

“实用生物信息技术”课程小组讨论总结报告

组： G4 次： R1 组长： G4A 高嘉禾 执笔： G4A 高嘉禾

1. 时间：2026年3月18日 18:30
2. 方式：线上（腾讯会议，微信）
3. 主题：熟悉 NCBI、EBI 和 CNCB 等网站，熟练进行文献检索，完成课外练习 1，巩固学习成果
4. 内容：
 - A. 熟悉 NCBI、EBI 和 CNCB 等数据库的使用方法；
 - B. 小组成员分别介绍如何使用 NCBI、EBI 和 CNCB 等网站中的不同功能搜索与自己研究方向相关的基因、蛋白和菌种；
 - C. 小组成员分别介绍*课外练习 1 中相应部分；
 - D. 小组成员分别说明在使用上出现的问题并讨论解决问题。

*课外练习 1:

（一）生物大分子月报：[PDB-101: Molecule of the Month: Lenacapavir](#)

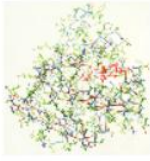
1. 按功能分类，该网站的生物大分子分为哪几大类？

- (1) Health and Disease;
- (2) Molecules of Life;
- (3) Biotech and Nanotech;
- (4) Structures and Structure Determination

The screenshot displays the PDB-101 website interface, showing four main categories of biological molecules and their associated articles:

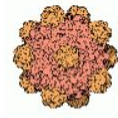
- Health and Disease:** Antimicrobial Resistance (10), Cancer (40), Coronavirus (7), Diabetes (10), Drug Action (36), Drugs and the Brain (13), HIV and AIDS (9), Immune System (26), Infectious Disease (14), Peak Performance (38), Toxins and Poisons (18).
- Molecules of Life:** Biological Energy (32), Biology of Plants (27), Bioluminescence and Fluorescence (3), Cellular Signaling (56), Central Dogma (13), Enzymes (72), Molecular Evolution (17), Molecular Infrastructure (14), Molecular Motors (8), Molecules for a Sustainable Future (13), Nucleic Acids (23).
- Biotech and Nanotech:** Biotechnology (18), Nanotechnology (20), Recombinant DNA (5), Renewable Energy (4).
- Structures and Structure Determination:** Biomolecular Structural Biology (19), Biomolecules (14), Integrative/Hybrid Methods (10), Nobel Prizes and PDB structures (31), PDB Data (1), Protein Structure Prediction, Design, and Computed Structure Models (3).

2. 2000 年 12 篇月报中描述的蛋白质分子，你熟悉的有哪些？



January 2000

- (1) **Myoglobin** 肌红蛋白：在肌肉细胞中储存氧气。



February 2000

- (2) **Bacteriophage phiX174** 噬菌体 phiX174：复制病毒。



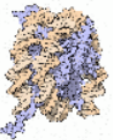
March 2000

- (3) **DNA Polymerase** DNA 聚合酶：能够准确复制细胞的基因组。



April 2000

- (4) **Collagen** 胶原蛋白：为身体提供结构。



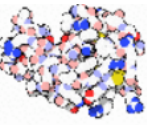
July 2000

- (5) **Nucleosome** 核小体：细胞的基因组由核小体储存和保护。



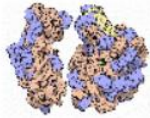
August 2000

- (6) **Restriction Enzymes** 限制性内切酶：分子剪刀，是基因工程的有用工具。



September 2000

- (7) **Lysozyme** 溶菌酶：攻击细菌的细胞壁。



October 2000

(8) **Ribosomal Subunits** 核糖体亚基：揭示了 RNA 在蛋白质合成中的核心作用。

3. 有关 DNA、rRNA、tRNA 的月报有几篇？

DNA: 39 篇

rRNA: 1 篇

tRNA: 15 篇

4. 该网站中和 DNA 复制相关的蛋白质分子有哪些？

- (1) DNA 聚合酶：DNA 聚合酶能够准确复制细胞的基因组；
- (2) 滑动夹：滑动夹沿 DNA 链滑动，在复制过程中保持 DNA 聚合酶的轨道；
- (3) DNA 连接酶：DNA 连接酶能重新连接断裂的 DNA 链，并用于工程重组 DNA 的制造
- (4) DNA 解旋酶：DNA 解旋酶通过 DNA 双螺旋分离两条链，由 ATP 驱动

5. 该网站中介绍的病毒有哪些？

- (1) Zika Virus (寨卡病毒)
- (2) Dengue Virus (登革热病毒)
- (3) Tobacco Mosaic Virus (烟草花叶病毒)
- (4) Simian Virus 40 (猿猴病毒 40)
- (5) Eastern Equine Encephalitis Virus (东部马脑炎病毒)
- (6) Adenovirus (腺病毒)
- (7) Poliovirus and Rhinovirus (脊髓灰质小儿麻痹症与鼻病毒)
- (8) Parvoviruses (细小病毒)
- (9) Human Papillomavirus (人类乳突病毒)
- (10) Measles Virus (麻疹病毒)
- (11) Ebola Virus (埃博拉病毒)
- (12) Hepatitis C Virus (丙型肝炎病毒)
- (13) Coronavirus (冠状病毒)

