

# Nat Imm成年人数据分析(不建议使用)

## 0 分析之前

Nat Imm中大部分人都是儿童患者/对照，而我们知道儿童的免疫系统本来就和成年人有较大区别。为了控制变量，先只分析了文章中的成年人队列，一共14个人，6位健康对照，8位SLE患者，均为女性，年龄分布在24-63岁，文章中提供的meta如下

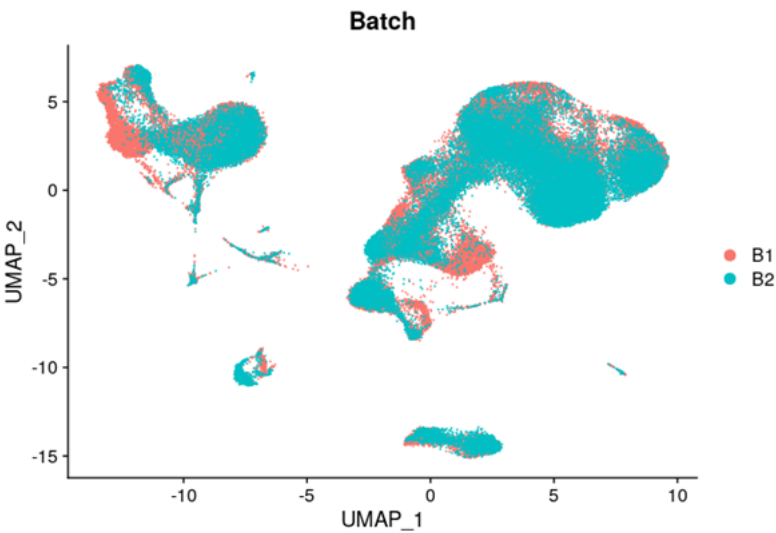
Names	Groups	Batch	Age	Gender	Race	Ethnicity	SLEDAI	Physician Notes	Collection_year	Symptoms	MDG	Arthritis	Rash	Proteinuria	Hematuria	Urinary Casts	Pyuria	dsDNA	low_com	fever	pleurisy	vasculitis	Alopecia	Leukopenia	MMF	OS	MTX	Plaque/nil	Neph_all	Proliferative_Neph	Membr_Neph
aHD1	aHD	B1	39	F	AA	NH	ND	NA	2018	NA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
aHD2	aHD	B1	50	F	C	NH	ND	NA	2018	NA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
aHD3	aHD	B1	43	F	H	H	ND	NA	2018	NA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
aHD4	aHD	B2	36	F	H	H	ND	NA	2018	NA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
aHD5	aHD	B2	46	F	C	NH	ND	NA	2018	NA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
aHD6	aHD	B2	47	F	AA	NH	ND	NA	2018	NA	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
aSLE1	aSLE	B1	37	F	AA	NH	15	NA	2018	NA	ND	0	0	4	0	0	0	2	0	0	0	8	0	1	0	0	1	1	1	ND	ND
aSLE2	aSLE	B1	63	F	not reported	H	14	NA	2018	NA	ND	0	0	4	4	0	4	2	0	0	0	0	0	0	1	0	1	0	ND	ND	ND
aSLE3	aSLE	B1	36	F	C	not reported	0	NA	2018	NA	ND	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	ND	ND	ND
aSLE4	aSLE	B1	58	F	not reported	H	2	NA	2018	NA	ND	0	0	0	0	0	0	2	0	0	0	0	0	1	1	0	1	0	ND	ND	ND
aSLE5	aSLE	B2	24	F	H	H	2	NA	2018	NA	ND	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	ND	ND	ND
aSLE6	aSLE	B2	27	F	H	H	4	NA	2018	NA	ND	0	2	0	0	0	0	0	0	0	0	2	0	0	1	0	0	0	ND	ND	ND
aSLE7	aSLE	B2	47	F	C	NH	6	NA	2018	NA	ND	4	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	ND	ND	ND
aSLE8	aSLE	B2	62	F	AA	NH	2	NA	2018	NA	ND	0	0	0	0	0	0	2	0	0	0	0	0	1	0	1	0	ND	ND	ND	ND

