

“实用生物信息技术”课程小组讨论总结 报告

组：G3

次：R4

组长：陈伟辉

执笔：陈伟辉

一、时间 2026 年 4 月 29 日星期三

二、方式 面对面交流

三、主题 课后复习

四、内容

(1) 系统进化树基础概念

(2) MEGA 12 软件的参数练习

(3) MEGA 12 软件的拓展功能挖掘

(4) 问题与建议

1. 参阅 ABC 提交的资料，并查阅相关文献，基础概念如下：

1) 分子演化和系统发生

- 分子演化：以 DNA、RNA、蛋白质序列为研究对象，分析生物大分子的突变、固定与扩散，揭示基因与蛋白质的进化规律。
- 系统发生：研究物种或基因的演化起源与亲缘关系，用树状结构表示分化路径与演化关系。

2) 序列相似性 (Similarity) 和序列同源性 (Homology)

- 序列相似性：可定量指标，指两条序列相同 / 相似位点的比例，反映序列匹配程度。
- 序列同源性：定性判断，指序列来自共同祖先的演化关系；同源序列通常相似性高，但高相似性不一定同源。

3) 直系同源 (Ortholog) 和旁系同源 (Paralog)

- 直系同源：不同物种中，因物种分化产生的同源基因，功能通常保守。
- 旁系同源：同一物种内，因基因复制产生的同源基因，功能易发生分化。

4) 核苷酸替换模型和氨基酸替换模型

- 核苷酸替换模型：描述 DNA 碱基相互替换概率的模型，如 Jukes-Cantor、Kimura 2-parameter、GTR。
- 氨基酸替换模型：描述氨基酸相互替换概率的矩阵，如 BLOSUM、PAM、GONNET，用于蛋白序列比对与建树。

5) 突变速率和分子钟

- 突变速率：单位时间内序列发生突变的概率。
- 分子钟：假定突变速率大致恒定，用序列差异估算物种 / 基因的分化时间。

6) 进化分支树 (Cladogram) 和系统发生树 (Phylogram)

- 进化分支树：仅展示分支拓扑关系，无分支长度。
- 系统发生树：同时展示拓扑关系与分支长度，分支长度代表进化差异。

7) 基因树和物种树

- 基因树：由单个基因序列构建，反映该基因的演化。

- 物种树：由全基因组或多基因整合构建，反映物种真实分化历史。

8) 无根树和有根树

- 无根树：仅显示类群间亲缘关系，无演化方向。
- 有根树：通过外类群等确定根节点，明确祖先→后代的演化方向。

9) 分支和节点

- 分支：连接节点的线段，代表进化路径。
- 节点：分支交汇点，代表分化事件。

10) 内部节点和外部节点

- 内部节点：树中间的节点，代表祖先类群。
- 外部节点：树末端的节点，代表现存类群。

11) 根节点和叶节点

- 根节点：有根树最顶端，代表所有类群的共同祖先。
- 叶节点：即外部节点，代表研究的现存样本。

12) 距离法和位点法

- 距离法：先计算序列间距离矩阵，再基于距离建树（如 NJ）。
- 位点法：直接利用每个位点信息建树（如 MP、ML）。

13) 最大简约法和最大似然法

- 最大简约法（MP）：选择突变总数最少的树为最优树。
- 最大似然法（ML）：基于替换模型，计算观测数据在该树结构下的最大概率。

2. 参阅 ABC 提交的资料，并查阅相关文献，回答以下问题

1) 构建系统发生树的基本步骤

1. 同源序列获取：检索并下载目标物种 / 基因的核酸或蛋白序列。
2. 多序列比对：用 ClustalW、MUSCLE 等进行序列对齐。
3. 替换模型选择：选择合适的核苷酸 / 氨基酸替换模型。
4. 建树方法选择：NJ、MP、ML、BI 等。

5. 系统发生树构建：使用 MEGA 等软件生成树。
6. 稳定性检验：Bootstrap 自举检验。
7. 树可视化与解读：标注分支、支持值、分化关系。

2) 选择核苷酸序列或氨基酸序列的原则

- 选用核苷酸序列
 - 近缘物种、分化时间短
 - 研究密码子偏好、选择压力、转录调控
 - 序列差异小、信息位点充足
- 选用氨基酸序列
 - 远缘物种、分化时间长
 - 避免同义突变干扰
 - 功能保守、序列差异大
- 通用原则：近缘用 DNA，远缘用蛋白；保守结构域优先用蛋白。

3) 自举法（Bootstrap）检验稳定性的原理

1. 从原始多序列比对中有放回重复抽样，生成与原长度相同的重采样数据集。
2. 对每个重采样数据集重新构建系统发生树。
3. 统计某一分支在所有重采样树中出现的频率，即为该分支的 Bootstrap 支持值。
4. 支持值越高，该分支拓扑结构越稳定可靠。

4) 确定无根树根节点的方法

1. 外类群法：加入亲缘关系明确的外类群，根位于外类群与内类群之间。
2. 中点 rooting：假设分子钟恒定，取最长路径中点为根。
3. 基因复制 rooting：利用旁系同源基因定位祖先分支。

5) 判断“先有物种”还是“先有基因”