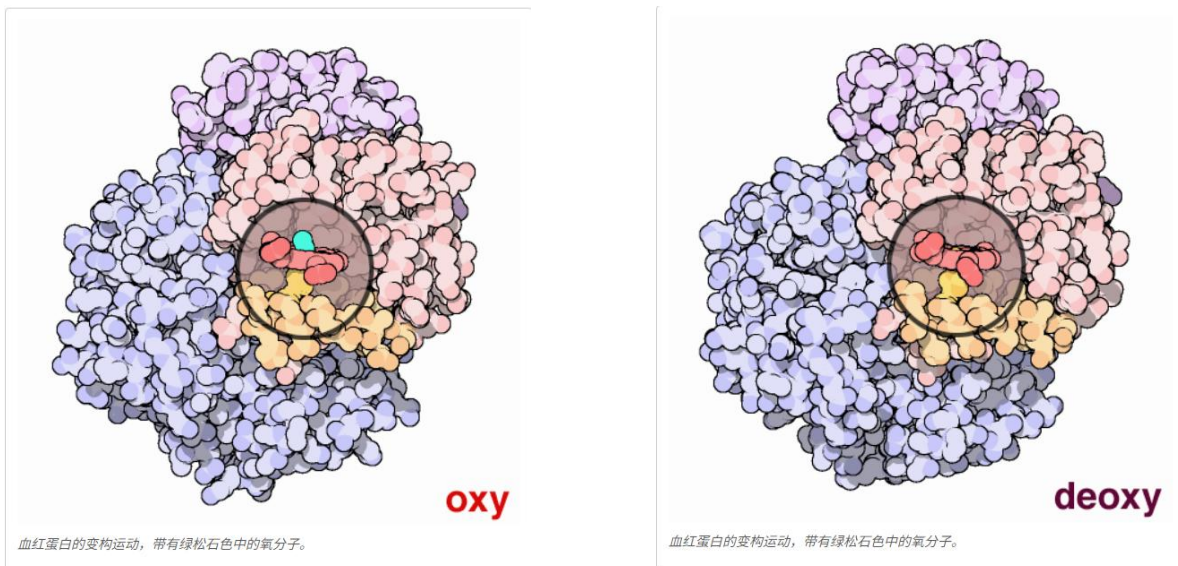
6. 该网站中你熟悉的酶分子有哪些？

- (1) Glycolytic Enzymes (糖解酶)：糖酵解的十种酶分解我们饮食中的糖分。
- (2) DNA Methyltransferases (DNA 甲基转移酶)：细胞会向 DNA 中添加甲基，以编码额外的表观遗传信息。
- (3) Pepsin (胃蛋白酶)：胃蛋白酶分解强胃酸中的蛋白质。
- (4) DNA Ligase (DNA 连接酶)：DNA 连接酶能重新连接断裂的 DNA 链，并用于工程重组 DNA 的制造。

- (5) Catalase (过氧化氢酶) 过氧化氢酶保护我们免受危险的反应性氧化分子侵害。
- (6) Acetylcholinesterase (乙酰胆碱酯酶): 阻断神经细胞与肌肉细胞之间的信号。
- (7) Alcohol Dehydrogenase (酒精脱氢酶): 解毒人体饮用的酒精。
- (8) Alpha-amylase (α -淀粉酶): 分解淀粉以产生葡萄糖。
- (9) Beta-galactosidase (β -半乳糖苷酶): 分解乳糖。
- (10) Restriction Enzymes (限制性内切酶): 分子剪刀, 精准切割特定 DNA 片段。
- (11) RNA Polymerase (RNA 聚合酶): DNA 转录关键酶。

7. 阅读该网站关于血红蛋白的短文, 说明其结合和释放氧气时不同亚基之间如何协同作用。

一个亚基的血红素基团 (小圆形窗口中显示) 被固定在一个位置, 以便你看到蛋白质在氧结合时如何绕行。氧分子以蓝绿色表示。当它与血红素中心的铁原子结合时, 会拉动血红素底部的组氨酸。这会改变整个 α 螺旋的位置, 这里以橙色显示在血红素下方。这种运动会沿着蛋白质链传递到其他链, 最终导致蓝色显示的两个亚基产生巨大的摇晃运动。



8. 阅读该网站关于流感病毒神经氨酸酶 (Influenza Neuraminidase) 的短文, 说明其结构特征和作用机制。

[PDB-101: Molecule of the Month: Influenza Neuraminidase](#)

结构特征: 神经氨酸酶由四个相同的亚基组成, 形成一个方形的四聚体。它通过一个长的蛋白质“茎”附着在病毒表面。酶的活性位点位于每个亚基上表面的一个深凹陷 (deep depression) 中, 这里是它与底物结合并催化反应的地方。酶的表面像人类细胞表面蛋白一样, 装饰着多条多糖链。

作用机制: 神经氨酸酶的主要作用发生在病毒生命周期后期, 确保新复制的病毒颗粒能够高效传播。

(1) 促进病毒释放 (核心功能): 当新的流感病毒从被感染的细胞出芽后, 会被细胞表面的唾液酸 (一种糖分子) 缠住。神经氨酸酶就像一个“分子剪刀”, 通过其活性位点剪断 (切割) 多糖链末端的糖分子, 从而切断病毒与受感染细胞之间的联系, 防止病毒聚集在细胞表面, 使它们能够自由地去感染新的细胞。

(2) 辅助病毒入侵: 此外, 神经氨酸酶还能帮助病毒穿过呼吸道表面的粘液层, 从而顺利到达并入侵目标细胞。

9. 列举该网站中与你研究方向相关的生物大分子。

DNA: [PDB-101: Molecule of the Month: DNA](#)

Insulin: [PDB-101: Molecule of the Month: Insulin](#)

10. 阅读你最感兴趣的生物大分子，阐述该分子的结构、功能特点。

DNA 的结构：双螺旋结构，由两条线性多核苷酸链组成，它们围绕一个共同的轴旋转，两条链反向平行，糖-磷酸骨架位于螺旋外侧，而碱基对在内部。双螺旋的直径约为 2 nm，相邻碱基对之间的垂直距离为 0.34nm，每圈螺旋大约包含 10.5 个碱基对，因此螺旋的螺距（完整一圈的轴向长度）约为 3.6 nm。遗传信息存储在碱基序列中（A、T、C、G）。通过一条链上的碱基与另一条链上的互补碱基形成氢键（A 与 T 配对，G 与 C 配对），确保了遗传信息的精确传递。在螺旋表面暴露的大沟和小沟中，不同的碱基对会呈现出特征性的化学基团排列模式。蛋白质无需解开双螺旋，就能通过识别这些模式来“读取”DNA 的遗传信息，从而调控基因的表达。

DNA 的功能：信息存储（只读存储器）、信息传递即中心法则、重组与修复。

（二）化学小分子月报（Molecular of the Month）：[The Molecule of the Month home page](#)



1. 与生物大分子月报相比，该网站的文章内容和写作风格有何特点？

生物大分子月报: [PDB-101: Molecule of the Month: Lenacapavir](#)

化学小分子月报: [The Molecule of the Month home page](#)

内容上:

生物大分子月报: 侧重结构决定机制。重点描述蛋白质折叠、结构域组装、活性位点构象变化，以及突变对功能的直接影响；化学小分子月报: 应侧重结合决定功能。重点描述小分子的化学结构特征（官能团、立体化学）、与大分子结合的相互作用力（氢键、疏水作用、范德华力），以及构效关系（Structure-Activity Relationship, SAR）分析。

写作风格: 生物大分子月报更学术严肃；化学小分子月报的风格更轻松跳脱。

2. 如何查看某个特定分子？

See [here](#) for a list of the awards and citations received by this page and a disclaimer.

Congo Red

Above is a drop-down alphabetical list of all the molecules. Click on your molecule of choice to take you directly to it in the chronological table below.

