

王翔宇 2023521039

## 1 实验名称

mtDNA 进化分析

## 2 实验目的

1. 了解线粒体 DNA (mtDNA) 的结构与功能
2. 掌握口腔上皮细胞基因组 DNA 的提取方法
3. 学习用软件对获得的 DNA 序列进行初步分析的方法
4. 对序列比对结果进行单核苷酸多态性的分析，从而追踪人类进化的轨迹，并进行亲缘关系的分析

## 3 实验原理

### 一、线粒体的基本结构与功能

线粒体是细胞中的重要细胞器，其结构与功能直接关联到 mtDNA 的遗传特性：

- **结构特征：**
  - 光镜下呈颗粒状，电镜下可见双层膜结构及嵴（内膜折叠形成），这为氧化磷酸化提供了巨大表面积。
  - 作为“能量工厂”，线粒体分布有三羧酸循环系统和电子传递系统，是细胞产生 ATP（腺苷三磷酸）的主要场所。
  - 线粒体是核外唯一含有遗传效应物质的细胞器，这意味着 mtDNA 独立于核基因组，具有独特的遗传规律。
- **功能角色：**mtDNA 编码部分线粒体蛋白，参与能量代谢，其突变可能影响细胞功能，如与 Leber 遗传性视神经病等疾病相关。

### 二、线粒体 DNA 的遗传学特征

mtDNA 的遗传方式具有特殊性，下列六个关键特征决定了其在进化分析和亲缘关系研究中的价值：

1. **母系遗传：**mtDNA 通常通过母系传递，但也有研究（如 Lu 等、Rius 等）探讨父系遗传的可能性，结论是双亲遗传并不常见，临床实践中仍以母系遗传为主。
2. **异质性：**个体内可能存在多种 mtDNA 序列变体，这源于突变或细胞分裂中的不均匀分布（即异质质/均质质现象）。
3. **阈值效应：**mtDNA 突变需积累到一定比例才会表现表型，这与线粒体的功能冗余相关。
4. **半自主性：**mtDNA 部分依赖核基因组进行复制和表达，但拥有独立的遗传系统。

5. **高突变率**: mtDNA 的突变率比核 DNA 高约 10 倍, 原因包括线粒体缺乏高效的 DNA 修复机制且暴露于氧化应激环境。
6. **多态性**: mtDNA 在人群中呈现高度变异, 为人类进化和疾病研究提供了分子标记。

### 三、mtDNA 的多态性类型及其机制

文档将 mtDNA 多态性分为两类, 这直接关联到实验中的序列分析方法:

- **点的多态性**: 指单个碱基的突变 (如碱基替换), 可能导致限制性内切酶位点的丢失或新增, 可通过 Southern 杂交或酶切图谱等技术检测。例如, Leber 遗传性视神经病与 nt11778 G>A 突变相关, 体现了点突变的功能影响。
- **序列多态性**: 由缺失、重复或插入事件引起, 或由于高变区内串联重复序列的拷贝数差异所致。这类多态性不一定改变酶切位点的碱基, 但会改变其在基因组中的相对位置, 常需通过 PCR 与测序来分析。序列多态性主要集中在 D-loop (控制区), 包括高变区 HV1 (nt16024–nt16365, 342 bp) 和 HV2 (nt73–nt340, 268 bp), 这些区域变异速率高, 是实验的重点分析靶点。

### 四、高变区 (HV1 和 HV2) 的关键作用

高变区 HV1 和 HV2 是 mtDNA 中变异最集中的区域, 碱基变异速率高, 适用于追踪人类进化轨迹和进行群体分析。实验原理中常用于 PCR 扩增的引物示例包括 HV1 引物 (如 L15997 和 H16391), 可用于扩增约 430 bp 的 HV1 片段; HV2 引物用于扩增约 400 bp 的 HV2 片段。通过扩增并测序这些区域, 结合参考序列 (如 Revised Cambridge Reference Sequence 或 RSRs) 进行比对与 SNP 分析, 可识别个体间的差异并初步判断 mtDNA 单倍群归属。

以上内容已插入实验原理节, 便于后续对 PCR、酶切及序列分析的原理与实验设计进行引用和展开。

## 4 实验试剂和耗材

1. 材料: 人口腔上皮细胞
2. 试剂: 灭菌水, 1 × TAE, Ezup 柱式口腔拭子基因组 DNA 抽提试剂盒, 2 × Taq Master Mix, Marker, loading buffer, 引物
3. 仪器: PCR 仪, 高速离心机, 金属浴, 电泳仪
4. 其他用品: 一次性纸杯, 1.5 ml 离心管, 0.2 ml PCR 反应管, 各种规格灭菌吸头, 各量程移液器

