

主要把课件看懂，下面的题可以看一下。

1. 某一原核细菌的蛋白质序列如下图所示，若将其编码基因克隆在大肠杆菌中表达，并希望最终获得序列纯净的重组蛋白，那么请：

- (1) 确定合适的工程菌构建策略，并说明依据；
- (2) 简述工程菌发酵后的下游分离纯化操作流程。

Met Tyr His Leu Gln Asp Asp His Phe Ser Ile Ile Val Tyr Arg Lys Ala Asn Thr Gln  
Glu Ala Pro Leu Phe Asp Asp His Gly Arg Trp Ser Gly Glu Gln Ala Leu Val Phe Thr  
Thr Asn Ser Tyr Met Leu Pro Ile Lys Lys His Trp Pro Trp Asp Val Val Gly His Arg  
Phe Ile Tyr Tyr Tyr Lys Gln Ser Phe Gly Ser Gln Ala Glu Arg Arg His Trp His Ala  
Gly Val Leu Leu Lys Ala Phe Ser Gly Glu Ala Asn Ser Asp Pro His Gly Ile Glu Gly  
Arg Ser Ile Val Tyr His Thr Met Phe Phe Lys Gly Gly Pro Asp Thr Ser Ser Ser Pro  
Leu Gln His Pro Glu Ile Val Arg Leu Thr Ala Gly Gln Trp Phe Phe Phe Ala Thr Ala  
Gly Ser His Leu Ala His Arg Trp Gly Leu Leu Leu Arg Glu Asp Phe Asn Ala Val Thr  
Trp Pro Met Trp Glu Ala Phe His His Ser Ala Gly Ser Val Val Ser Gln Lys Pro Pro  
Val Glu His His Tyr Phe Gln Ser Ile Ile His Tyr Asn Leu Ala Leu Ser Thr Trp Glu  
Val Ala Gly Pro Val Trp Trp Phe Leu Thr Val Lys Lys Asn Leu Gly Asp Asp His Met  
Ala Ala Phe Asp His Trp Arg His Thr Gly Leu Tyr Tyr Val Glu Pro Gly Glu His Tyr  
Leu Gly Pro His Arg Val Gly Ala Gln Asp Thr Ser Val His Ala Val Trp Arg Gln Ala  
Asn Leu Pro Ile Ser Asn

关键词：外源表达序列纯净的蛋白 下游分离纯化

分析：

该蛋白有两个特点：1，分子量比较小；2，没有 Cys。这说明该蛋白容易变性和复性。为了表达序列纯净的蛋白而不引入 tag，可以采取包涵体法表达并纯化该蛋白。

为了得到包涵体，必须使该蛋白高表达，以致占细胞总蛋白量的 20% 以上。所以使用 pCP3 即温度可诱导型的载体。（低温时该载体为低拷贝数，对细胞生长的抑止作用较小；升高温度不但有利于该质粒的扩增，同时亦有助于包涵体形成）

由于该蛋白没有 Cys，而且才不到 300 个氨基酸，所以它比较容易变性和复性，即对变性复性的条件要求不高。

答：

(1) a. 根据氨基酸序列，合成相应的核苷酸序列（以及终止密码子）并在两端加上合适的酶切位点（可采用分段合成法），并注意大肠杆菌的密码子偏好。

该蛋白对应的核苷酸序列没有给出，而且长度不是很长，所以决定人工合成，同时密码子可以采用最适合大肠杆菌的密码子。

b. 将 pCP3 载体用相应的酶切开，并将合成的序列加入，用 DNA 连接酶连接。

pCP3 是温度可诱导型载体，在 28 度培养可以迅速扩增菌量，在 42 度培养可以迅速获得高的质粒拷贝数以及蛋白表达量。而且 42 度容易得到包涵体。

并不是所有大肠杆菌菌株都适合用来获取蛋白，所以应该使用专门的菌株。

c. 转化适于表达蛋白的大肠杆菌菌株，并挑取若干个克隆，用 PCR 法鉴定得到阳性的克隆（最好多选取几个待用，以得到表达量最高的克隆）。

不可能所有长出的克隆都是目的克隆，而且各个克隆的蛋白表达能力不尽相同，

可以采用 His-tag，容易分离（Ni 柱），且切取容易（Xa 切断即可）

所以应该根据合成的 DNA 序列设计 PCR 引物并鉴定出阳性克隆,再选取高表达的克隆。

(2) a. 纯化包涵体。

将得到的发酵产物离心收菌,用水重悬,用超声裂解菌体,高速离心即可得到较为纯净的包涵体。可进行多次重悬离心操作以提高纯度。

b. 变性包涵体。

用尿素, 醋酸, DMSO 混合溶剂溶解包涵体。

c. 复性。

将上述混合液稀释到一定浓度, 蛋白即可自动折叠到正确的构相。

d. 纯化。

将上述混合液进行透析以除去不必要的离子。

e. 浓缩。

将上述溶液进行超滤以缩小体积。如果必要再加入甘油等保护剂,低温保存。取样检验。

2. 海藻糖是一种由两个葡萄糖分子以 , (1, 1) 糖苷键构成的非还原性双糖,存在于多种微生物和植物的细胞内,难以分泌至胞外。海藻糖在高温、低温、干燥等极端环境中具有保护生物大分子和细胞的重要功能,因此被广泛应用于医药、食品、保健品、化妆品的制造以及抗寒抗旱转基因植物的构建。相关研究表明,利用来自原核生物节杆菌 (*Arthrobacter*) 的麦芽寡糖基海藻糖合酶(其编码基因为 *mts*) 以及来自大肠杆菌的新型淀粉酶(其编码基因为 *amy*) 联合转化淀粉,即可实现海藻糖的大规模产业化,然而上述两种酶在野生型菌株中的含量甚微。已知节杆菌中 *mts* 基因的部分序列为:

