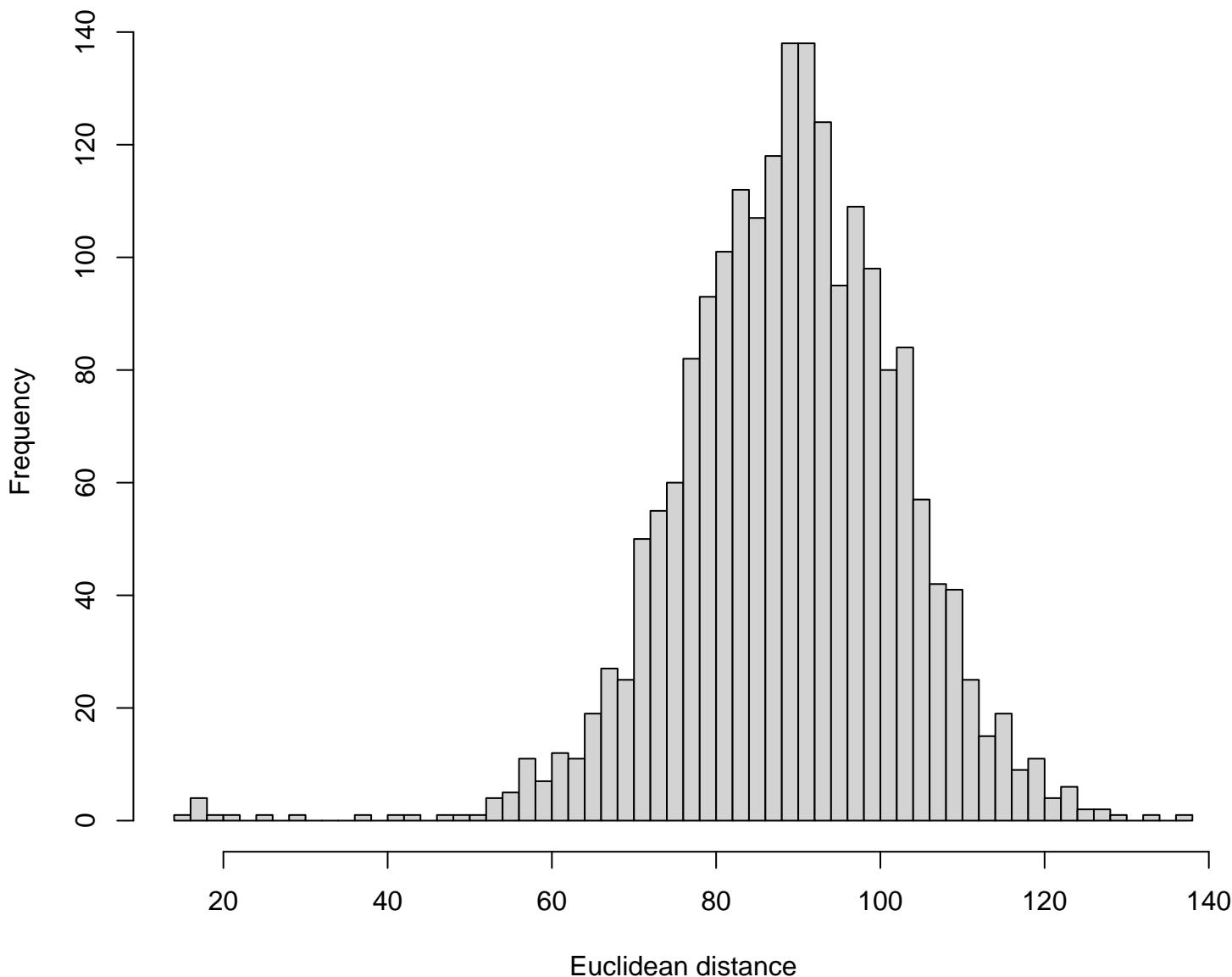
Pairwise distances between samples on reduced data with Z_75



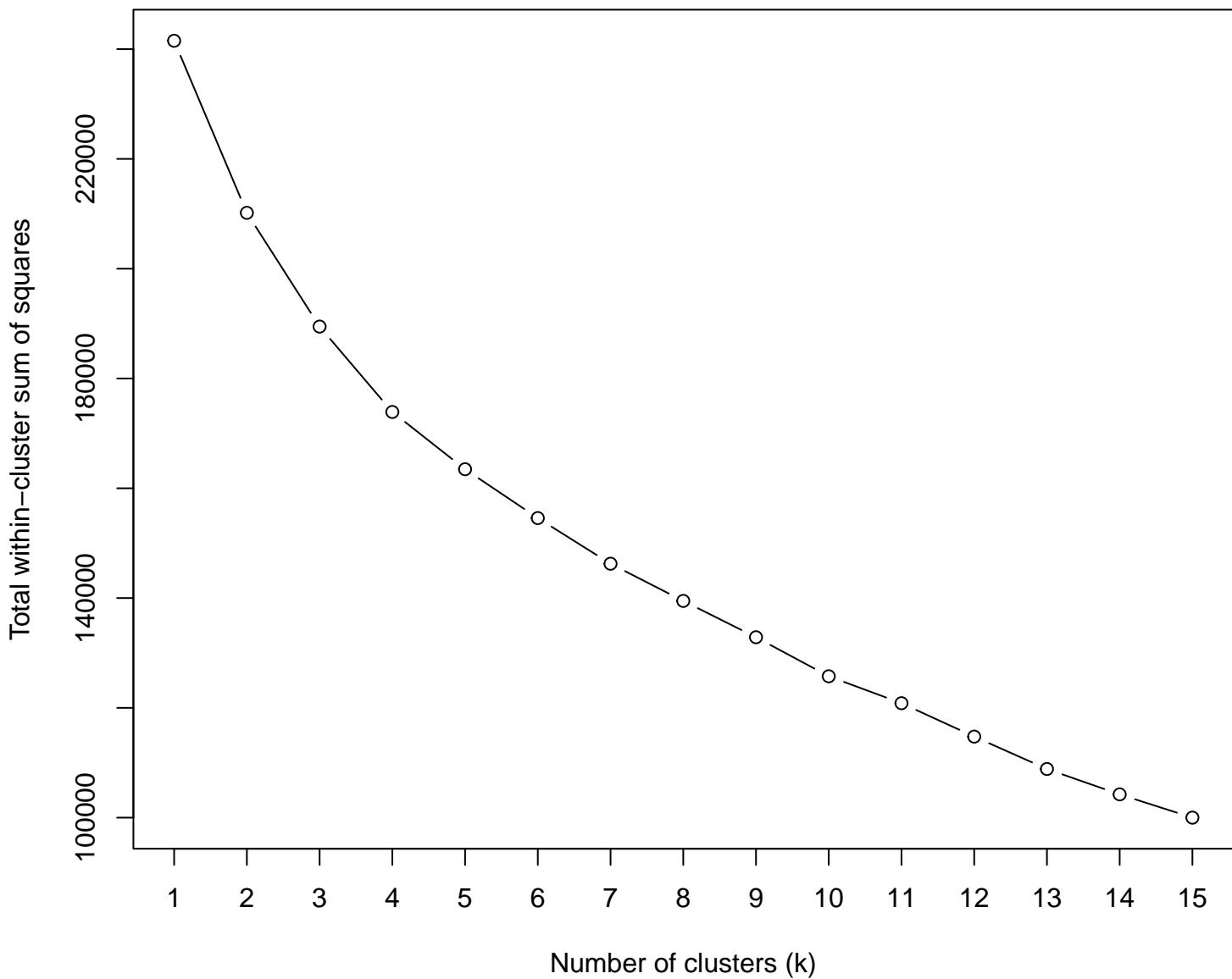
Pairwise distances between samples on reduced data with Z_90



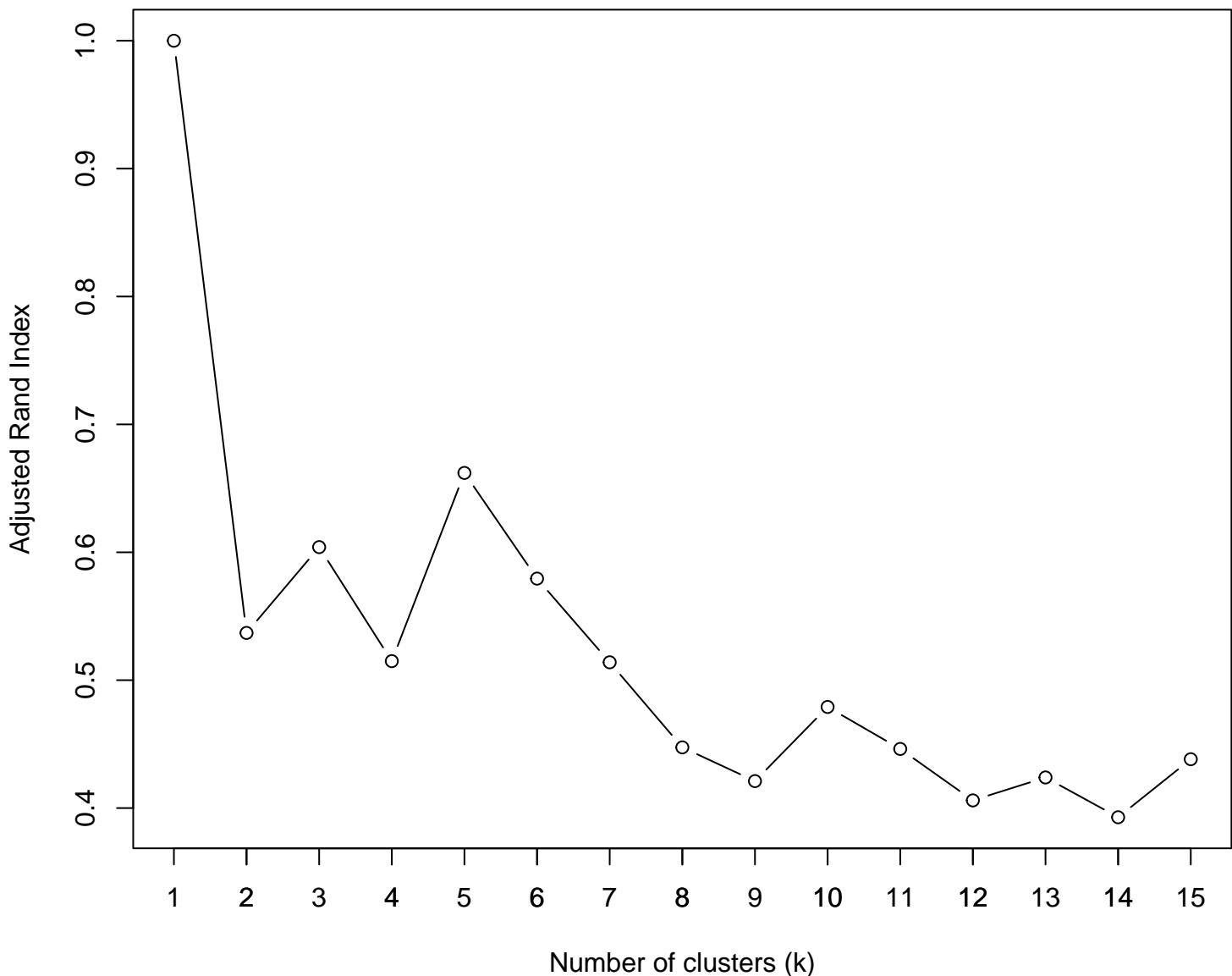
Pairwise distances between samples on reduced data with Z_95



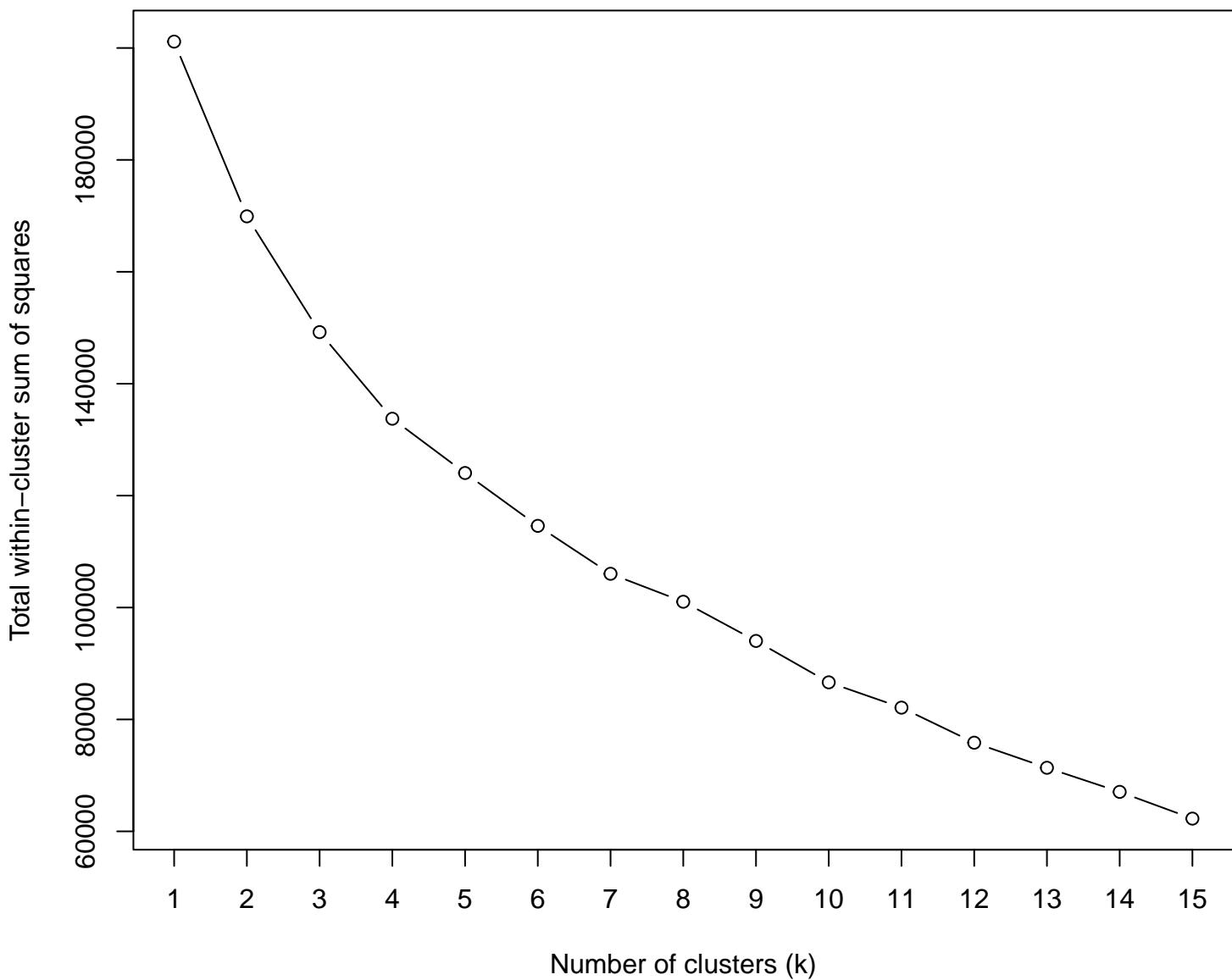
Elbow plot for k-means clustering on reduced data Z