在下拉界面（按 A-Z 排序）找到需要的分子，可跳转到相应的条目，点击 html 即可跳转到相应的界面；或者直接拉动页面下滑按年份找到需要的分子。

3. 该网站收录的维生素有哪些？各有什么功能？

目前有维生素 A B1 B2 B3 B12 C D E K,共 9 种：

<https://www.chm.bris.ac.uk/motm/vitaminA/vitaminah.htm>

<https://www.chm.bris.ac.uk/motm/vitaminB1/vitaminb1h.htm>

<https://www.chm.bris.ac.uk/motm/vitaminB2/vitaminb2h.htm>

<https://www.chm.bris.ac.uk/motm/vitaminB3/vitaminb3h.htm>

<https://www.chm.bris.ac.uk/motm/vitaminb12/>

<https://www.chm.bris.ac.uk/motm/vitaminD/vitaminsdh.htm>

<https://www.chm.bris.ac.uk/motm/vitaminC/vitaminsch.htm>

<https://www.chm.bris.ac.uk/motm/vitaminE/vitaminsch.htm>

<https://www.chm.bris.ac.uk/motm/vitaminK/vitaminskh.htm>

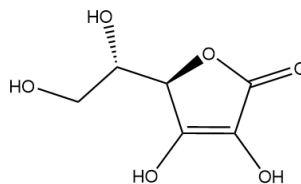
- 维生素 A：构成视觉色素（视紫红质）的核心成分。它在视网膜中作为光敏开关，通过吸收光能后发生构型变化（从 11-顺式变为全反式），触发视蛋白构象改变，从而产生神经信号，使人眼能够感知光线（尤其是夜视）。此外，其衍生物视黄酸还参与调节皮肤免疫功能和抗炎反应。
- 维生素 B1 (硫胺素)：作为关键辅酶（常转化为焦磷酸硫胺素，TPP）参与能量代谢，核心是催化脱羧反应。它帮助分解糖类和碳水化合物，释放能量供神经功能、肌肉张力等生命活动使用，对维持神经细胞髓鞘的构建也至关重要。
- 维生素 B2 (核黄素)：是两种重要辅酶——黄素单核苷酸（FMN）和黄素腺嘌呤二核苷酸（FAD）的前体。FAD 能灵活接受或提供电子，作为“ helper 分子”广泛催化体内的氧化还原反应，在脂肪、碳水化合物和蛋白质的能量代谢中起核心作用。
- 维生素 B3 (烟酸)：是两种关键辅酶——烟酰胺腺嘌呤二核苷酸（NAD⁺）和烟酰胺腺嘌呤二核苷酸磷酸（NADP⁺）的前体。NAD 主要参与分解代谢（如糖、脂肪、蛋白质分解以释放能量），而 NADP 则参与合成代谢（如合成脂肪酸和胆固醇），两者都是氧化还原反应的核心电子载体。
- 维生素 B12：（基于页面信息和通用知识）作为辅酶参与两类重要反应：一是催化分子内重排（如将甲基丙二酰辅酶 A 转化为琥珀酰辅酶 A，参与氨基酸和脂肪酸代谢）；二是在蛋氨酸合成循环中作为甲基载体，对维持神经系统健康和红细胞生成至关重要。其分子核心的钴原子是催化活性的关键。
- 维生素 C (抗坏血酸)：具有双重核心功能：1) 强效抗氧化剂，通过稳定地提供电子清除有害的自由基；2) 辅助酶促反应，作为辅因子维持胶原蛋白合成酶（脯氨酰羟化酶）的活性，确保结缔组织（皮肤、骨骼、血管）的完整性，并参与神经递质（如多巴胺）的合成。
- 维生素 D：作为激素前体，在体内经肝脏和肾脏两步羟化后转化为活性形式（骨化三醇）。它作为激素调节钙磷代谢，通过促进肠道对钙的吸收、维持血钙平衡，从而支持骨骼的正常矿化与重塑。同时，它也参与调节免疫功能。
- 维生素 E (α -生育酚)：最主要的角色是脂溶性抗氧化剂。它嵌入细胞膜中，利用其酚羟基捕获并

淬灭攻击膜脂质的自由基（过氧化物），中断脂质过氧化链式反应，保护细胞膜免受氧化损伤。其抗氧化功能可被维生素 C 等协同再生。

- 维生素 K： 作为辅酶参与“维生素 K 循环”，驱动特定蛋白质（如凝血因子凝血酶原）的谷氨酸残基发生 γ -羧化。这一修饰使这些蛋白能够结合钙离子，从而锚定在磷脂膜表面，启动并维持血液凝固级联反应。同时，此过程也参与骨骼代谢。

4. 阅读维生素 C 的文章，说明其分子结构、功能，及其发现过程、后续研究。

[Vitamin C - Molecule of the Month - July 2017 - HTML-only version](#)



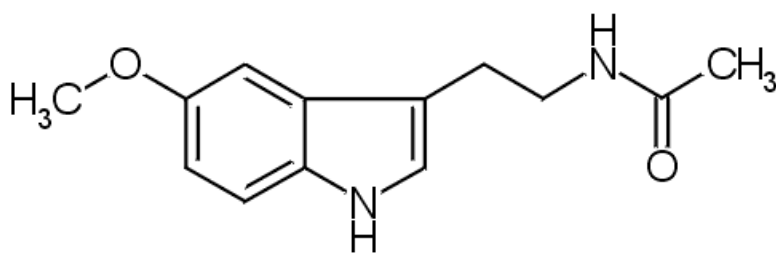
结构：Vitamin C (L-ascorbic acid, L-抗坏血酸)

功能：维生素 C 是一种抗氧化剂，能与自由基反应并清除自由基；另外辅助酶促反应，作为辅因子维持胶原蛋白合成酶（脯氨酰羟化酶）的活性，确保结缔组织（皮肤、骨骼、血管）的完整性，并参与神经递质（如多巴胺）的合成。

发现过程和后续研究：坏血病是一种导致海上死亡人数多于海盗、沉船及其他所有疾病总和的疾病。它会导致痛苦的死亡，起初表现为牙龈出血、牙齿松动和旧伤口重新裂开。当时没有人知道确切的原因，但当人们长时间无法食用新鲜植物食物或动物肉类时，就会发生这种情况。1747 年 5 月，船医詹姆斯·林德进行了一个实验，给部分水手提供橘子和青柠，进行对照实验。发现有橘子和青柠的水手没有得坏血病。1932 年，分离出抗坏血病因子，最终在红辣椒中分离出来。

