尊敬的各位老师大家好

我是来自生物科学系的赵洵，我的科研实践II的内容是胚胎器官发育中单细胞测序分析的应用，指导老师是北京基因组研究所的刘江老师。

我将从研究背景、研究方法、实验结果以及未来计划四个方面进行汇报。

首先是研究背景，核酸测序技术已经被广泛应用在各种研究领域中，但是过去只能对细胞群体进行测序，这样并不能很好的体现出细胞与细胞之间的异质性，所以出现了基于高通量测序技术，又具有高分辨率特点的单细胞测序技术，它能获得各个细胞的各个特征的值，例如RNA测序中的各个基因表达量、ATAC测序中染色质区域的开放程度等等，最终以矩阵的形式表达。在胚胎器官发育的过程中，细胞类型的分布处于动态变化的过程中，并且由于实验样品的特殊性，往往样品量缺乏，因为在少量组织中就存在很多种不同的细胞类型，所以在这个时候单细胞测序技术就体现出了优越性。

我所做的主要工作是对测序数据进行分析，利用的工具是一个名为Seurat的R语言函数包。它可以读取测序后的矩阵数据，进行过滤、降维、聚类等统计分析。在聚类后，一方面可以直观的看到聚类的可视化结果，另一方面也可以通过热图的形式了解不同类别所具有的特征，经过查阅文献或查询其他数据库，将特征和细胞类型对应起来，综合后得到每个类群的细胞类型。

在这里我以小鼠着床13.5天的胚胎肝脏数据为例，经过比较不同类群的特异表达或者高表达的基因，我将肝脏数据分为这样一些类型。基于这样的分类结果，可以有更进一步的发现。例如，已经有文献证明在13.5天时，肝脏是胚胎中一个主要的造血器官。在这幅图当中，大部分细胞被注释为血细胞类群，而与成熟肝脏相关的肝细胞则只占很少一部分。另一方面，这样的数据也可以表明细胞类型变化的过程，例如这样一些细胞被注释为巨核细胞、红细胞共有的前体细胞，这样一些细胞被注释为红细胞前体细胞，这样一些细胞被注释为红细胞。这样造血过程中细胞分化路径与已有文献中描述的假设模型相吻合。至于为什么这些血细胞类型在RNA测序得到的转录组上有区别，但是在ATAC测序得到的染色质开放程度却很相近，这个问题还有待进一步研究。

除了对单个器官分析以外，还可以对不同器官中的相同细胞类型进行分析。例如，从眼、胰腺、胃、肠四个非中枢神经系统器官中选取的神经细胞分布显示出，消化系统器官的神经元有着高度一致性，并且与眼的神经元有明显差异。

这些就是我在科研实践II中所作的部分工作，通过和导师的交流我意识到，目前我所做的只是对数据的处理和对已有现象的描述。之后我需要综合多种组学的数据，观察不同器官、不同时期、不同细胞类型之间的差异，将数据的变化展示位生物学过程的变化，将着重点放在生物学意义上而不是代码算法上。