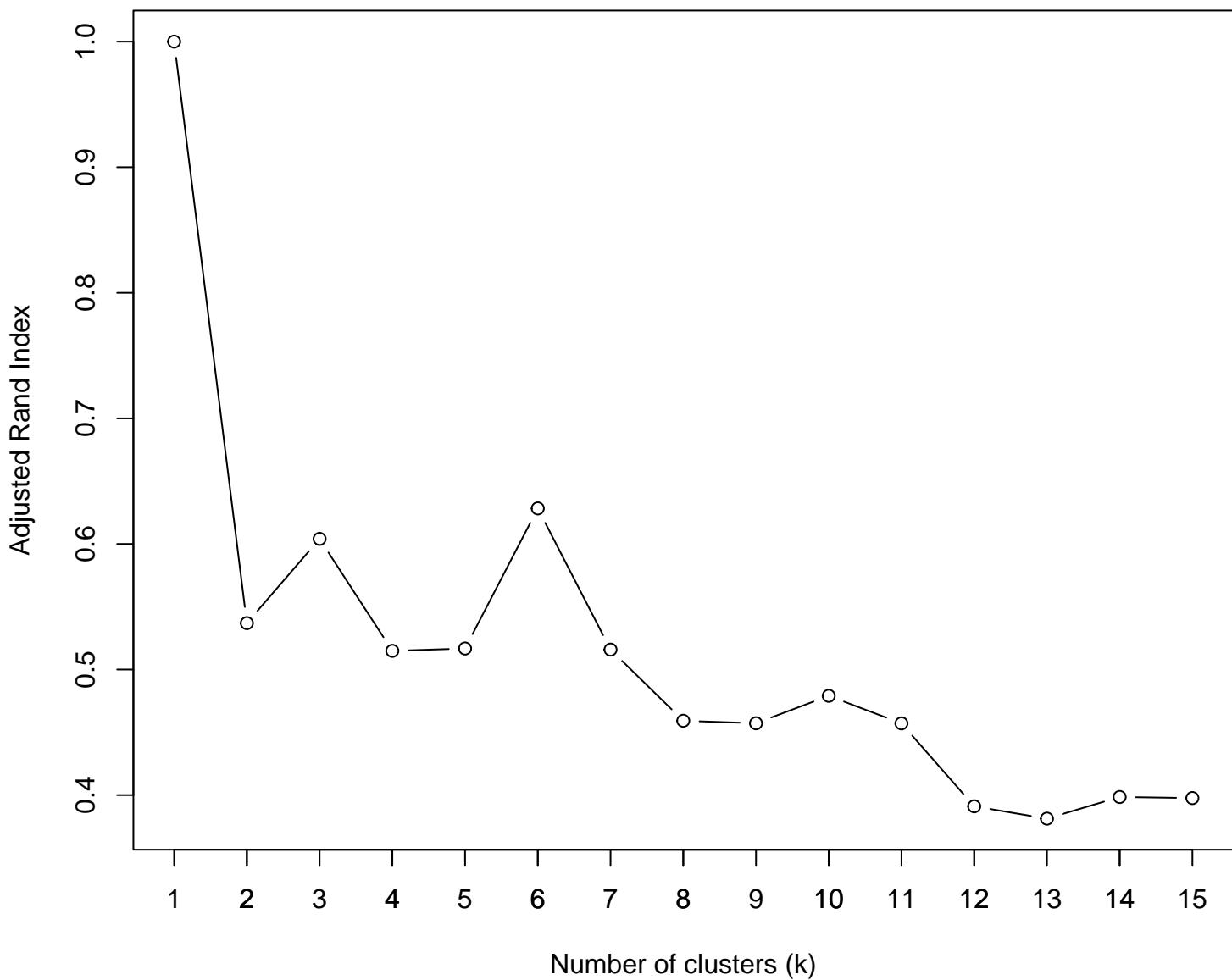
ARI vs k clusters on reduced data Z



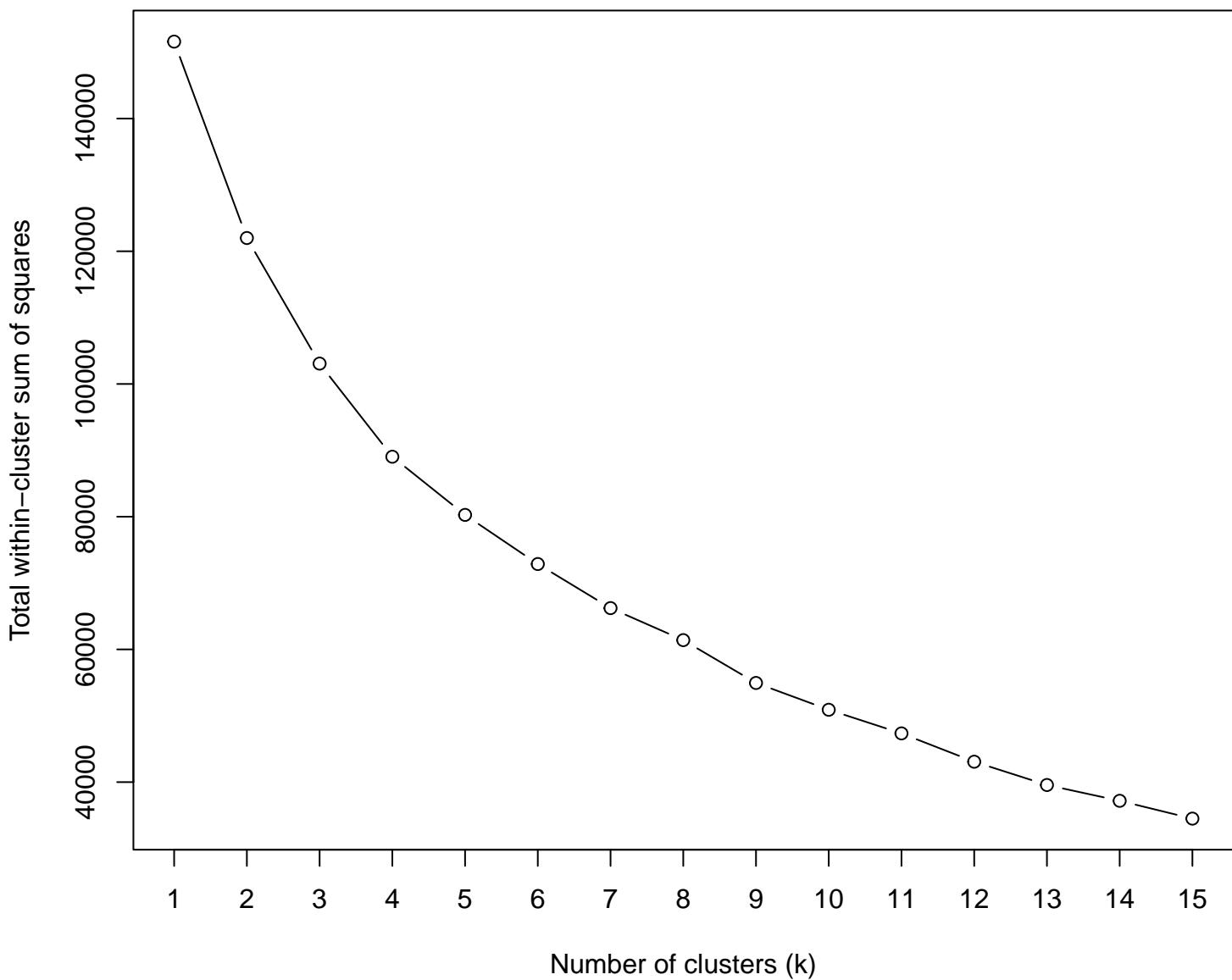
Elbow plot for k-means clustering on reduced data Z_75



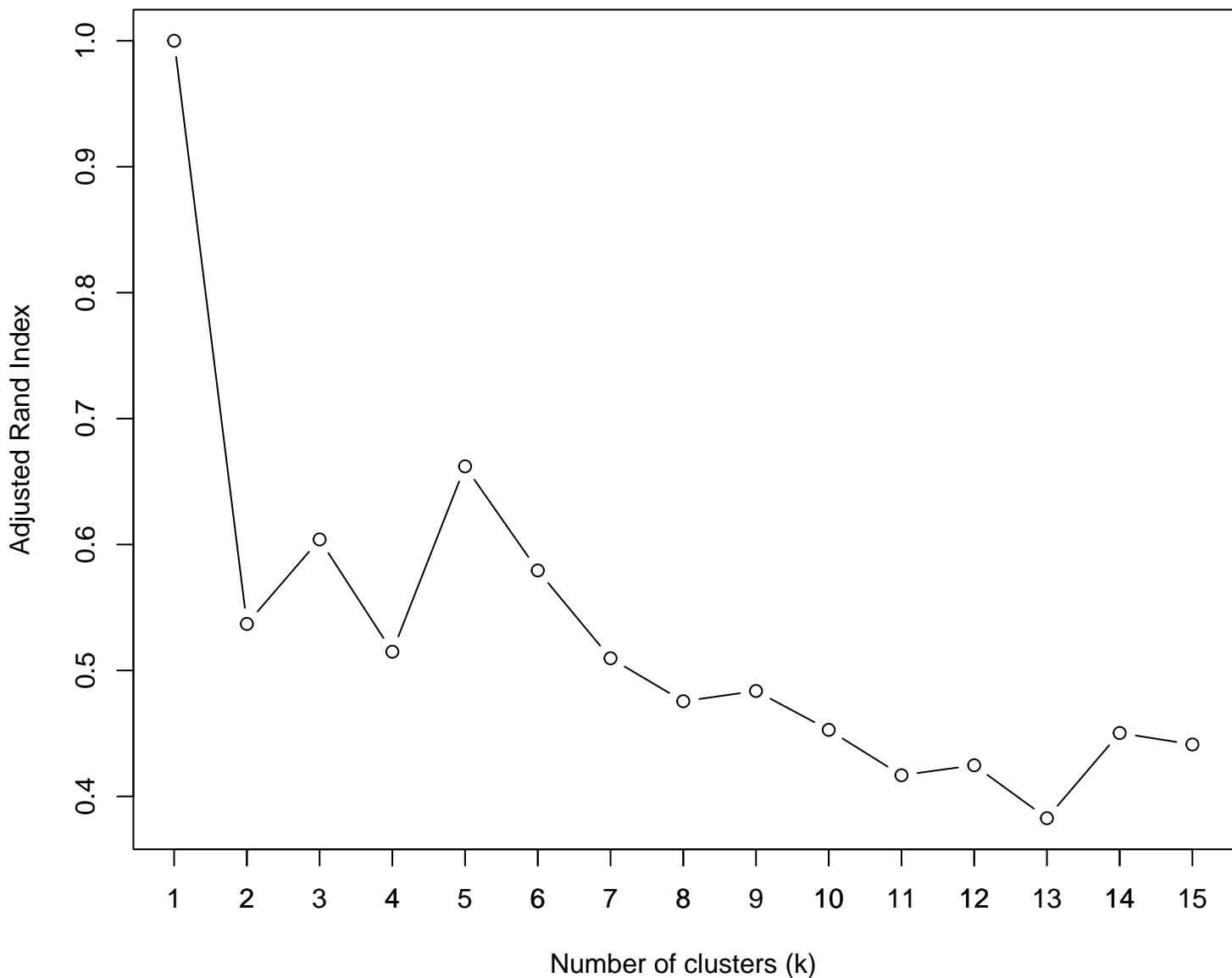
ARI vs k clusters on reduced data Z_75



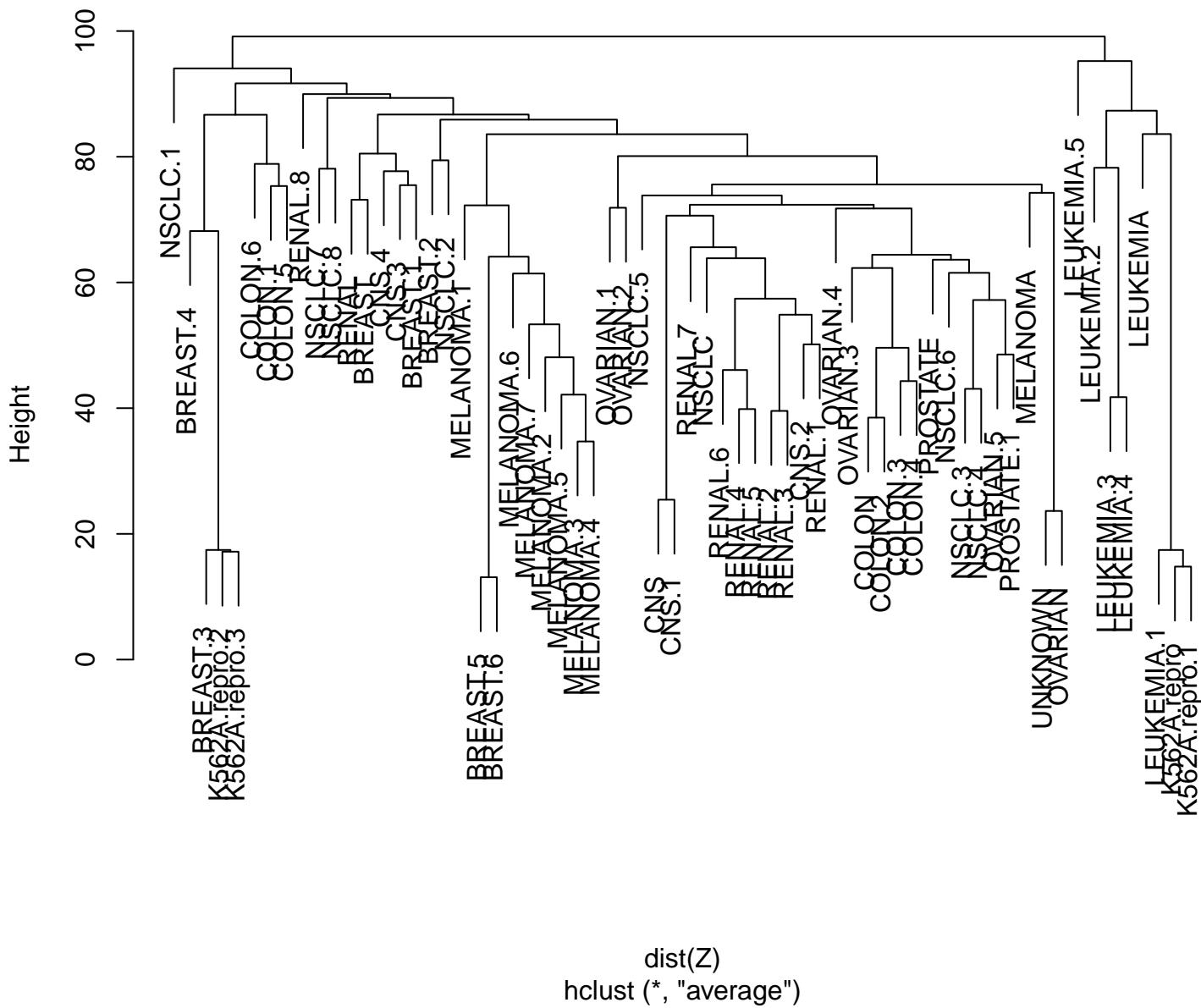
Elbow plot for k-means clustering on reduced data Z_sp