## 5 实验步骤

### 5.1 样本采集与 DNA 提取

#### 实验前准备

- **自备材料**: 小型高速离心机 (最大离心力  $\geq 12,000 \times g$ )、水浴锅、2 ml 离心管、无水乙醇、RNase A 溶液 (10 mg/ml)。
- **试剂盒预处理**:
  - 检查 Buffer ACL 和 Buffer CL 是否有沉淀, 如有则在 65°C 溶解后使用。
  - 在 Wash Solution 中加入规定量的无水乙醇 (15 ml Wash Solution 加 45 ml 无水乙醇)。

## 标准操作流程

### 1. 样本采集与预处理

- 使用口腔拭子在口腔内壁擦拭 6–10 次，晾干 2 小时后保存。
- 注意事项：取样前 30 分钟内勿进食或饮水，避免样本污染。

### 2. 细胞裂解

- 1) 将拭子棉签部分剪入 2 ml 离心管，加入 400  $\mu\text{l}$  Buffer PBS。
- 2) 加入 400  $\mu\text{l}$  Buffer ACL、200  $\mu\text{l}$  Buffer CL 和 20  $\mu\text{l}$  Proteinase K。
- 3) 震荡混匀后，置 65°C 水浴 10 分钟，期间间歇混匀。

### 3. DNA 结合与纯化

- 1) 加入 400  $\mu\text{l}$  无水乙醇，充分混匀。
- 2) 将全部溶液转移至硅胶膜吸附柱（吸附柱容量为 750  $\mu\text{l}$ ，若体积过大需分两次过柱）。
- 3) 10,000 rpm 离心 1 分钟，弃废液。

### 4. 洗涤步骤

- 1) 加入 500  $\mu\text{l}$  Wash Solution，10,000 rpm 离心 1 分钟，弃废液。
- 2) 重复洗涤一次，确保彻底去除杂质。

### 5. DNA 洗脱

- 1) 12,000 rpm 离心 2 分钟（关键步骤：去除残留乙醇）。
- 2) 将吸附柱转入新 1.5 ml 离心管，在膜中央加入 50  $\mu\text{l}$  TE Buffer。
- 3) 静置 3 分钟后，12,000 rpm 离心 2 分钟，收集洗脱液。
- 4) 获得的 DNA 溶液可于 -20°C 保存或立即使用。

## 质量保证

- 预期  $\text{OD}_{260}/\text{OD}_{280}$  比值为 1.7–1.9。
- 无需苯酚/氯仿等有毒试剂，符合安全标准。

## 5.2 PCR 扩增（与原流程一致）

### 反应体系配制（50 $\mu\text{l}$ 体系）

| 组分                      | 实验组              | 对照组                                |
|-------------------------|------------------|------------------------------------|
| 模板 DNA                  | 5 $\mu\text{l}$  | 5 $\mu\text{l}$ ddH <sub>2</sub> O |
| 引物 1（10 $\mu\text{M}$ ） | 2 $\mu\text{l}$  | 2 $\mu\text{l}$                    |
| 引物 2（10 $\mu\text{M}$ ） | 2 $\mu\text{l}$  | 2 $\mu\text{l}$                    |
| 2× Taq Master Mix       | 25 $\mu\text{l}$ | 25 $\mu\text{l}$                   |
| ddH <sub>2</sub> O      | 16 $\mu\text{l}$ | 16 $\mu\text{l}$                   |

## 扩增程序

- 预变性：94°C 5 分钟。
- 35 个循环：94°C 30 秒 → 55–60°C 30 秒 → 72°C 1 分钟。
- 最终延伸：72°C 10 分钟。

## 5.3 电泳验证与测序

1. 琼脂糖凝胶电泳：取 10  $\mu$ l PCR 产物进行电泳分析，预期条带大小：HV1 区约 430 bp，HV2 区约 400 bp。
2. 样本送测：有条带的样本移交测序分析。

## 5.4 序列分析与 SNP 鉴定

1. 软件分析：使用 MEGA/PHYLIP 等软件比对 rCRS/RSRS 参考序列，识别 HV1/HV2 区单核苷酸多态性 (SNP)。
2. 进化追踪：与 RSRS 序列对比以确定单倍群，并分析个体间差异位点用于进化或亲缘关系推断。

## 5.5 常见问题解决方案

1. 得率低：检查拭子类型（推荐植绒拭子）、确认 Wash Solution 中乙醇添加量、检查洗脱液 pH（建议 >7.5）。
2. DNA 降解：避免样本反复冻融，裂解时间不超过 30 分钟。
3. 杂质残留：离心后开盖风干或轻轻孵育以去除乙醇残留。
4. 电泳异常：建议使用 0.8% 琼脂糖凝胶，适当减少上样量并检查 Marker/缓冲液浓度。

