```
> library(PLIER)
```

1 Vaccination Dataset

Load data

```
> data(bloodCellMarkersIRISDMAP)
> data(svmMarkers)
> data(canonicalPathways)
> data(vacData)
```

Construct a joint pathway matrix by merging canonicalPathways, bloodCellMarkersIRISDMAP and svmMarkers and select genes appearing in both gene expression profile and the joint pathway matrix.

```
> allPaths=combinePaths(bloodCellMarkersIRISDMAP, svmMarkers,canonicalPathways)
> cm.genes=commonRows(allPaths, vacData)
```

Normalize the data and count the number of latent variables in the data by num.pc(). The result is 24. Then set max.iter = 250, k = 24 and all other parameters to be default.

```
> vacDataN=rowNorm(vacData)
> num.pc(vacDataN[cm.genes,])
```

```
  [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1
 [38] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 [75] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[112] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[149] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[1] 24
```

```
> plierResult=PLIER(vacDataN[cm.genes,], allPaths[cm.genes,],k=24, trace=T, max.iter=150)
```

```
[1] "L2 is set to 82.0573187891353"
[1] "L1 is set to 41.0286593945677"
[1] "L3 is set to 0.0111182399988254"
```

We correlate the decomposition result with SPVs from CellCODE. We have nice one-to-one correspondence, though the "DendriticCell" signature from CellCODE is more closely related to the Type-I interferon transcriptional response so it is probably not cell-type induced variation.

```
> data(SPVs)
> plotMat(cor(t(plierResult$B), SPVs))
```

Visualize the cross-validation results

```
> plotMat(plierResult$Uauc)
```

Plot all of U and visualize the top genes

```
> plotU(plierResult,auc.cutoff = 0.5, pval.cutoff = 1)
> plotTopZ(plierResult, vacDataN, allPaths, top = 10)
```

The "$\mathrm{PID}_{A}TF2_{P}ATHWAY"looksalittletenuousandwecancheckitsstatistics.$

```
> plierResult$summary[which(plierResult$summary$`LV index`==3),]
\end{Sinput}
\begin{Soutput}
           pathway LV index       AUC   p-value
5 PID_ATF2_PATHWAY        3 0.5718064 0.0536596
\end{Soutput}
\end{Schunk}
The association with "PID_ATF2_PATHWAY" is not significant: this pathway has only 42 genes t
\begin{Schunk}
\begin{Sinput}
> plotTopZ(plierResult, vacDataN, allPaths, index=c(3), top=50)
\end{Sinput}
\end{Schunk}
\section{HCC Dataset}
Load data
\begin{Schunk}
\begin{Sinput}
> data(HCCdataTumor)
> data(canonicalPathways)
> data(chemgenPathways)
> data(oncogenicPathways)
\end{Sinput}
\end{Schunk}
Construct a joint pathway matrix by merging canonicalPathways, chemgenPathways and oncogenic
\begin{Schunk}
\begin{Sinput}
> CancerPath=combinePaths(canonicalPathways, chemgenPathways, oncogenicPathways)
> cmHCC=commonRows(HCCdataTumor, CancerPath)
>
\end{Sinput}
\end{Schunk}
Remove small pathways, not strictly necessary but saves computation time by making the pathw
\begin{Schunk}
\begin{Sinput}
> ii=which(colSums(CancerPath[cmHCC,])<20)
> HCCpath=CancerPath[, -ii]
\end{Sinput}
\end{Schunk}
Prescale the data
\begin{Schunk}
```

```
\begin{Sinput}
> HCCdataN=rowNorm(HCCdataTumor[cmHCC,])
\end{Sinput}
\end{Schunk}
Precompute Chat, which is used to define active pathways and is expensive for large pathway
\begin{Schunk}
\begin{Sinput}
> HCCchat=computeChat(CancerPath[cmHCC,])
\end{Sinput}
\end{Schunk}
Compute the number of latent variables by num.pc(HCCdataUse) and the result is 52. Then set
\begin{Schunk}
\begin{Sinput}
> plierResultHCC=PLIER(HCCdataN, CancerPath[cmHCC,], k = 52, Chat = HCCchat, trace=T)
\end{Sinput}
\begin{Soutput}
[1] "L2 is set to 132.670764279278"
[1] "L1 is set to 66.3353821396389"
[1] "L3 is set to 0.00807273868235351"
\end{Soutput}
\end{Schunk}
Plot the result with a high AUC cutoff so it is not too busy
\begin{Schunk}
\begin{Sinput}
> plotU(plierResultHCC, auc.cutoff = 0.9)
\end{Sinput}
\end{Schunk}
We found two immune components, interferon alpha and genes related to interferon gamma/CD8/T
\begin{Schunk}
\begin{Sinput}
> plotTopZ(plierResultHCC, HCCdataN, CancerPath, index=c(26, 40), top = 20)
\end{Sinput}
```