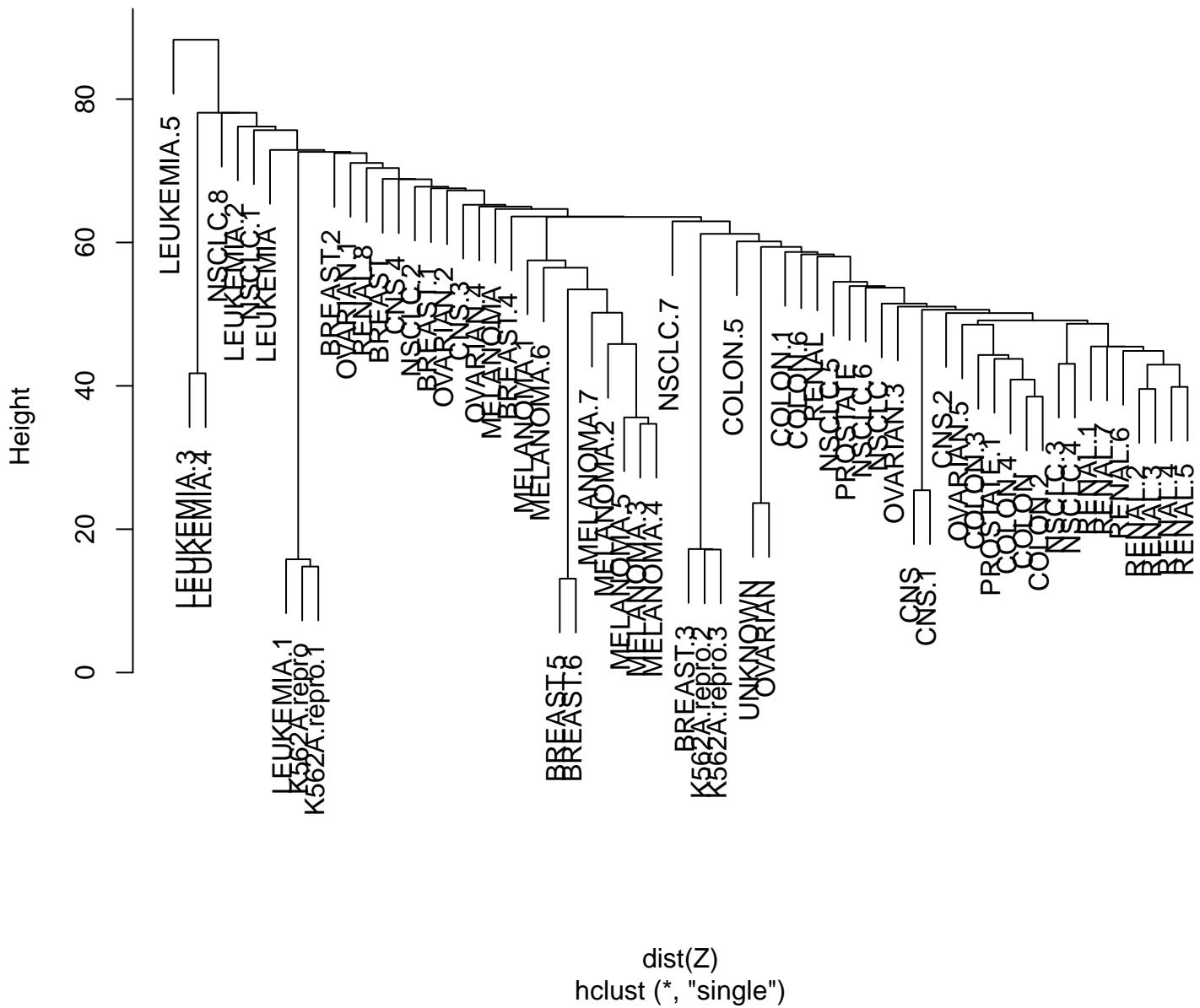
ARI vs k clusters on reduced data Z_sp



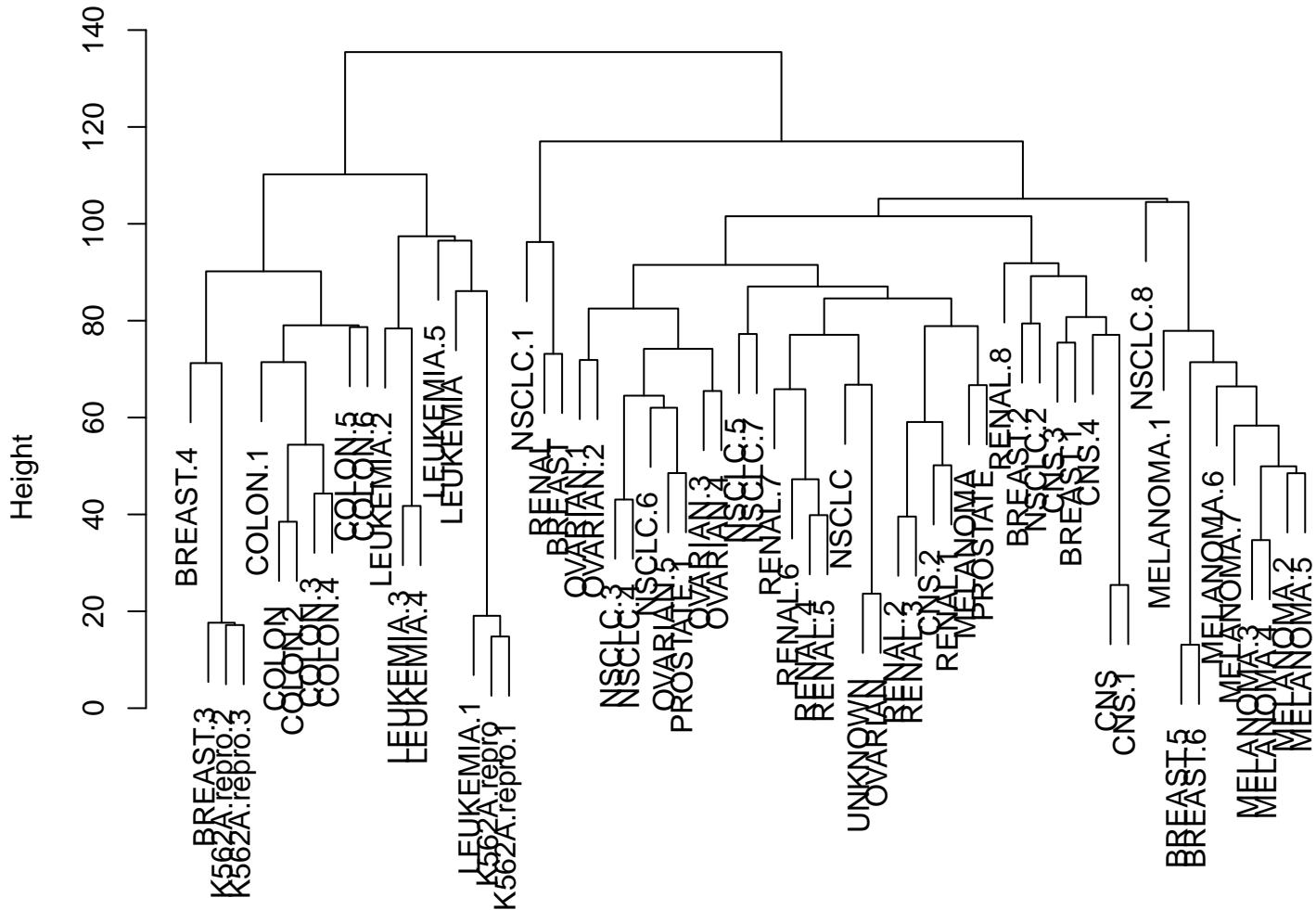
Average linkage on reduced data Z



Single linkage on reduced data Z

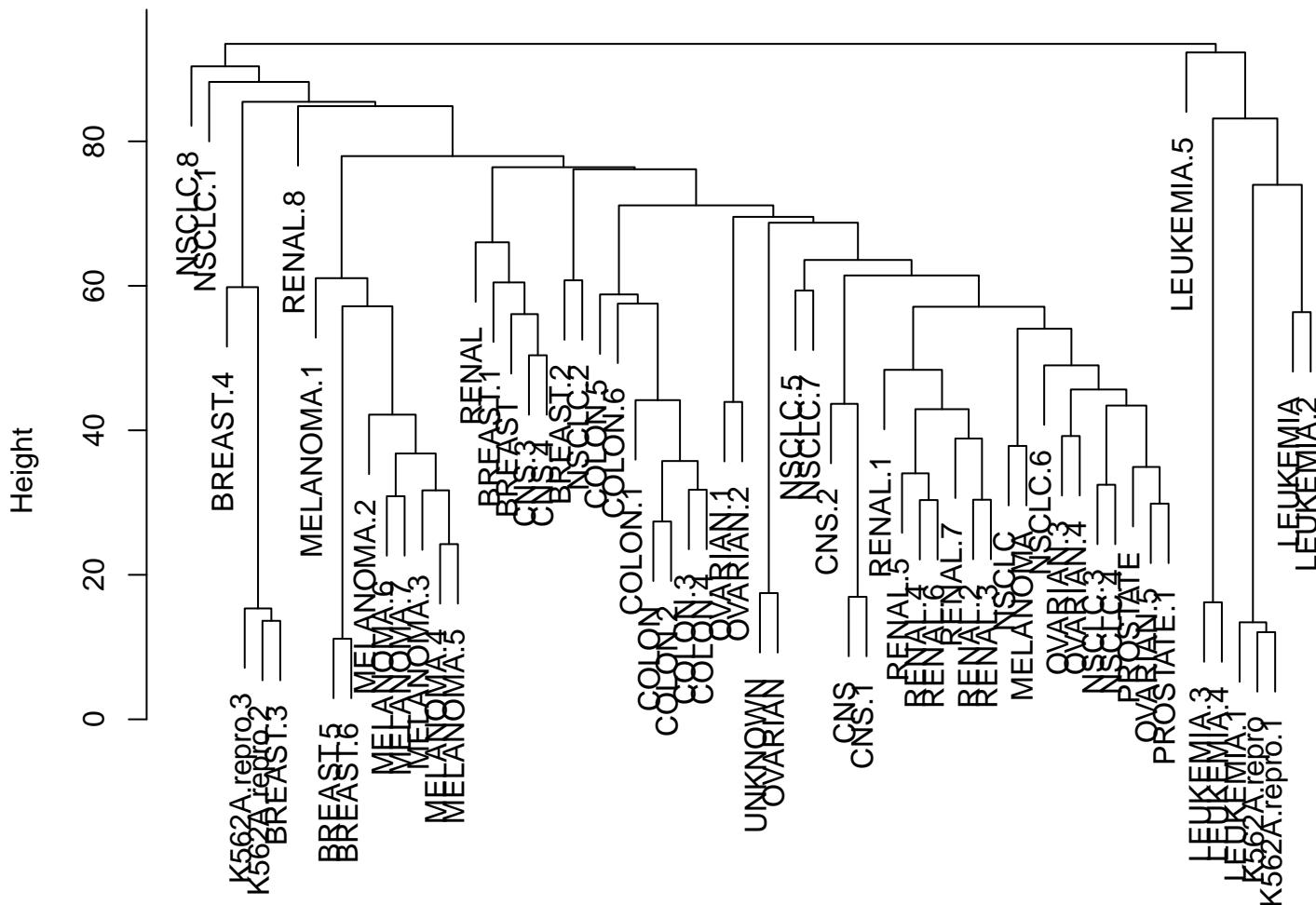


Complete linkage on reduced data Z



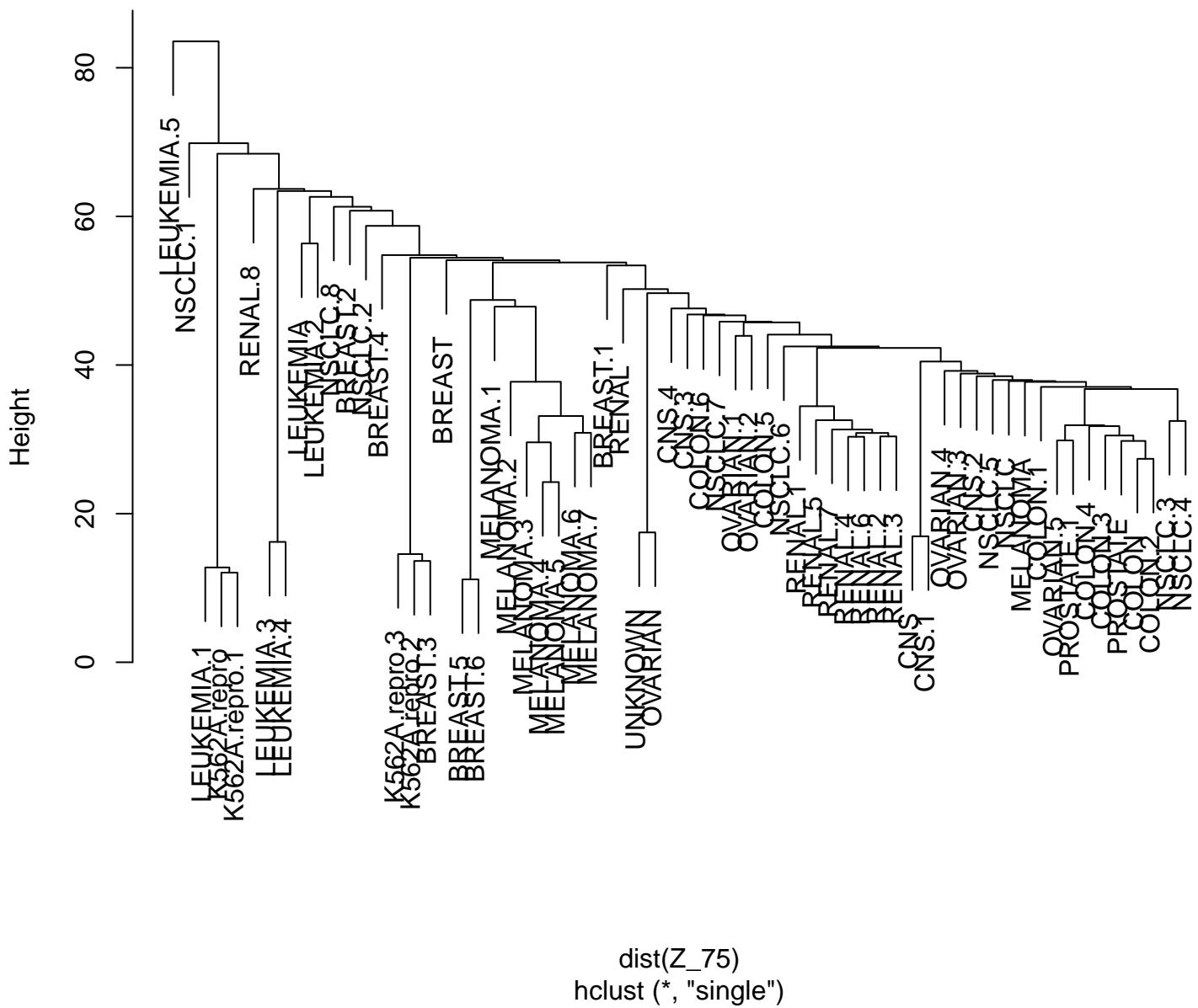
```
dist(Z)  
hclust (*, "complete")
```

Average linkage on reduced data Z_75

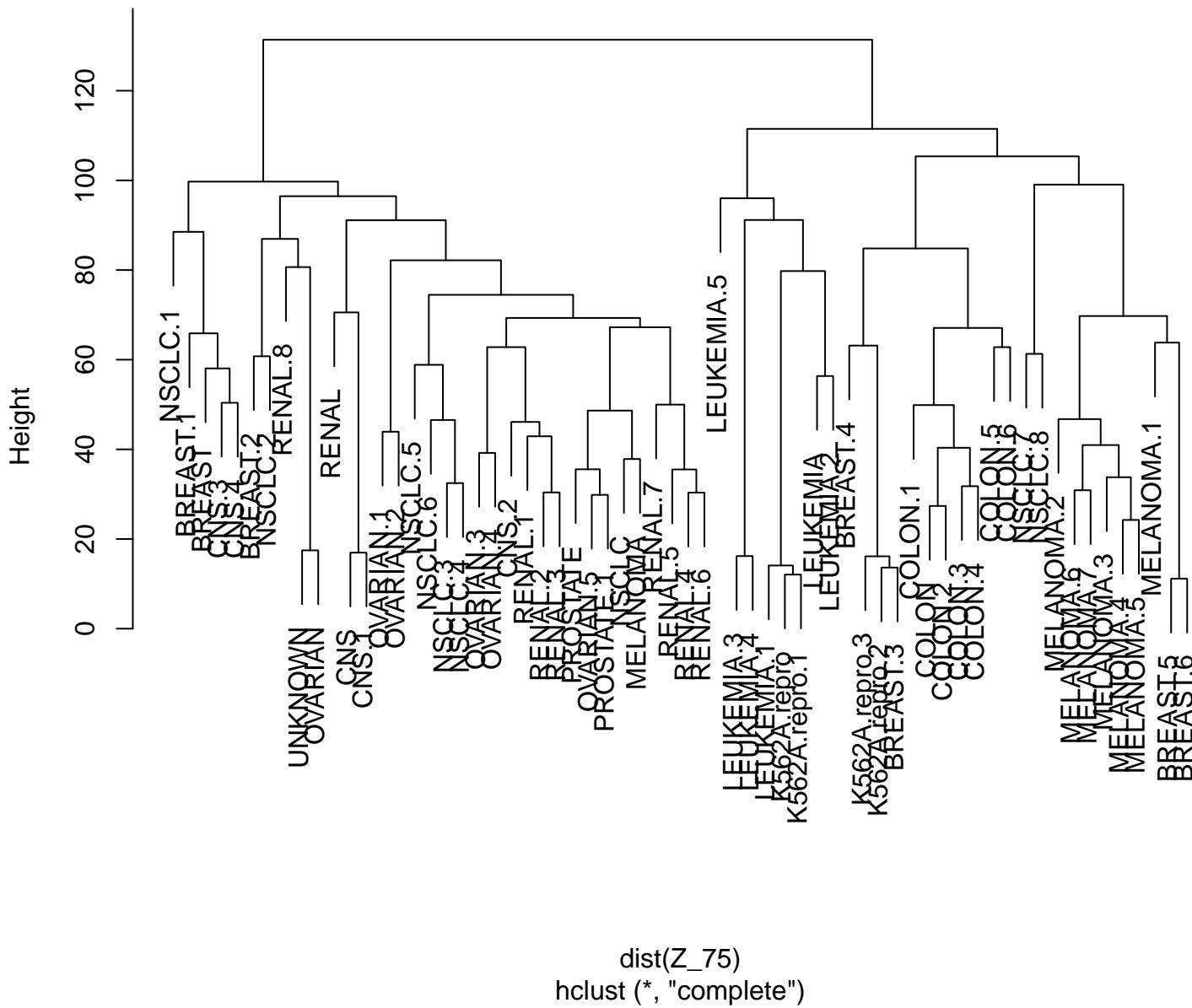


dist(Z_75)
hclust (*, "average")