# 6 数据记录

## 6.1 电泳结果

电泳结果如图 1 所示，我的样品编号为 6。（关于 1 号样品出现这种现象的解释：胶没凝好就拔孔导致第一个孔漏了，于是我们又在第八个空补了一点，但是已经跑了一会儿了，所以这条带比其他条带要靠后一点。）

## 6.2 测序数据

HV1 区测序正常，可以直接从测序公司获取拼接后的序列数据。下图截取了 HV1，F 方向引物的部分测序结果（图 2），可以看到信号较好，碱基识别清晰。

HV2 去测序结果不理想，公司分别提供了 F 和 R 两个方向的测序数据，F 方向出现套峰，R 方向被报告为中断，但使用 SnapGene 尝试拼接，仍获得一段可用序列。人工查看 F 方向和 R 测序结果，峰还比较清晰，可以继续分析。F 和 R 方向的部分测序结果以及最后的拼接结果示意如图 3 所示。

HV1 区测序结果如下：

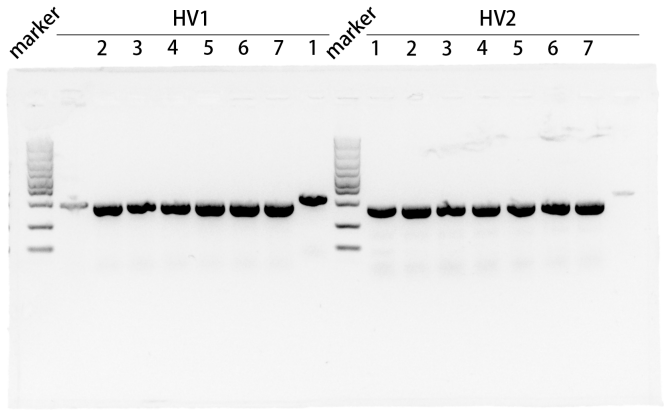


Figure 1: 电泳结果图

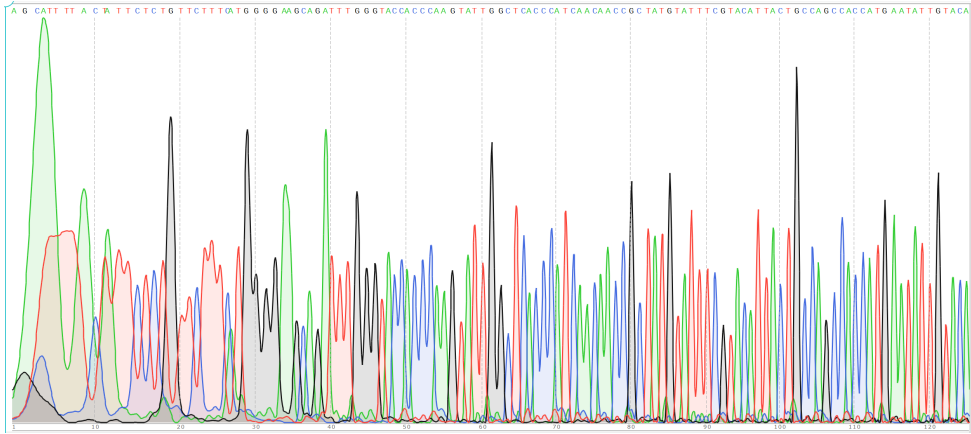
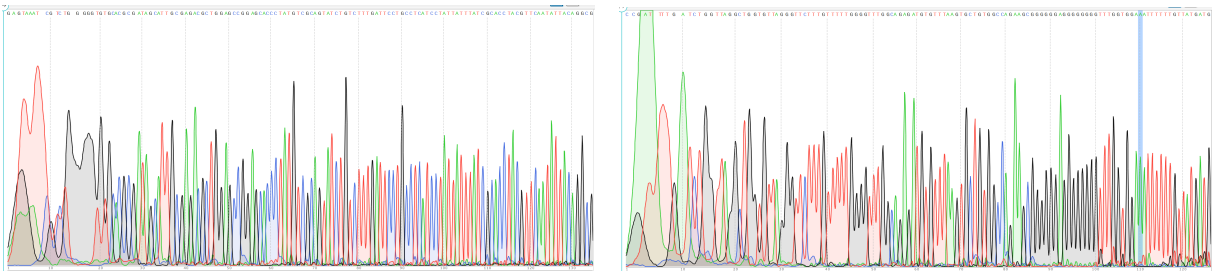
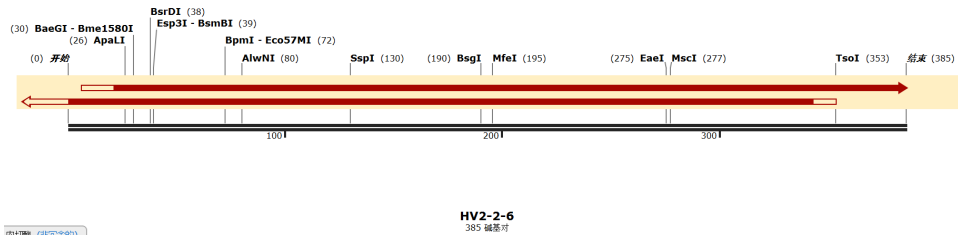