5' TCACCGATCCCTCCGCCGTCGACCCCGAACGCGGCGGGCCGGAGGGC 3'

请详细设计一株能高产分泌型麦芽寡糖基海藻糖合酶的基因工程菌的构建路线,内容包括目的基因的克隆、重组、表达以及海藻糖规模化生产的策略。

关键词: 分泌型表达 目的基因的克隆

分析: 由于该蛋白的基因尚未得到克隆,因此需要从 cDNA 文库中筛选。给出的序列即是作为探针使用的序列。探针需要单链结构,至少 12bp, 内部不含互补区, 并有放射性或者生物素标记。筛选到该基因后,由于要求分泌型表达,所以应该用含有细菌素释放蛋白的质粒,和在 *mts* 的 N 端加上了信号肽的质粒共转染大肠杆菌,这样就可以得到能够分泌麦芽寡糖基海藻糖合酶的菌株了。根据题目所给出的信息,可以知道如果想生产海藻糖,似乎并不一定要将酶提纯出来,而是将可以分泌表达新型淀粉酶的菌株和分泌表达麦芽寡糖基海藻糖合酶的菌株一起培养在含有淀粉的培养基中,发酵完毕后只需要进行一步纯化,就可以得到海藻糖。

答:

基因的克隆:

a. 构建 cDNA 文库。先将节杆菌的 mRNA 抽提出来,再使用引导合成法合成序列完整的 cDNA。将 cDNA 补成平末端插入非表达型载体,并转入大肠杆菌制成节杆菌的 cDNA 文库。

b. *mts* 基因的筛选。用菌落原位杂交法筛选 *mts* 基因。将节杆菌 cDNA 文库的大肠杆菌用影印平板法影印到硝酸纤维素薄膜上,再用 0.4M NaOH 浸泡以裂解,用 SSC 溶液浸泡清洗,再在 80 度烘干以固定 DNA。用人工合成的含有放射性标记末

端的序列 TCACCGATCCCTCCGCCGTC 进行杂交。用 SSC-SDS 溶液洗涤杂交膜并进行放射自显影。得到的阳性克隆挑出进行测序以进一步验证，并用 PCR 法得到该序列的 DNA。

基因的重组：

将得到的 mts 基因的 DNA（用 PCR 法时在两端加上了适合的酶切位点）**连接到 N 端带有大肠杆菌信号肽的分泌型表达质粒上，并和带有细菌素释放蛋白的质粒共转染大肠杆菌**（上述两个基因使用相同性质的启动子以使之同步表达），并挑选出阳性克隆，保存菌株。

蛋白的表达：

在适当的培养基中培养该菌株到合适的时期，麦芽寡糖基海藻糖合酶即可分泌到培养基中。**加入适当底物以检测表达的蛋白的活性，并探索能得到蛋白最高表达量的条件。**



海藻糖规模化生产：

先进行小规模试验：将上述菌株诱导蛋白表达后，再将分泌表达有新型淀粉酶菌株的培养液两者以一定比例混合，再加入淀粉作为反应底物，反应一定时间后检测海藻糖的产量。针对两种菌的培养液的比例，底物浓度，反应时间等条件进行改进，以得到最佳反应条件。再进行中等规模试验，最后进行大规模试验并投入生产。

概要：麦芽糖结合蛋白（MBP）可促进表达分泌型蛋白，构建 MBP-mts 融合蛋白，用支链淀粉-琼脂糖凝胶亲和层析可以得到纯化的融合蛋白，凝血因子 Xa 裂解后，过分子筛可以得到纯化的麦芽寡糖基海藻糖合酶。

在 ncbi 中，BLAST 找出 mts 基因序列如下：

```
ATGAGAACGCCAGTCTCCACGTACAGGCTGCAGATCAGGAAGGGATTACACTCTTCGACG
CGGCCAAAACCGTTCCGTACCTGCACTCGCTCGGCGTCGACTGGGTCTACCTTTCTCCGGTCCTGACTG
CCGAGCAGGGCTCCGACCACGGGTACGACGTCACCGATCCCTCCGCCGTCGACCCCGAACGCGGCGGGC
CGGAGGGCCTCGCGGCGGTTTCCAAGGCGGCCCGCGCCGCGGGCATGGGCGTGCTGATCGACATCGTG
CCAACCACGTGGGCGTCGCGACGCCGGCGCAGAACCCCTGGTGGTGGTCGCTGCTCAAGGAGGGACGCC
AGTCCCGTTACGCGGAGGCGTTTCGACGTCGATTGGGACCTCGCCGGGGGACGCATCCGGCTGCCGGTGC
TCGGCAGCGACGATGACCTCGACCAGCTCGAAATCAGGGACGGGGAGCTGCGGTACTACGACCACCGAT
TCCCGCTCGCCGAGGGAACCTACGCCGAAGGCGACGCCCCGCGGGATGTCCACGCCCGGCAGCACTACG
AGCTCATCGGCTGGCGCCGCGCGGACAACGAGCTGAACTACCGCCGCTTTTTCGCGGTGAACACGCTCG
CCGGCGTCCGCGTGAAATCCCCGCCGTCTTCGACGAGGCACACCAGGAGGTGGTGCCTGTTCCGCG
AGGACCTTGCGGACGGCCTGCGGATCGACACCCGGACGGCCTCGCTGACCCCGAGGGGTACCTGAAGC
GACTCCGGAAGTCACCGGCGGCGCTTACCTGCTGATC
```