5. 阅读你感兴趣的化学小分子，阐述该分子的结构、功能特点。

褪黑素：[褪黑素 --- Melatonin](#)



5-甲氧基-N-乙酰色胺，是一种重要的激素，在调节神经内分泌系统中发挥作用。它在调节昼夜睡眠-觉醒周期中扮演重要角色。它还控制着新陈代谢、性欲、生殖、食欲、平衡、肌肉协调和免疫系统等基本功能，以抵抗由细菌、病毒、化学污染物和过量自由基活动引发的疾病。它通常在夜间由**松果体**分泌，松果体是哺乳动物眼后脑深部的一个微小腺体，以响应光照水平的环境变化。松果体本身受下丘脑中视交叉上方的成对神经细胞群控制。这些细胞被称为视交叉上核（SCN），其中包含昼夜节律振荡器。每晚，SCN 通过下丘脑和脊髓中的一系列神经元向松果体发送冲动，以刺激褪黑素分泌。SCN 中的计时机制受阳光控制，阳光通过视网膜和视-下丘脑通路到达 SCN。血液中循环的褪黑素含量已被证明在一天中会上升和下降。

(三) 蛋白质分子精选 (Protein Spotlight)

1. 和生物大分子月报相比, 该网站在文章内容、写作风格、数据库链接、文献引用等方面有什么特点?

Protein Spotlight: 聚焦单一蛋白质/家族, 以发现史、科学故事、人物轶事串联功能与结构, 生动幽默、叙事性强, 侧重科普叙事+科学里程碑, 每篇讲透一个蛋白的“前世今生”, 偏好引用里程碑论文 + 经典综述, 聚焦发现、测序、合成、应用等关键节点, 数量精简, 突出历史意义。

生物大分子月报: 按功能/结构分类 (如酶、结构蛋白、能量代谢), 侧重三维结构解析、PDB 数据、分子机制, 严谨客观、结构导向, 语言偏学术化, 聚焦结构细节与机制, 偏百科式、结构导向。引用结构解析、机制研究的前沿文献, 数量更多, 侧重结构生物学最新进展。

2. 阅读该网站关于胰岛素分子的短文, 说明作者将其称作“20 世纪蛋白质分子的代表”的原因。

胰岛素分子开创了蛋白质研究的多个“第一次”, 是现代生物化学与生物技术的里程碑。是 1926 年最早以纯形式结晶的蛋白质之一、是 1955 年首个完全测序的蛋白质, 1958 年首次通过化学合成合成的蛋白质、1979 年则是通过生物技术制造的第一个人类蛋白质。胰岛素测序是自动蛋白质测序的起点, 也是蛋白质测序数据库日益增长的需求的开端。

3. 阅读该网站关于绿色荧光蛋白 (GFP) 短文 (The greenest of us all) 以及相关文献, 说明该蛋白质分子的结构功能特点。

GFP 是一种自发荧光蛋白, 它拥有一种新颖的三维结构: β 罐。整体形状为桶形, 直径 30Å, 长度 40Å, 十一条紧密贴合的 β 板条构成枪管本体, 轴线上有一条 α 螺旋。后者形成一个支架, 蛋白质的发色团被固定在 其上。绿色荧光蛋白能利用水母发光蛋白发出的蓝光产生绿光, 其生色团是一段经过修饰的丝氨酸—酪氨酸—甘氨酸 (Ser Tyr Gly) 序列。只要这段短三肽在少量氧气的作用下形成环状构象, 绿色荧光蛋白就可以发光, 且整个过程既不需要辅助因子, 也不依赖酶类组分。

4. 该网站中你最感兴趣的蛋白质分子有哪些? 简述它们的生物学功能。

Chameleonin 变色蛋白, 是一种球状蛋白质, 末端像天线一样突出。能根据不同的环境条件呈现不同的色调, 低于 18°C 时, 蛋白质的球状实体变形; “天线”承受结构应力, 其形状相应改变。正是这种形状的变化导致变色龙脊柱的颜色从黄褐色变为明亮的橙色。这一现象类似于龙虾的壳在煮沸时从蓝色变成红色。将这种蛋白质涂抹于墙纸上, 超过 18°C 时, 墙纸会发黄光; 低于 18 度时, 会出现鲜艳的橙色。

5. 该网站中与你研究方向相关的蛋白质分子有哪些? 简述它们的生物学功能。

Rubisco, 固定大气中的二氧化碳, 并将其提供给 RuDP, 然后将新链切入两个相同的磷酸甘油酸分子。

甲硫氨酸合酶还原酶 (Mtrr), 它是叶酸代谢与甲硫氨酸循环之间真正的连接纽带, 换句话说: 它与发育密切相关。事实上, 叶酸所携带的甲基会向下游传递, 并最终用于 DNA 甲基化。Mtrr 利用来自叶酸循环的甲基合成甲硫氨酸, 而甲硫氨酸又是 S 腺苷甲硫氨酸 (S adenosylmethionine, 简称 SAM) 的前体。SAM 是众多细胞底物的甲基供体, 底物包括蛋白质、RNA 和 DNA。因此, Mtrr 对叶酸循环与甲硫氨酸循环的正常进行至关重要; 若缺少该酶, 基因表达会受到严重阻碍。

（四）演化生物学网络教程（Evolution 101）

1. 该网站内容和写作风格有什么特点？

网站内容特点：系统完整、由浅入深，从演化基本概念、自然选择、遗传变异、物种形成、系统发育、化石证据到人类演化，层层递进；强调科学证据，用化石、比较解剖、分子生物学、实验观察等多方面证据讲解演化；纠正常见误区，专门解释“演化不是越来越高级”“人不是猴子变的”“演化无目标”等误解；配有大量示意图，用进化树、对比图、流程图降低理解难度；紧扣现代演化理论，包含种群遗传学、中性演化、趋同演化、共演化等核心内容。

写作风格特点：通俗简洁、无专业壁垒，适合中学生与入门者，语言直白；用达尔文雀、桦尺蛾、鲸鱼演化等经典例子故事化讲解；逻辑清晰、重点突出，每部分只讲一个核心知识点；客观严谨，强调演化是事实+理论，不是假说；