实际在作者公布的数据中只有7个SLE和5个对照（未提供样本的在表中标红），每个样本的细胞数量如下，SLE的患者的细胞

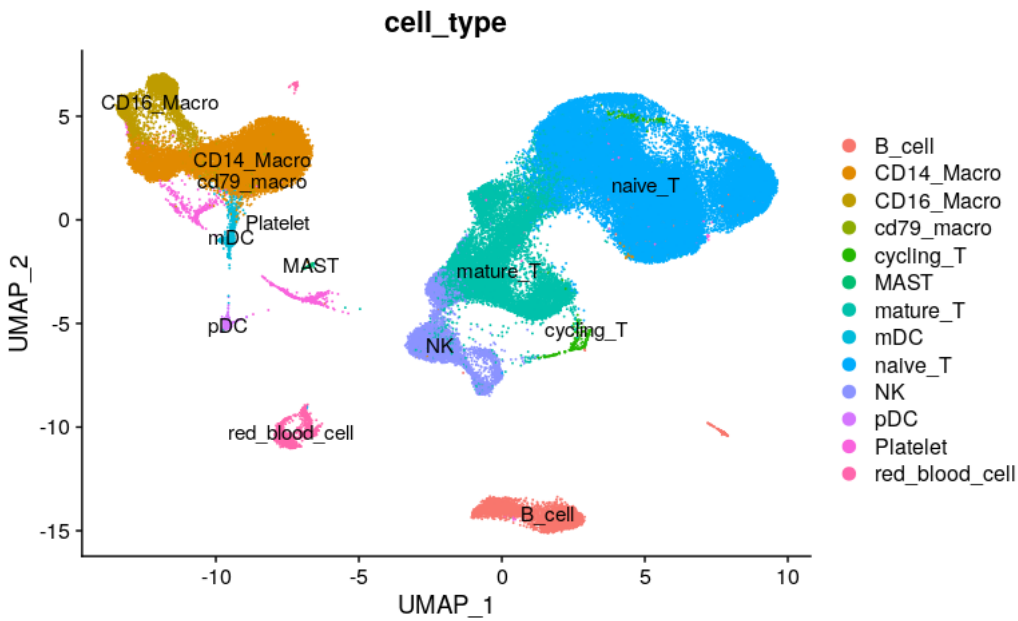
	aHD1	aHD3	aHD4	aHD5	aHD6	aSLE1	aSLE2	aSLE3	aSLE4	aSLE5	aSLE6	aSLE7
HD	7179	7174	6606	8116	12249	0	0	0	0	0	0	0
SLE	0	0	0	0	0	4298	4505	7442	6193	4578	3725	4543