- 先有基因（基因先分化，物种后分化）

同一基因家族的不同类群先聚为一支，再按物种分开。

- 先有物种（物种先分化，基因后复制）

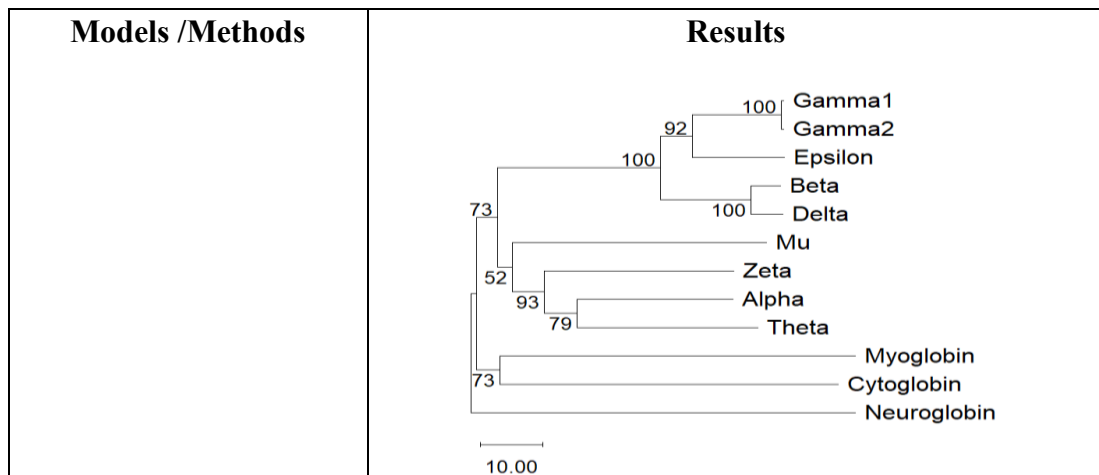
同一物种的序列先聚为一支，再按基因家族类型分开。

6) 不同建树方法的基本原理和特点

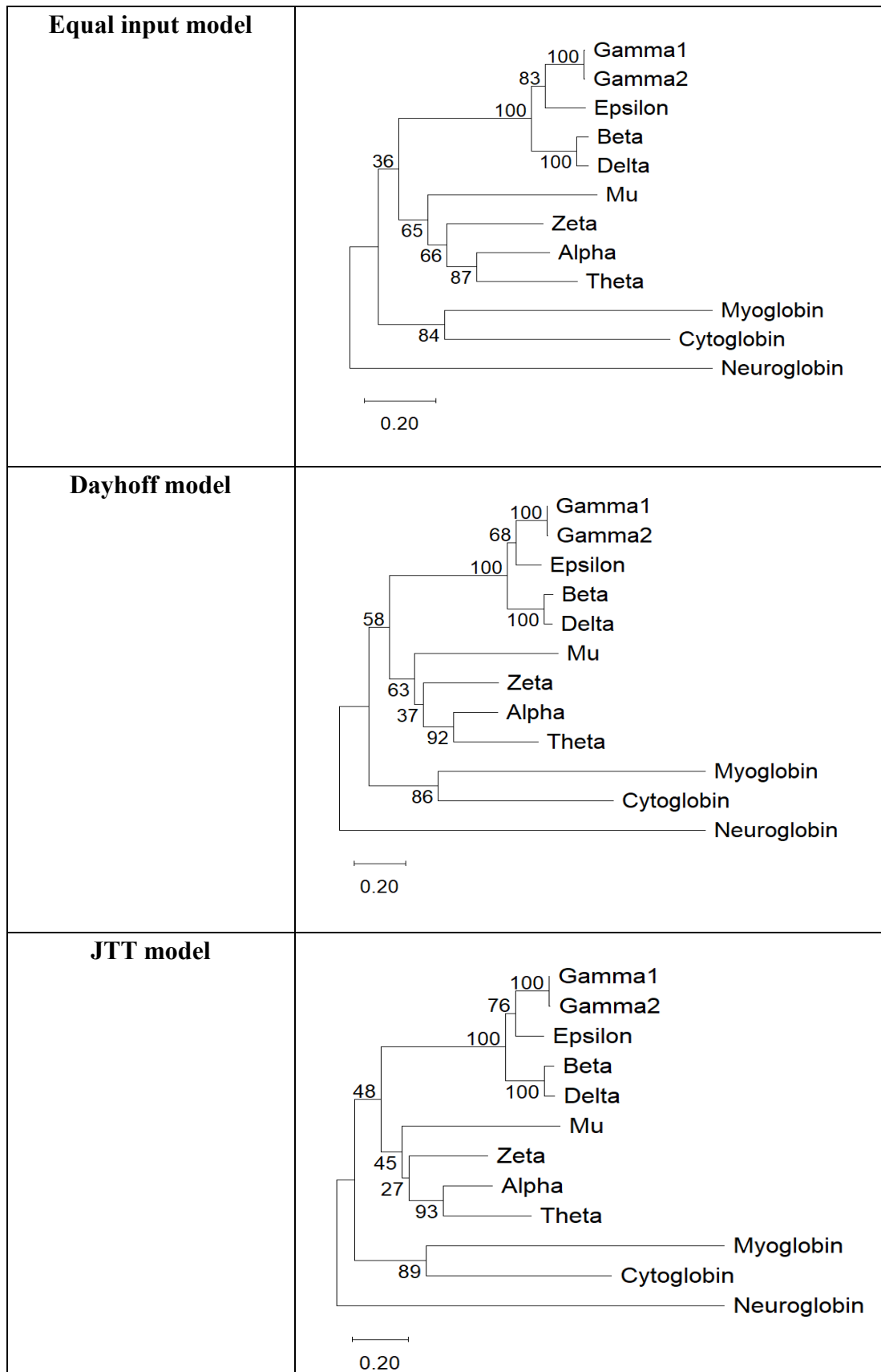
方法	原理	优点	缺点
邻接法 (NJ)	基于距离矩阵，迭代寻找最近邻居	速度快、适合大数据集	对长枝吸引较敏感
最大简约法 (MP)	选择突变次数最少的树	直观、无模型依赖	易长枝吸引、大数据慢
最大似然法 (ML)	计算序列在给定树的似然值	准确、稳健、结果可靠	计算量大、速度慢
贝叶斯法 (BI)	基于后验概率评估树	支持值直观、结果稳健	计算极慢、参数复杂

3. 人珠蛋白基因家族系统发生树实例

1) 以人珠蛋白基因家族 12 个成员蛋白质序列;用 MEGA 邻接法构建系统发生树; 选择不同氨基酸替换模型 (Substitution Model), 比较所构建的系统发生树的拓扑结构和稳定性值 (Bootstrap value), 说明不同替换模型对结果的影响。



<p>No.of differences</p>	<p>Phylogenetic tree based on the number of differences. The tree shows relationships between Gamma1, Gamma2, Epsilon, Beta, Delta, Mu, Zeta, Alpha, Theta, Myoglobin, Cytoglobin, and Neuroglobin. Bootstrap values are shown at the nodes: 94, 100, 89, 100, 100, 90, 88, 77, 80. A scale bar of 0.10 is provided.</p>
<p>p-distance</p>	<p>Phylogenetic tree based on p-distance. The tree shows relationships between Gamma1, Gamma2, Epsilon, Beta, Delta, Mu, Zeta, Alpha, Theta, Myoglobin, Cytoglobin, and Neuroglobin. Bootstrap values are shown at the nodes: 94, 100, 89, 100, 100, 90, 88, 77, 80. A scale bar of 0.10 is provided.</p>
<p>Poisson model</p>	<p>Phylogenetic tree based on the Poisson model. The tree shows relationships between Gamma1, Gamma2, Epsilon, Beta, Delta, Mu, Zeta, Alpha, Theta, Myoglobin, Cytoglobin, and Neuroglobin. Bootstrap values are shown at the nodes: 71, 100, 88, 100, 100, 80, 68, 89, 85. A scale bar of 0.20 is provided.</p>



