Supp 1

The aim of this document is to investigate the correlation of standard error and sample size, and to show how the presence of small sample sizes and large standard errors in biologically 'unique' tissues drives incompatibilities between the fold-size sharing heatmap and significance sharing heatmap.

```
## Warning: package 'knitr' was built under R version 3.2.5
```

Look at the ordering of Sample Size and see how it is almost identical to that of standard error, though no sample sizes differ by more than about 4 fold.

Now we look at the median posterior variances:

```
marginal.var=read.table("../../Data_vhat/withvhatmarginal.var.txt")[,-1]
```

Now let's plot effective sample size. Recall:

$$n_{jeff} = \frac{s_j^2}{\tilde{s}_j^2}$$

Let's plot again with order by original sample size:

```
original.var=as.matrix(standard.error.from.z)^2
#original.var=(standard.error.from.z/standard.error.from.z)^2
size=as.matrix(exp.sort)
post.var=as.matrix(marginal.var)*standard.error.from.z^2
njeffective=size*original.var/post.var

missing.tissues=c(7,8,19,20,24,25,31,34,37)

gtex.colors=read.table('../../Data/GTExColors.txt', sep = '\t', comment.char = '')[-missing.tissues,2]

##ask why some of the origianl variancres are smaller than posterior variances, even though calculation
###15093  44
increase=njeffective/size


par(mfrow=c(1,2))
samplesize=apply(size,2,function(x){unique(x)})
sampleorder=order(samplesize,decreasing = T)
median.nj.effective=apply(njeffective,2,median)
median.nj.increase=apply(increase,2,median)

par(mar=c(5.1,8,4.1,0.1))
barplot(median.nj.effective[sampleorder],cex.names=0.4,las=2,col=as.character(gtex.colors[sampleorder])
title("MedianNj_effective",cex.main=0.8)
par(mar=c(5.1,2,4.1,6))
barplot(median.nj.increase[sampleorder],cex.names=0.4,las=2,col=as.character(gtex.colors[sampleorder]),
title("Median(Nj_effective/Nj_original)",cex.main=0.8)
```
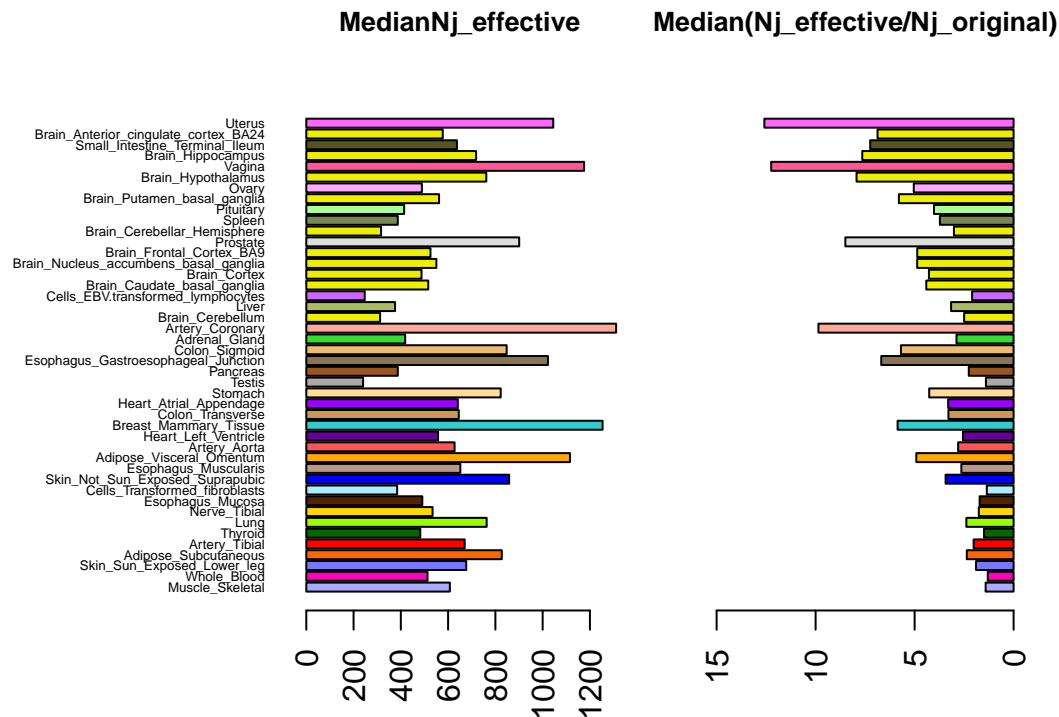
MedianNj_effective   **Median(Nj_effective/Nj_original)**

```r
par(mfrow=c(1,2))
samplesize=apply(size,2,function(x){unique(x)})
sampleorder=order(samplesize,decreasing = T)
median.nj.effective=apply(njeffective,2,median)
median.nj.increase=apply(increase,2,median)


par(mar=c(5.1,8,1.1,0.1))
barplot(samplesize[sampleorder],cex.names=0.4,las=2,col=as.character(gtex.colors[sampleorder]),horiz = T
title("Sample Size",cex.main=0.8)

par(mar=c(5.1,2,1.1,6))
barplot(median.nj.effective[sampleorder],cex.names=0.4,las=2,col=as.character(gtex.colors[sampleorder])

title("Effective Sample Size",cex.main=0.8)
```

Sample Size   **Effective Sample Size**

Uterus
Brain_Anterior_cingulate_cortex_BA24
Small_Intestine_Terminal_Ileum
Brain_Hippocampus
Vagina
Brain_Hypothalamus
Ovary
Brain_Putamen_basal_ganglia
Pituitary
Spleen
Brain_Cerebellar_Hemisphere
Prostate
Brain_Frontal_Cortex_BA9
Brain_Nucleus_accumbens_basal_ganglia
Brain_Cortex
Brain_Caudate_basal_ganglia
Cells_EBV.transformed_lymphocytes
Liver
Brain_Cerebellum
Artery_Coronary
Adrenal_Gland
Colon_Sigmoid
Esophagus_Gastroesophageal_Junction
Pancreas
Testis
Stomach
Heart_Atrial_Appendage
Colon_Transverse
Breast_Mammary_Tissue
Heart_Left_Ventricle
Artery_Aorta
Adipose_Visceral_Omentum
Esophagus_Muscularis
Skin_Not_Sun_Exposed_Suprapubic
Cells_Transformed_fibroblasts
Esophagus_Mucosa
Nerve_Tibial
Lung
Thyroid
Artery_Tibial
Adipose_Subcutaneous
Skin_Sun_Exposed_Lower_leg
Whole_Blood
Muscle_Skeletal

0   500   1000   1500   2000   0   500   1000   1500   2000