

02_Compositional_Analysis

January 24, 2023

```
[1]: suppressPackageStartupMessages({  
      library(Seurat, quietly = T)  
  
      library(dplyr, quietly = T)  
      library(reshape2, quietly = T)  
  
      library(ggplot2, quietly = T)  
    })
```

```
data_path = '/data3/hratch/norcross_abc/'
```

```
[2]: abc.integrated<-readRDS(paste0(data_path, 'processed/abc_annotated.RDS'))  
     abc.tcells<-readRDS(paste0(data_path, 'processed/abc_tcells.RDS'))
```

Get the cell type frequencies:

Level 1:

```
[3]: md<-abc.integrated@meta.data
```

```
[27]: context.counts<-dplyr::count(md, orig.ident, Cell.Type.Level1)  
      context.counts<-reshape2::dcast(context.counts, orig.ident ~ Cell.Type.Level1,   
      ↪ value.var="n")  
      rownames(context.counts)<-context.counts$orig.ident  
      context.counts<-context.counts[names(context.counts) != 'orig.ident']  
      context.counts[is.na(context.counts)]<-0  
      context.counts
```

		Pre-B cells	Naive B cells	T cells	Natural killer cells	Conventional De
		<int>	<dbl>	<int>	<int>	<int>
A data.frame: 5 × 13	UNTR	907	0	1853	20	15
	ABC	13	0	2156	62	264
	DT_Veh	57	27	788	23	1651
	DT_ABC	873	41	3740	97	2091
	aCD4_ABC	2008	1	5677	132	654

```
[28]: context.freq<-prop.table(data.matrix(context.counts),2)  
      context.freq
```

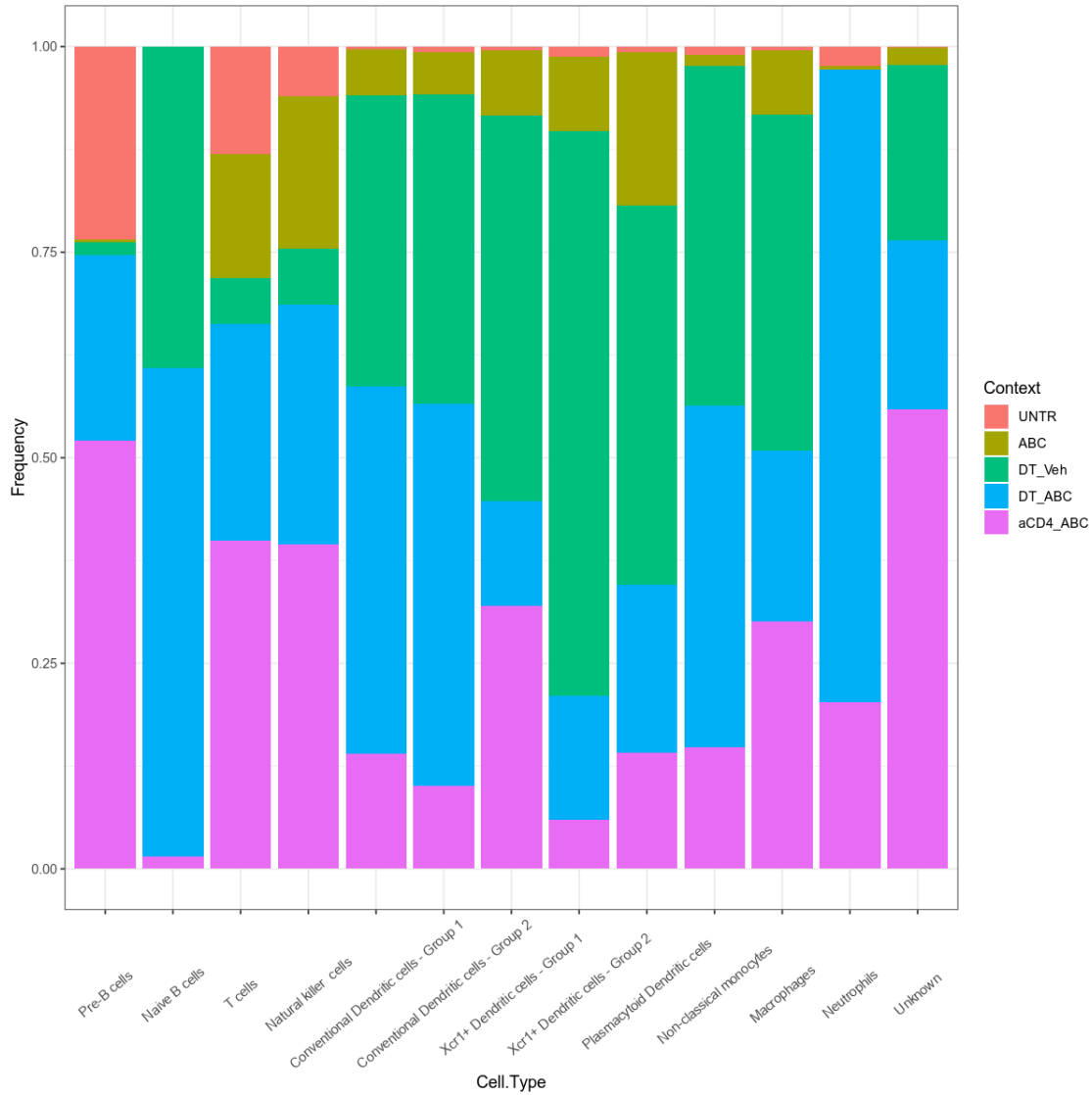
		Pre-B cells	Naive B cells	T cells	Natural killer cells	Conve
A matrix: 5 × 13 of type dbl	UNTR	0.235095905	0.000000000	0.1303644	0.05988024	0.003
	ABC	0.003369622	0.000000000	0.1516814	0.18562874	0.056
	DT_Veh	0.014774495	0.39130435	0.0554383	0.06886228	0.353
	DT_ABC	0.226283048	0.59420290	0.2631209	0.29041916	0.447
	aCD4_ABC	0.520476931	0.01449275	0.3993950	0.39520958	0.139

```
[42]: viz.df<-melt(context.freq)
names(viz.df)<-c('Context', 'Cell.Type', 'Frequency')

h_ = 10
w_ = 10
options(repr.plot.height=h_, repr.plot.width=w_)

g1<-ggplot(viz.df, aes(fill=Context, y=Frequency, x=Cell.Type)) +
  geom_bar(position="stack", stat="identity") + theme_bw() +
  theme(axis.text.x = element_text(angle = 40, vjust = 0.5, hjust=0.5))
g1

for (ext in c('.svg', '.png', '.pdf')){ggsave(paste0(data_path, 'figures/',
                                                    ↵
                                                    'celltype_level1_composition', ext), g1,
                                                    height = h_, width = w_)}
```