附：默认重复次数 100

拓扑结构就是这棵树的分支“骨架”，也就是各个类群之间的亲缘关系和分组方式，不考虑分支长度，只看谁和谁聚在一起。

Bootstrap 值是用来衡量树里每个分支的可靠性的，数字越大，说明这个分支结构越稳定、越可信。它代表在 100 次自举重采样建树中，这个分支被重复构建出来的次数。比如图里的 100，表示 100 次重采样中，这个分支每次都出现，说明结构非常稳定。一般来说：

Bootstrap 值 ≥ 70 ：这个分支的拓扑结构是可靠的；

Bootstrap 值在 50-70 间：可信度中等，需谨慎解读；

Bootstrap 值 < 50 ：这个分支结构不稳定，不能作为结论依据；

选择不同氨基酸替换模型（Substitution Model），比较所构建的系统发生树的拓扑结构和稳定性值（Bootstrap value），可以得出结论：

上述 6 种模型的拓扑结构一致：

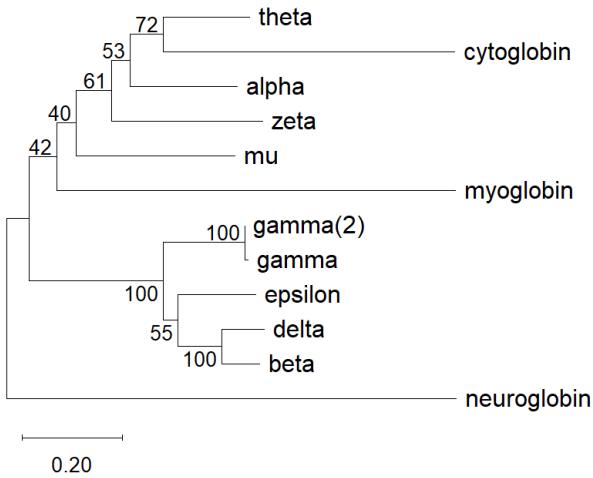
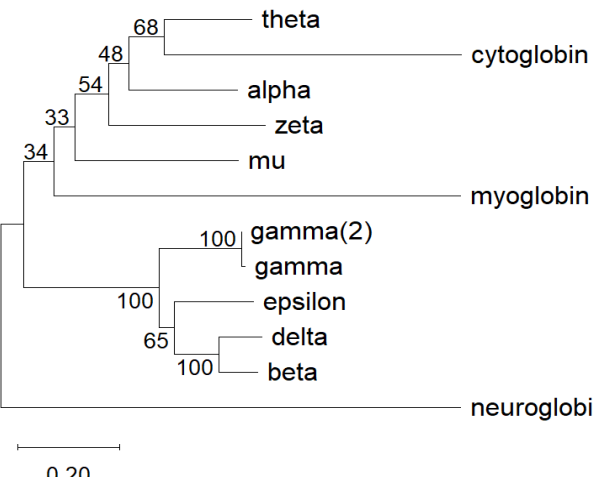
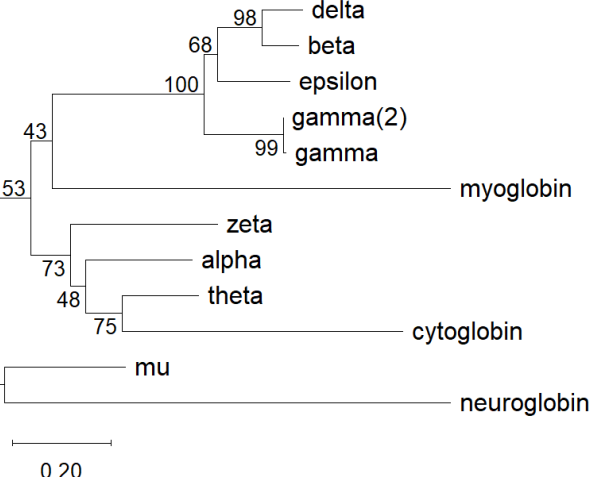
- 第一大分支：Gamma1/Gamma2/Epsilon/Beta/Delta（胎儿和成体血红蛋白亚基）
- 第二大分支：Mu/Zeta/Alpha/Theta（另一类血红蛋白亚基）
- 第三大分支：Myoglobin/Cytoglobin（肌红蛋白 / 胞红蛋白）
- 最外侧：Neuroglobin（神经红蛋白，作为外类群）

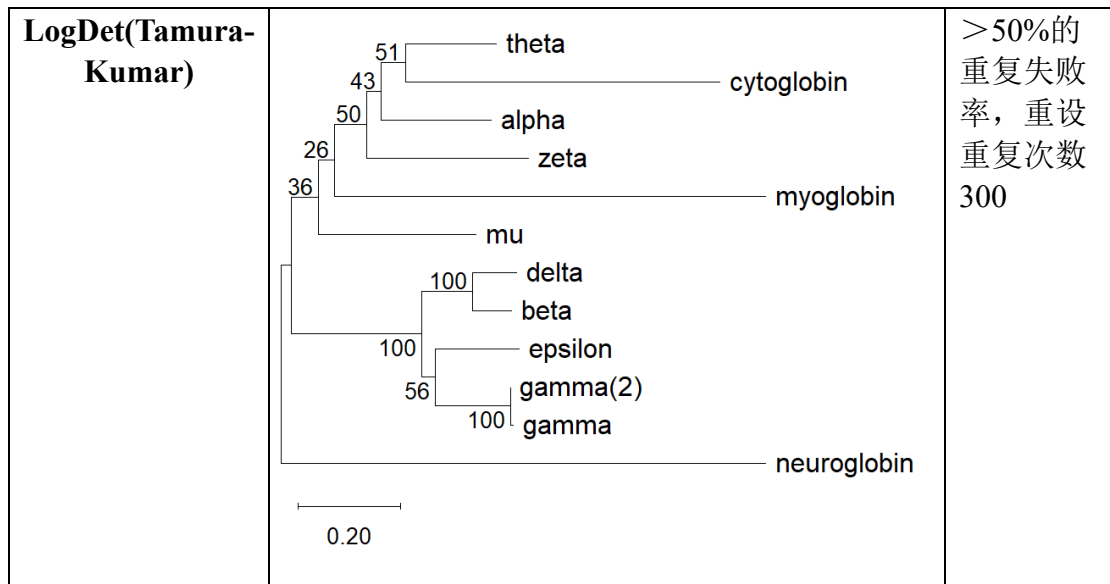
稳定性值：上述 6 种模型中 p-distance 与 Poisson model 的分支结构最稳定，结果最为可靠，相反，Equal input model、Dayhoff model 与 JTT model 这三种分支的稳定性值低于 50，分支结构不稳定，不能作为结论依据。