Figure 2: HV1 区 F 方向引物测序结果截图



(a) HV2 区 F 方向引物测序结果截图

(b) HV2 区 R 方向引物测序结果截图



(c) HV2 区 F 和 R 方向引物测序结果拼接示意图

Figure 3: HV2 区测序结果

TATTCTCTGTTCTTTTCATGGGGAAGCAGATTGGGTACCACCCAAGTATTGGCTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAG  
CCACCATGAATATTGTACAGTACCATAAATACTTGACCACCTGTAGTACATGAAAACCCAACCCACATCAAAACCCCTCCCATGCTTACAAGCA  
AGTACAGCAATCAACCCTCAACTATCACACATCAACTGCAACTCGAAAGCCACCCCTCACCCACTAGGATACCAACAAACCTACCCACCCCTTAACA  
GCACATAGTACATAAAGCCATTACCGTACATAGCACATTACAGTCAAATCCCTTCTCGTCCCATGGATGACCCCTCAGATAGGGGTCCCTTG  
AACC

HV2 区测序结果如下：

CATTTGGTATTTTCGTCTGGGGGTGTGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCCCTATGTCGCAGTATCTGTCTTTGATTCTG  
CCTCATCTATTATTTATCGCACCTACGTTCAATATTACAGGCGAACATACTTACTAAAGTGTGTTAATTAATTAATGCTTGTAGGACATAATAAT  
AACAATTGATGTCTGCACAGCGCTTTCCACACAGACATCATAACAAAAAATTTCCACCAAAACCCCTCCCTCCGCTTCTGGCCACAGCACTT  
AAACACATCTCTGCCAAACCCCAAAACAAAGAACCCTAACACCAGCCTAACAGATTTCAAATTTTATCTTTTGGCGGTATGCCTTTTGAACAG  
A

## 7 结果分析

### 7.1 SNP 鉴定

使用 Mega12，分别将我的 HV1 区和 HV2 区测序结果与班级中所有测序成功的同学的测序结果进行比对，鉴定出 SNP 位点。HV1 区共鉴定出 26 个 SNP 位点，详细结果见附表 1；HV2 区共鉴定出 7 个 SNP 位点，详细结果见附表 2。

### 7.2 进化树构建

使用 MEGA12 软件，根据 HV1 区和 HV2 区的序列，构建了最大似然法（Maximum Likelihood）进化树。进化树如图 4 所示。

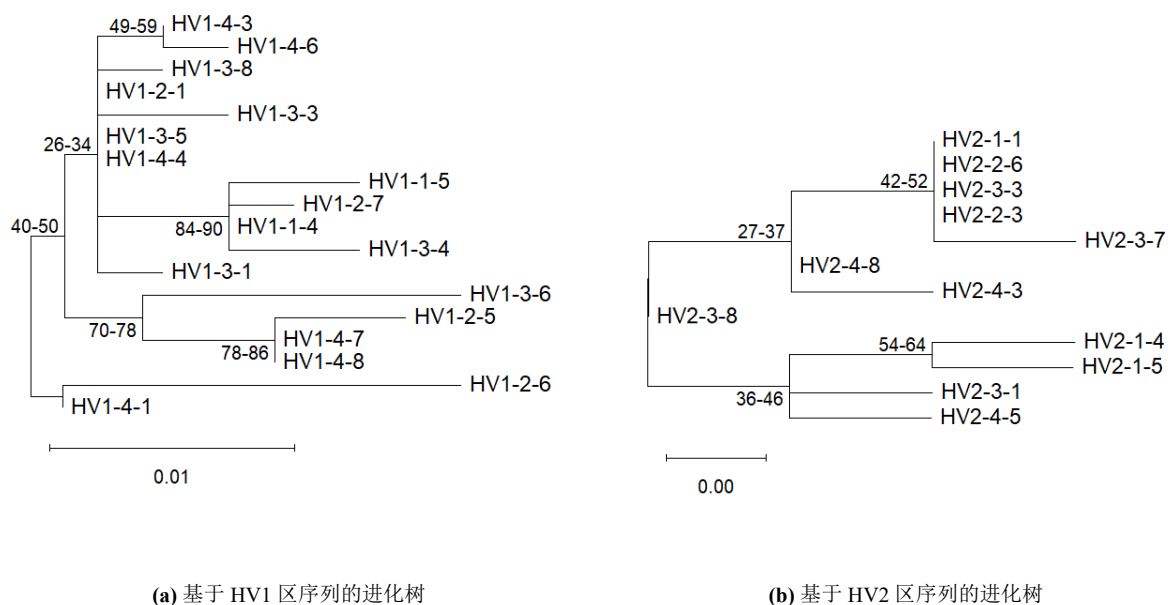


Figure 4: 基于 HV1 区和 HV2 区序列的最大似然法进化树

### 7.3 单倍群分析

为追踪人类进化轨迹，将我的测序结果与从 phylotree.org 下载的 Reconstructed Sapiens Reference Sequence (RSRS) 进行比对，共检测到 96 个 SNP 位点，详细结果见附表 3。