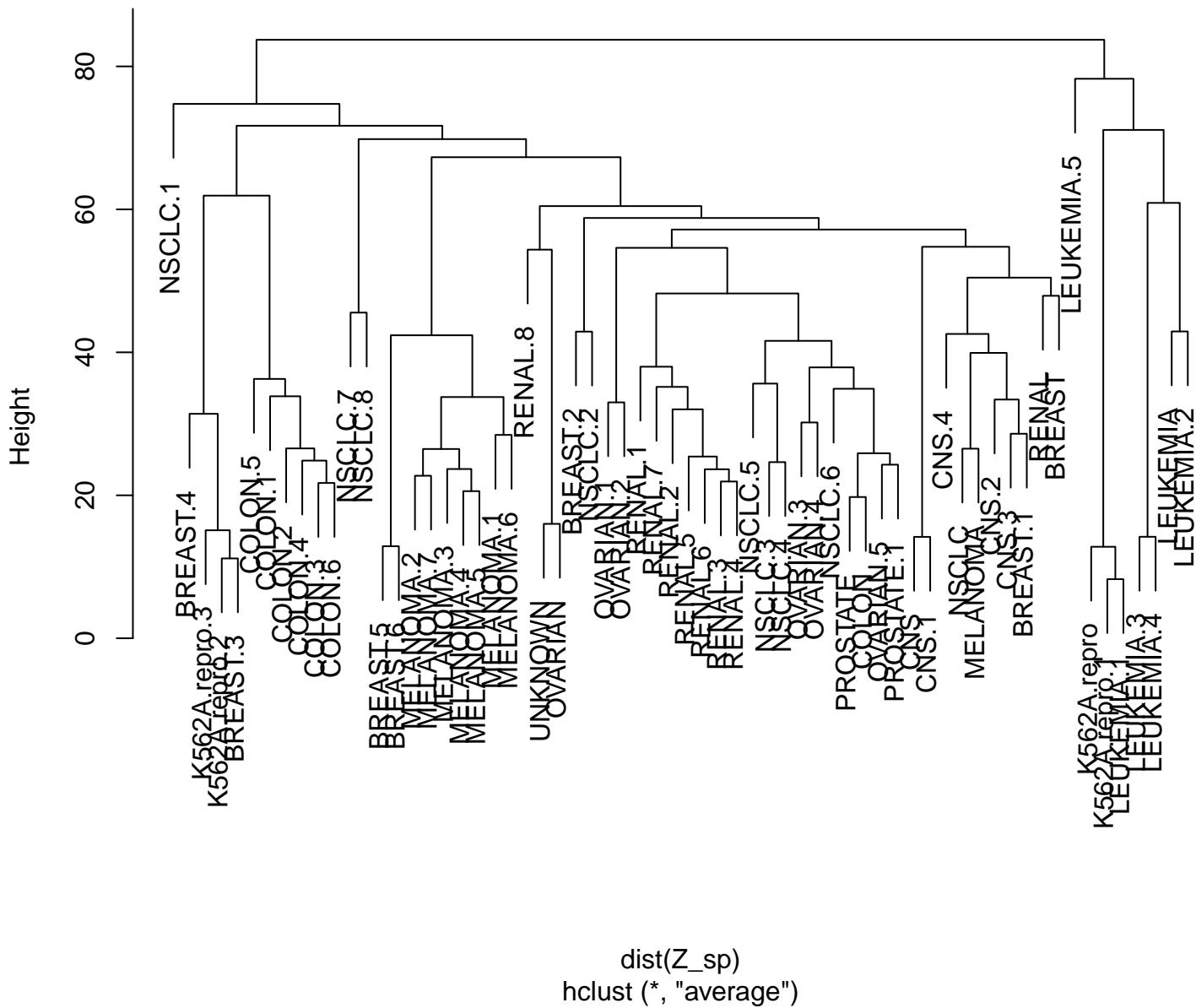
Single linkage on reduced data Z_75



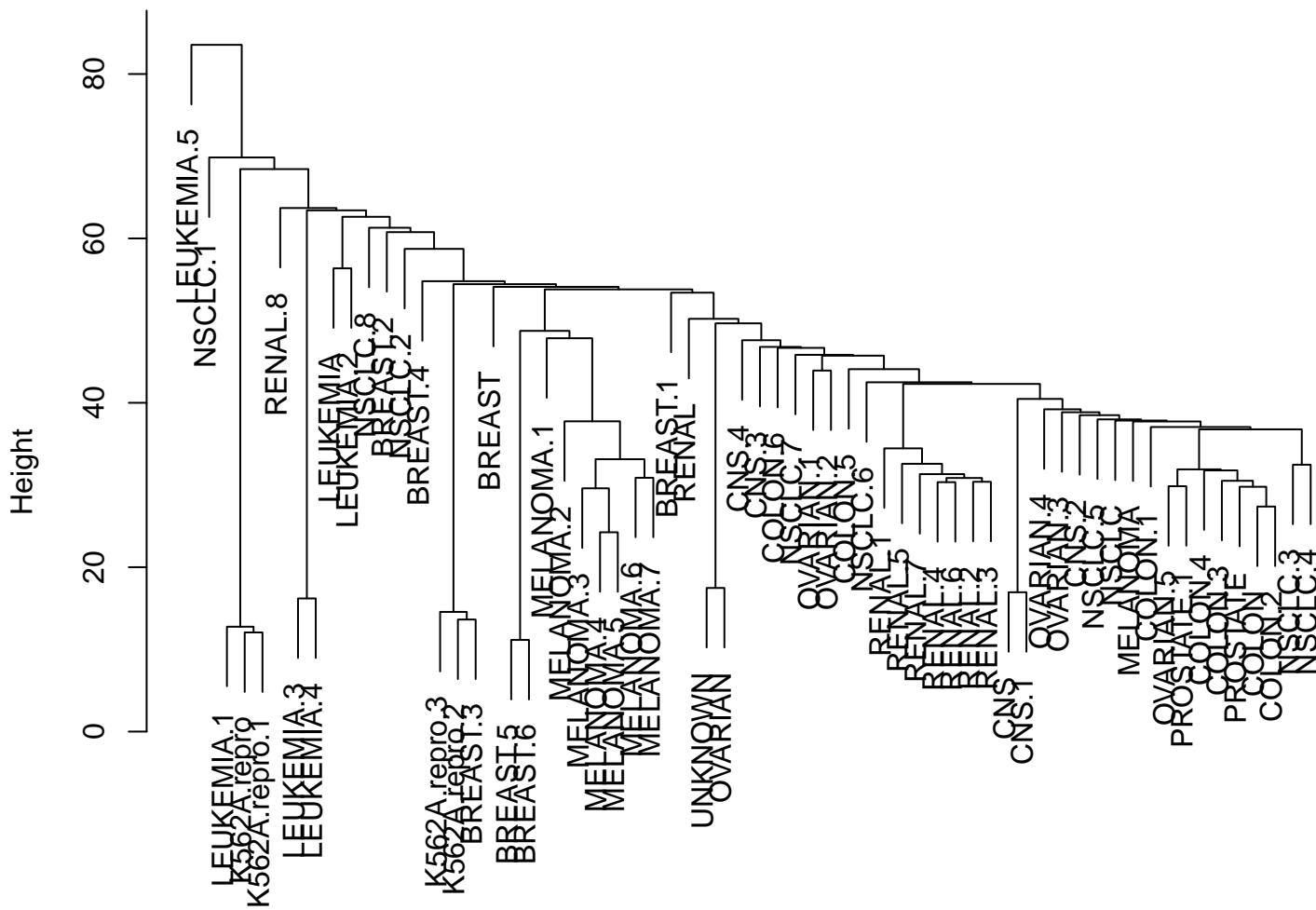
Complete linkage on reduced data Z_75



Average linkage on reduced data Z_sp

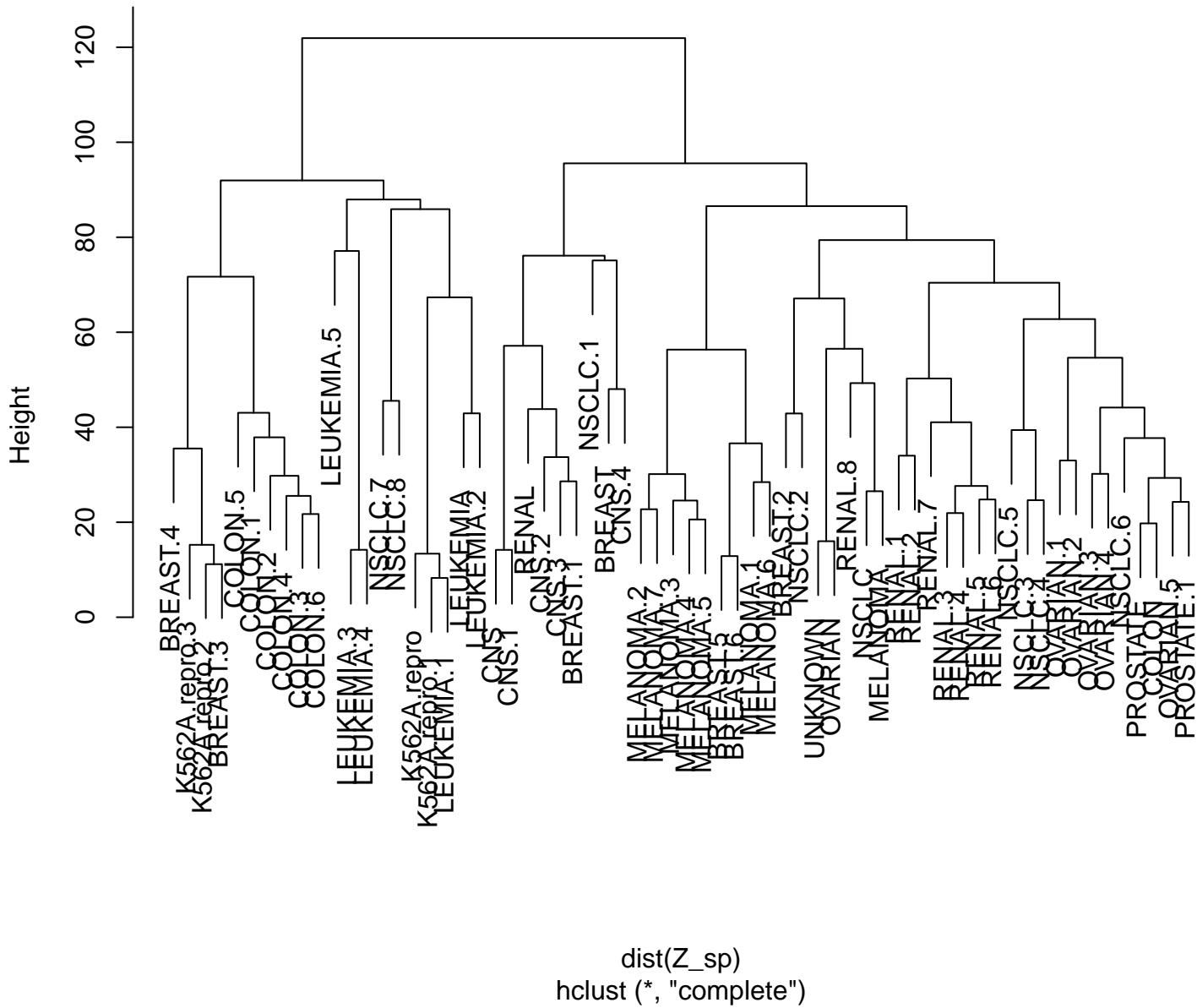


Single linkage on reduced data Z_sp

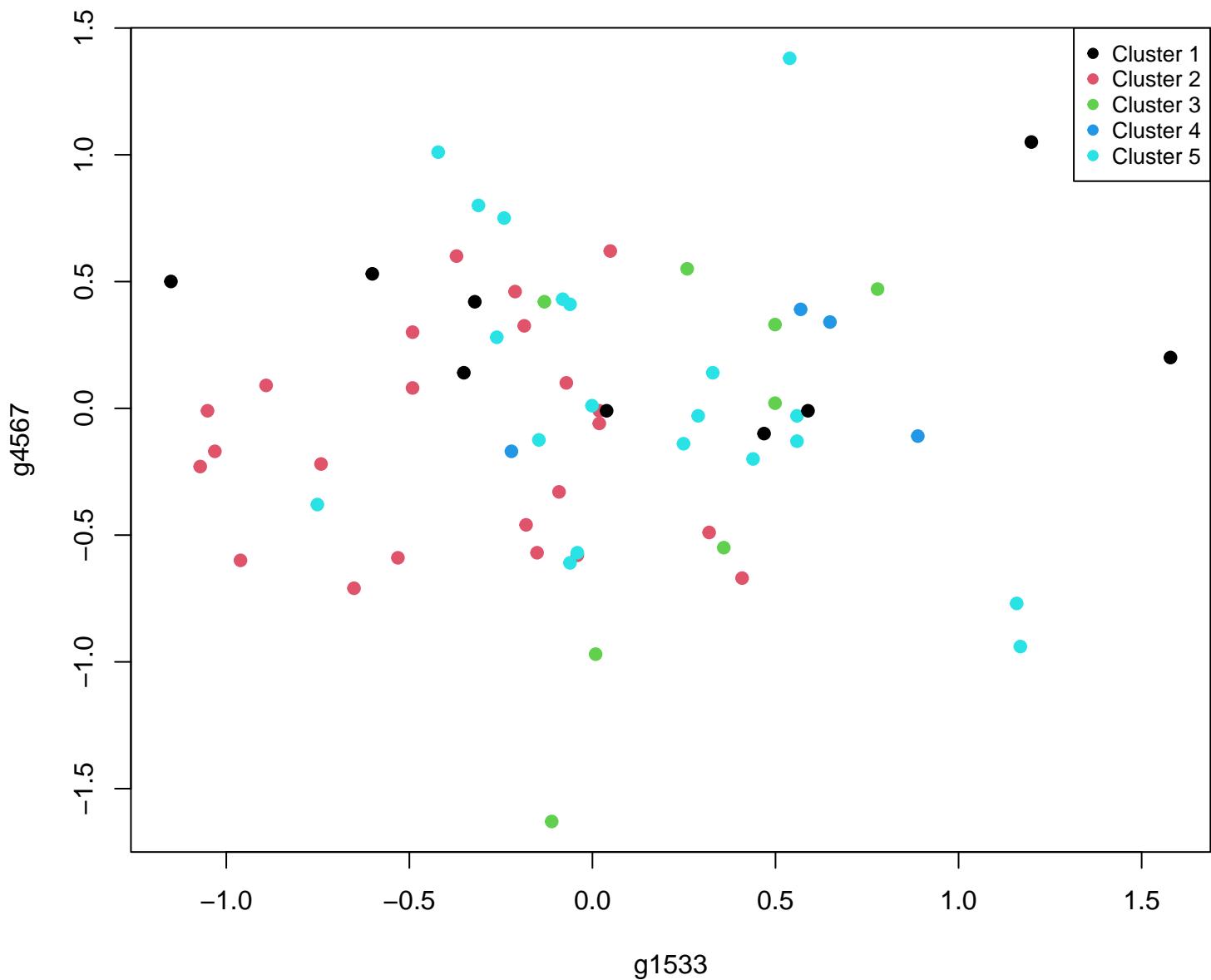


```
dist(Z_75)  
hclust (*, "single")
```

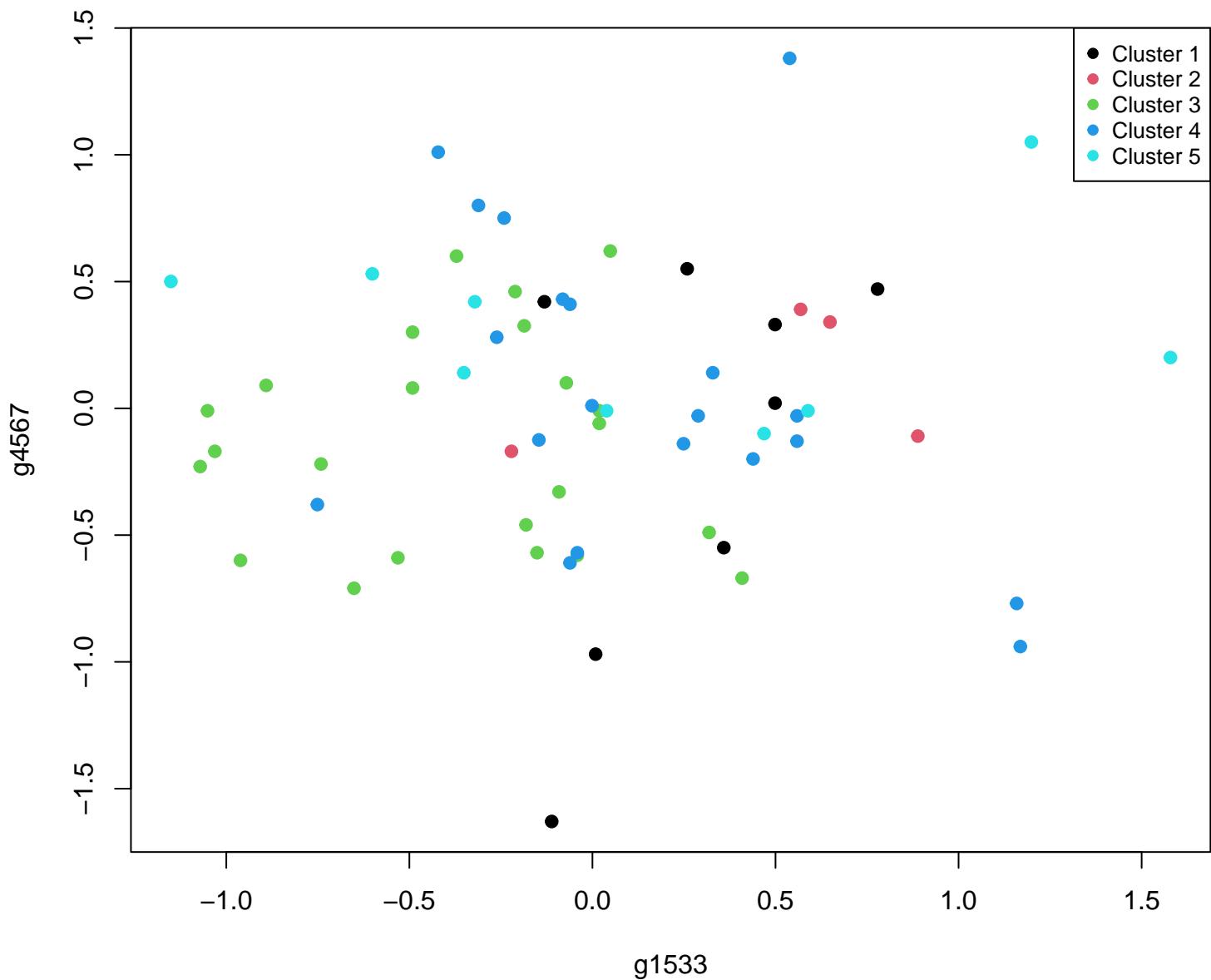
Complete linkage on reduced data Z_sp



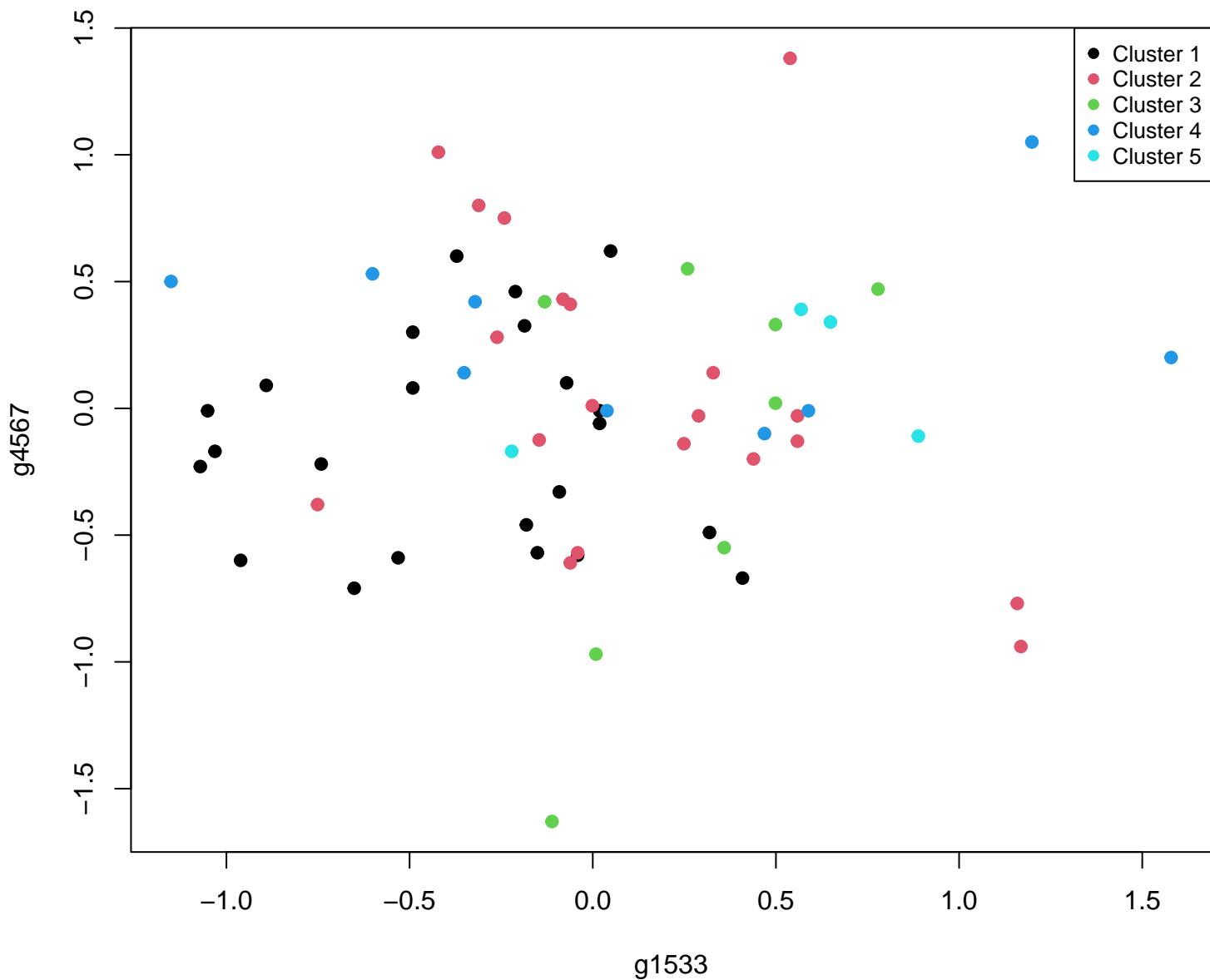
Reduced data Z k-means clusters with 5 clusters (using 2 randomly sampled genes)



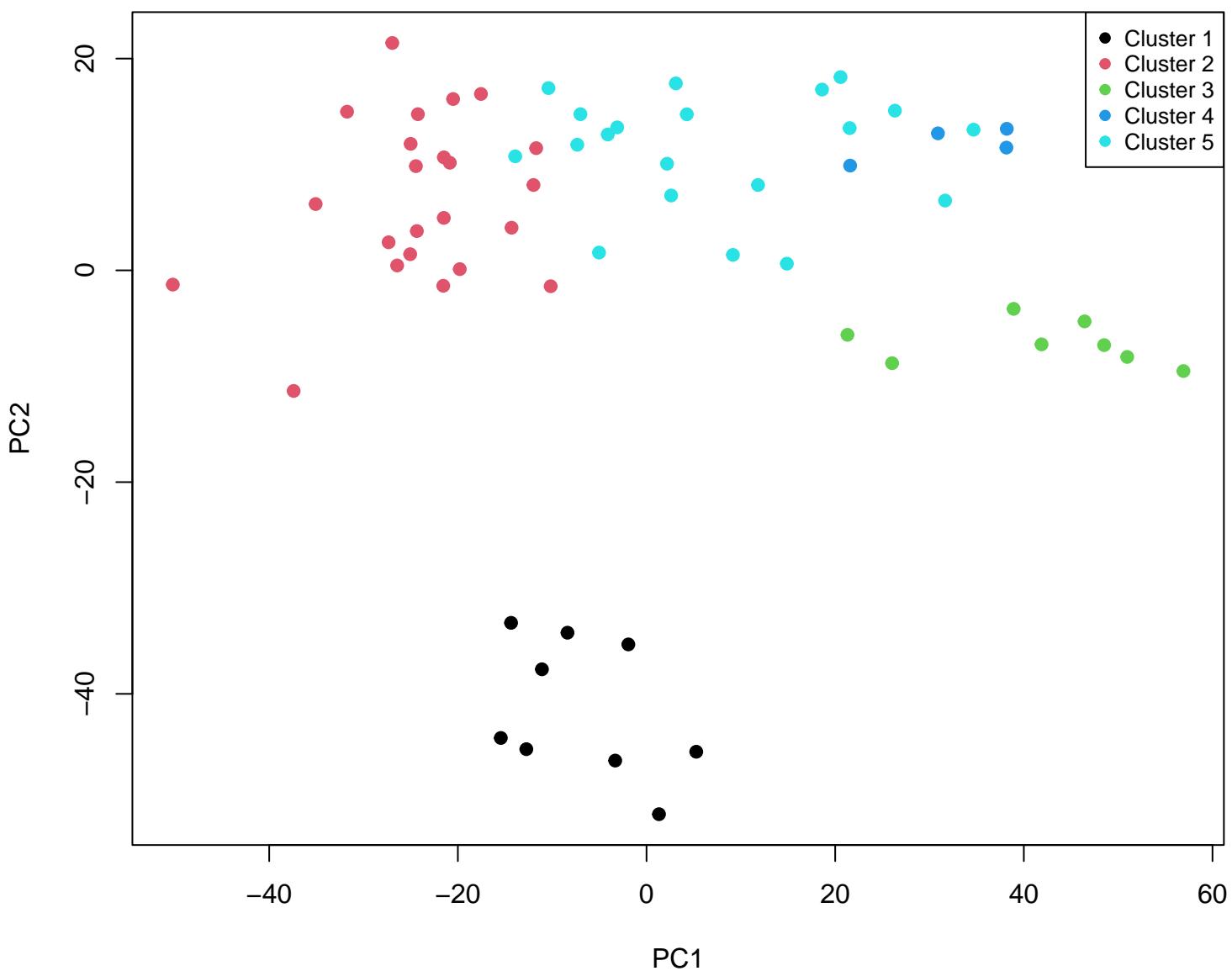
Reduced data Z_75 k-means clusters with 5 clusters (using 2 randomly sampled genes)



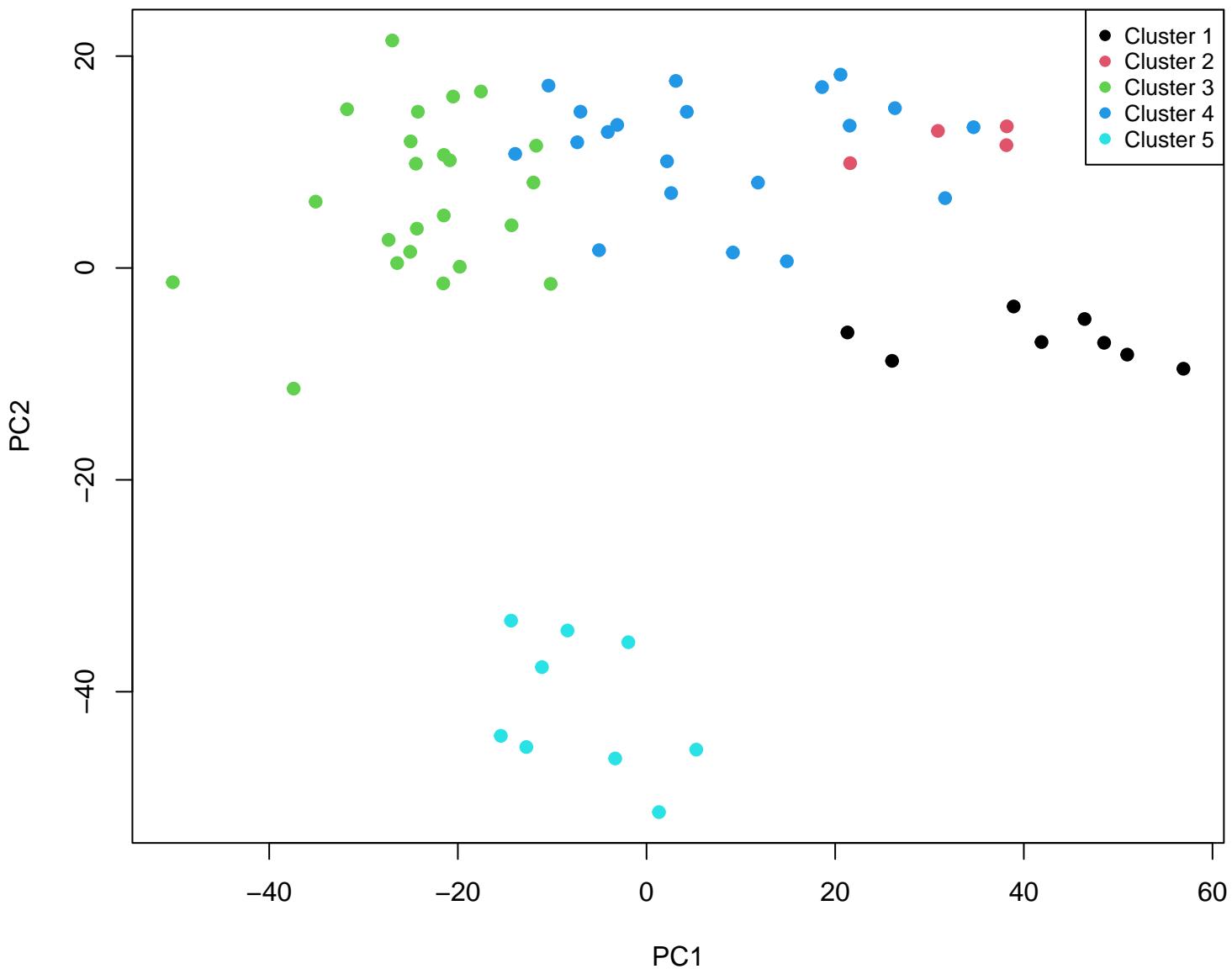
Reduced data Z_sp k-means clusters with 5 clusters (using 2 randomly sampled genes)