2) 以人珠蛋白基因家族 12 个成员编码区序列，用 MEGA 邻接法构建系统发生树；，选择不同核苷酸替换模型，比较所构建的系统发生树的拓扑结构和稳定性值（Bootstrap value），说明不同替换模型对结果的影响。

Models	Results	补充
p-distance		

<p>Jukes-Cantor model</p>	<p>Phylogenetic tree showing relationships between globin genes (delta, beta, epsilon, gamma(2), gamma, myoglobin, zeta, alpha, theta, cytoglobin, mu, neuroglobin) under the Jukes-Cantor model. Bootstrap values are indicated at nodes. A scale bar of 0.20 is shown.</p>	
<p>Kimura 2-parameter model</p>	<p>Phylogenetic tree showing relationships between globin genes (delta, beta, epsilon, gamma(2), gamma, myoglobin, zeta, alpha, theta, cytoglobin, mu, neuroglobin) under the Kimura 2-parameter model. Bootstrap values are indicated at nodes. A scale bar of 0.20 is shown.</p>	
<p>Tajima-Nei model</p>	<p>Phylogenetic tree showing relationships between globin genes (delta, beta, epsilon, gamma(2), gamma, myoglobin, alpha, zeta, theta, cytoglobin, mu, neuroglobin) under the Tajima-Nei model. Bootstrap values are indicated at nodes. A scale bar of 0.20 is shown.</p>	

<p>Tamura 3-parameter model</p>	 <p>Phylogenetic tree showing relationships between cytoglobin, myoglobin, and neuroglobin. The tree is rooted at the bottom left. A scale bar of 0.20 is shown at the bottom left. Bootstrap values are indicated at the nodes: 72 (theta/alpha), 53 (theta/alpha/zyta), 61 (alpha/zyta), 40 (alpha/zyta/mu), 42 (mu/myoglobin), 100 (gamma(2)/gamma), 100 (epsilon/delta/beta), 55 (epsilon/delta/beta), 100 (epsilon/delta/beta).</p>	
<p>Tamura-Nei model</p>	 <p>Phylogenetic tree showing relationships between cytoglobin, myoglobin, and neuroglobin. The tree is rooted at the bottom left. A scale bar of 0.20 is shown at the bottom left. Bootstrap values are indicated at the nodes: 68 (theta/alpha), 48 (theta/alpha/zyta), 54 (alpha/zyta), 33 (alpha/zyta/mu), 34 (mu/myoglobin), 100 (gamma(2)/gamma), 100 (epsilon/delta/beta), 65 (epsilon/delta/beta), 100 (epsilon/delta/beta).</p>	<p>20%的重复失败率，重设重复次数300</p>
<p>Maximum Composite Likelihood</p>	 <p>Phylogenetic tree showing relationships between cytoglobin, myoglobin, and neuroglobin. The tree is rooted at the bottom left. A scale bar of 0.20 is shown at the bottom left. Bootstrap values are indicated at the nodes: 98 (delta/beta), 68 (delta/beta/epsilon), 100 (delta/beta/epsilon/gamma(2)), 43 (delta/beta/epsilon/gamma(2)/gamma), 53 (gamma/myoglobin), 73 (zeta/alpha), 48 (zeta/alpha/theta), 75 (zeta/alpha/theta/cytoglobin), mu/neuroglobin.</p>	



附：默认重复次数 100

选择不同核苷酸替换模型，比较所构建的系统发生树的拓扑结构和稳定性值明显差异，Jukes-Cantor model、Kimura 2-parameter model、Tajima-Nei model 与 Maximum Composite Likelihood 这 4 个模型所得拓扑结构相近，参考蛋白质序列建树结果及课程文献，与实际进化规律不符；p-distance、Tamura 3-parameter model、Tamura-Nei model 及 LogDet(Tamura-Kumar)这 4 个模型所得拓扑结构相近，参考蛋白质序列建树结果及课程文献与实际进化规律相符，但稳定性值较差，综合蛋白质建树结果，p-distance 模型为较为稳定的模型。

建树报错时，树的拓扑结构本身不一定是错的，但 **Bootstrap** 支持值的可信度会下降，需要谨慎解读。

3) 根据所构建的系统发生树，参阅 **Burmester** 和 **Hardison** 论文，说明人珠蛋白基因家族 12 个成员之间的演化关系。

人珠蛋白基因家族 12 个成员的演化关系可简略总结为：

最古老分支：神经珠蛋白（Neuroglobin）最先分化，与其他成员亲缘关系最远。

早期分化分支：肌红蛋白（Myoglobin）与胞红蛋白（Cytoglobin）聚为一支，是次早分化的谱系。

核心血红蛋白两大支：其余成员分化为 **β-** 样与 **α-** 样两大分支：

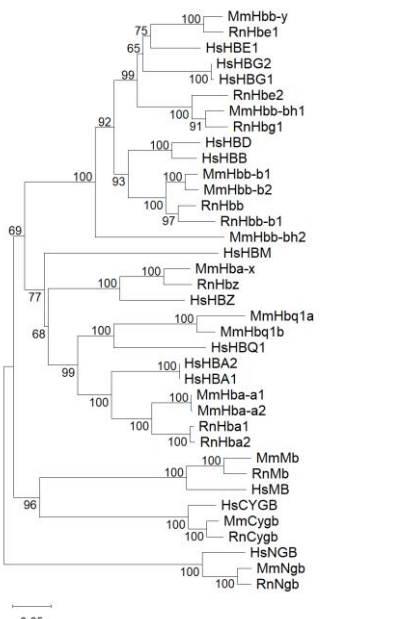
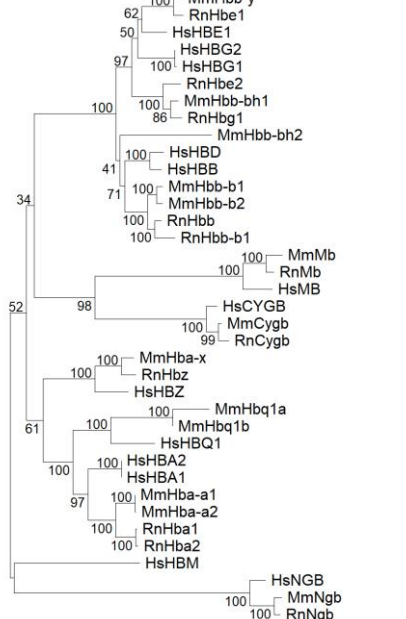
β- 样分支：Gamma1/Gamma2、Epsilon、Beta、Delta 亲缘关系最近，按胚胎→胎儿→成人的表达模式，通过多次复制形成层级清晰的演化关系。