根据序列设计合适引物从节杆菌（*Arthrobacter*）基因组中 PCR 扩增出 mts 基因，PCR 引物中 5' 端加入 XbaI 和 3' 端加入 HindIII 酶切位点。

使用 pMAL-C2X 作为载体，双酶切后将 mts 基因连入载体。

转入 DH5 α，Amp<sup>r</sup>抗性筛选，扩增，双酶切鉴定，测序验证。

IPTG 诱导表达。（为获得较高产量可摸索不同 IPTG 浓度梯度）

收菌，超声裂解

支链淀粉-琼脂糖凝胶亲和层析纯化的 MBP-mts 融合蛋白

凝血因子 Xa 裂解 MBP-mts 融合蛋白

分子筛分离（MBP 约为 42KD，mts 约 30KD），得到纯化的 mts

主要是记，下面的题都可能考到。

1. 在进行小鼠基因剔除实验中，常用 129 小鼠的 ES 细胞和 C57 小鼠的囊胚，请说明有什么好处和缺点？

好处：毛色不同，好筛选

目前成熟的ES细胞建立方法均基于129小鼠，而C57小鼠基因组较为稳定，同时两者毛色不同，便于表型区分。

缺点：遗传背景不纯

2. 请简述目前在生物医药研究中常用的模式生物有哪些（试举 3 种以上不同进化层次的生物）？它们各自的研究优势是什么？

常用的模式生物有：大肠杆菌、酵母、线虫、果蝇、斑马鱼、小鼠

(1) 大肠杆菌：是分子生物学，遗传学和生物化学研究中的重要模式生物，也用于人类感染性疾病和其他微生物研究的模式生物。其研究优势包括：

① 适合实验室培养（可以液体培养和固体培养），培养条件简单，生长迅速。

② 安全，克隆方便，基因操作简便。操作体系成熟

③ 1997 已完成测序，用 4.3M，4400 基因。

(2) 酵母：是分子生物学，遗传学，生物化学，真核细胞周期调控，信号转导等研究的重要模式生物，其研究优势包括：

① 它是单细胞的真核生物，适合实验室培养（可以液体培养和固体培养）。

② 安全性好，克隆方便，基因操作简单 操作体系成熟

③ 可以有丝分裂，也可以出芽生殖；可以以单倍体和双倍体生长，且可在实验条件下较方便地控制单倍体和二倍体之间的相互转换，这点对其基因功能的研究十分有利

④ 完整基因组已测序，有 16 条染色体，12M，约 6000 个蛋白编码基因，基因结构紧凑

(3) 线虫：是研究发育（细胞命运决定）、细胞凋亡、细胞迁移、神经系统的重要模式生物，其研究优势包括：

① 成虫大小约 1mm，实验室饲养简单，以细菌为食，生命周期短，保种方便，大部分为雌雄同体（雄虫占 0.05%），全身透明，在显微镜下不需经过染色，其体内的器官如肠道、生殖腺等均可一览无余，且不具寄生性，适合实验室饲养

② 成虫由 959 个细胞构成，其中 302 个细胞为神经细胞，成虫过程中有 131 个细胞经历凋亡；所有的细胞谱系清楚，人们已经建立了完整的线虫从受精卵到所有成体细胞的谱系图。

③ 转基因和 RNAi 等遗传操作成熟

④**已完成基因组测序**，有 5 对常染色体，1 对性染色体，97M，约 20000 个基因

(4) 果蝇：**是遗传学，发育生物学理想的模式生物**，其研究优势包括：

①**个体小，3mm，适合实验室饲养，生命周期短（2 周），繁殖力强（数百个卵）**

②**具有成熟的遗传操作的方法，具备众多的突变品系，有特殊的染色体结构（多线染色体，可分成 102 条区带）**

③**有比较复杂的行为能力**

④**测序已完成**，3 对常染色体，1 对性染色体（X，Y）；165M，14000 基因

(5) 斑马鱼：**主要用于遗传和发育学研究**，其研究优势有：

①**个体小，淡水饲养，繁殖周期短，产仔频繁且量高，适合实验室养殖**

②**脊椎动物，胚胎发育周期短，遗传操作方法成熟**，其个体发育是在全透明状态下完成的，这不仅使人们很容易得到胚胎，而且还可以在显微镜下直接观察斑马鱼胚胎发育的过程，是进行胚胎发育机理和基因组研究的好材料