## 1 全PBMC分析

这批数据对批次效应控制较好，在作者声明的两个Batch中，未出现明显的batch effect效应。

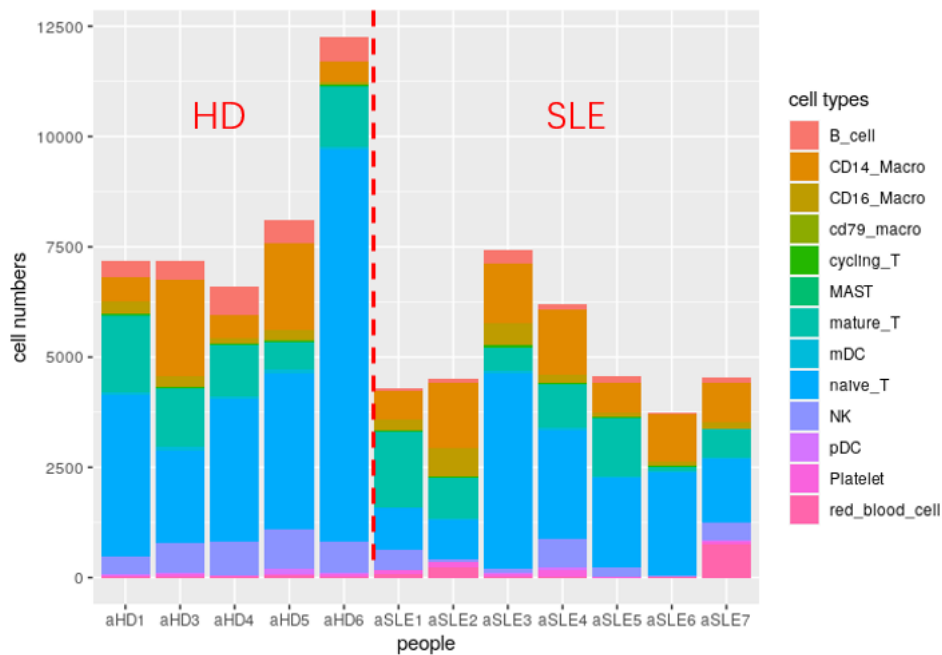


对于主要细胞类型的注释如下：

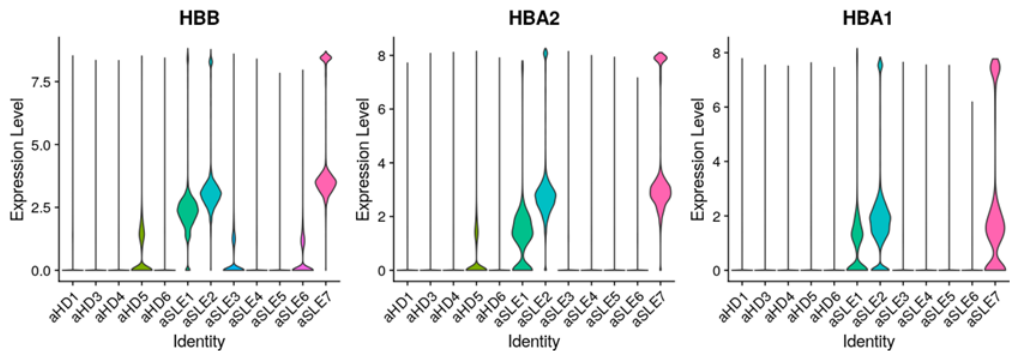


同时发现了他们数据中的一些质量问题：

1. 没有免疫球蛋白相关基因，按理来说10x cellranger的结果肯定是有相关基因的，最可能的解释是作者有意删除了，拿到原始数据之后重新运行上游步骤就清楚了
2. 不同患者测序数目不平衡，可以明显发现，大部分SLE病人所测序的细胞数量要低于健康对照

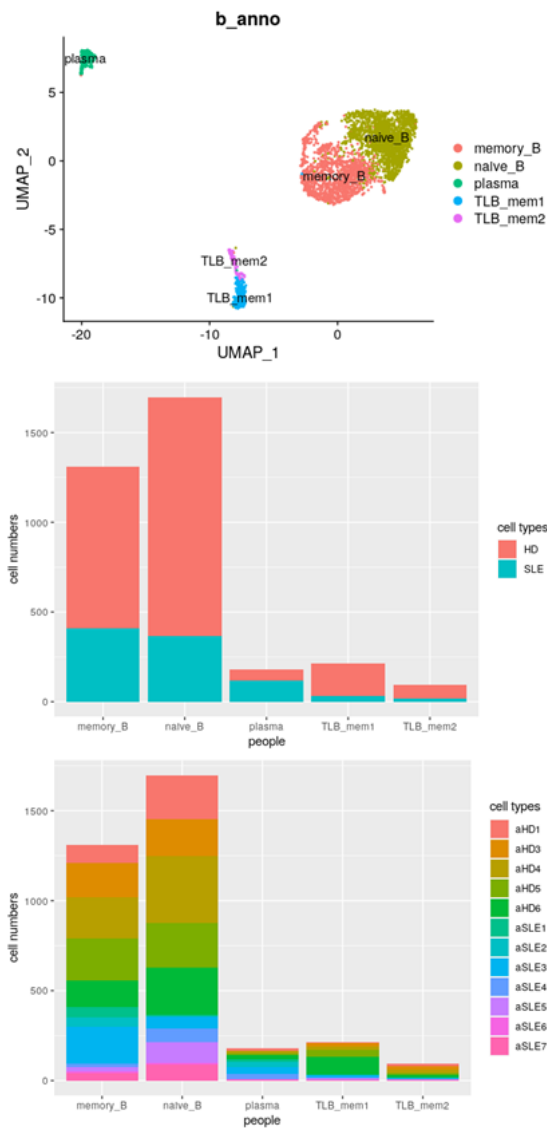


3. 数据质量不高，有一定比例的红细胞，这说明PBMC解离的步骤不标准，并且在一些样本（如下图）的各个细胞类型中都检测到了红细胞marker gene的表达，可能是mRNA污染。在1，2，7这3个SLE患者样本中这个现象尤其严重