α- 样分支：Mu、Zeta、Alpha、Theta 聚为一支，其中 Mu、Zeta 为胚胎型，Alpha 为成人型，Theta 为假基因，反映了复制与功能丢失事件。

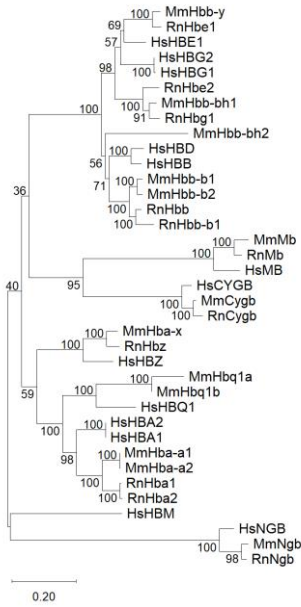
4. 人、小鼠和大鼠三个物种珠蛋白家族系统发生树实例

1) 以人、小鼠和大鼠三个物种珠蛋白家族 37 个成员编码区序列，采用邻接法、最大简约法和最大似然法构建系统发育树，选择适当的替换模型和参数，比较采用不同方法、不同模型和不同参数时所构建的系统发生树的拓扑结构和稳定性值。

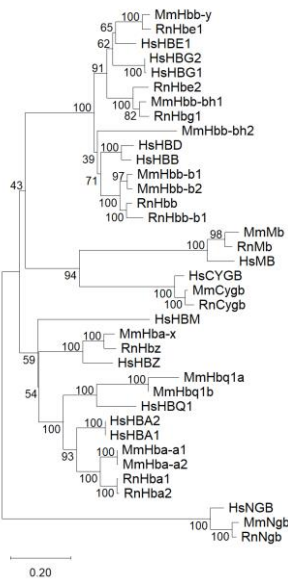
(1) 邻接法

Modle/methods	结果	
<p>P-distance</p>		<p>最佳拓扑结构与稳定性值</p>
<p>Jukes-cantor</p>		

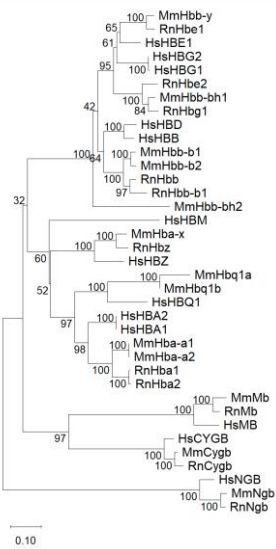
Kimura 2-parameter



Tajima-nei



Kimura 3-parameter

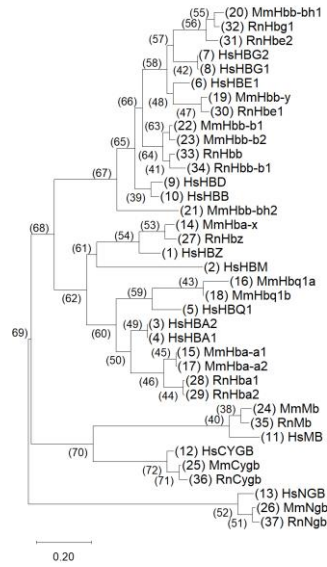


<p>Kimura -nei</p>		
<p>Maximum composite</p>		
<p>Logdet</p>	<p>报错</p>	

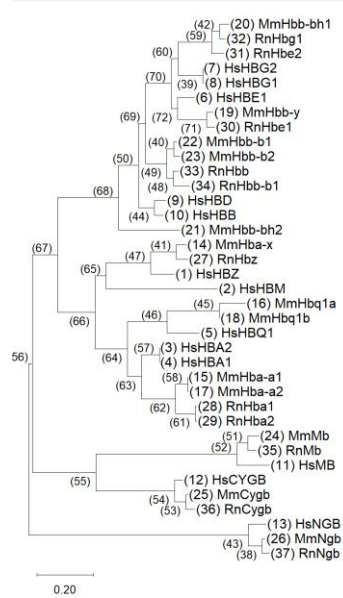
(2) 最大似然法

Modle/methods	结果	
<p>Jukes-cantor</p>		

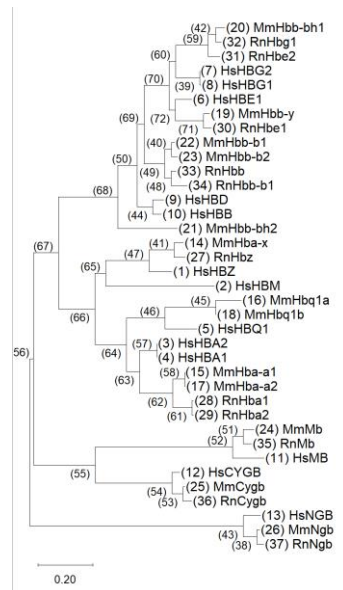
Kimura 2-parameter



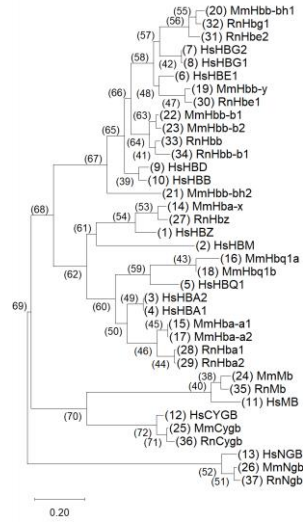
Tamura 3-parameter



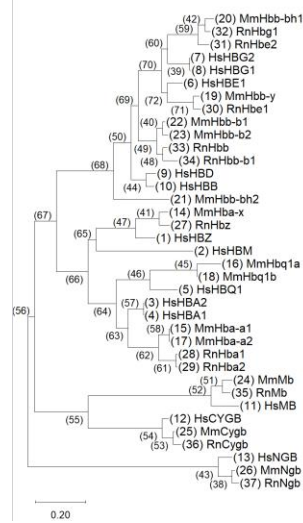
Hasegawa-Kishino



Tamura-nei

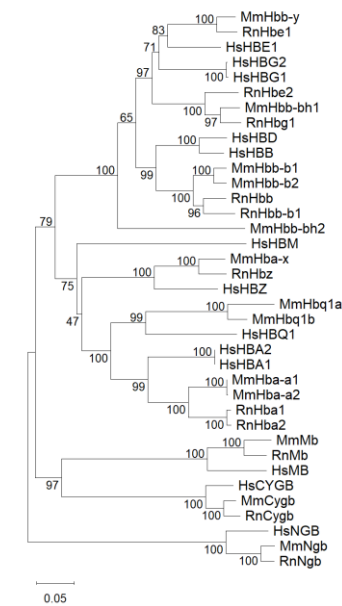


General time reversible

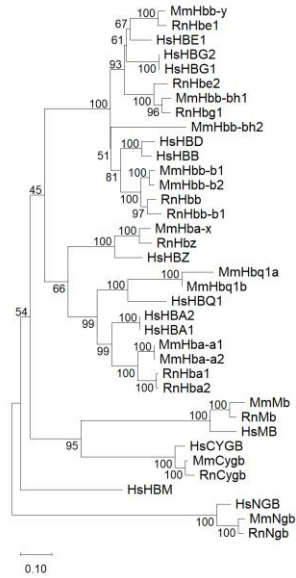


(3) 最大简约法