使用 Haplogrep 3，初步判断我的 mtDNA 单倍群归属于 F (图 5)。根据上一步检测到的 SNP 位点，在 Phylotree 上进一步确认，验证 Haplogrep 3 的预测无误。然而受测序长度的限制，我的 mtDNA 单倍群无法进一步精确。Haplogrep 3 预测结果链接：  
<https://haplogrep.i-med.ac.at/jobs/PA27AU5CVWFQ7ZLlgLWpu8KE1zg6MMUYdtWRP9zHPeFs2DX6In>

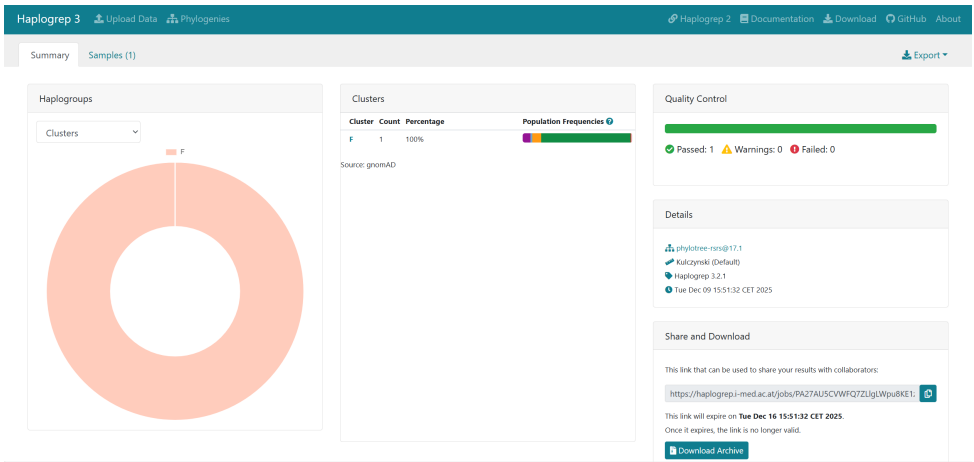


Figure 5: Haplogrep 3 结果

8 实验结论

本实验通过口腔上皮细胞基因组 DNA 提取、PCR 扩增、测序及生物信息学分析，系统研究了 mtDNA 高变区（HV1 和 HV2）的遗传变异特征，得出以下结论：

8.1 DNA 提取与扩增

- 1. 成功利用 Ezup 柱式试剂盒从口腔拭子样本中提取基因组 DNA，质量符合 OD<sub>260</sub>/OD<sub>280</sub> 比值 1.7–1.9 的标准，避免了传统苯酚/氯仿法的毒性风险。
- 2. PCR 成功扩增 mtDNA 高变区 HV1（约 430 bp）和 HV2（约 400 bp）片段，电泳验证显示目的条带清晰，表明引物设计合理且扩增特异性良好。
- 3. HV1 区测序质量较高，信号清晰可直接使用；HV2 区出现套峰现象，通过正反向测序拼接获得可用序列，说明双向测序策略对提高序列准确性的重要性。

8.2 群体遗传变异分析

- 1. 班级群体内 HV1 区检测到 26 个 SNP 位点，HV2 区检测到 7 个 SNP 位点，反映出 mtDNA 控制区的高度多态性特征，验证了 D-loop 区域作为分子标记的有效性。
- 2. SNP 位点分布不均匀，集中于特定区域，符合 mtDNA 高变区的突变热点特征，这些变异为群体遗传结构分析和个体识别提供了丰富的遗传标记。
- 3. 通过最大似然法构建的系统发育树显示，不同个体间存在明显的分支结构，表明样本间的进化关系和遗传距离差异，HV1 和 HV2 区域的进化树拓扑结构基本一致，相互验证了分析结果的可靠性。