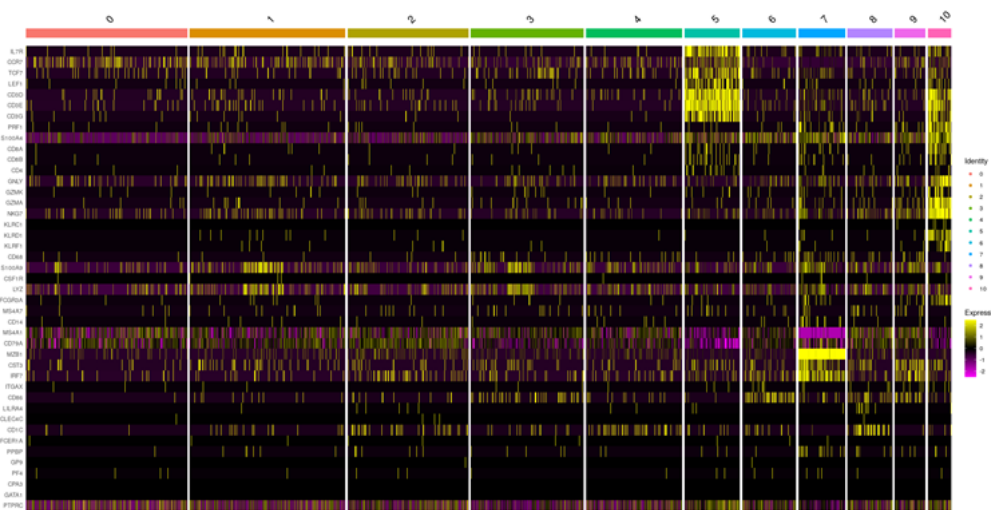


## 2 B cell

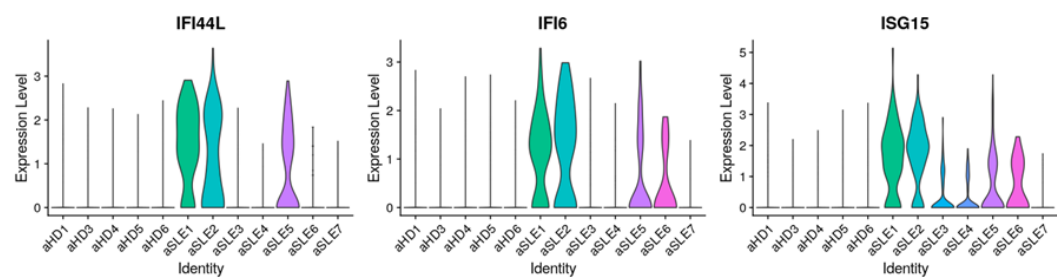
B cell在SLE和患者中分布差异不大，并且没有发现和我们自己数据中表达多种细胞marker的细胞，也没有发现TLB\_plasma (abnormal)，但是发现了TLB\_mem1和TLB\_mem2