Level 2: only within total T cells

```
[43]: md<-abc.tcells@meta.data
```

```
[45]: context.counts<-dplyr::count(md, orig.ident, Cell.Type.Level2)
context.counts<-reshape2::dcast(context.counts, orig.ident ~ Cell.Type.Level2,
  value.var="n")
rownames(context.counts)<-context.counts$orig.ident
context.counts<-context.counts[names(context.counts) != 'orig.ident']
context.counts[is.na(context.counts)]<-0
context.counts
```

		Naive CD8+ T cells <int>	Naive/Early Activted ISG CD8+ T cells <int>	Early Act <int>
A data.frame: 5 × 10	UNTR	726	54	157
	ABC	746	139	205
	DT_Veh	217	10	57
	DT_ABC	1102	141	207
	aCD4_ABC	2177	164	612

```
[46]: context.freq<-prop.table(data.matrix(context.counts),2)
context.freq
```

		Naive CD8+ T cells	Naive/Early Activted ISG CD8+ T cells	Ear
A matrix: 5 × 10 of type dbl	UNTR	0.14613527	0.10629921	0.1
	ABC	0.15016103	0.27362205	0.1
	DT_Veh	0.04367955	0.01968504	0.0
	DT_ABC	0.22181965	0.27755906	0.1
	aCD4_ABC	0.43820451	0.32283465	0.4

```
[49]: viz.df<-melt(context.freq)
names(viz.df)<-c('Context', 'Cell.Type', 'Frequency')

h_ = 10
w_ = 10
options(repr.plot.height=h_, repr.plot.width=w_)

g2<-ggplot(viz.df, aes(fill=Context, y=Frequency, x=Cell.Type)) +
  geom_bar(position="stack", stat="identity") + theme_bw() +
  theme(axis.text.x = element_text(angle = 40, vjust = 0.5, hjust=0.5))
g2

suppressWarnings({
  for (ext in c('.svg', '.png', '.pdf')){ggsave(paste0(data_path, 'figures/',
  ↵
    'celltype_level2_composition', ext), g2,
                                height = h_, width = w_)}
})
```