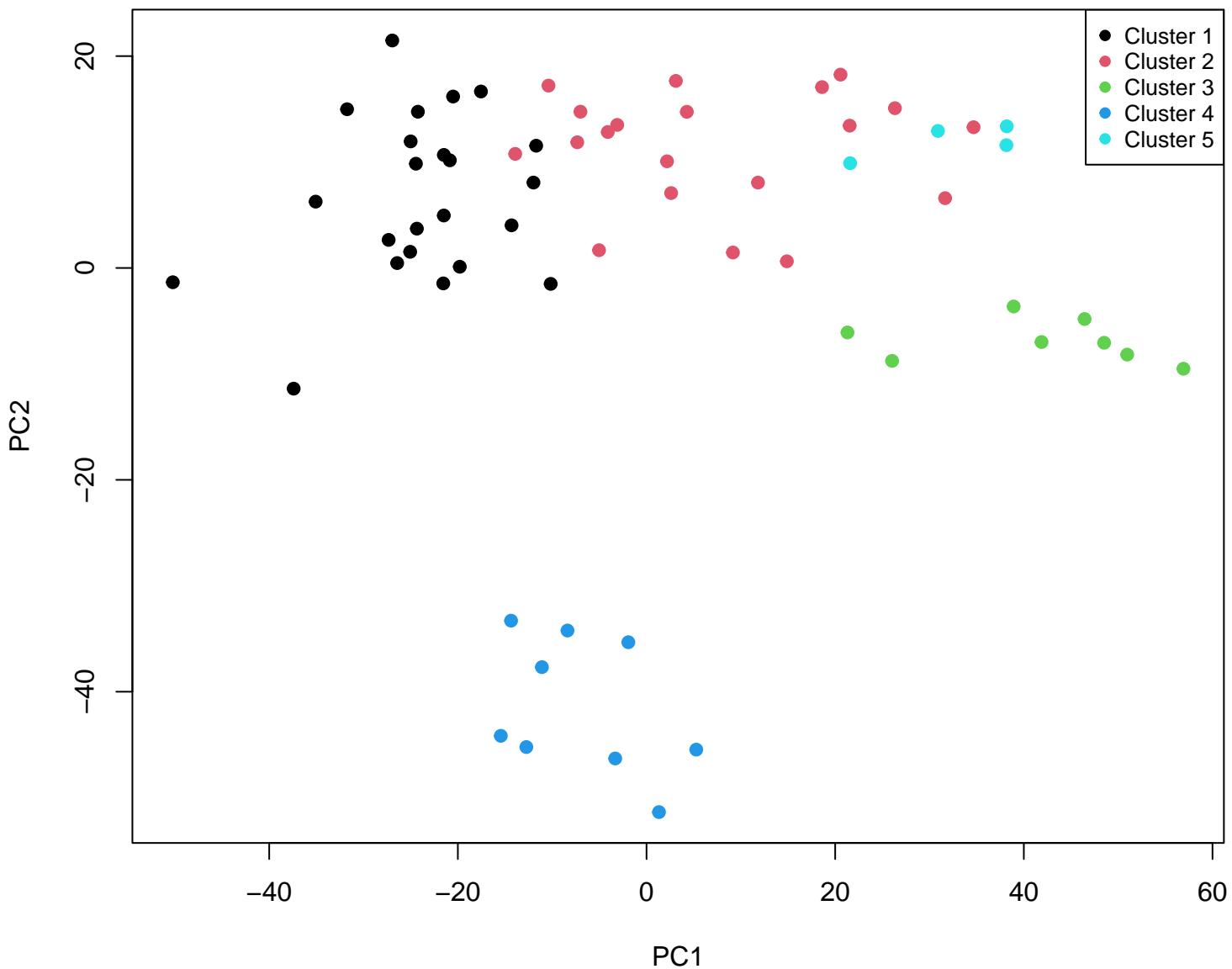
K-means clustering (k=5) on reduced data Z



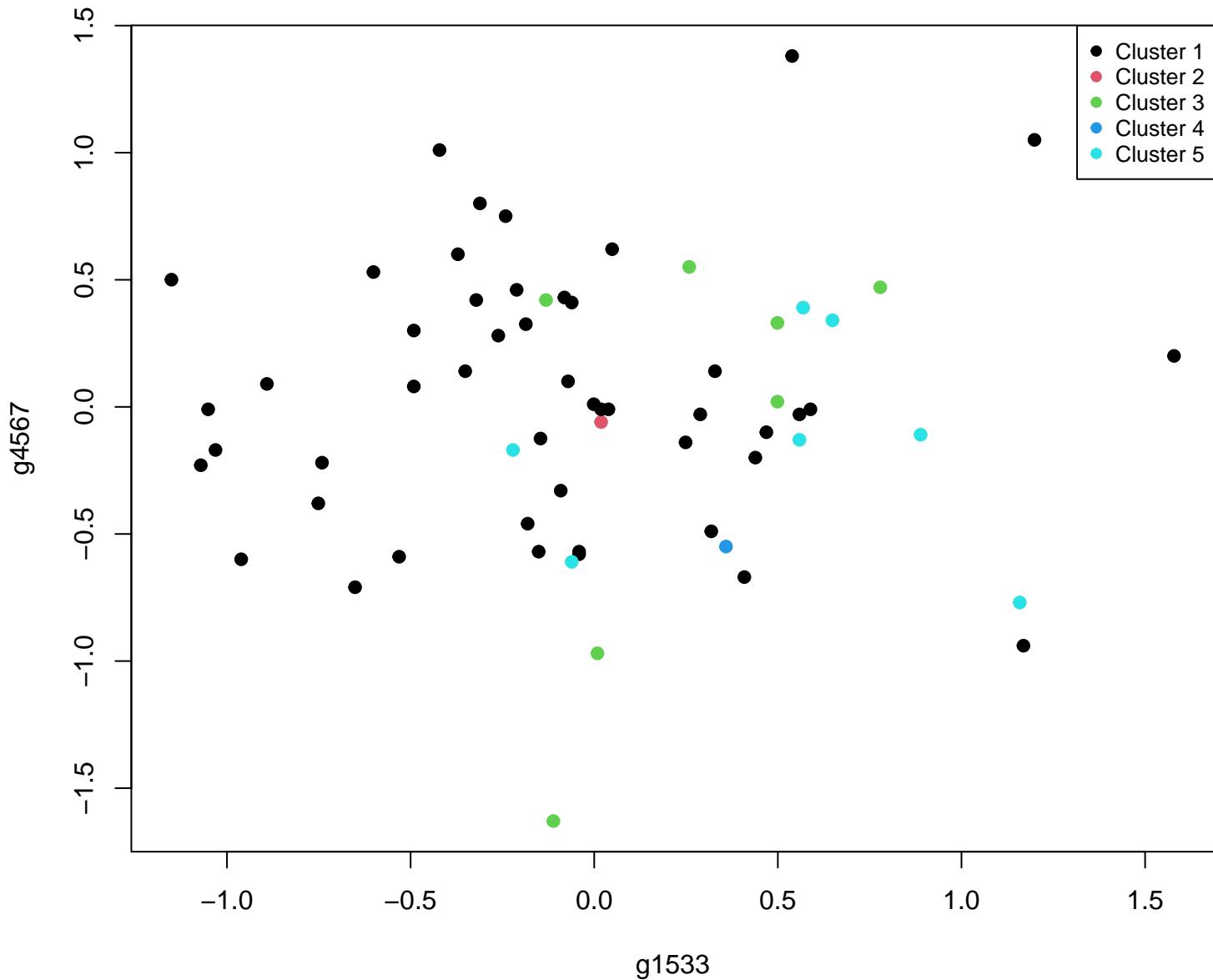
K-means clustering (k=5) on reduced data Z_75



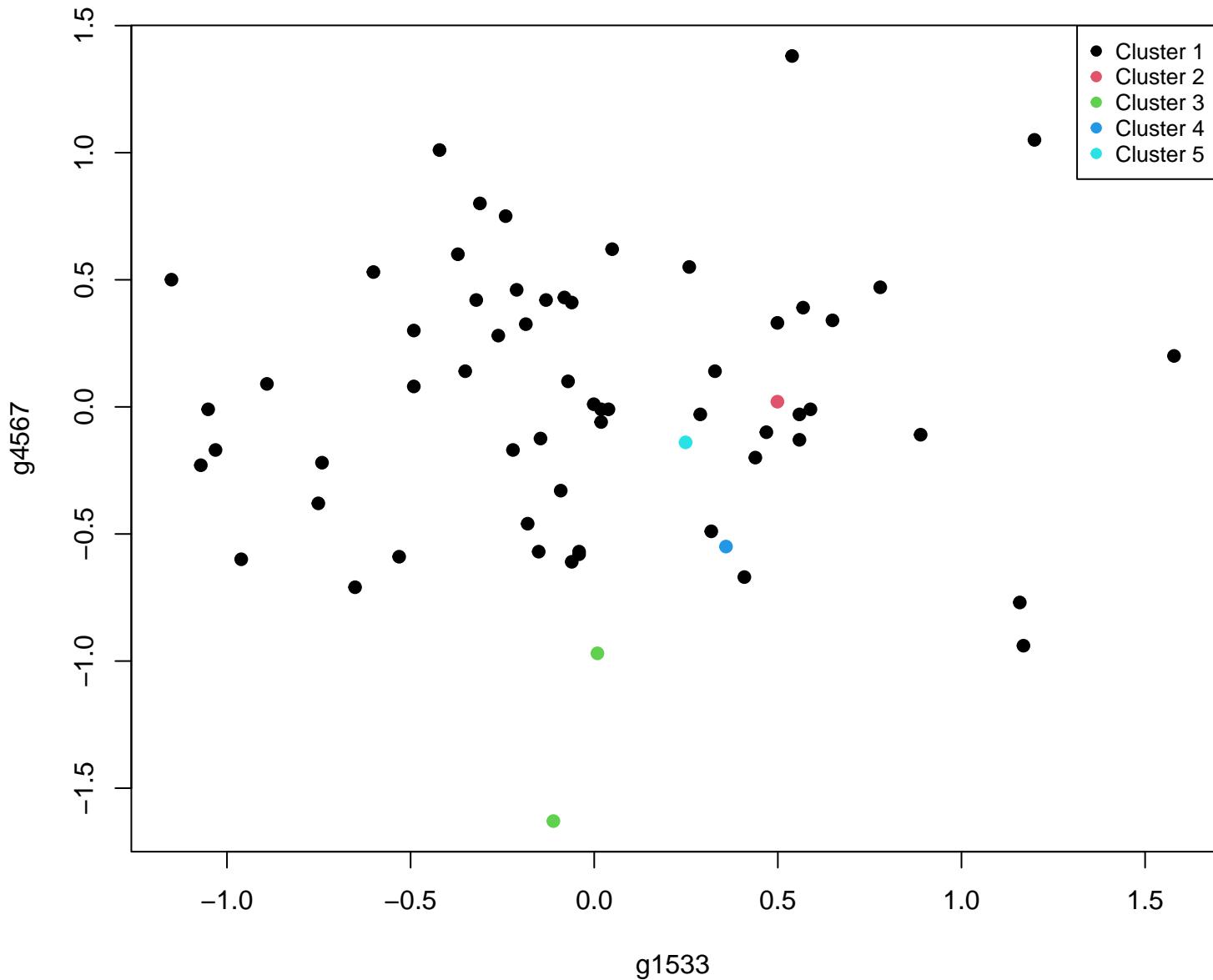
K-means clustering (k=5) on reduced data Z_sp



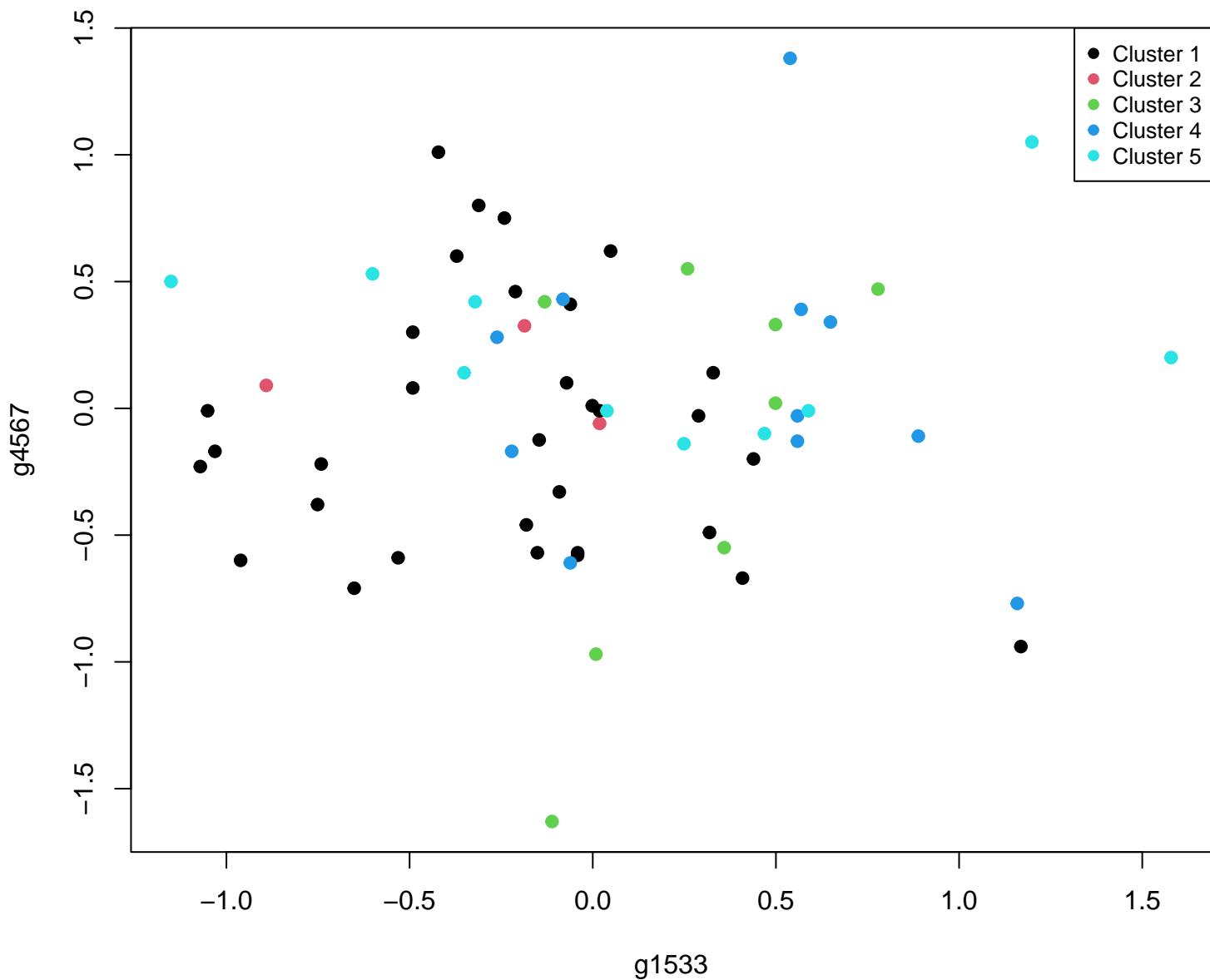
Average linkage with 5 cluster cutoff on Z (sample)



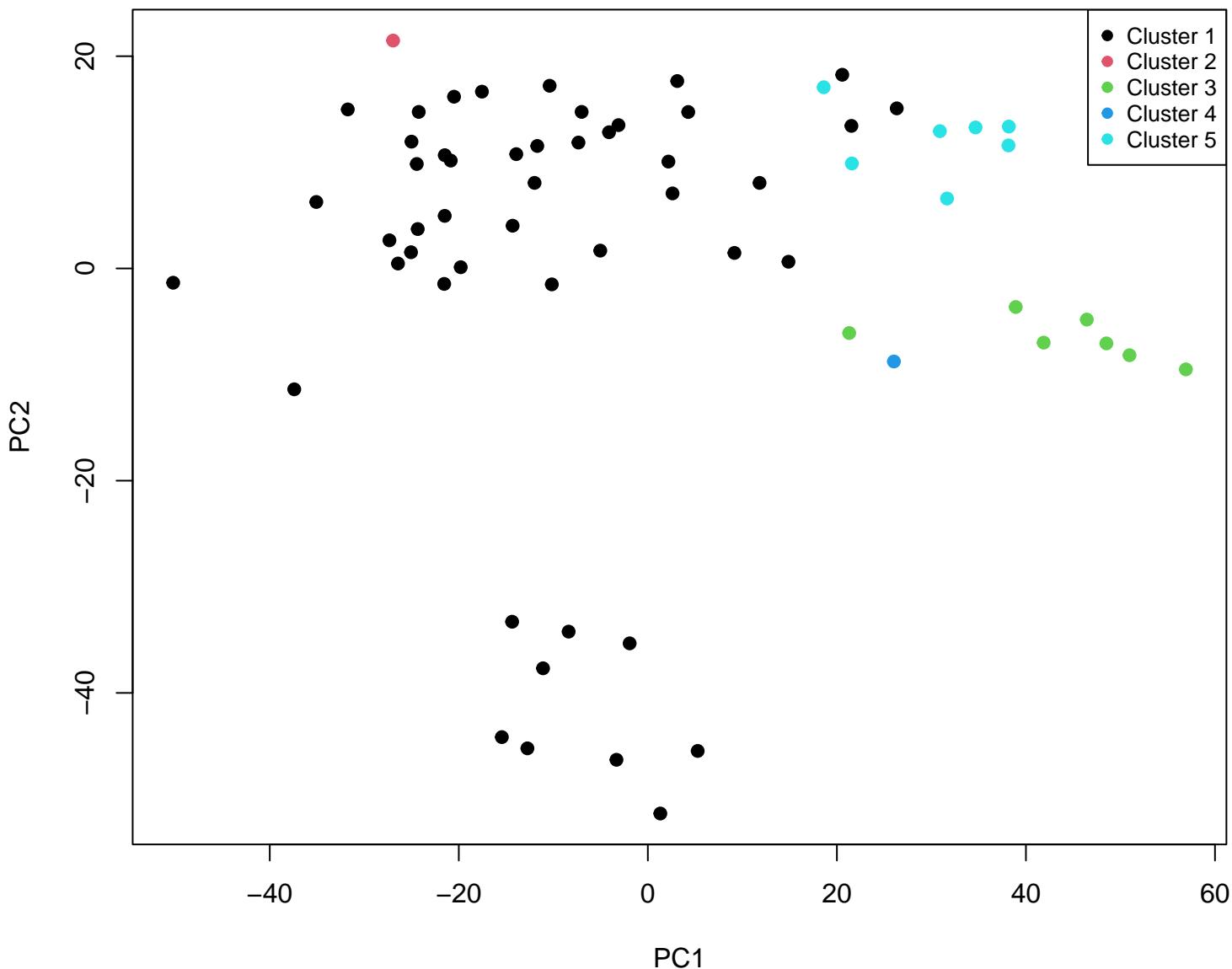
Single linkage with 5 cluster cutoff on Z (sample)