但是可能是这个数据集质量的问题，背景有点高，体现在B cell发现其他细胞marker也有星星点点的表达，



干扰素相关基因在部分患者中有高表达,提示不能忽略了患者之间的异质性

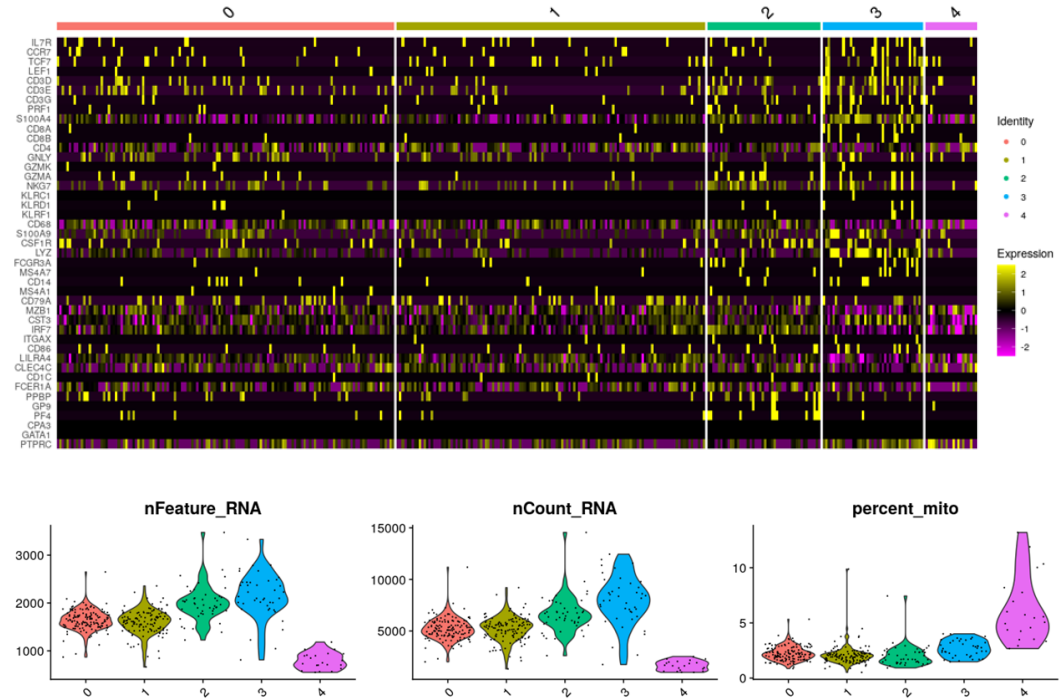


此外，在这一组成年人的B cell中没有发现CD11C+ B cell，但是文章在child SLE队列中报道了，在我们的数据中也发现了。CD11C+ B cell被认为可能是在生发中心之外成熟，因而可能与自身免疫有较强的关系。

### 3 pDC

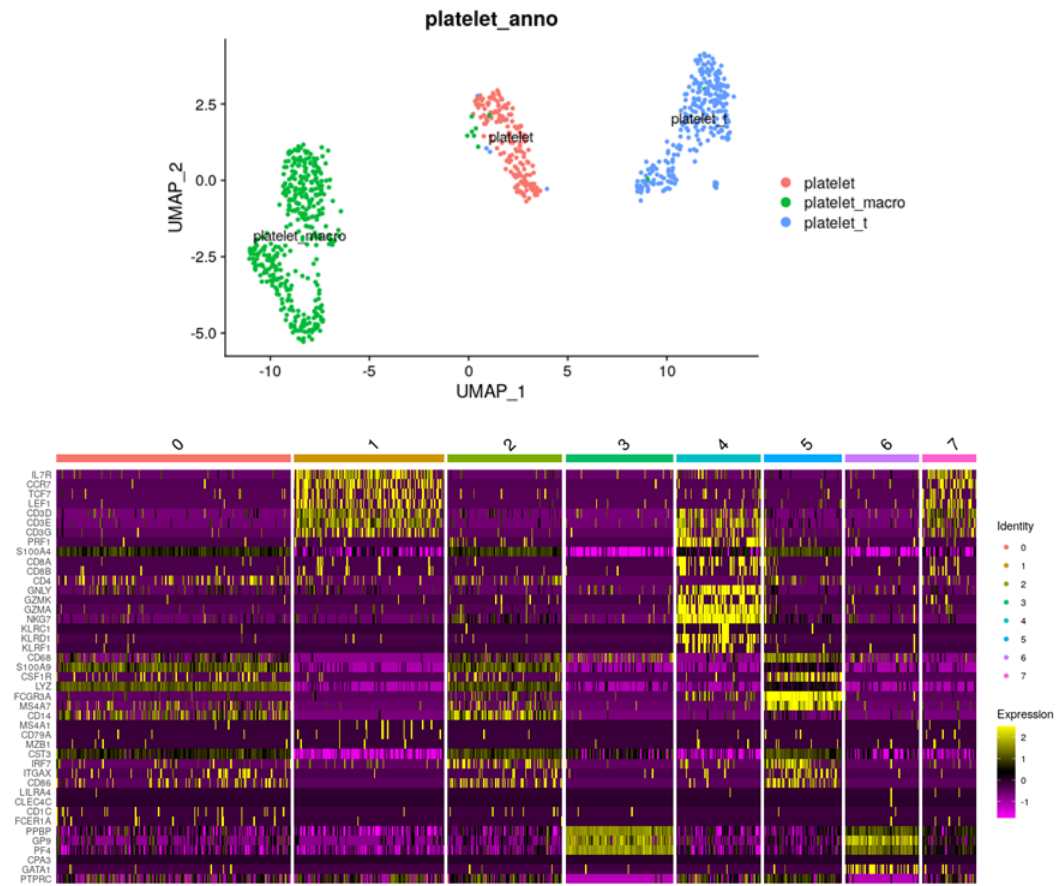
pDC中出现了更加明显“混合marker”表达的细胞，尤其是cluster3，但是pDC的总数很小，8个病人只有372个细胞，可能随机因素比较大一点。