2. 下载该网站中 7 个代表性脊椎动物的图片，写出它们的名称和特点。

文昌鱼（Amphioxus/Lancelet）：最原始的脊索动物代表，无真正脊椎，是脊椎动物共同祖先的“活化石”。

七鳃鳗（Lamprey）：无颌、无成对附肢，最原始的脊椎动物之一，保留古老特征。

鲨鱼（Shark）：软骨鱼类，骨骼为软骨，有颌，是最早出现颌骨的脊椎动物类群。

鲈鱼 / 硬骨鱼（Bony fish / Perch）：硬骨鱼类，骨骼钙化，有鳔，是多数现代鱼类的代表。

青蛙（Frog）：两栖动物，水生到陆生过渡，幼体水生、成体可登陆。

蜥蜴（Lizard）：爬行动物，羊膜卵，完全适应陆地生活。

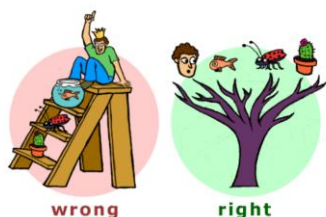
人类（Human）：哺乳类→灵长类，大脑高度发达，直立行走，是演化末端的一个分支，而非“最高级”。

3. 认真阅读分析该网站最后 Big Issue 图片，说明为什么关于演化地位的梯子模型不正确，而树枝模型正确。

梯子模型错误地把演化看成从低等到高等、从简单到复杂、一条直线向上，比如：鱼→两栖→爬行→鸟→哺乳→人，把人放在最顶端，视为“最进化、最高级”。但实际上演化不是直线进行，而是不断分支，现存的生物都是现代物种，不是谁进化成谁。演化也没有“目标”，不是朝着“更高级、更聪明、更大”前进。简单≠低级，很多简单生物高度适应环境，生存极成功。

树枝模型又为什么正确呢？所有生物来自共同祖先，像树枝从树干长出。演化是不断分支，形成树式结构。每个分支都是独立演化，各自适应环境。现存物种都位于树枝顶端，没有谁在谁的“中间阶段”。亲缘关系用共同祖先判断，而不是“高低级”。

总之，梯子模型把演化看作线性、有目标、从低到高等的过程，不符合真实演化历史；树枝（进化树）模型强调共同祖先、分支演化、无高低等级、所有现代生物平等，符合化石、解剖与分子生物学证据，因此是正确的。



（五）PubMed 文献摘要数据库

1. PubMed 生物医学文献摘要数据库和 PubMed Central 全文数据库有哪些主要区别？

PubMed 是包含了生物医学和生命科学文献，收录了 4000 万篇文献，文献没有全文有链接：

PubMed Central 是美国国立卫生研究院国家医学图书馆（NIH/NLM）的免费生物医学和生命科学期刊文献，收录了 1000 万篇有全文的文献。

2. PubMed 医学主题词（Medical Subject Heading, MeSH）包括哪些大类？

进入方式：PubMed – MeSH - NLM MeSH Homepage - MeSH Browser – Tree view

分析、诊断和治疗技术和设备类别、解剖学类别、人类学、教育学、社会学和社会现象类、检查标签类别、化学品和药物类别、学科和职业类别、疾病类别、地理位置类别、医疗保健类别、人文类、信息科学类、生物分类、人员分类、药理作用类别、现象和过程类别、精神病学和心理学类、发布类型类别、副标题类别、技术和食品饮料类。

- Anatomy [A] +
- Organisms [B] +
- Diseases [C] +
- Chemicals and Drugs [D] +
- Analytical, Diagnostic and Therapeutic Techniques, and Equipment [E] +
- Psychiatry and Psychology [F] +
- Phenomena and Processes [G] +
- Disciplines and Occupations [H] +
- Anthropology, Education, Sociology, and Social Phenomena [I] +
- Technology, Industry, and Agriculture [J] +
- Humanities [K] +
- Information Science [L] +
- Named Groups [M] +
- Health Care [N] +
- Publication Characteristics [V] +
- Geographicals [Z] +

3. 生物领域（Biological Science Discipline, H01.158）包括哪些学科？

解剖学、生物化学、生物学、生物物理学、生物技术、时间生物学、老年科学、神经科学、药理学、生理学、毒理学、转化生物科学和生物医学。

- Natural Science Disciplines [H01]
- Biological Science Disciplines [H01.158] -**
- Anatomy [H01.158.100] +
- Biochemistry [H01.158.201] +
- Biology [H01.158.273] +
- Biophysics [H01.158.344] +
- Biotechnology [H01.158.550] +
- Chronobiology Discipline [H01.158.580]
- Geroscience [H01.158.595]
- Neurosciences [H01.158.610] +
- Pharmacology [H01.158.703] +
- Physiology [H01.158.782] +
- Toxicology [H01.158.891] +
- Translational Science, Biomedical [H01.158.945]

4. 生物学学科（Biology, H01.158.273）包括哪些分支？

植物学、细胞生物学、计算生物学、低温生物学、细胞学、发育生物学、生态学、外空生物学、遗传学、糖生物学、实验动物科学、微生物学、自然历史、神经生物学、寄生虫学、光生物学、放射生物学、

社会生物学、合成生物学。

```
Natural Science Disciplines [H01]
  Biological Science Disciplines [H01.158]
    Anatomy [H01.158.100]
    Biochemistry [H01.158.201]
    Biology [H01.158.273]
      Botany [H01.158.273.118]
      Cell Biology [H01.158.273.160]
      Computational Biology [H01.158.273.180]
      Cryobiology [H01.158.273.195]
      Cytology [H01.158.273.198]
      Developmental Biology [H01.158.273.200]
      Ecology [H01.158.273.248]
      Exobiology [H01.158.273.295]
      Genetics [H01.158.273.343]
      Glycobiology [H01.158.273.356]
      Laboratory Animal Science [H01.158.273.368]
      Microbiology [H01.158.273.540]
      Natural History [H01.158.273.602]
      Neurobiology [H01.158.273.610]
      Parasitology [H01.158.273.688]
      Photobiology [H01.158.273.738]
      Radiobiology [H01.158.273.789]
      Sociobiology [H01.158.273.866]
      Synthetic Biology [H01.158.273.904]
      Zoology [H01.158.273.943]
```