### 8.3 进化追踪与单倍群鉴定

1. 将个人测序结果与 RSRS (Reconstructed Sapiens Reference Sequence) 比对, 共检测到 96 个 SNP 位点, 其中 HV2 区 84 个位点 (位置 146–433), HV1 区 12 个位点 (位置 16066–16402), 这些差异位点记录了现代人类从非洲起源后的进化轨迹。
2. 利用 Haplogrep 3 和 Phylotree 数据库分析, 成功鉴定个人 mtDNA 单倍群归属为 **F 单倍群**, 该单倍群主要分布于东亚和东南亚地区, 与中国人群的遗传背景相符, 验证了 mtDNA 单倍群分析在人类起源与迁徙研究中的应用价值。
3. 受测序长度限制, 无法进一步细分至 F 单倍群的亚型, 提示全长 mtDNA 测序对精确单倍群鉴定的必要性, 但现有数据已足以进行宏观的群体归属和地理来源推断。

### 8.4 实验技术与方法

1. 掌握了从样本采集、DNA 提取、PCR 扩增、电泳检测到测序分析的完整实验流程, 建立了 mtDNA 分析的标准操作规范。
2. 熟练应用 MEGA 软件进行序列比对、SNP 检测和系统发育树构建, 运用 Haplogrep 和 Phylotree 等在线工具进行单倍群分析, 提升了生物信息学数据处理能力。
3. 理解了 mtDNA 母系遗传、高突变率、无重组等遗传特性, 以及 D-loop 区域 HV1 和 HV2 作为分子标记在群体遗传学、法医学和进化生物学中的应用原理。

综上所述, 本实验系统验证了 mtDNA 高变区在人类进化分析和群体遗传研究中的重要价值, 成功通过 SNP 分析追踪了个体的母系遗传谱系, 鉴定出 F 单倍群归属, 为理解东亚人群的遗传结构和迁徙历史提供了分子证据。实验结果表明, mtDNA D-loop 区域的序列变异分析是研究人类起源、迁徙和群体分化的有效手段。

## 9 思考题

### 1. mtDNA 相比于核 DNA 高突变率的原因?

mtDNA 的突变率约为核 DNA 的 10 倍, 主要原因包括:

1. **氧化应激环境:** 线粒体是细胞进行氧化磷酸化的主要场所, 在 ATP 合成过程中会产生大量活性氧自由基 (ROS), 这些自由基对 mtDNA 造成持续的氧化损伤, 导致碱基修饰和 DNA 链断裂。由于 mtDNA 紧邻电子传递链, 直接暴露于高浓度 ROS 环境中, 受损伤程度远高于细胞核内的核 DNA。
2. **DNA 修复机制不完善:** mtDNA 缺乏组蛋白保护, 呈裸露状态, 更易受到损伤。同时, 线粒体内的 DNA 修复系统相对简单, 缺乏核苷酸切除修复 (NER) 等重要修复途径, 虽具备碱基切除修复 (BER) 和错配修复 (MMR) 机制, 但修复效率远低于核 DNA, 导致突变更容易积累和固定。
3. **复制保真度较低:** 线粒体 DNA 聚合酶  $\gamma$  (Pol  $\gamma$ ) 虽然具有 3'→5' 外切酶活性进行校对, 但其整体保真度仍低于核 DNA 复制酶复合体。此外, mtDNA 复制不依赖细胞周期, 持续进行高频率复制, 增加了复制错误的累积机会。
4. **缺乏内含子和非编码区保护:** mtDNA 基因组极为紧凑 (约 16.5 kb), 编码序列占比高达 93%, 几乎没有内含子和保护性非编码区。这意味着突变更容易直接影响功能基因, 虽然有害突变会被选择淘汰, 但中性或轻微有害突变更易固定, 导致观察到的高突变率。
5. **母系遗传与瓶颈效应:** mtDNA 通过卵细胞母系遗传, 在生殖细胞发育过程中经历瓶颈效应, 导致突变的随机漂变和快速固定, 加速了 mtDNA 的进化速率。



## 2. mtDNA 多态性有哪些应用？

mtDNA 多态性因其母系遗传、高突变率和无重组等特性，在多个领域具有重要应用：

### 1. 人类进化与起源研究：

- 通过 mtDNA 单倍群分析，追溯现代人类“走出非洲”的迁徙路线和时间，重建人类祖先的地理分布和扩散历史。
- 线粒体夏娃理论：利用 mtDNA 系统发育分析，推测所有现代人类 mtDNA 均可追溯到约 20 万年前生活在非洲的共同女性祖先。
- 研究不同人群的遗传关系和分化时间，揭示古代人群迁徙、混合和隔离事件。

### 2. 法医学与个体识别：

- 高度降解样本（如骨骼、牙齿、毛发）的个体鉴定，因 mtDNA 拷贝数高（每个细胞数百至数千份），在核 DNA 严重降解时仍能成功提取和分析。
- 母系亲缘关系鉴定：当无法获取直系父母样本时，可通过比对母系亲属（如外祖母、姨母）的 mtDNA 确认身份。
- 灾难遇难者身份识别和历史人物遗骸鉴定（如末代沙皇家族遗骸鉴定）。