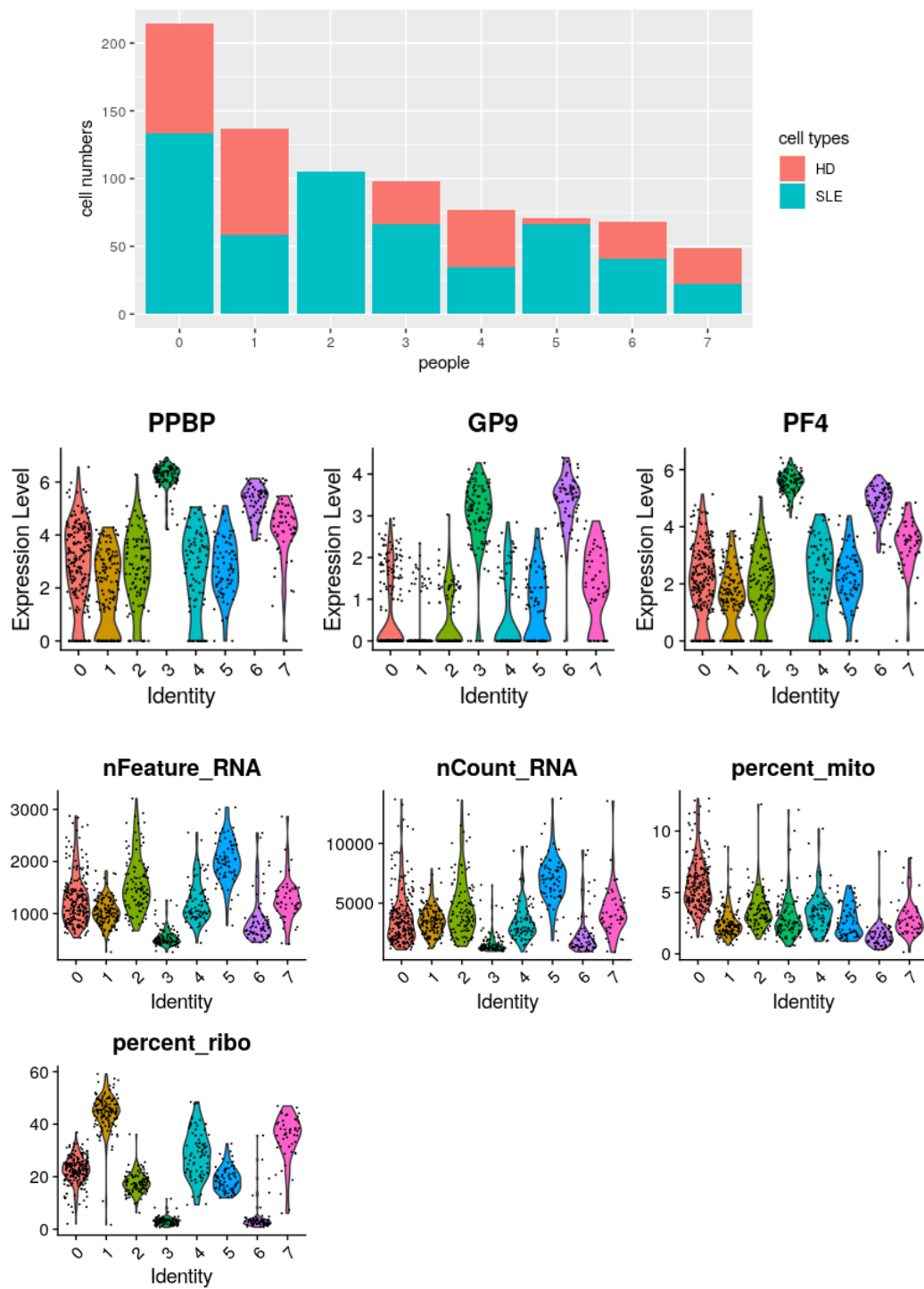
需要注意的是，这里“混合marker”表达的细胞的UMI和Gene的数量并没有显著小于正常细胞，反而比正常的pDC要更高，暂时无法解释