5. 遗传学分支（Genetics, H01.158.273.343）包括哪些研究方向？

细胞遗传学、遗传研究、遗传学，行为学、遗传学，微生物学、遗传，种群、基因组学、人类遗传学免疫遗传学、分子生物学、遗传药理学、遗传药理学、表型组学、辐射遗传学、毒理遗传学

```
Natural Science Disciplines [H01]
  Biological Science Disciplines [H01.158]
    Biology [H01.158.273]
      Botany [H01.158.273.118]
      Cell Biology [H01.158.273.160]
      Computational Biology [H01.158.273.180]
      Cryobiology [H01.158.273.195]
      Cytology [H01.158.273.198]
      Developmental Biology [H01.158.273.200]
      Ecology [H01.158.273.248]
      Exobiology [H01.158.273.295]
      Genetics [H01.158.273.343]
        Cytogenetics [H01.158.273.343.180]
        Genetic Research [H01.158.273.343.249]
        Genetics, Behavioral [H01.158.273.343.290]
        Genetics, Microbial [H01.158.273.343.330]
        Genetics, Population [H01.158.273.343.335]
        Genomics [H01.158.273.343.350]
        Human Genetics [H01.158.273.343.385]
        Immunogenetics [H01.158.273.343.420]
        Molecular Biology [H01.158.273.343.595]
        Pharmacogenetics [H01.158.273.343.750]
        Phenomics [H01.158.273.343.775]
        Radiation Genetics [H01.158.273.343.800]
        Toxicogenetics [H01.158.273.343.900]
```

6. 你的研究方向属于哪个大类？哪个领域？哪个学科？哪个分支？

Disciplines and Occupations [H] - Natural Science Disciplines [H01] - Biological Science Disciplines [H01.158] - Biology [H01.158.273] - Computational Biology [H01.158.273.180]

7. 分子生物学（Molecular Biology）同时属于哪几个学科？哪几个分支？包括什么专题研究方向？

[MeSH Browser](#)

同时属于 3 个学科，分别是：

- H01.158.201.636：属于自然科学—生命科学--Biochemistry [H01.158.201]生物化学
- H01.158.273.343.595：属于自然科学—生命科学—生物学 Genetics [H01.158.273.343]遗传学
- H01.181.122.650：属于自然科学—化学—Biochemistry [H01.181.122]生物化学

8. 生物信息学 (Bioinformatics) 在 MeSH 数据库中归哪个主题词? 其研究方向和范围包括哪些?

Bioinformatics, 在 MeSH 数据库中属于 Computational Biology, 归属于自然科学和信息学。

有 2 个研究方向, 分别是:

- H01.158.273.180: 基因组学、免疫信息学、代谢组学、表型组学、系统生物学;
- L01.313.124: 生物信息整理、免疫信息学、多组学

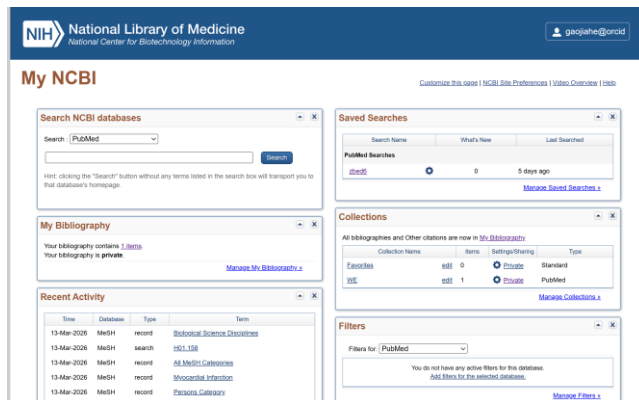
9. 何谓 PubMed 检索 Stop 词? 列举 3-5 个常见 Stop 词。

Stop Words: 也称停用词, 指的是那些本身缺乏特定检索意义, 或因出现过于频繁而被系统自动忽略的单词。如: a, an, the, in, on, at, for, of, and 等

10. 结合 PubMed 网站帮助文档和用户指南, 总结 PubMed 文献检索的策略和经验。

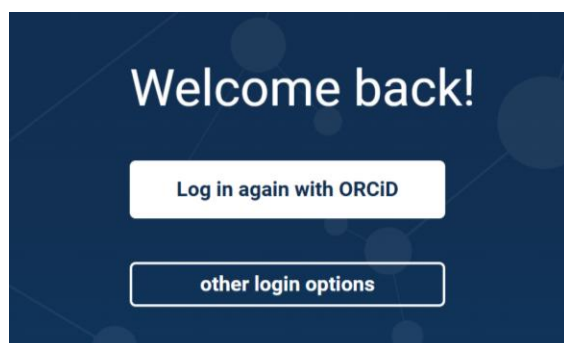
PubMed 文献检索是将关键词与 MeSH 词相结合: 关键词用于初步探索和捕捉最新未标引文献, MeSH 词通过匹配同义词提高查准率, 两者并用可实现查全与查准。在精确检索时, 可用英文双引号固定词组、利用字段标签限定检索范围 (如作者、标题), 并借助 AND、OR、NOT 进行逻辑组合。检索结果出来后, 可利用侧边栏筛选器按文章类型、年份、免费全文等进行过滤, 还可通过检索历史组合多次检索结果以实现更复杂的检索需求。注册 MyNCBI 账号可以保存检索策略、设置自动更新提醒以便持续追踪新文献, 并能创建个人文献集、生成共享链接, 实现个性化文献管理与协作。

(六) MyNCBI



1. 如何注册 MyNCBI?

使用 ORCID 登录



2. 如何保存检索结果？

文献--sent to-collection 或直接文献--collection

3. 如何导出检索结果？

文献--sent to--citations manager

4. 如何自动获取更新文献？

点击搜索框下的 create alert

5. 如何保存文献集（Collection）？

点击 send to 选择 collection

6. 如何将个人文献集与他人共享？

my ncbi--collection--齿轮图标--private 改为 public--link

7. 如何设置文献属性筛选条件（Filter）？

搜索框左侧可以看到

8. 根据以上具体用法，结合帮助文档，总结 MyNCBI 的特点和用途。

MyNCBI 是 NCBI 提供的免费个人账户服务，旨在帮助科研人员高效管理文献和学术产出，核心功能包括文献追踪与自动化管理：可以保存使用过的检索策略并且设置自动更新提醒，之后系统会按设定频率将最新相关文献发送至邮箱，从而便于长期跟踪特定研究领域的文献。

在文献整理方面，可以将感兴趣的文献保存到个人收藏夹，并且创建不同的文献集进行分类管理。可以从 NCBI 导出至 EndNote、Zotero 等文献管理软件方便引用。此外，MyNCBI 还有资源共享与协作的功能，可以将自己的文献集设为公开，生成共享链接，与他人分享收藏的文献。同时可根据常用筛选需求自定义页面过滤器，提升日常工作效率。