### 3. 群体遗传学与生物地理学：

- 研究种群遗传结构、基因流动和遗传多样性，评估濒危物种的保护遗传学问题。
- 分析地理隔离群体的分化程度和迁徙历史，揭示环境因素对种群分布的影响。
- 构建生物地理分布模式，推断物种扩散和定居的历史过程。

### 4. 疾病遗传学与医学研究：

- mtDNA 突变与线粒体疾病的关联研究：如 Leber 遗传性视神经病（LHON）与 nt11778 G>A 突变、MELAS 综合征与 nt3243 A>G 突变等。
- 研究 mtDNA 多态性对复杂疾病（如帕金森病、阿尔茨海默病、糖尿病、癌症）易感性的影响。
- 线粒体单倍群与疾病风险、药物代谢、运动能力等表型关联分析。
- mtDNA 异质性（heteroplasmy）在疾病表达和遗传咨询中的应用。

### 5. 古 DNA 研究与考古学：

- 从古代样本（如尼安德特人、丹尼索瓦人化石）中提取 mtDNA，研究古人类与现代人类的遗传关系。
- 追溯农业起源、家养动植物驯化历史和古代贸易路线。
- 重建古代文明的人口迁移和文化交流模式。

### 6. 物种鉴定与系统发育分析：

- 利用 mtDNA 条形码（如 COI 基因）进行物种快速鉴定，应用于生物多样性调查、食品安全检测和濒危物种保护。
- 构建动物系统发育树，解析物种间的进化关系和分类地位。

综上所述，mtDNA 多态性已成为遗传学、进化生物学、法医学和医学研究中的重要分子标记，为理解人类起源、疾病机制和生物多样性提供了强有力的工具。

### 9.1 3. 分析比较你和同学之间序列的差异。

详见章节 [7.1](#)和 [7.2](#)

## 10 收获体会

学会使用 MEGA 软件，进行序列对齐，SNP 分析，绘制进化树，学会使用 [phylotree.org](http://phylotree.org) 查找单倍群，同时了解到 Haplogrep 3 等快捷查找单倍群的工具。

A 附表

Table 1: HV1 区 SNP 分析结果

| 位置  | 1-4 | 1-5 | 2-1 | 2-5 | 2-6 | 2-7 | 3-1 | 3-3 | 3-4 | 3-5 | 3-6 | 3-8 | 4-1 | 4-3 | 4-4 | 4-6 | 4-7 | 4-8 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 50  | A   | A   | A   | A   | G   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   |
| 77  | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | C   | T   | T   |
| 95  | C   | C   | C   | T   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   |
| 108 | T   | C   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   |
| 110 | T   | T   | T   | T   | T   | C   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   |
| 113 | G   | G   | G   | A   | A   | G   | A   | G   | G   | G   | G   | G   | G   | G   | G   | G   | G   | G   |
| 129 | G   | G   | G   | G   | G   | G   | G   | G   | G   | G   | A   | G   | G   | G   | G   | G   | G   | G   |
| 142 | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | G   | A   | A   | A   | A   | A   | A   |
| 146 | A   | A   | A   | A   | G   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   |
| 152 | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | T   | C   | C   | C   | C   | C   | C   | C   |
| 156 | T   | T   | T   | T   | C   | T   | T   | T   | T   | T   | T   | T   | C   | T   | T   | T   | T   | T   |
| 163 | C   | C   | C   | C   | C   | C   | C   | C   | T   | C   | C   | C   | C   | C   | C   | C   | C   | C   |
| 172 | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | T   | C   | C   | C   | C   | C   | C   | C   |
| 176 | C   | C   | C   | C   | C   | C   | C   | C   | T   | C   | C   | C   | C   | C   | C   | C   | C   | C   |
| 207 | T   | T   | T   | T   | C   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   |
| 218 | C   | C   | C   | C   | C   | C   | C   | T   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   |
| 240 | C   | T   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   |
| 241 | C   | C   | C   | A   | C   | C   | C   | C   | C   | C   | T   | C   | C   | C   | C   | C   | A   | A   |
| 245 | C   | C   | C   | T   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | T   | T   |
| 274 | T   | T   | C   | C   | C   | T   | C   | C   | T   | C   | C   | C   | C   | C   | C   | C   | C   | C   |
| 288 | T   | T   | T   | T   | C   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   |
| 295 | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | C   | T   | T   | T   | T   | T   | T   | T   |
| 300 | A   | A   | A   | A   | A   | A   | A   | G   | A   | A   | A   | A   | A   | A   | A   | A   | A   | A   |
| 303 | A   | A   | G   | G   | G   | A   | G   | G   | A   | G   | G   | G   | G   | G   | G   | G   | G   | G   |
| 309 | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | T   | C   | T   | C   | T   | T   |
| 346 | C   | C   | C   | T   | T   | C   | C   | C   | C   | C   | T   | C   | C   | C   | C   | C   | T   | T   |