**胚胎透明  
便于观测**

③**拥有大量突变体，可以模拟人类疾病模型。通过对斑马鱼的突变体进行遗传疾病模型的研究，不仅能够明确一些已知基因的功能，而且还可以发现一些新的基因并提供这些基因的功能。**

④**测序完成**，25 对染色体，168M，30000 基因

(6) 小鼠：**是生物医学研究中广泛使用的模式生物**，也是当今世界上研究最详尽的哺乳类实验动物，它作为模式生物的优势有：

①**是最小的哺乳动物之一**，时代周期较短，**饲养管理费用低**

②**生物进化上与人类接近**（60-75 百万年）

③**胎盘形成和早期胚胎发育与人类相近，组织器官结构和细胞功能与人类相近**

④**有高级神经活动**

⑤**小鼠基因测序计划已经完成**，人类 99% 的基因存在于小鼠，基因同源性高达 78.5%；基因组 93% 的区域基因排列顺序与人类相同

⑥**基因组改造技术成熟**，经长期人工饲养选择培育，已育成多达千余个独立的远交群和近交系，利用小鼠可以进行基因体内功能及基因组功能的研究，以及建立人类疾病动物模型。

因此，小鼠可以代替人体进行安全性和毒性试验，可以用于生物效应测定和药物效价比较、药物的筛选、放射学研究，以及各种生理病理研究，是研究哺乳动物和人类基因功能的模式生物。这将对人类了解自身、预防和治疗各类疾病产生重要和深远的影响。



#### 4. 小鼠的常规基因剔除和条件性基因剔除有什么不同，后者主要通过什么手段来达到条件剔除的目的？

基因剔除是利用基因同源重组 (gene homologous recombination) 又称基因打靶的原理，用外源片段整合到活体细胞 DNA 的同源序列中，使某个基因被取代或破坏而失活。

常规基因剔除是使用一个带有正负筛选标记的打靶载体转至 ES 细胞，然后将筛选成功的 ES 克隆导入囊胚中以进行发育。常规方法的缺陷在于：它使得实验个体的所有细胞都被改变，往往引起严重的发育缺陷和早期胚胎死亡；在培养纯系小鼠的过程中，有些基因的互补效应也增加了表型分析的复杂性；另外，由于广泛应用 neo 和 HSVtk 作为正负选择系统，发生同源重组的细胞基因组中总留有外源的选择标记 (neo) 基因；该基因可能影响相邻基因的表达，不利于对突变表型的精确分析。

条件基因剔除则可以克服以上缺陷，它与常规基因剔除的不同之处就在于条件剔除对于基因的敲除是具有组织特异性的，可以在特定的发育阶段诱导敲除，因此可以避免早期胚胎的死亡，发生同源重组的细胞基因组中不含选择标记 (neo) 基因；

而条件打靶通常采用两个载体，其中一个载体带有目的基因，另一个带有调控基因（带有组织特异性的 promoter 和 Cre 等可诱导基因）。将这两个载体分别转入两系小鼠中，将这两系小鼠杂交后就可以得到在特定组织特定时间 gene knockout 的小鼠。

双载体+双品系：亦是已有的模型动物，如 flox 小鼠

利用 DNA 同源重组技术和胚胎干细胞 (ES 细胞) 发育全能性的原理，先将待剔除的基因制成缺失突变型，在缺失的位置上插入一个选择基因如新霉素抗性基因 (neo)，同时再接上另一个选择基因如胸苷激酶基因 (tk)。将这一段带有已失去原有功能的待剔除基因的 DNA 片段装在载体上，通过一定的方式 (电穿孔法或显微注射) 转入在体外培养的小鼠胚胎干细胞 (embryonic stem cells, ES cells)。在细胞内，通过同源重组将基因组里有功能的待剔除基因剔除掉。然后通过有效的筛选方法，通常用 PCR 技术，southern blot 技术或 PNS (positive-negative selection) 筛选出目的基因被敲出的干细胞，注射进小鼠早期胚胎，将产生的缺失杂合体小鼠同野生型小鼠杂交，将获得的杂合体后代交配，从子代中选出该基因缺失的纯合体，其概率为四分之一。

基因敲除技术的局限性：

胚胎致死

所有细胞都改变（条件性敲除和诱导性敲除可改善该缺点）

功能代偿


冗余

#### 5. 转基因小鼠技术

- 受精卵单细胞原核注射
- 精子介导的转基因技术
- 病毒介导的转基因技术（逆转录病毒、Lentivirus）
- 穿膜肽（核定位蛋白）介导的基因转移



- **转座酶介导的基因转移**

用受精卵单细胞原核注射导入基因的特点：

- **随机插入**
  - 单位点和多位点插入并存
  - **多拷贝串连**
- 插入位点及整合方式对转入基因表达的影响
  - **插入位点影响**
  - 插入序列引发的 de novo **甲基化**
  - **同源性依赖的基因沉默** 
  - **使用大片段的 DNA**，如 BAC 等
  - 使用隔离子 (Insulator 如小鼠 H19 基因) 来减少插入位点对转基因表达的影响
- **方法技术要求高，设备投入大**