5. 问题

- i. 课程中提到的一些国外网站（如 NCBI）打开不方便，校园网无法访问，需要附加条件（已解决：热点+梯子）；
- ii. 不熟悉 EBI 网站中的工具在哪（已解决：在工具栏“service”中可以找到需要的工具）
- iii. 掌握了使用 NCBI、EBI 和 CNCB 等数据库查找基因和蛋白的方法，发现在 EBI 中可以进行蛋白质数据分析，并且在查找基因时 EBI 可以直接链接并进行深入的蛋白质序列分析，相较于其他两个数据库更加方便快捷。
- iv. 学习使用文献管理软件，如 EndNote、Zotero 等。

6. 组员个人总结

A. G4A 高嘉禾个人总结:

课程第一部分对 NCBI、EBI、CNCB 等数据库进行了系统学习，使我对生物信息资源在功能基因位点挖掘中的应用有了更深入的理解，收获如下：

一、通过学习掌握了 3 个数据库的基本使用方法：

1. NCBI: 通过基因数据库可以定位猪的目标基因；利用 PubMed 跟踪与某些性状或基因的最新的相关文章；可以使用 BLAST 功能进行序列的同源性比对。

2. EBI: 可以查询已经发布的猪的参考基因组；借助 Ensembl 基因组浏览器获取猪的基因组注释信息，其中“Service”入口可直接链接到 InterPro、UniProt 等工具，便于进行蛋白质结构域和功能位点的深度分析，便于挖掘潜在的功能突变位点。

3. CNCB: 作为国内的备份数据库，可以在网络不稳定时可作为重要补充。另外随着 CNCB 的逐渐发展，功能也逐步完善。

二、拓展了对从分子结构到功能位点的认知

通过阅读 PDB-101 的生物大分子月报和 UniProt 的蛋白质分子精选以及化学小分子月报进一步学习了不同生物大分子的功能和结构，通过模式图对蛋白质的结构有了更深刻的认识。通过胰岛素、生长激素受体为例子练习，学习了这两种蛋白的发现历程和结构解析过程，使我意识到在研究中要将候选位点与蛋白质功能验证相结合，也提示我在进行基因位点挖掘时，不能仅关注 SNP 本身，还应结合蛋白质三维结构，分析突变是否位于关键结构域（如 DNA 结合域、活性中心等），从而预测功能性影响。

三、提升了在 NCBI 中管理文献信息的能力

学习了 PubMed 与 PubMed Central 的区别，掌握了利用 MeSH 词进行精准检索的方法。如在搜索猪泛基因组分析的相关文献时，将“Sus scrofa”与“pangenomes”结合 MeSH 词进行组合检索，可以提高查准率，而且此方法在任何数据库查询文献时都适用。

学会了使用 ORCID 账号登陆 NCBI 进入 MyNCBI，可以设置自己感兴趣的内容的自动更新提醒，并且可以利用 Collection 功能分类保存重要文献，从而对文献长期追踪和管理。

在 MyNCBI 中构建了自身需要的文献数据库，为今后的论文写作和课题设计等科研工作打下基础。

四、问题与反思

1. 网络问题：部分国外数据库访问不稳定，需通过校园网优化或使用 VPN 工具解决，才能够确保研究连续性，这也说明了发展我国自主的数据库的必要性和重要性。

2. EBI 平台可以将基因查找与蛋白质分析衔接起来，不需要查询多个网站，更加便利，可以更多利用 EBI 进行序列的深度功能注释。同样给国内数据库的发展提供了参考，从而可以让我们在使用过程中更加流畅。

B. G4B 赵雪倩个人总结:

一、数据库资源为抗菌肽研究提供核心支撑

抗菌肽作为一类具有广谱抗菌活性的小分子多肽，其结构与功能关系的研究高度依赖生物信息学工具。

NCBI: 可检索抗菌肽相关基因序列、蛋白序列及文献，通过 PubMed 获取最新研究进展，利用 BLAST 工具进行同源性比对，帮助发现新型抗菌肽及其同源物。

EBI: 其提供的蛋白质数据分析工具（如 InterPro、UniProt）可直接用于抗菌肽的结构域预测、功能注释和序列分析，相较于其他数据库，EBI 在蛋白质层面的分析衔接更为流畅。

CNCB: 作为国内生物信息数据平台，在访问稳定性方面具有优势，适合作为国外数据库的补充，用于序列存储与本地化分析。

二、分子结构与功能的学习深化了对活性肽作用机制的理解

课外练习中对 PDB-101 “生物大分子月报”的学习，重点涉及多种与活性肽研究相关的分子结构与功能机制。

血红蛋白的协同作用: 展示了蛋白质亚基之间如何通过构象变化实现功能调控，这一机制在理解抗菌肽与膜受体相互作用、多肽协同抗菌效应方面具有借鉴意义。

流感病毒神经氨酸酶: 其四聚体结构和活性位点特征，为理解酶活性中心构象与功能的关系提供了范例。对于研究抗菌肽是否通过酶抑制机制发挥作用，具有启发价值。

此外，化学小分子月报中对维生素 C、褪黑素等分子的学习，展示了小分子结构与其功能之间的构效关系（SAR），为抗菌肽的结构优化与改造提供了思路借鉴。

三、蛋白质专题网站的学习为活性肽研究提供历史视角

在 Protein Spotlight 网站的学习中，重点关注了胰岛素和绿色荧光蛋白（GFP）的发现与发展历程 ‘

胰岛素: 作为“20 世纪蛋白质分子的代表”，其从结晶、测序到合成、基因工程的历程，展示了蛋白质研究的技术演进。对于活性肽研究而言，这一路径提示了从天然产物中分离、结构鉴定到人工合成与改造的完整技术链条。

GFP: 其独特的 β 桶结构与自发荧光特性，为抗菌肽的标记与示踪研究提供了技术参考。未来可将抗菌肽与荧光蛋白融合表达，用于研究其在细胞内的定位与作用机制。

四、演化生物学网络教程与 PubMed 功能的学习拓展了研究视野

演化教育网站的学习帮助树立了正确的生物演化观，强调了“树枝模型”而非“梯子模型”的科学认知。在抗菌肽的进化研究中，这一视角有助于理解抗菌肽在不同物种中的保守性、多样性及适应性演化规律。

PubMed 与 MyNCBI 的功能学习则直接服务于日常科研管理：掌握了检索策略保存与自动更新功能，可长期跟踪抗菌肽领域的最新文献；学会建立文献集并与他人共享，便于课题组内资料整合；明确自身研究方向在 MeSH 分类中的归属（Computational Biology），为精准检索提供支持。