Table 2: HV2 区 SNP 分析结果

| 位置  | 1-1 | 1-4 | 1-5 | 2-3 | 2-6 | 3-1 | 3-3 | 3-7 | 3-8 | 4-3 | 4-5 | 4-8 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 82  | C   | C   | C   | C   | C   | C   | C   | C   | T   | T   | C   | T   |
| 84  | T   | C   | C   | T   | T   | C   | T   | T   | C   | T   | C   | T   |
| 126 | C   | C   | C   | C   | C   | C   | C   | T   | C   | C   | C   | C   |
| 132 | A   | A   | G   | A   | A   | A   | A   | A   | A   | G   | A   | A   |
| 167 | A   | G   | G   | A   | A   | A   | A   | A   | A   | A   | A   | A   |
| 195 | G   | A   | G   | G   | G   | T   | G   | G   | G   | G   | G   | G   |
| 230 | C   | C   | C   | C   | C   | C   | C   | C   | C   | C   | T   | C   |

Table 3: 基于 RSRS 参考序列的 SNP 分析结果

| 位置  | RSRS | HV1 | HV2 | 位置    | RSRS | HV1 | HV2 |
|-----|------|-----|-----|-------|------|-----|-----|
| 146 | C    | -   | T   | 382   | C    | -   | G   |
| 152 | C    | -   | T   | 383   | T    | -   | C   |
| 195 | C    | -   | T   | 384   | A    | -   | C   |
| 247 | A    | -   | G   | 385   | A    | -   | T   |
| 310 | T    | -   | C   | 386   | C    | -   | A   |
| 311 | C    | -   | T   | 387   | C    | -   | A   |
| 316 | G    | -   | C   | 388   | A    | -   | C   |
| 318 | T    | -   | G   | 389   | G    | -   | C   |
| 319 | T    | -   | C   | 391   | T    | -   | G   |
| 320 | C    | -   | T   | 392   | T    | -   | A   |
| 322 | G    | -   | C   | 394   | C    | -   | T   |
| 323 | G    | -   | T   | 395   | A    | -   | T   |
| 324 | C    | -   | G   | 396   | A    | -   | C   |
| 325 | C    | -   | G   | 398   | T    | -   | A   |
| 326 | A    | -   | C   | 399   | T    | -   | A   |
| 329 | G    | -   | C   | 402   | A    | -   | T   |
| 330 | C    | -   | A   | 404   | C    | -   | A   |
| 331 | A    | -   | G   | 406   | T    | -   | C   |
| 333 | T    | -   | A   | 409   | G    | -   | T   |
| 334 | T    | -   | C   | 410   | G    | -   | T   |
| 335 | A    | -   | T   | 411   | C    | -   | G   |
| 336 | A    | -   | T   | 413   | G    | -   | C   |
| 338 | C    | -   | A   | 414   | T    | -   | G   |
| 342 | T    | -   | C   | 415   | A    | -   | G   |
| 343 | C    | -   | A   | 417   | G    | -   | A   |
| 347 | G    | -   | C   | 418   | C    | -   | T   |
| 348 | C    | -   | T   | 419   | A    | -   | G   |
| 349 | C    | -   | G   | 421   | T    | -   | C   |
| 350 | A    | -   | C   | 425   | A    | -   | T   |
| 351 | A    | -   | C   | 426   | A    | -   | T   |
| 353 | C    | -   | A   | 427   | C    | -   | G   |
| 354 | C    | -   | A   | 429   | G    | -   | A   |
| 357 | A    | -   | C   | 430   | T    | -   | C   |
| 358 | A    | -   | C   | 431   | C    | -   | A   |
| 362 | C    | -   | A   | 432   | A    | -   | G   |
| 364 | A    | -   | C   | 433   | C    | -   | A   |
| 366 | G    | -   | A   | 16066 | A    | G   | -   |
| 368 | A    | -   | G   | 16162 | A    | G   | -   |
| 369 | C    | -   | A   | 16172 | T    | C   | -   |
| 370 | C    | -   | A   | 16187 | T    | C   | -   |
| 372 | T    | -   | C   | 16189 | C    | T   | -   |
| 373 | A    | -   | C   | 16223 | T    | C   | -   |
| 374 | A    | -   | T   | 16230 | G    | A   | -   |
| 375 | C    | -   | A   | 16278 | T    | C   | -   |
| 378 | C    | -   | A   | 16304 | T    | C   | -   |
| 379 | A    | -   | C   | 16311 | C    | T   | -   |
| 380 | G    | -   | C   | 16400 | C    | A   | -   |
| 381 | C    | -   | A   | 16402 | A    | C   | -   |