### 4 platelet

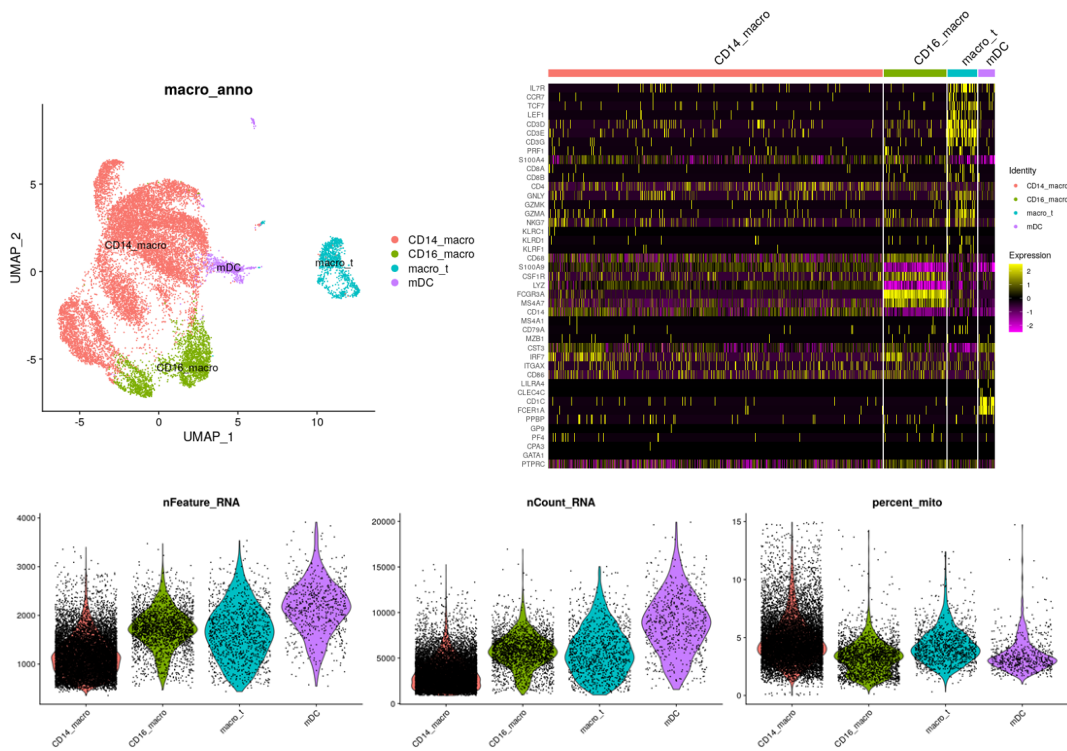
有趣的是在platelet中观察到了T cell marker（cluster 1, 4, 7）以及macrophage marker（0, 2, 5）同时表达的情况，当然也有正常的platelet（cluster3, 6）





## 5 Macro and cDC

发现混杂表达T cell 和macrophage marker的细胞，但是这类细胞（标记为macro\_t）并没有UMI和Gene的显著下降，而是在正常的区间范围