C. G4C 张俊鑫个人总结:

水稻作为全球主要粮食作物和单子叶植物研究的模式植物，其种质资源的收集、鉴定、挖掘与创新是保障粮食安全和推动育种进步的核心前提。随着多组学技术的快速发展，海量水稻种质资源相关数据应运而生，NCBI、EBI、CNCB 等综合性生物信息数据库，已成为水稻作物种质资源学研究中不可或缺的工具。结合自身专业学习与研究实践，我对这些数据库在本领域的应用、价值及现存问题形成了一些思考与总结。

一、三大核心数据库在水稻种质资源学中的应用实践与核心价值

NCBI、EBI、CNCB 三大数据库虽分属不同地区、侧重不同领域，但均以“数据共享、工具支撑、助力研究”为核心，在水稻种质资源的多样性分析、基因挖掘、资源创新等环节发挥着不可替代的作用，且各有侧重、互补共生，构成了水稻种质资源研究的数字化支撑体系。

（一）NCBI：全球水稻种质资源的“数据枢纽”，侧重序列与功能关联

NCBI 作为全球最具影响力的生物信息数据库，其核心优势在于数据的全面性和开放性，尤其在水稻种质资源的基因组序列、遗传变异及功能注释方面提供了海量基础数据。NCBI 是获取水稻种质资源测序数据的核心渠道，其中收录了全球不同地区、不同生态型水稻（包括栽培稻、野生稻）的重测序、转录组测序数据，为研究水稻种质资源的遗传多样性提供了原始素材。

此外，NCBI 收录了大量水稻基因序列、EST 序列及线粒体、叶绿体基因组序列，为水稻种质资源的功能基因挖掘提供了便利。在研究水稻抗逆、高产等优异性状相关种质时，可检索目标性状相关基因的序列信息，结合种质资源的表型数据，实现“基因型-表型”的关联分析，为筛选携带优异基因的水稻种质提供依据。同时，NCBI 收录了大量水稻种质资源相关的研究文献，可辅助我们追踪领域研究热点，了解不同数据库在水稻种质资源研究中的应用案例，为自身研究提供思路借鉴。其内置的在线分析工具，还能实现遗传变异的初步分析，降低了初学者的数据分析门槛。

（二）EBI：水稻种质资源的“多组学整合平台”，侧重数据分析与可视化

EBI 的核心特色的是多组学数据的整合与深度分析工具的提供，其数据库资源更注重数据的标准化和可操作性，这与水稻种质资源学中“多维度解析种质价值”的研究需求高度契合。在实践中，EBI 的 Ensembl 是专业常用的工具之一，该数据库提供了水稻（包括日本晴、9311 等参考品种）的基因组注释、遗传变异、基因表达等多维度数据，且支持基因组浏览器可视化，能够直观呈现水稻种质资源的基因分布、变异位点及表达模式。

此外，EBI 提供的 EMBOSS 在线分析工具，可实现水稻种质资源的序列比对、进化树构建、基因序列分析等功能，无需本地安装复杂软件，极大提升了研究效率。相较于 NCBI，EBI 的工具更侧重多组学数据的整合分析，能够帮助研究者从基因组、转录组等多维度解析水稻种质资源的遗传特征，推动研究向更深更广的方向拓展。

（三）CNCB：水稻种质资源的“本土化支撑平台”，侧重特色资源与产业适配，使用便捷迅速

CNCB 作为我国自主建设的国家级生物信息中心，其数据库资源更贴合我国水稻种质资源的研究与产业需求，弥补了国际数据库在我国特色水稻种质资源收录、注释方面的不足，是我国水稻种质资源研究的重要支撑。CNCB 收录了大量我国自主培育的水稻品种、地方品种及野生稻种质的测序数据，包括稻属不同物种的基因组数据、遗传变异数据等，尤其注重我国特有水稻种质资源的保存与共享。

此外，CNCB 整合了全国农作物种质资源信息平台的相关数据，涵盖水稻等 26 种作物 30 余万份种质

资源的来源、性状特点、共享状态等信息，实现了种质资源“数据查询-申请获取”的全流程服务，为科研人员精准获取所需水稻种质资源提供了便利。同时，CNCB 提供的在线分析工具更贴合我国研究者的使用习惯，且支持中文检索，降低了专业学习和研究的门槛。在实践中，通过 CNCB 可快速获取我国地方水稻品种的基因组数据，结合其提供的变异分析工具，能够解析我国水稻种质资源的遗传多样性特征，为地方品种的保护与利用提供数据支撑，助力我国水稻种业的自主创新。

二、数据库应用中的现存问题与个人反思

在学习三大数据库开展水稻种质资源学相关学习和研究的过程中，我也发现了一些现存问题，这些问题不仅影响了研究效率，也反映出当前数据库应用与水稻种质资源研究需求之间的差距。

首先，数据整合度不足。三大数据库及各类水稻专用数据库之间的数据缺乏有效的联动，同一水稻种质的序列数据、表型数据、变异数据可能分散在不同数据库中，研究者需要在多个数据库之间切换查询、下载，增加了研究工作量。

其次，数据分析门槛较高，工具适配性有待提升。虽然三大数据库都提供了在线分析工具，但部分工具的操作难度较大，对于初学者而言，需要花费大量时间学习工具的使用方法；同时，部分工具的分析功能针对性不强，难以满足水稻种质资源研究中“精准挖掘优异基因”“解析种质遗传分化”等具体需求。

最后，种质资源数据的更新与共享机制不够完善。水稻种质资源的研究处于快速发展中，新的测序技术、新的种质资源不断涌现，但部分数据库的数据更新不及时，导致研究者无法获取最新的种质资源数据；同时，部分特色水稻种质资源的数据共享存在壁垒，尤其是一些自主创制的优异种质资源，其测序数据和表型数据未能及时公开共享，限制了种质资源的高效利用。

三、总结

NCBI、EBI、CNCB 等数据库作为水稻作物种质资源学研究的核心数字化工具，为水稻种质资源的收集、鉴定、挖掘与创新提供了强大的支撑，其应用贯穿于水稻种质资源研究的各个环节通过对这些数据库的学习，我认识到，数据库不仅是数据的存储载体，更是推动水稻种质资源学发展的重要动力，熟练掌握数据库的应用方法，是水稻种质资源研究者必备的专业能力。

同时，我也认识到，当前数据库应用中存在数据整合不足、分析门槛较高、共享机制不完善等问题，这不仅需要数据库建设者不断优化完善，也需要我们这些普通的研究人者不断提升自身能力，积极进行数据共享，而不仅仅作为数据的使用者，更好地发挥数据库的价值。