显微注射、电穿孔、磷酸钙沉淀及逆转录病毒载体感染等基因转移技术能够有效地将外源基因导入靶细胞内，但这些导入的外源基因在靶细胞基因组中整合的位点一般是随机的，可能导致下面几种结果：(1) 外源基因整合位点在某一正常基因的中部，导致该基因表达的缺失；(2) 外源基因导入位点在正常基因的侧翼序列，影响周围正常基因的活性；(3) 由于外源基因整合位点不恰当，在细胞内不表达或表达难以控制；(4) 外源基因的导入可能激活细胞内的原癌基因。为避免上述情况，最好的途径就是将

转基因小鼠的繁殖

1. 决定所需要的小鼠品系背景；首建者分别和选定的野生型品系进行配种繁殖；首建者小鼠之间不能进行相互交配。 
2. 做好出生小鼠的标牌（出生日期，亲本信息）；小鼠编号（1-2 周）；同时采集小鼠尾部组织或者耳轮组织；提取 DNA，进行基因型鉴定 (gene typing)；对阳性小鼠进行断奶、雌雄分笼。 

转基因载体的构建


1. 获得所需要表达蛋白的 cDNA
2. 选择 **合适的基因表达启动子**
3. 保证转录后的正确加工

-  基因调控序列的研究及报告基因
  - 启动子；增强子；负调控元件 等等

- 常用的报告基因

- E. coli  $\beta$ -半乳糖苷酶 (lac Z)
  - 底物: ONPG; MUG; X-gal
- E. coli 葡萄糖苷酸酶 (gusA)
  - 底物: X-gluc
- 荧光素酶 (luc)
  - 底物: luciferin
- 绿色荧光蛋白 (GFP) 等

-----可诱导的基因表达表达方法

- 利用内源性可诱导的启动子:
  - 热休克蛋白, hsp70 基因
  - 金属硫蛋白基因
- 重组型的可诱导启动子
  - LacI; tet-on or tet-off; 蜕皮激素; 化合物诱导的二聚化 (CID)
- 蛋白水平的诱导
  - ER 受体融合蛋白/Tamoxifen
- 重组酶系统 

## 6. 基因打靶

---该技术通过外缘 DNA 与染色体 DNA 之间的同源重组, 进行精确的定位修饰和基因改造, 具有专一性强、染色体 DNA 可与目的片段共同稳定遗传等特点。

---基因敲除技术分为完全基因敲除和条件性基因敲除。

- 完全基因敲除是指通过同源重组完全消除细胞或动物个体中的靶基因的活性。

在 ES 细胞中进行基因打靶最常用的策略是置换型载体的正-负双选择 (PNS)。正向选择基因 neo 通常插入载体靶 DNA 功能最关键的外显子中, 或通过同源重组法置换靶基因的功能区。Neo 基因有双重作用, 一方面形成八位点的插入突变, 同时可作为正向筛选标记。负向选择基因 HSV-tk 置于目的片段外侧, 含有该基因的重组细胞不能在选择培养基上生长。如果细胞中发生了随即重组, 负向选择基因就可能被整合到基因组中, 导致细胞死亡。见现代分子生物学 p161

最后将筛选成功的 ES 克隆导入早期胚胎的囊胚腔中进行发育。

常规方法的缺陷在于:



它使得实验个体的所有细胞都被改变，往往引起严重的发育缺陷和早期胚胎死亡；在培养纯系小鼠的过程中，有些基因的互补效应也增加了表型分析的复杂性；另外，由于广泛应用 neo 和 HSVtk 作为正负选择系统，发生同源重组的细胞基因组中总留有外源的选择标记(neo)基因；该基因可能影响相邻基因的表达，不利于对突变表型的精确分析。

- 条件性基因敲除是指通过定位重组系统实现特定时间和空间的基因敲除。(噬菌体的 Cre/Loxp 系统, Gin/Gix 系统, FLP/FRT 系统, R/RS 系统实现阶段重用的死重定位重组系统, 尤为第一种系统应用最为广泛。)

条件基因剔除则可以克服以上缺陷，它与常规基因剔除的不同之处就在于条件剔除对于基因的敲除是具有组织特异性的，可以在特定的发育阶段诱导敲除，因此可以避免早期胚胎的死亡，发生同源重组的细胞基因组中不含选择标记(neo)基因。

而条件打靶通常采用两个载体，其中一个载体带有目的基因，另一个带有调控基因(带有组织特异性的 promoter 和 Cre 等可诱导基因)。将这两个载体分别转入两系小鼠中，将这两系小鼠杂交后就可以得到在特定组织特定时间 gene knockout 的小鼠。

把下面内容看懂，有些需要去记的。

1. 蛋白质 A 是细胞内的一个信号分子，可以结合酪氨酸磷酸化蛋白质，它有三个结构域：  
一个结合酪氨酸磷酸化蛋白质结构域  
一个磷酸化位点结构域  
一个结合蛋白质 B 的结构域，并且能独立表达出有功能的结构域。

注意前提！

蛋白质 A 本身可以被受体酪氨酸激酶磷酸化，磷酸化的蛋白质 A 与蛋白质 B 的结合能力下降。蛋白质 B 也能被酪氨酸磷酸化，但其只有在 A 存在时才磷酸化，而 A 本身在任何时候都不具有激酶活性。磷酸化位点都为已知。