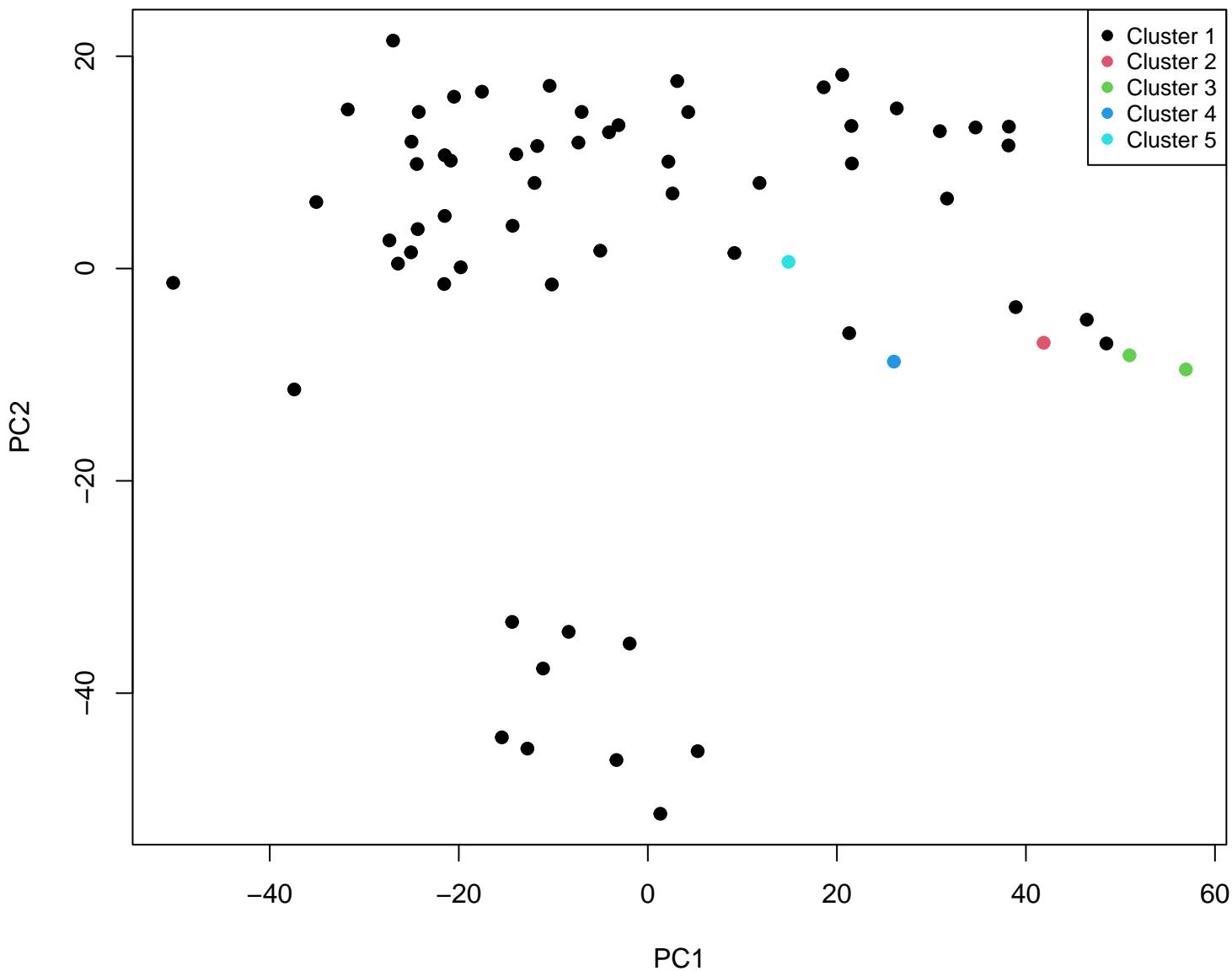
Complete linkage with 5 cluster cutoff on Z (sample)



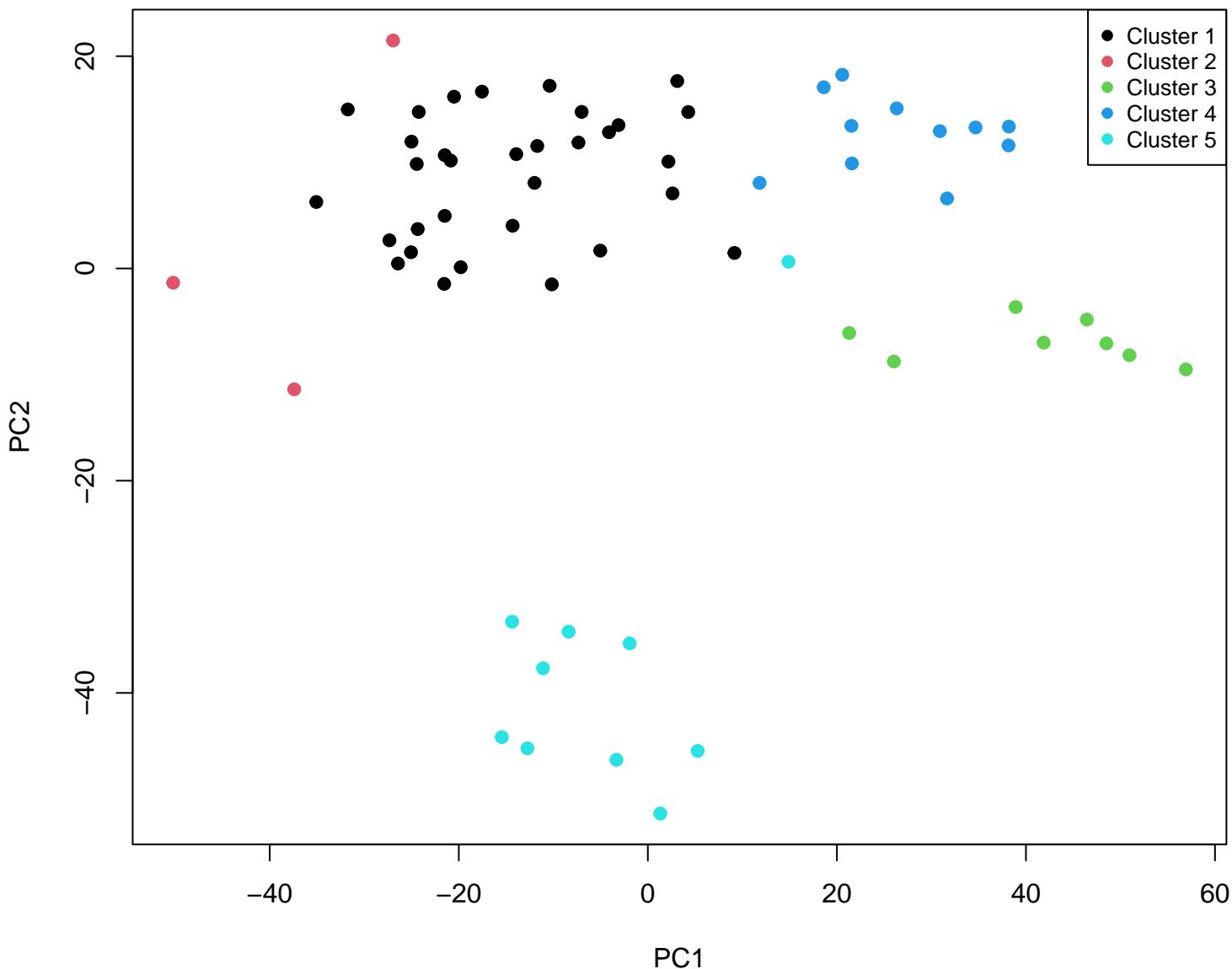
Average linkage clusters (first 2 PCs, 5 cluster cut-off) on Z



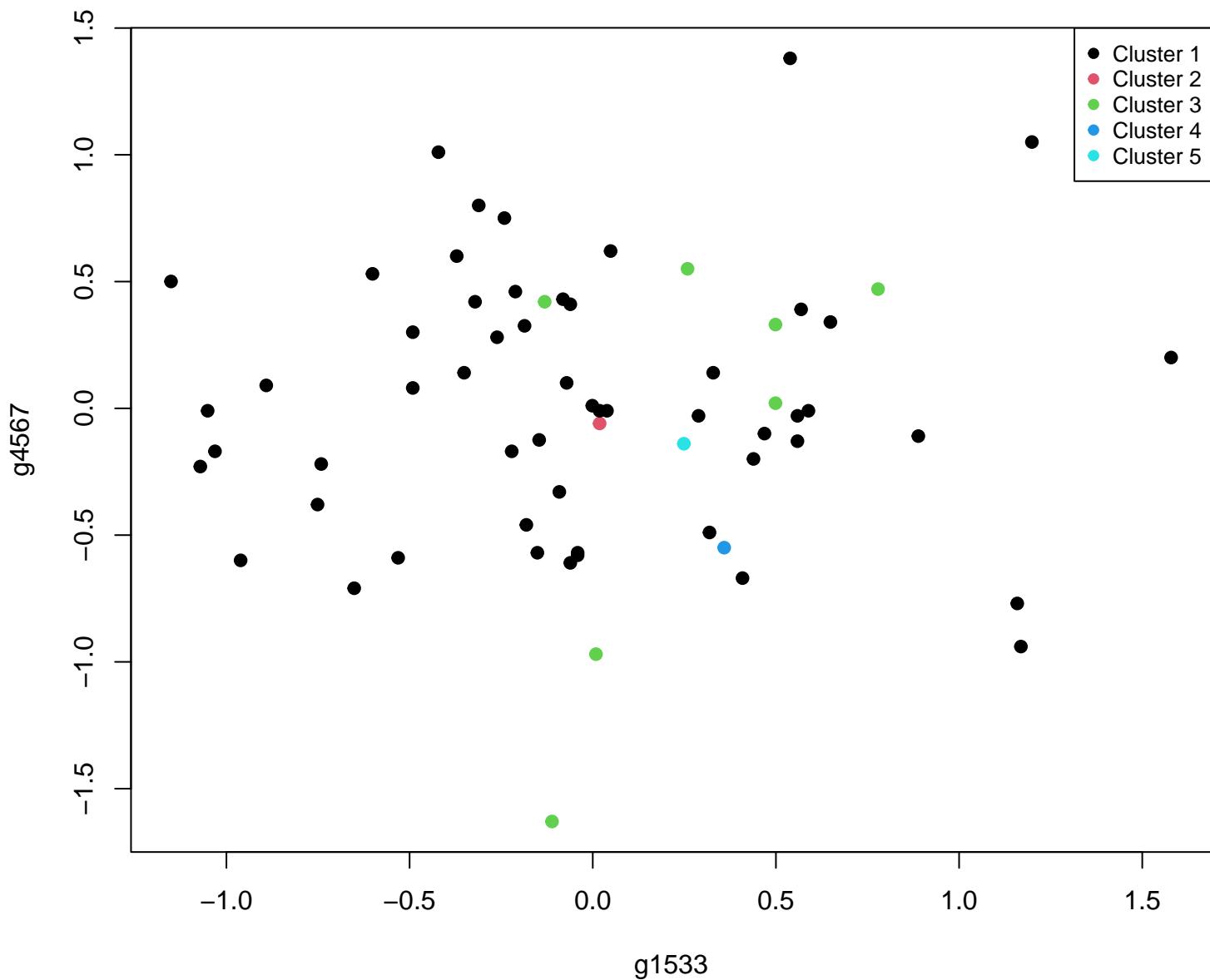
Single linkage clusters (first 2 PCs, 5 cluster cut-off) on Z



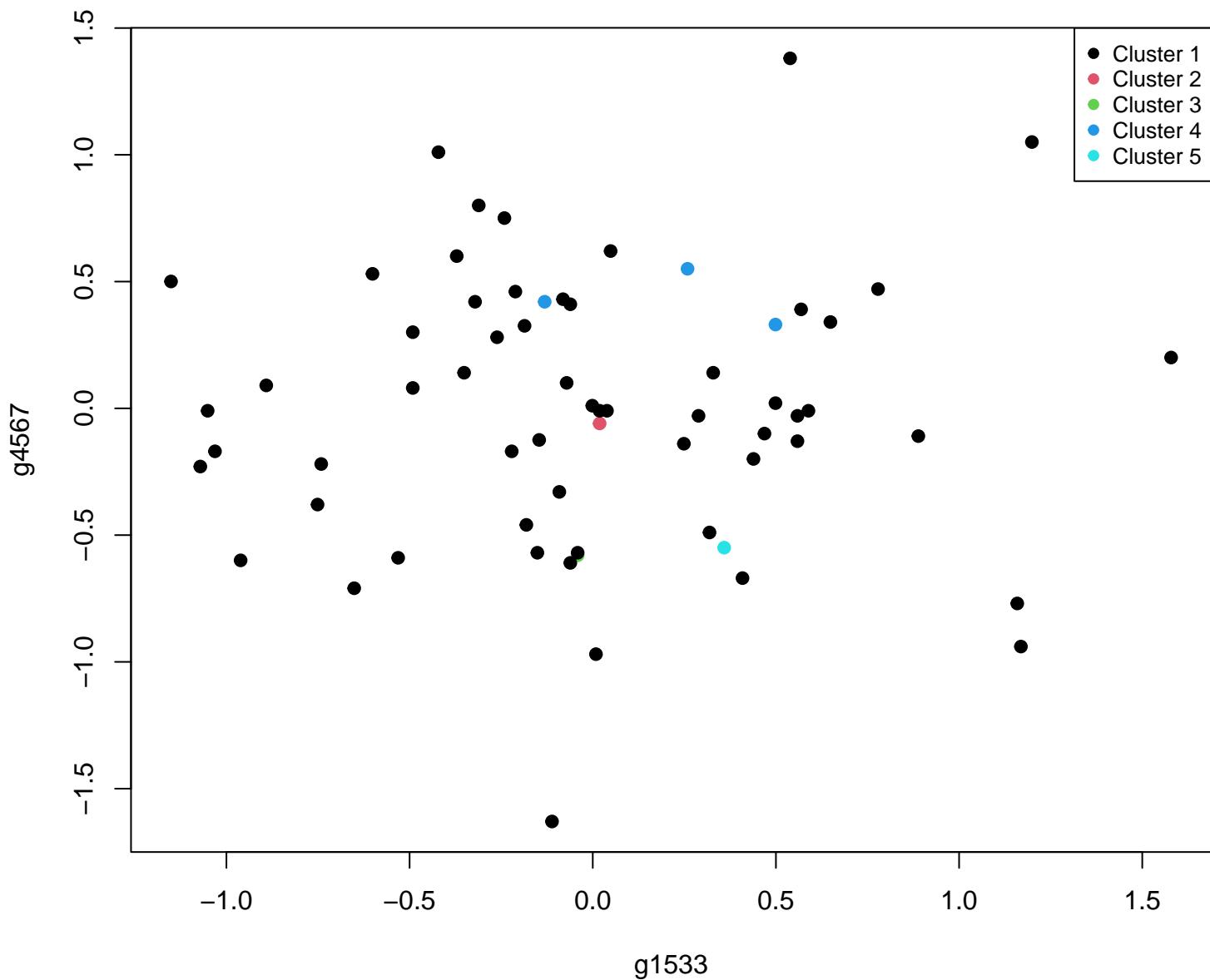
Complete linkage clusters (first 2 PCs, 5 cluster cut-off) on Z



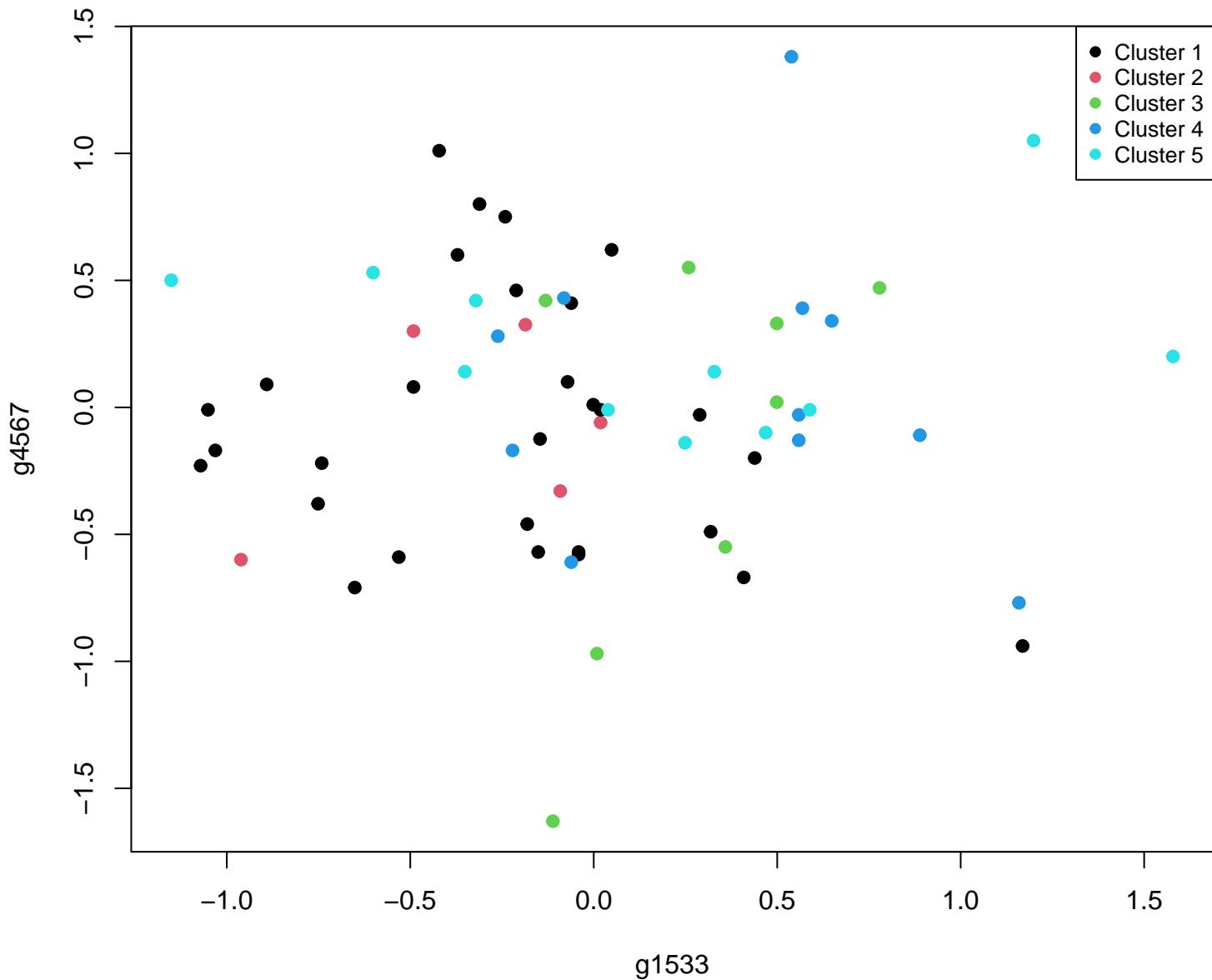
Average linkage with 5 cluster cutoff on Z_75