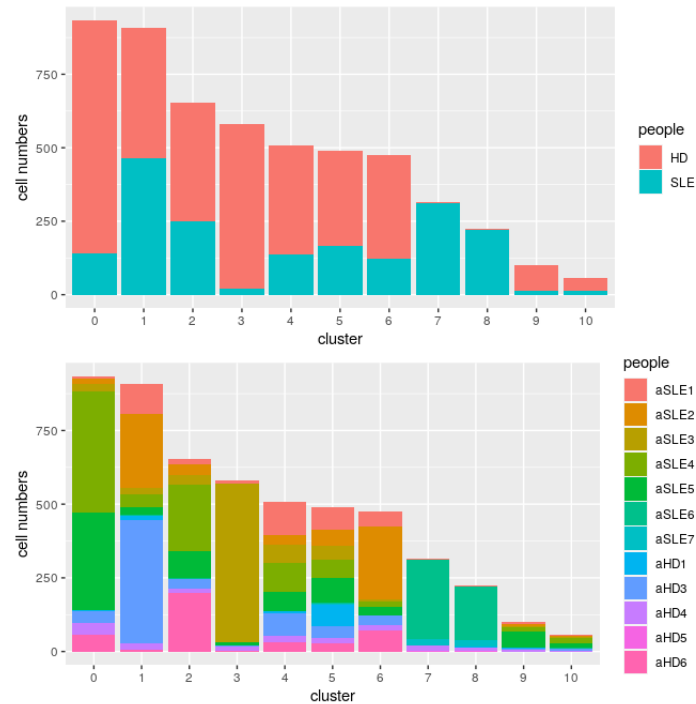


## 6 T cell

被红细胞的基因污染，在想解决的办法

## 7 NK

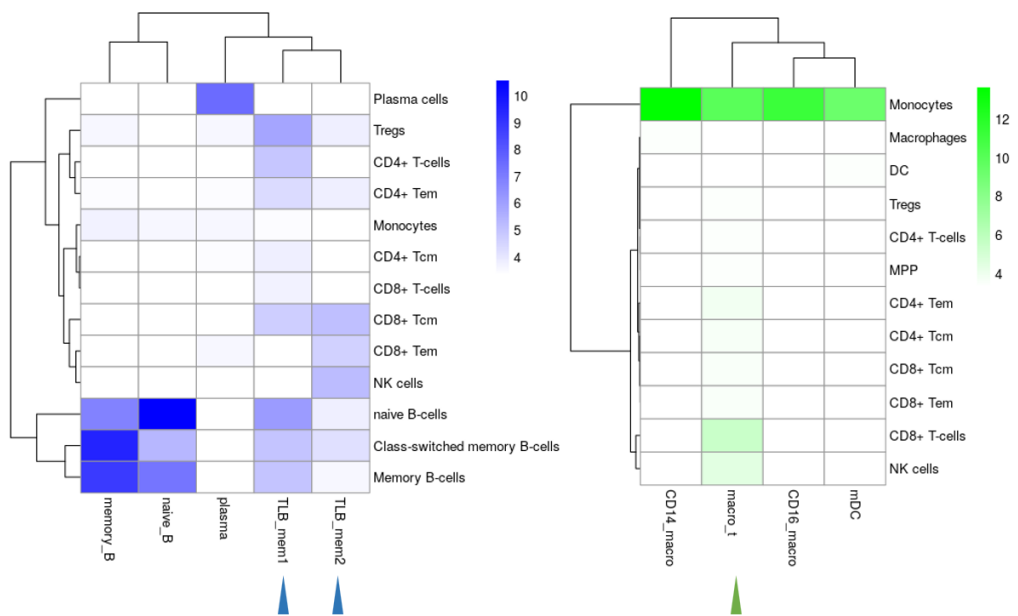
NK细胞按照疾病状态和患者来源如下，发现个体异质性比较大，



## 8 其他发现

TLB细胞与macro\_t之间的可能区别，张泽民老师在发表在Cell上的髓样细胞图谱的工作中也报道过一些macrophage细胞表达T cellmaker，他们认为是doublet直接扔掉了。但是他们在之前遇到同时表达T cell和 B cell marker 的细胞的时候引用了 [T1D的文献](#)，而没有直接说是doublet。虽然不知道张老师组如何定义doublet，但是他们肯定有深入调研过这类表达混合marker细胞。

利用singleR从全转录组相似度的角度看，TLB比macro\_t在两种细胞类型之间的得分更加平均。简单的说，TLB像是T cell和B cell等比例的融合，而macro\_t像是非等比例的融合



## 9 总结

以上结论仅供参考，我总感觉数据有点奇怪，并且作者貌似删除了IG（免疫球蛋白）相关的基因，不敢确认目前发现的生物学的现象的真实性。

建议等丁阳师兄的原始数据，我们重新从cellranger的步骤分析