根据以上条件提出蛋白质间相互作用以及被磷酸化调控的模型。模型不得自己增加条件，并且必须符合以上所有条件。根据你的模型设计实验并加以证明。

要求：1、不包括蛋白质单分子技术类的方法。2、每一个证明实验必须用两种或两种以上的不同方法加以证明（如分子筛、超离心、SDS 电泳为三种不同的测量分子量的方法）。3、模型用示意图表示并附文字说明。

注释：

A-I：结合酪氨酸磷酸化蛋白质结构域

A-II：磷酸化位点结构域

A-III：结合蛋白质 B 的结构域

A/I：没有 A-I 的 A

A/II：没有 A-II 的 A

A/III：没有 A-III 的 A

TyrK：酪氨酸激酶

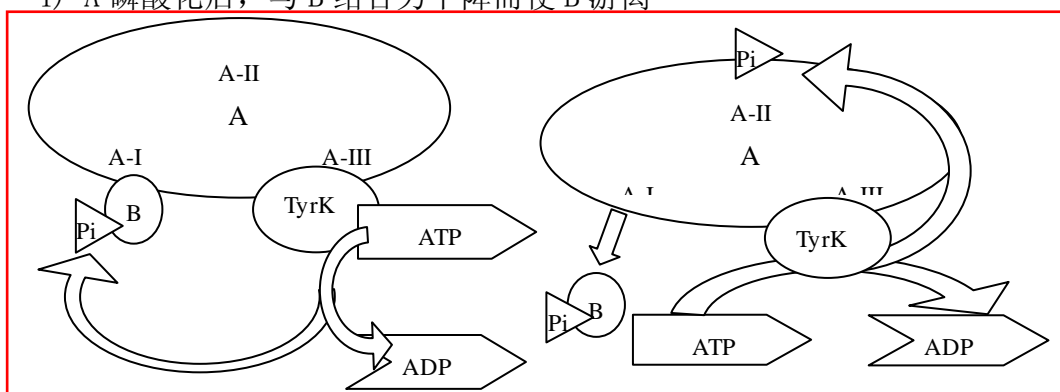
条件：

$A-B \rightarrow A-Pi + B : TyrK$

$B \rightarrow B-Pi : A + TyrK$

模型：

- 1) 没有磷酸化的 A 是与 B 结合的
- 2) A 是否结合 B 不依赖于 B 的磷酸化
- 3) A 无论是否被磷酸化，都可以接合 TyrK，磷酸化 B 或自身
- 4) A 磷酸化后，与 B 结合力下降而使 B 游离



实验：

1. 采用酵母双杂交和胞内 Co-IP，证明 A/A-I 与 B 结合
2. 采用胞外 Co-IP 和 GST-pull-down，证明体外 A/A-I 与 B-Pi 结合，而 A-Pi 与 B/B-Pi 结合能力下降 **WB+CoIP+FRET——双杂交则可用于结构域的分析**
3. 采用酵母双杂交和胞内 Co-IP 的方法证明 A-III 与 TyrK 结合
4. 将 A 和 TyrK 共转，提取蛋白做针对 A-Pi 的 Western Blots 或对 ATP 标记看是否 IP 的 A 含有放射性，证明 TyrK 磷酸化 A；通过 CoIP 的方法证明，TyrK 与 A-Pi

该模型有一个问题 --  
--即A不一定由K来磷酸化 -- --因为可能是其他负调控的酶参与；对于第2点可以予以考虑，但对于3，则暂不考虑，可以在问题与进一步分析中说明，即：  
(1) A可以结合p-K  
(2) A结合B  
(3) p-A结合B能力下降，造成B磷酸化下降（注意：题中所说B仅在A存在是磷酸化，且A本身无激酶活性）  
(4) A结合B后促进B磷酸化  
此外，对于已给定的题目条件不必再作证明！而对于p-B与A的结合性、A的负调控可再做进一步考虑

结合

5. 将 A/III 和 TyrK 共转，提取蛋白做针对 A-Pi 的 Western Blots 或 ATP 标记看是否 IP 的 A/III 含有放射性，证明 TyrK 磷酸化 A 要与 A 相互作用
6. 将 B 和 TyrK 共转，提取蛋白做针对 B-Pi 的 Western Blots 或 ATP 标记看是否 IP 的 B 含有放射性，证明 TyrK 不能单独磷酸化 B
7. 将 B, A/I 和 TyrK 共转，提取蛋白做针对 B-Pi 的 Western Blots 或用 ATP 标记看是否 IP 的 B 含有放射性，证明 TyrK 磷酸化 B 要 A, B 相互作用
8. 将 B, A/III 和 TyrK 共转，提取蛋白做针对 B-Pi 的 Western Blots 或 ATP 标记看是否 IP 的 B 含放射性，证明 TyrK 磷酸化 B 要 A, TyrK 相互作用
9. 将 B, A/II 和 TyrK 共转，提取蛋白做针对 B-Pi 的 Western Blots 或对 ATP 标记看是否 IP 的 B 含放射性，证明 TyrK 磷酸化 B 与 A 磷酸化无关；同时，提取蛋白做 CoIP 和 Western Blots，证明 A 与 B-Pi 是结合的
10. 将 B, A 和 TyrK 共转，提取蛋白做针对 B-Pi 的 Western Blots 或对 B 进行 IP 后 SDS-PAGE 是否有 Gel Shift，证明 TyrK 在 A 存在下可以磷酸化 B；同时提取蛋白做 CoIP 和 Western Blots，证明 A-Pi 没有结合 B-Pi