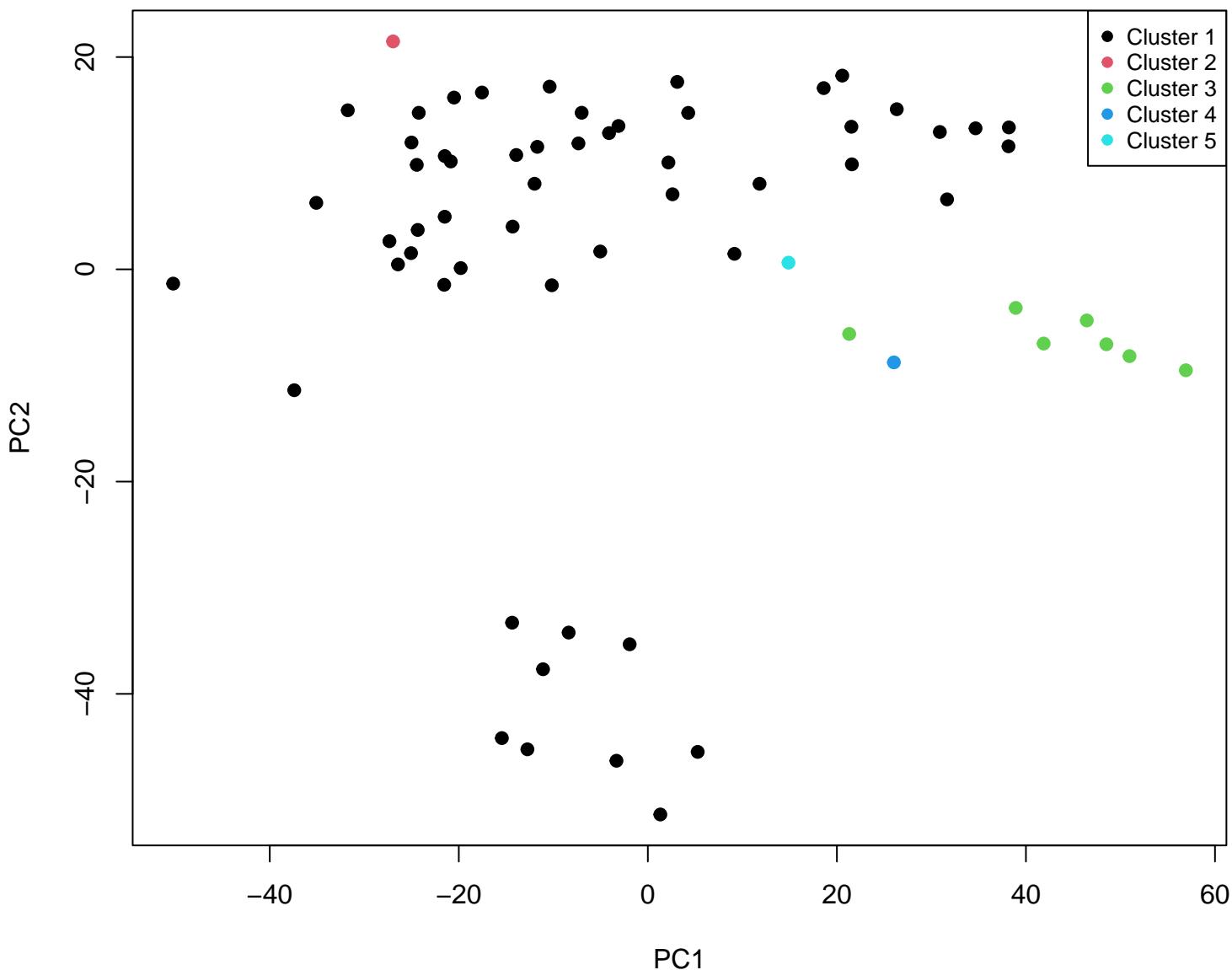
Single linkage with 5 cluster cutoff on Z_75



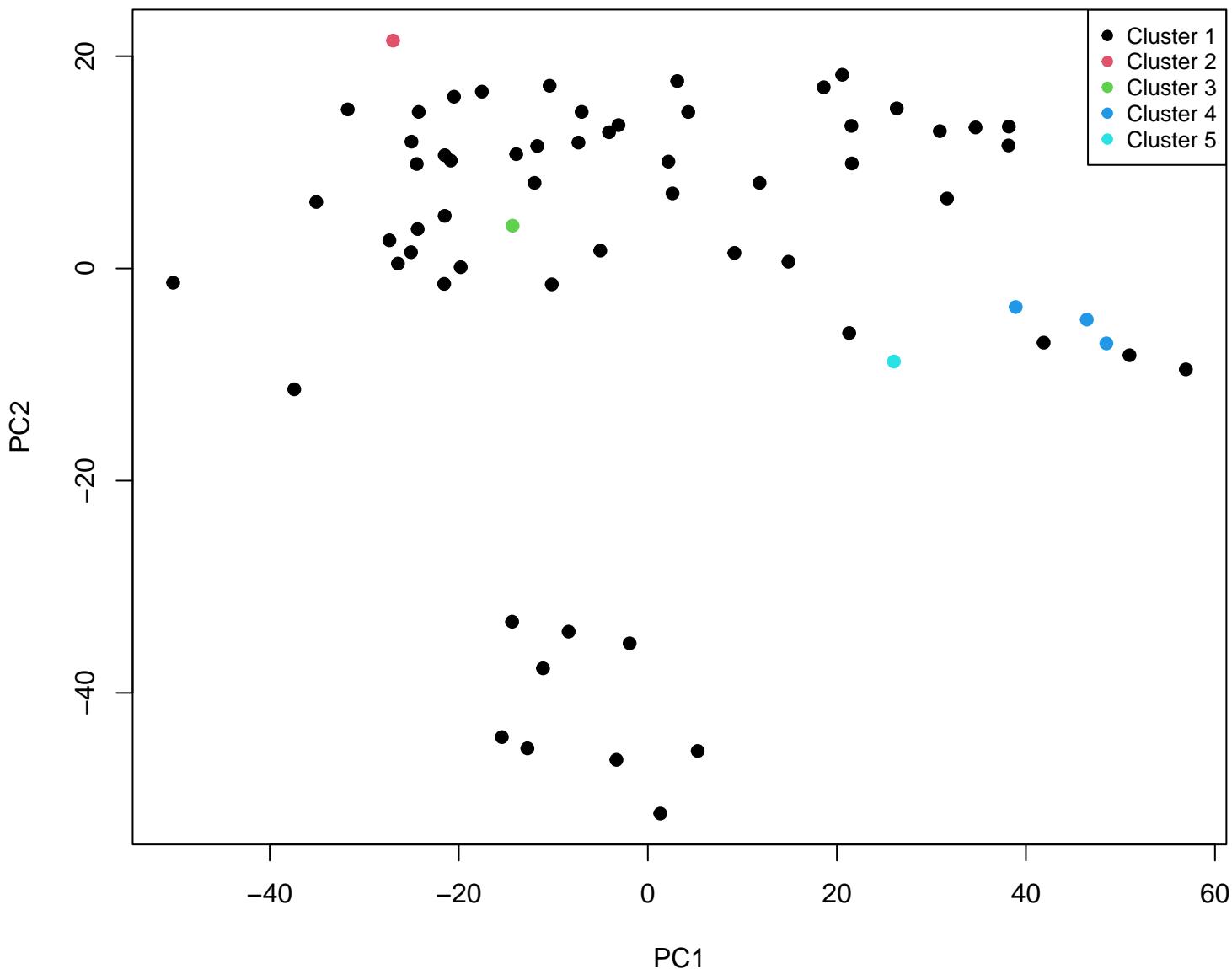
Complete linkage with 5 cluster cutoff on Z_75



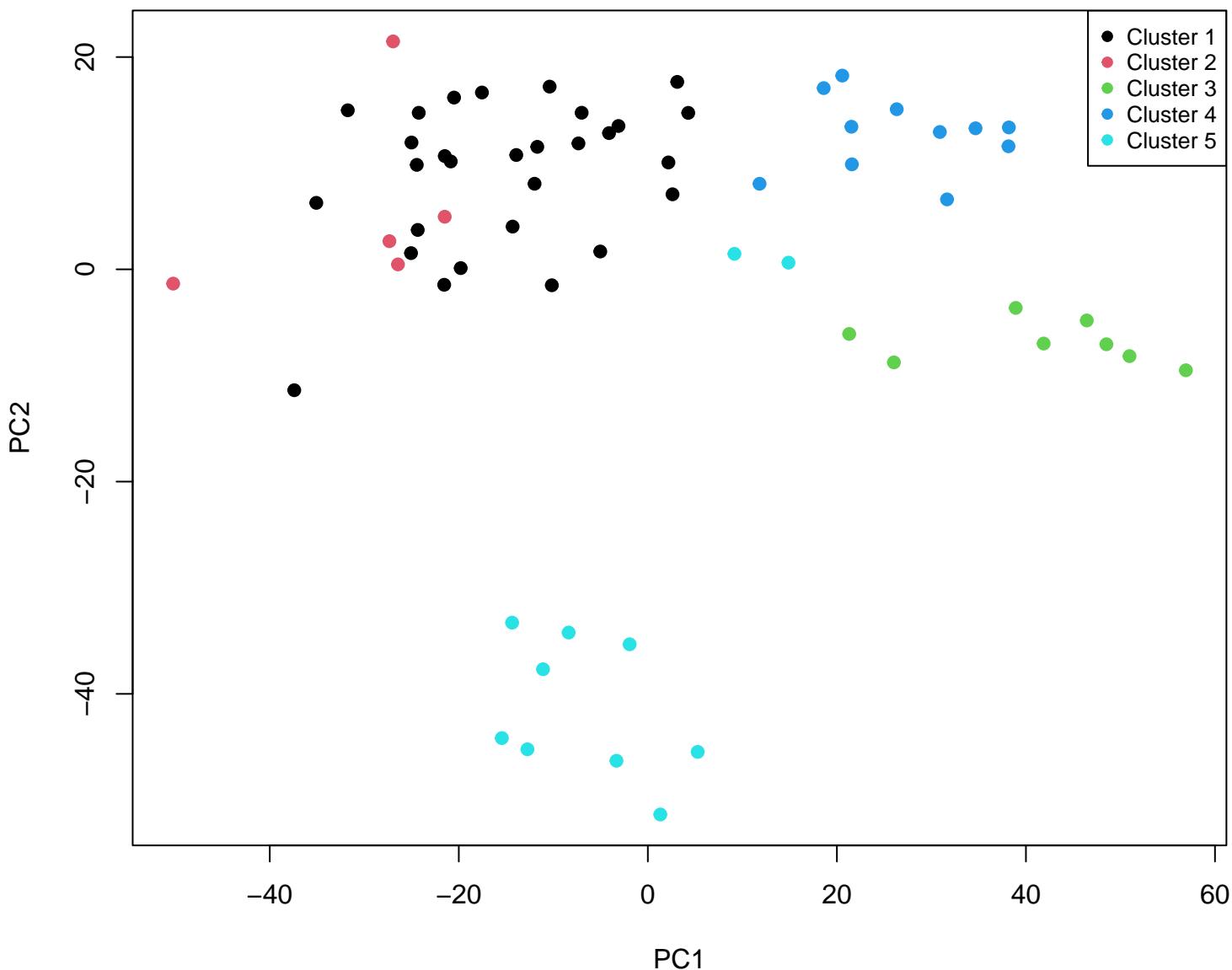
Average linkage clusters (first 2 PCs, 5 cluster cut-off) on Z_75



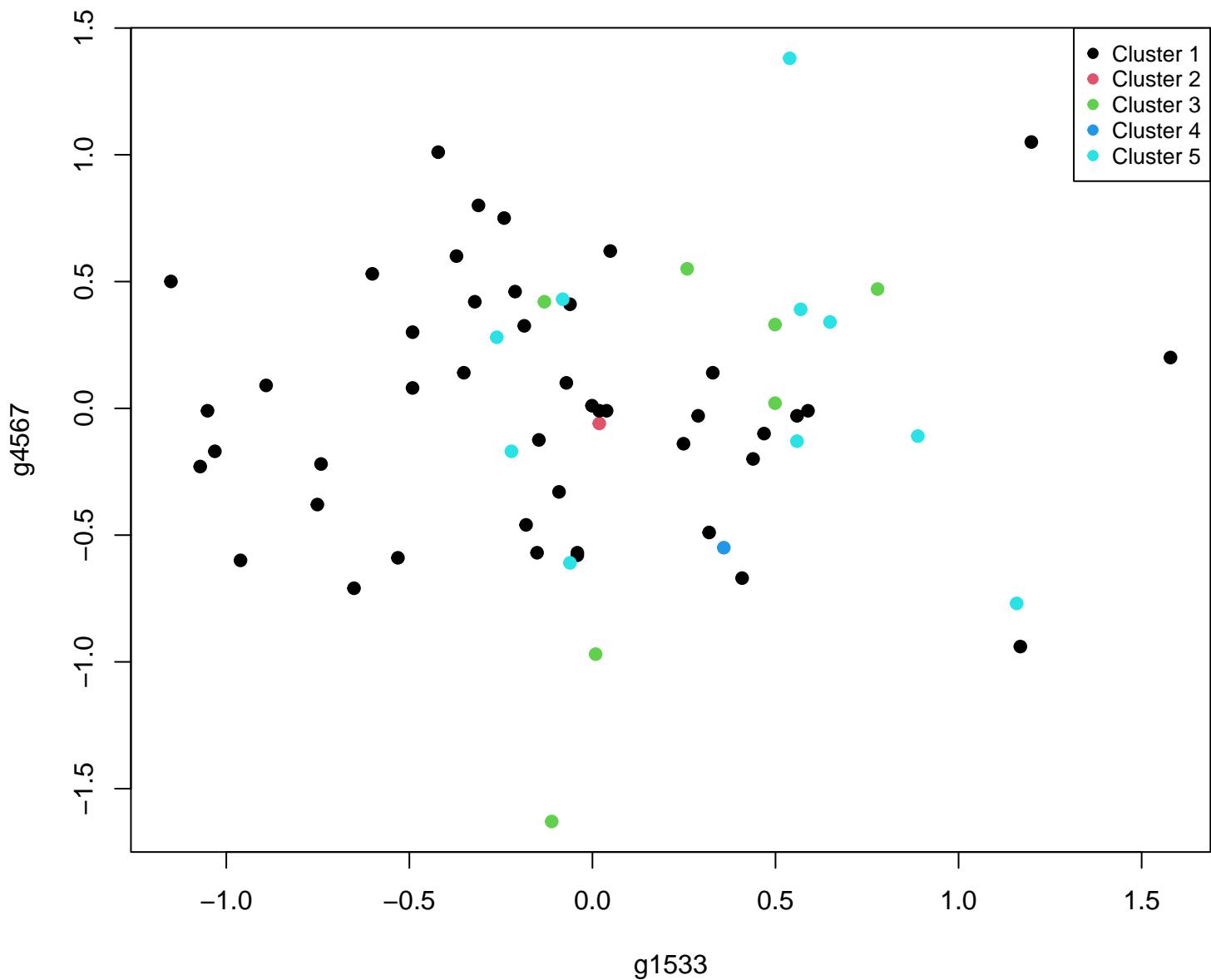
Single linkage clusters (first 2 PCs, 5 cluster cut-off) on Z_75



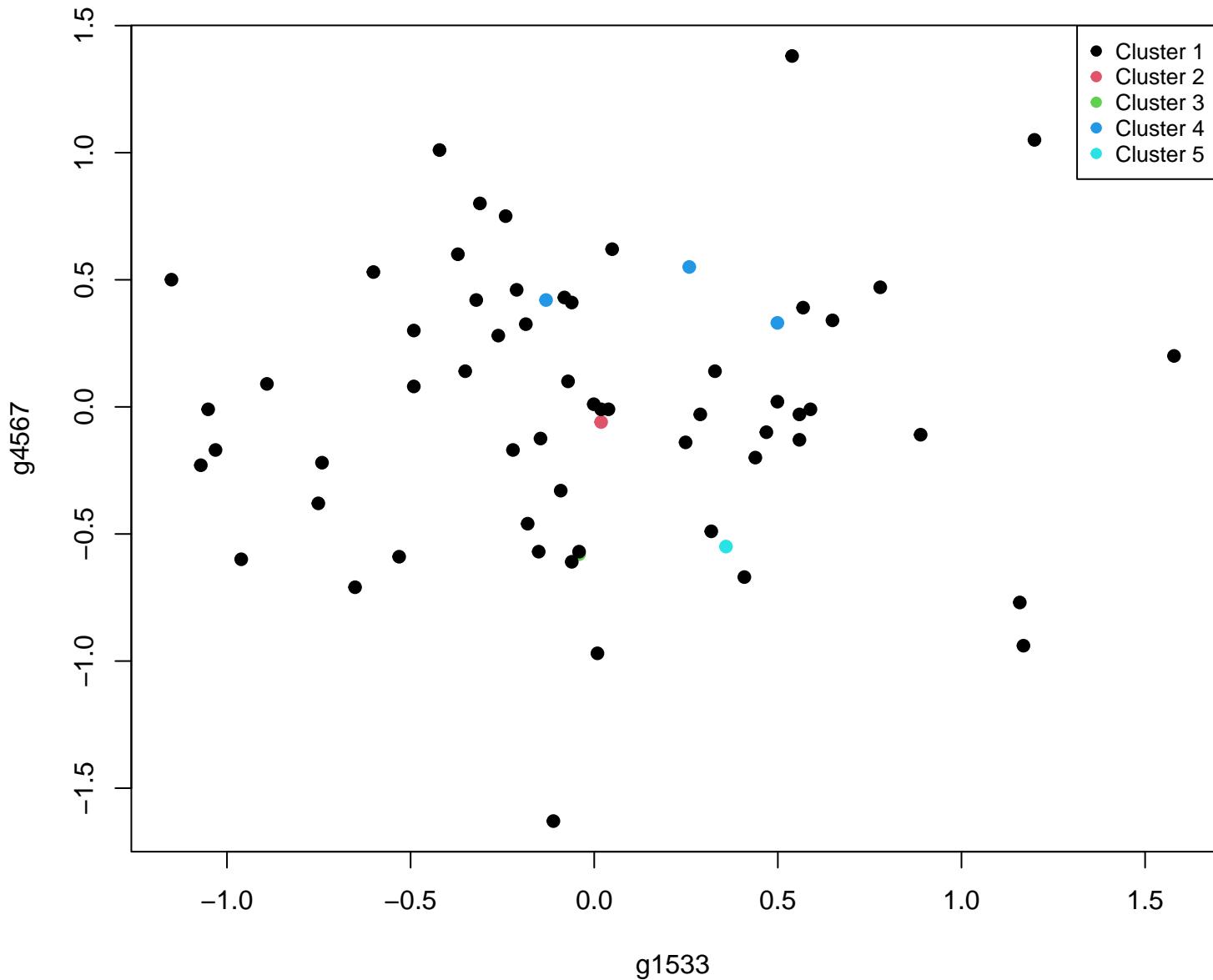
Complete linkage clusters (first 2 PCs, 5 cluster cut-off) on Z_75