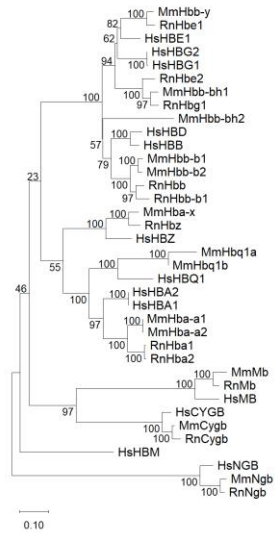
P-distance



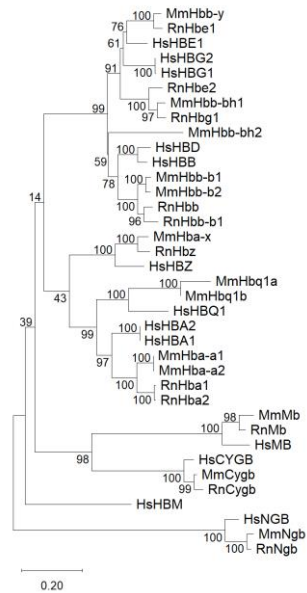
Jukes-cantor



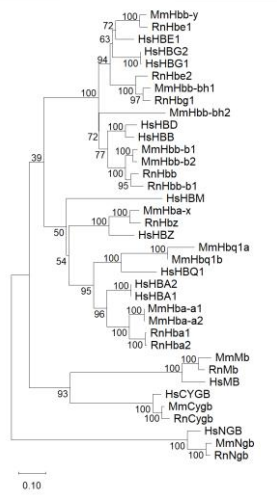
Kimura 2-parameter



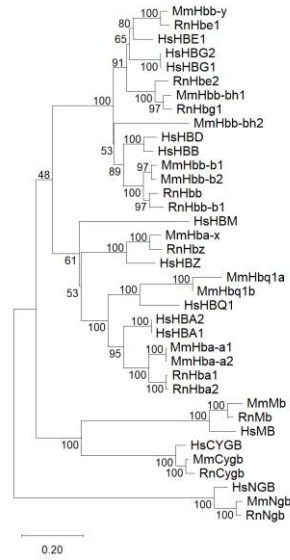
Tajima-nei



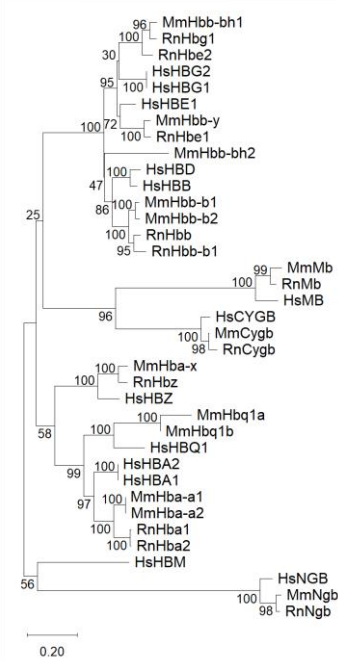
Kimura 3-parameter



Kimura -nei



Maximum composite



2) 根据上述人、小鼠和大鼠三个物种珠蛋白家族 37 个成员编码区序列系统发生树，参阅相关文献，说明珠蛋白基因家族的起源和演化。

(1) 系统发生树的基本结构

肌红蛋白 (Myoglobin, MB) 分支：最古老，与其它珠蛋白分化最早。

神经珠蛋白 (Neuroglobin, NGB) 与细胞珠蛋白 (Cytooglobin, CYGB) 分支：二者互为姐妹群，分化时间早于血红蛋白家族。

血红蛋白 (Hemoglobin) 分支：进一步分为 α -类 (包括 α 、 ζ 、 θ 等) 和 β -类 (包括 β 、 γ 、 δ 、 ϵ 等)，分别形成两个亚簇。

各物种的直系同源基因 (如人的 HBA1 与小鼠的 Hba-a1) 在树上彼此相邻，而旁系同源基因 (如人的 HBA 与 HBB) 则分别位于不同亚簇，表明基因复制事件发生在物种分化之前。

(2) 物种间对比

每个物种的 α -类成员与 β -类成员均各自聚类，表明 α/β 分化早于哺乳动物辐射。

同一簇内的直系同源基因 (如人 HBA1 vs 小鼠 Hba-a1) 的序列差异小于旁系同源基因 (如人 HBA1 vs 人 HBB)，符合物种演化时间。

小鼠和大鼠的珠蛋白成员数量相近，且存在一些啮齿类特有的复制 (如大鼠的 β -珠蛋白假基因)，反映了谱系特异性演化。

Mega 软件进一步功能探索

图标名称	功能说明
ALIGN	多序列比对工具，内置 ClustalW 和 MUSCLE 算法，支持 DNA / 蛋白序列比对，也能手动编辑比对结果，是后续所有分析的基础。