## 2. 检测蛋白质之间相互作用的实验方法有哪些？这些检测方法各有什么缺点？

### 1. 生化方法 (Biochemical approaches)

- 1.1 共纯化、共沉淀 Traditional co-purification (chromatography co-purification and co-sedimentation) 在不同基质上进行色谱层析
- 1.2 蛋白质亲和层析 (Affinity chromatography) 将一种蛋白质固定于某种基质上（如 Sepharose），当细胞抽提液经过改基质时，可与改固定蛋白相互作用的配体蛋白被吸附，而没有吸附的非目标蛋白则随洗脱液流出。被吸附的蛋白可以通过改变洗脱液或者洗脱条件而回收下来。

**GST pull down:** 为了更有效的利用蛋白质亲和色谱，可以将待纯化的蛋白以融合蛋白的形式表达，即将“诱饵”蛋白与一种易于纯化的配体蛋白融合。例如与 GST 融合的蛋白再经过 GSH 的色谱柱时，就可以通过 GST 和 GSH 的相互作用而被吸附。当载有细胞抽提物经过柱时，就可以得到能够与“诱饵”蛋白相互作用的目标蛋白了。

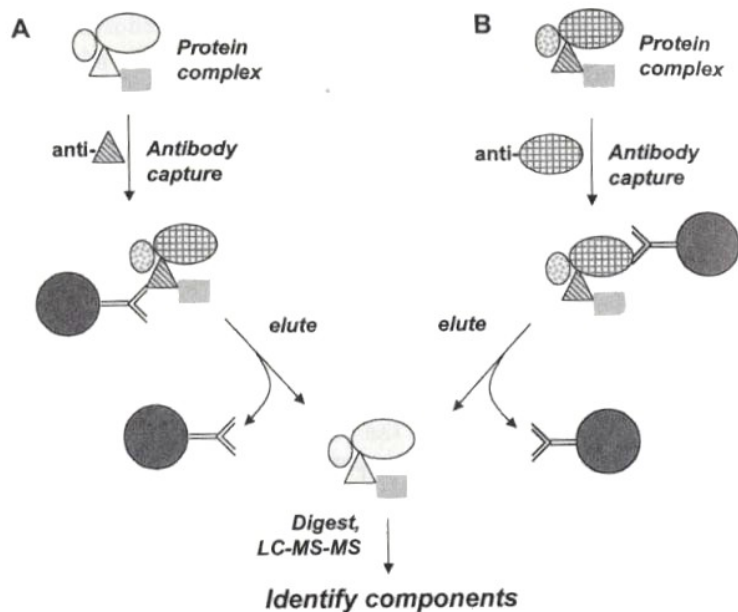
**Epitope-tag:** 表位附加标记技术 就是将附加的抗原融合到目的蛋白以检测目的蛋白的表达，同时还可以通过亲和层析法来纯化目的蛋白。缺点：表位附加标记可能会使融合蛋白不稳定，改变或使融合蛋白功能丧失。

以上两种方法都要共同的缺点：假阳性。实验所检测到的相互作用可能由蛋白质所带电荷引起的，并不是生理性的相互作用；蛋白的相互作用可能并不是直接的，可是由第三者作为中介的；有时会检测到两种在细胞中不可能相遇却有极强亲和力的蛋白。因此实验结果还应经其他方法验证。

### 1.3 免疫共沉淀 (Immunoprecipitation)

免疫共沉淀是以抗体和抗原之间的专一性作用为基础的用于研究蛋白质相互作用的经典方法。改法的优点是蛋白处于天然状态，蛋白的相互作用可以在天然状态下进行，可以避免认为影响；可以分离得到天然状态下相互作用的蛋白复合体。

缺点：免疫共沉淀同样不能保证沉淀的蛋白复合物时候为直接相互作用的两种蛋白。另外灵敏度不如亲和色谱高。



#### 1.4 Far-Western

又叫做亲和印记。将 PAGE 胶上分离好的凡百样品转移到硝酸纤维膜上，然后检测哪种蛋白能与标记了同位素的诱饵蛋白发生作用，最后显影。

缺点是转膜前需要将蛋白复性。

#### 2. 表面等离子共振 (Surface plasmon resonance)

该技术是将诱饵蛋白结合于葡聚糖表面，葡聚糖层固定于几十纳米厚的技术膜表面。当有蛋白质混合物经过时，如果有蛋白质同“诱饵”蛋白发生相互作用，那么两者的结合将使金属膜表面的折射绿上升，从而导致共振角度的改变。而共振角度的改变与该处的蛋白质浓度成线性关系，由此可以检测蛋白质之间的相互作用。该技术不需要标记物和染料，安全灵敏快速，还可定量分析。

缺点：需要专门的等离子表面共振检测仪器。芯片比较昂贵，如果重复使用可能导致结果不好。

#### 3. 遗传学方法 (Genetic approaches)

##### 3.1 基因外抑制子

基因外抑制子是通过一个基因的突变来弥补原有基因的突变。比如相互作用的蛋白 A 和 B，如果 A 发生了突变使两者不再相互作用，此时 B 如果再发生弥补性突变就可以使两者的相互作用恢复，那么 B 就是 A 的基因外抑制子。缺点：需要知道基因，要有表型，筛选抑制子比较费时。