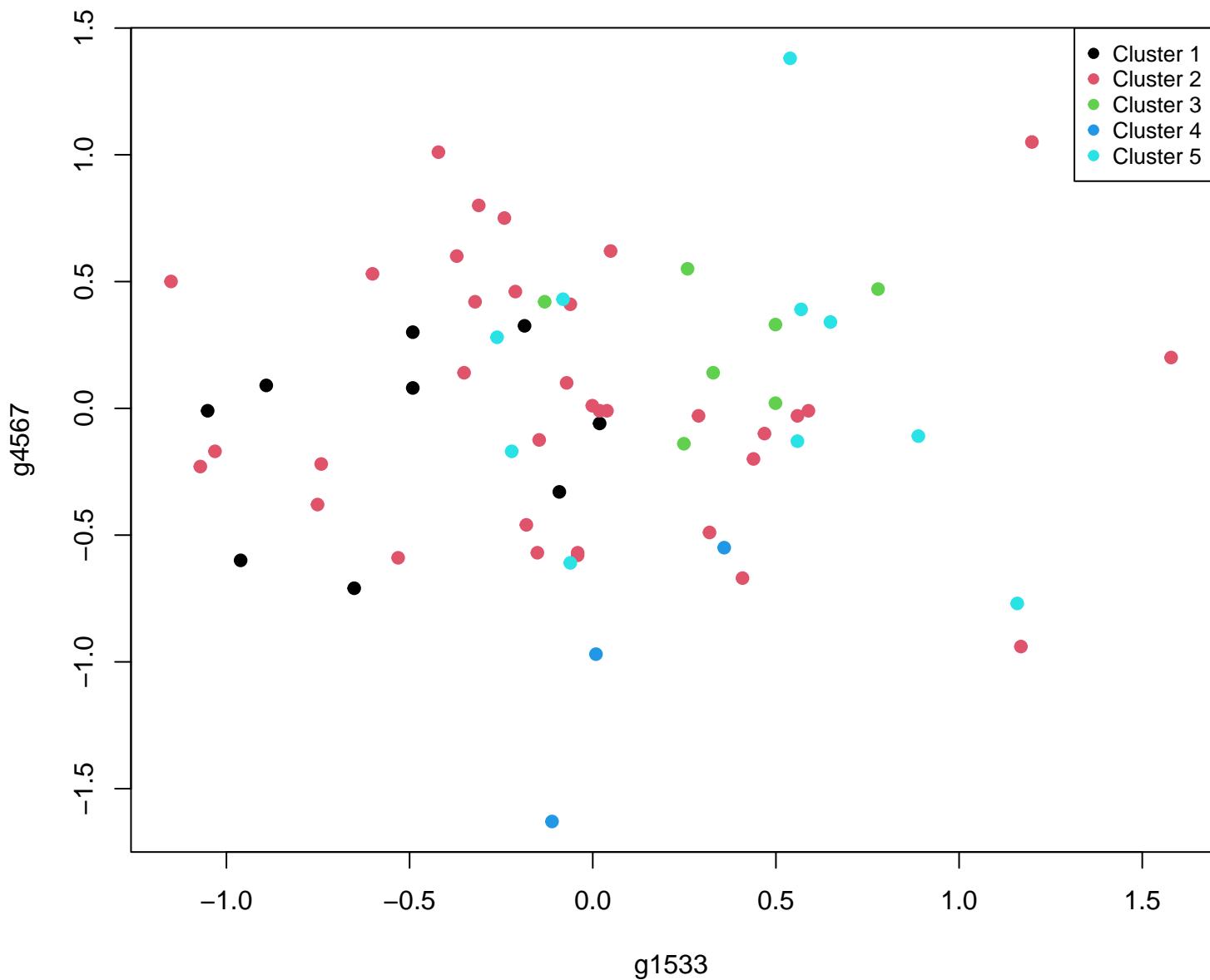
Average linkage with 5 cluster cutoff on Z_sp



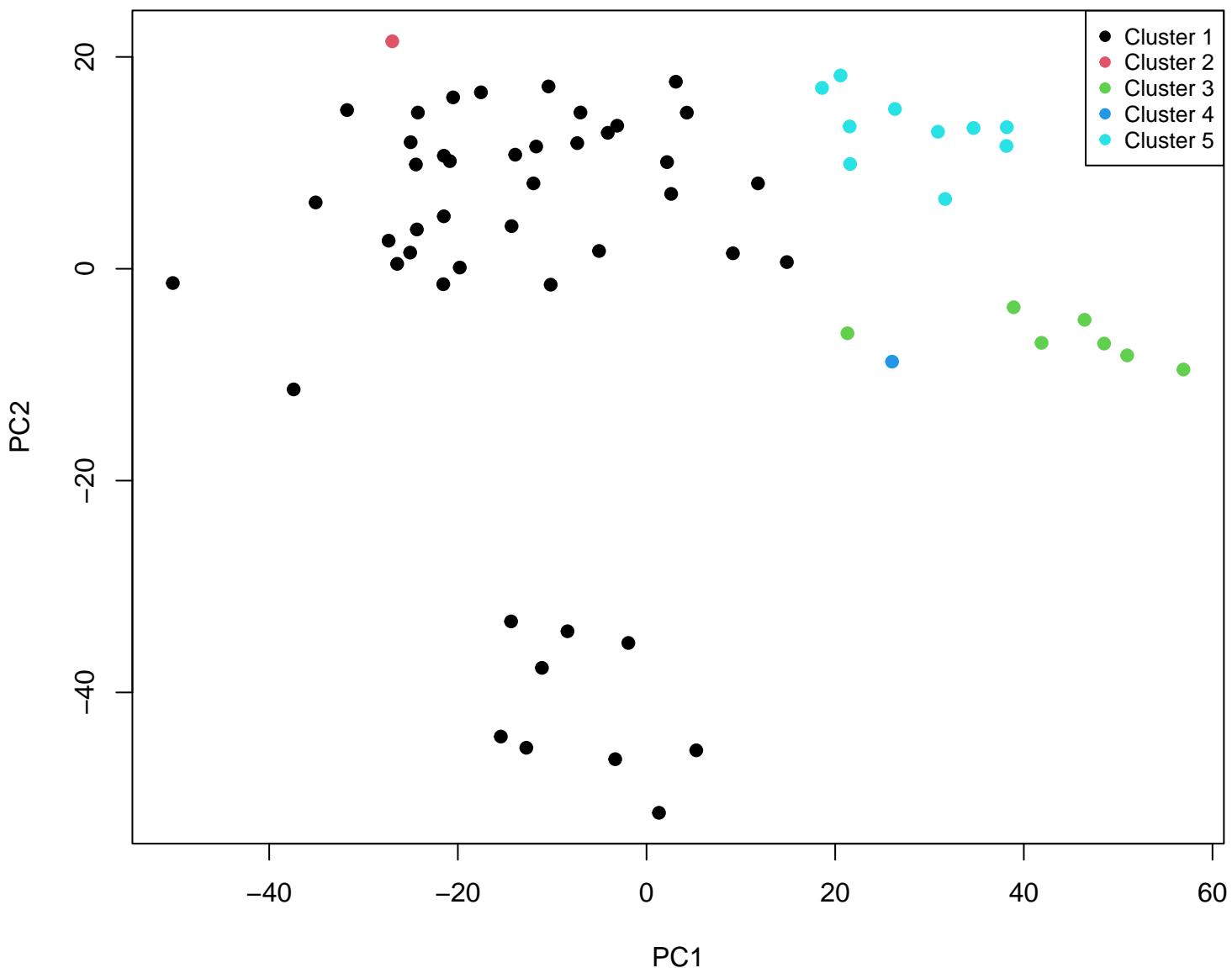
Single linkage with 5 cluster cutoff on Z_sp



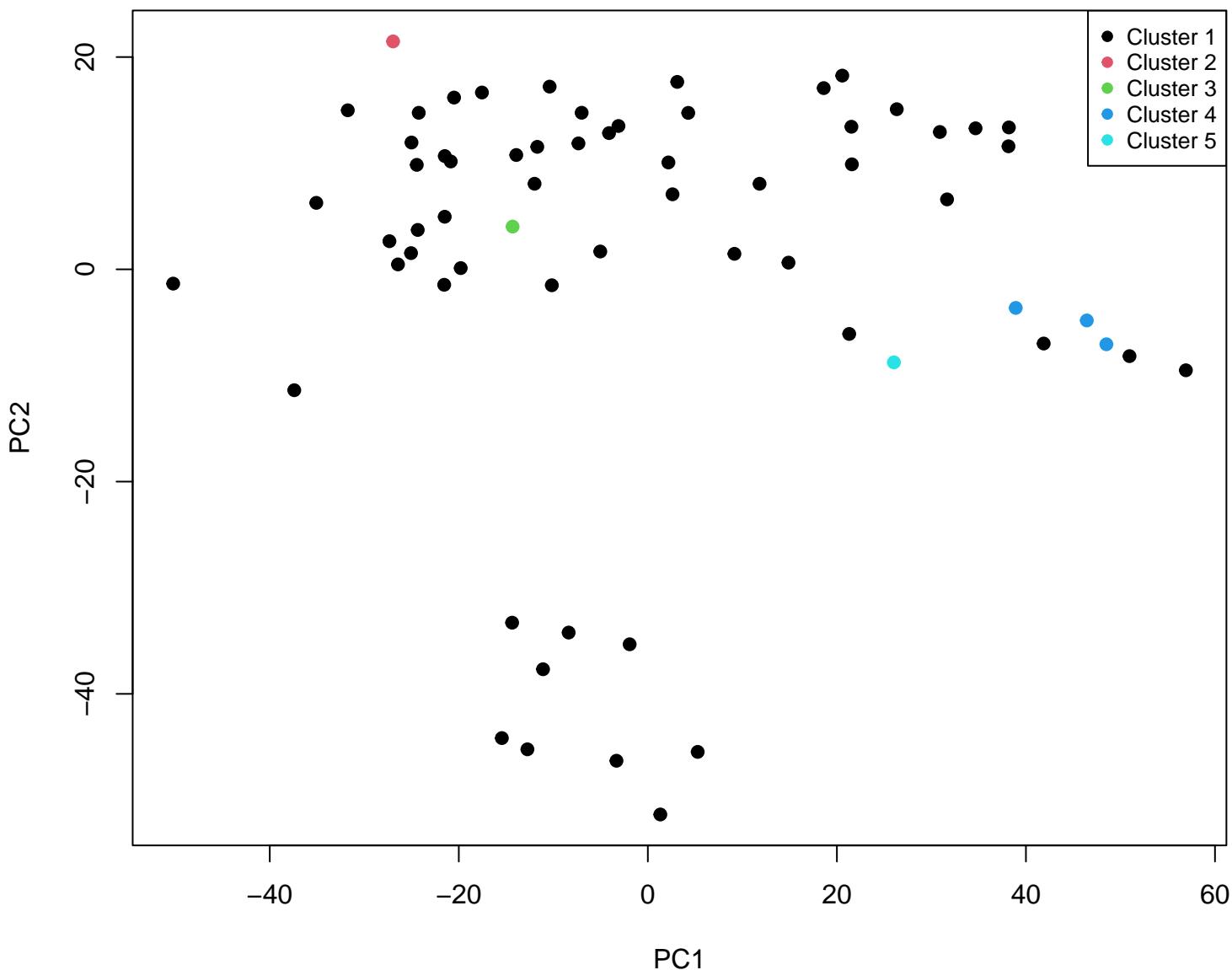
Complete linkage with 5 cluster cutoff on Z_sp



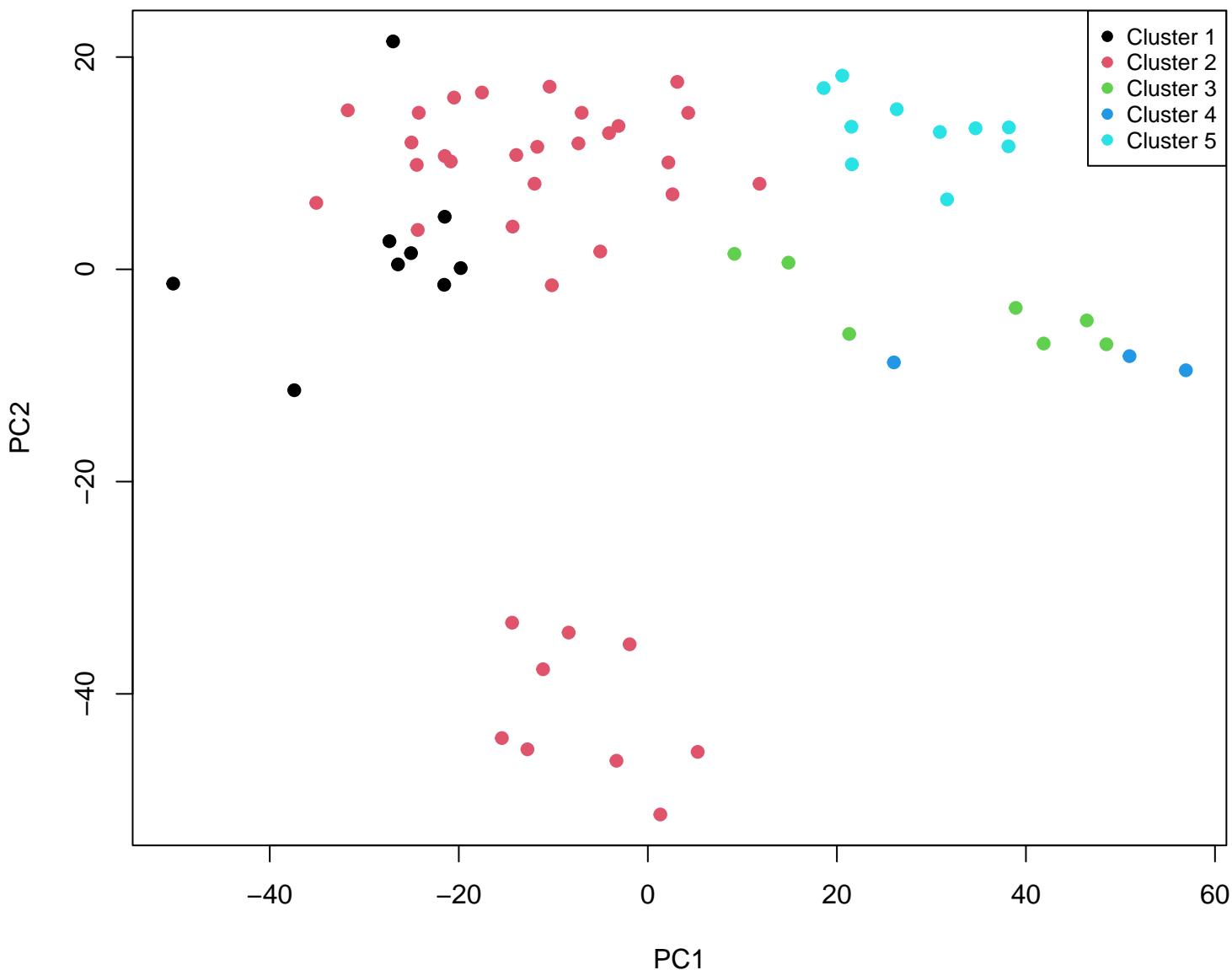
Average linkage clusters (first 2 PCs, 5 cluster cut-off) on Z_sp



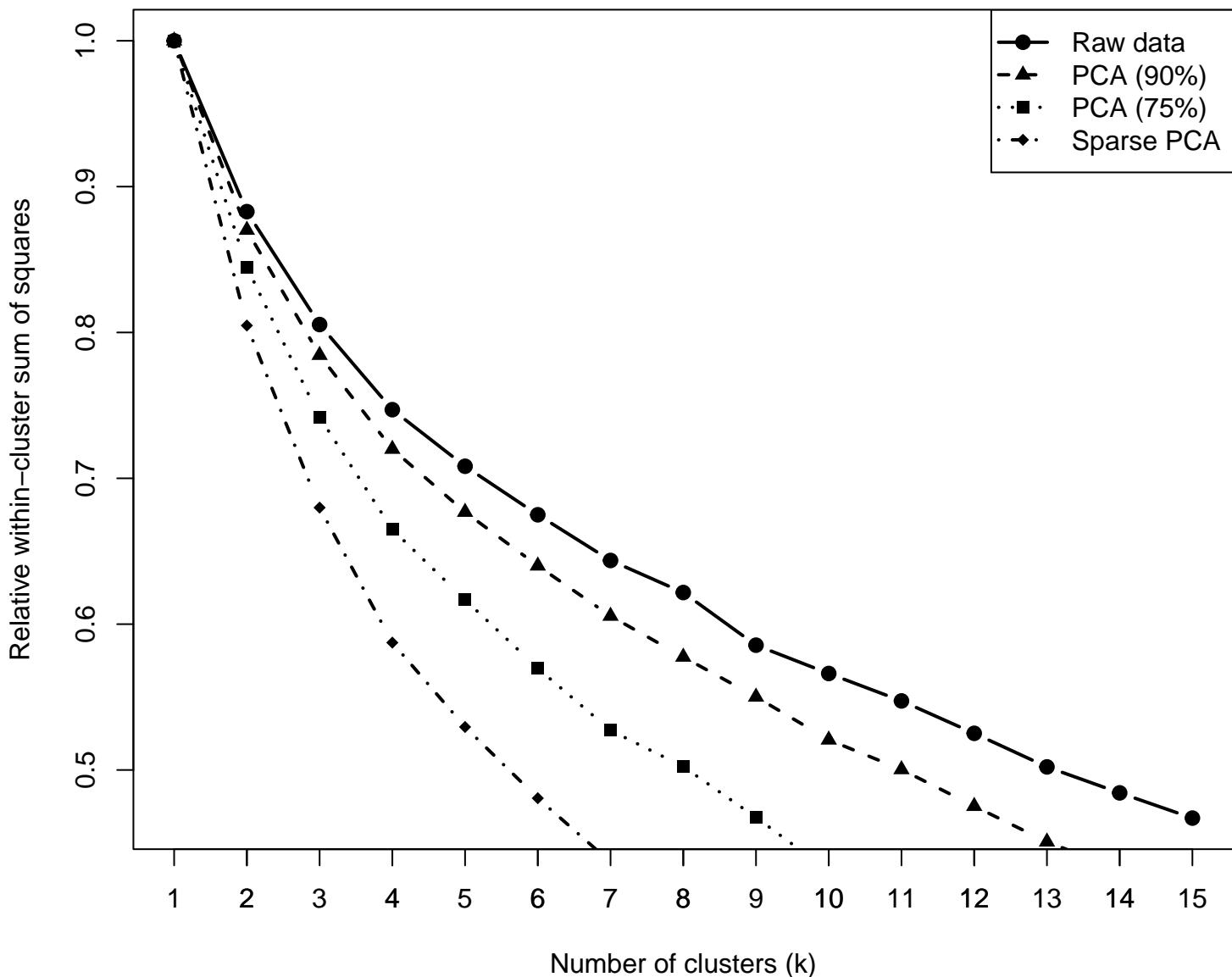
Single linkage clusters (first 2 PCs, 5 cluster cut-off) on Z_sp



Complete linkage clusters (first 2 PCs, 5 cluster cut-off) on Z_sp



Normalised elbow plots for k-means



ARI vs number of clusters for k-means