图标名称	功能说明
DATA	数据管理中心，可导入 / 导出序列文件、筛选序列、处理数据集（比如去除冗余序列、定义分组），也支持序列格式转换。
MODELS	进化模型选择工具，能根据序列数据自动筛选最优的核苷酸 / 氨基酸替换模型（如 GTR、HKY、JTT 等），为后续建树和距离计算提供参数基础。
DISTANCE	遗传距离计算模块，可计算序列间的进化距离（如 p-distance、Kimura 2-parameter 距离等），也支持构建距离矩阵，是邻接法建树的前置步骤。
DIVERSITY	遗传多样性分析工具，用于计算核苷酸多态性 (π)、单倍型多样性、群体内 / 群体间遗传分化等群体遗传学指标。
PHYLOGENY	系统发育树构建模块，支持邻接法 (NJ)、最大似然法 (ML)、最大简约法 (MP) 等多种建树算法，还能进行 bootstrap 检验评估树的可靠性。
USER TREE	用户树分析工具，可导入外部 Newick 格式的系统发育树，基于已有树进行后续分析（如祖先序列重建、选择压力分析）。
ANCESTORS	祖先序列重建工具，利用最大似然法或简约法推断序列的共同祖先序列，可用于研究基因 / 蛋白的进化历史。
SELECTION	选择压力分析模块，通过计算非同义替换率 (dN) 与同义替换率 (dS) 的比值 ($\omega=dN/dS$)，检测序列中是否存在正选择、负选择或中性进化位点。
RATES	进化速率分析工具，可计算不同分支、不同位点的进化速率差异，也支持检测分子进化速率的异质性。
CLOCKS	分子钟分析模块，支持严格分子钟、宽松分子钟模型，可结合化石校准数据估算物种分化时间（如 RelTime 方法）。
DIAGNOSE	数据诊断工具，用于检测序列数据中的潜在问题（如重组、碱基组成偏倚、序列饱和等），帮助评估数据质量。

1.序列预处理与管理

多序列比对（支持 ClustalW/MUSCLE 算法）及手动编辑

序列格式转换（FASTA、MEGA、PHYLIP 等）、数据筛选与分组

支持从 ABI 测序文件中直接读取序列，也能对接 BLAST 进行同源序列搜索

2.进化模型选择与距离计算

自动筛选最优替换模型，为后续分析提供可靠参数

计算多种遗传距离矩阵，支持距离矩阵的导出与可视化

3.群体遗传学分析

计算核苷酸多态性 (π)、单倍型多样性、Tajima's D 检验等群体遗传指标

分析群体内 / 群体间的遗传分化与基因流

4.选择压力检测

计算 dN/dS 比值，检测正选择位点（如分支位点模型、位点模型）

识别适应性进化的基因 / 位点，为功能进化研究提供依据

5.祖先序列重建与进化历史推断

重建基因 / 蛋白的祖先序列，追溯序列的进化轨迹

结合分子钟模型估算物种 / 基因的分化时间

6.数据质量诊断

检测序列数据中的重组事件、碱基组成偏倚、序列饱和等问题

评估数据是否适合后续的系统发育分析

5. 血红蛋白 alpha 亚基系统发生树实例

1) 从脊椎动物中选取若干代表性物种，根据传统分类学知识，描述它们之间的系统发生关系，用 MEGA 软件中 User Tree/Display Newick trees 绘制系统发生树。

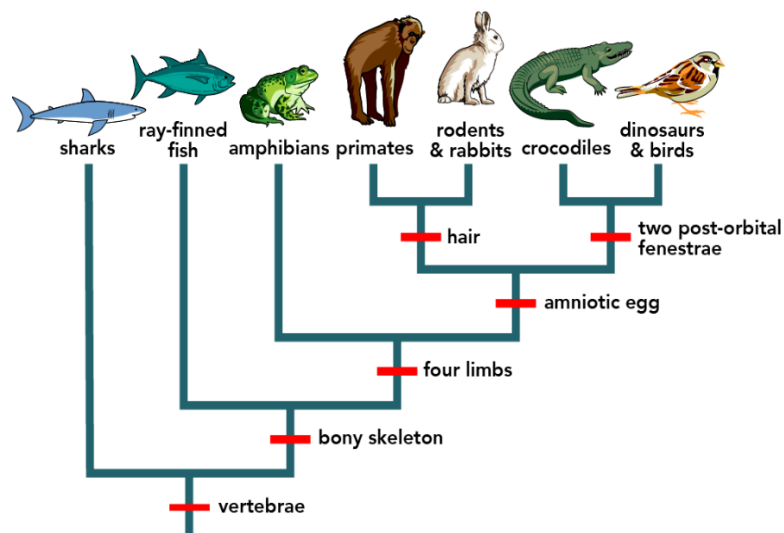
2)从 UniProt 中提取上述代表性物种的血红蛋白 alpha 亚基氨基酸序列，用

MEGA 构建系统发生树； 选择适当的方法、模型和参数，以获得稳定性较好的系统发生树。比较所构建的系统发生树与基于传统分类学的系统发生树的异同。

①EHA101, 网址: <https://evolution.berkeley.edu/evolution-101/>

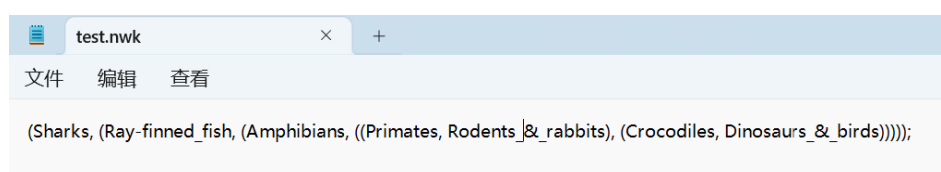
找到 The history of life: looking at the patterns 版块，为七种代表性的脊椎动物，从左至右依次为：鲨鱼、辐鳍鱼类（硬骨鱼）、两栖动物、灵长类、啮齿类与兔类、鳄鱼、恐龙与鸟类。

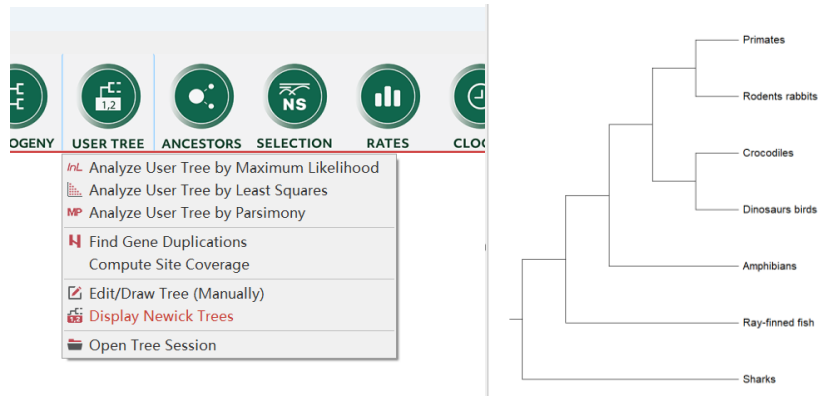
从系统发育树基部梳理这些脊椎动物的演化关系：所有类群共享脊椎特征，构成脊椎动物亚门，软骨鱼纲的鲨鱼是该演化支最早分化的类群，保留软骨骨骼，与后续硬骨脊椎动物形成明确分支。硬骨骨骼出现后，辐鳍鱼纲率先分化，作为多样性最高的水生类群，以鳍条支撑的鱼鳍为运动器官，代表硬骨脊椎动物早期水生演化分支。随后四肢结构演化形成四足总纲，两栖纲是此支最早分化类群，初步适应陆地却依赖水环境繁殖，幼体需水中发育，是水生向陆生过渡的代表。两栖类之后，羊膜卵的出现让脊椎动物摆脱繁殖对水环境的依赖，形成羊膜动物并分化为两大分支：其一为以毛发为关键特征的合弓纲，包含灵长类、啮齿类与兔形目在内的哺乳动物均属此支，具备恒温、胎生、哺乳等特征；其二是以双颞孔为特征的蜥形纲，鳄鱼与恐龙、鸟类同属此支，鸟类是兽脚类恐龙后代，系统发生上属于恐龙演化支延续，与鳄鱼亲缘关系更近。整体来看，这些脊椎动物演化呈现从水生到陆生、从低等到高等的脉络，亲缘关系随演化分支分化愈发清晰。



②用 MEGA 软件中 User Tree/Display Newick trees 绘制系统发生树

Newick 文件





③从 UniProt 数据库中找到上述 7 个代表性脊椎动物 alpha 血红蛋白。

④方法、模型与参数选择

My Basket			
UniProtKB (7)			
Entry	Entry Name	Organism	
P69905	HBA_HUMAN	Homo sapiens (Human)	
P01942	HBA_MOUSE	Mus musculus (Mouse)	
P01994	HBA_CHICK	Gallus gallus (Chicken)	
Q90487	HBA_DANRE	Danio rerio (Zebrafish) (Brachydanio rerio)	
P07428	HBA_XENTR	Xenopus tropicalis (Western clawed frog) (Silurana tropicalis)	
P01999	HBA_ALLMI	Alligator mississippiensis (American alligator)	
Q9YGW2	HBA_MUSGR	Mustelus griseus (Spotless smooth-hound)	

建树方法：邻接法（Neighbor-Joining, NJ）

进化模型：软件默认蛋白质替换模型

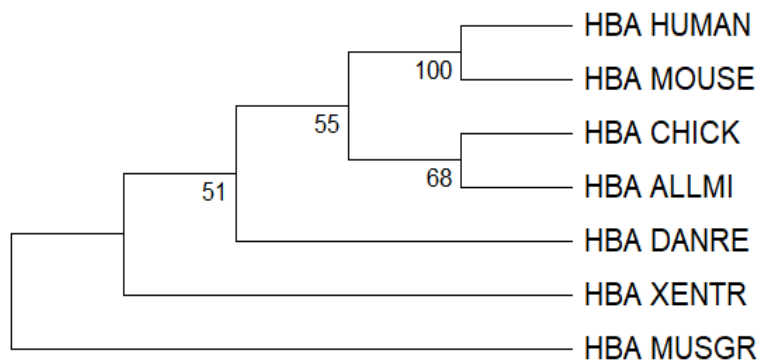
稳定性检验：Bootstrap 自举法

重复次数：100 次

空位与缺失数据：默认处理方式

序列类型：蛋白质序列

⑤MEGA 构建系统发生树



⑥比较所构建的系统发生树与基于传统分类学的系统发生树的异同。

本次基于 α - 血红蛋白氨基酸序列构建的分子系统发生树，与传统分类学系统发生树整体拓扑结构高度一致，均能准确反映 7 种代表性脊椎动物的演化关系。二者均呈现出从原始到高等、从水生到陆生的分化顺序，即无颌类位于基部，硬骨鱼、两栖类、鸟类、哺乳类依次分化，人、小鼠、大鼠等哺乳动物聚为单系群，分支格局与传统分类确立的脊椎动物演化框架完全吻合。

主要差异体现在构建依据与分析精度上。传统分类学以形态解剖、胚胎发育等表型特征为依据，侧重宏观层级划分；分子系统发生树以基因序列差异为基础，客观性更强，可通过自举值量化分支可靠性，对近缘物种的亲缘关系判定更精细，分支长度能直观反映遗传分化程度。

整体而言，两棵树相互印证、彼此支持，分子树为传统分类学提供了直接的分子演化证据，也进一步验证了本次建树方法与参数的合理性。

Q1: 问题:当使用 MEGA 构建系统发生树时，应该怎样选择模型得到可用结果？若以经验选择模型后所得树的分支结构不稳定是应当放弃数据还是更改参数？

建议:当经验值不足时，模型选择可尝试 MEGA 自带的「Model Selection」工具（具体操作可询问 AI）

Q2: MEGA 支持同时构建多个不同的系统进化树吗？

尝试:在已打开第一组数据的基础上打开第二个数据时 MEGA 弹窗默认保留到本地，不同时兼容 2 组数据同时运行。

建议:当 MEGA 操作内容或界面使用存疑时，可在线点击右上 Help→Ask MEGA-GPT→展开详细询问（英语），但需注意该闻讯界面需要在线登录微软

账户，国内 AI 模型在精准语令下可基本代替该功能

The image shows the MEGA-GPT AI chat interface. On the left is a navigation sidebar with options: 新聊天, 搜索聊天, 图片, 应用, 深度研究, 查看套餐和定价, 设置, 帮助. Below these is a section for personalized responses with a '登录' button. The main area features the MEGA-GPT logo and a description: 'MEGA-GPT provides assistance to users of MEGA (Molecular Evolutionary Genetics Analysis) Software. If you use MEGA-G your research please consider citing our paper: Allard JB, Kun Mol Biol Evol. 2025;42(6) https://doi.org/10.1093/molbev/ms...'. A login overlay is active, showing the user 'weihui chen' and a '继续以“weihui”的身份登录' button. Below the overlay are four question prompts: 'I want to create a Maximum Likelihood tree', 'How can I align protein sequences?', 'What is new in MEGA 12.1?', and 'How can I estimate divergence times?'. At the bottom is a text input field with a '+' icon and a '语音' button. A footer note states: '向 AI 聊天机器人 ChatGPT 发送消息即表示，你同意我们的条款并阅读我们的隐私政策。'