##### 3.2 合成致死筛选

指两个基因同时发生突变会产生致死效应，而当每个基因单独发生突变时则无致死效应。用于分析两个具有相同重要蛋白之间的相互作用。

#### 4. 酵母双杂交 (Two-hybrid)

原理基于真核细胞转录因子的结构特殊性，这些转录因子通常需要两个或以上相互独立的结构域组成。分别使结合域和激活域同诱饵蛋白和猎物蛋白形成融合蛋白，在真核细胞中表达，如果两种蛋白可以发生相互作用，则可使结合域和激活域在空间上充分接近，从而激活报告基因。优点是酵母细胞内的 *in vivo* 相互作用，只需要 cDNA，简单，弱的相互作用也能检测到。

缺点：自身有转录功能的蛋白或者有其他蛋白插足介导或者自身激活报告基因会造成假阳性。融合蛋白会影响蛋白的真实结构和功能；不利于核外蛋白研究，会导致假隐性。

另外还有酵母的翻译后修饰不尽相同。尤其是蛋白质的调控性修饰；对基因库的要求比较高，单向 1/3 是 in frame；蛋白质毒性等。

## 5. 荧光共振能量转移技术 (Fluorescence resonance energy transfer)

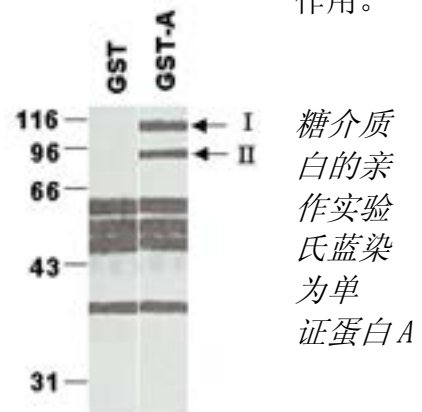
指两个荧光发色基团在足够近 (<100 埃) 时，它们之间可发生能量转移的现象。荧光共振能量转移技术可以研究分子内部对某些刺激发生的构象变化，也能研究分子间的相互作用。它可以 **在活体中检测，非常灵敏，分辨率高，能够检测大分子的构象变化，能够定性定量的检测相互作用的强度。**

缺点：此项技术要求发色基团的距离 <100 埃。另外 **设备昂贵，还需要融合 GFP 给蛋白标记。**

此外还有交联技术 (cross-linking)，**蛋白质探针技术，噬菌体展示技术 (Phage display)** 以及生物信息学的方法来检测蛋白质之间相互作用。

### 2004—周金秋

某实验室把纯化的重组融合蛋白 GST-A 共价交联到琼脂上，并以此对哺乳动物细胞的抽提物做亲和层析。GST 蛋白和层析样品作对照，而 GST-A 融合蛋白的亲和层析样品组。亲和层析洗脱后的样品用 SDS-PAGE 分离，并用考马斯。结果如图所示 (图左边的指示是分子量大小，以 kDa 位)。你将如何克隆编码蛋白 I 或 II 的 cDNA？你将如何验证与蛋白 I 或 II 的相互作用？



GST Pull Down 已经初步判定蛋白 A 与蛋白 I、II 可能有相互作用。

### 1. 克隆编码蛋白 I 或 II 的 cDNA

- 1.1 将 PAGE 胶上的两个条带挖下来打质谱；
- 1.2 **由质谱结果分析确定两个蛋白及其编码的基因 I、II；**
- 1.3 设计引物，收集细胞抽 RNA，RT-PCR 得到目的基因 I、II 的 cDNA。

**说明**：两个蛋白均较大，分别为 95KD 和 105KD，编码基因分别为 2400bp 和 2500bp 左右。

### 2. 验证相互作用 (以蛋白 I 为例说明与蛋白 A 的相互作用，蛋白 II 如法进行)

#### 2.1 coIP

- 2.1.1 设计合适引物，使用 **Flag-CMV4 载体**，构建 Flag-I 的重组表达质粒，GST-A 看有没有合适的酶切位点插入到 **Flag 载体**，如没有则重新设计引物构建 Flag-A 质粒；
- 2.1.2 将 Flag-A，Flag-I 共转染哺乳动物细胞 (如 COS-7，CHO)；

**说明**：以下需要抗体 anti-A，anti-I，anti-Flag，假设全部为鼠抗；

- 2.1.3 **用 anti-A 做 coIP，anti-Flag，anti-mouse，input 对照，使用 anti-I Western 检测，看是否能检测到蛋白 I；**
- 2.1.4 **用 anti-I 做 coIP，anti-Flag，anti-mouse，input 对照，使用 anti-A Western 检测，看是否能检测到蛋白 A；**

**说明**：anti-mouse 指示抗体轻重链位置，input 指示目的条带位置，anti-Flag confirm；

如果顺式和反式的 coIP 均有阳性结果则证明蛋白 A 和蛋白 I 有相互作用。

## 2.2 酵母双杂交

2.2.1 构建 pBD-A 与 pAD-I 两种重组表达质粒，这样使得待研究蛋白可融合表达 GAL4 DNA 结合结构域和 GAL4 激活结构域；

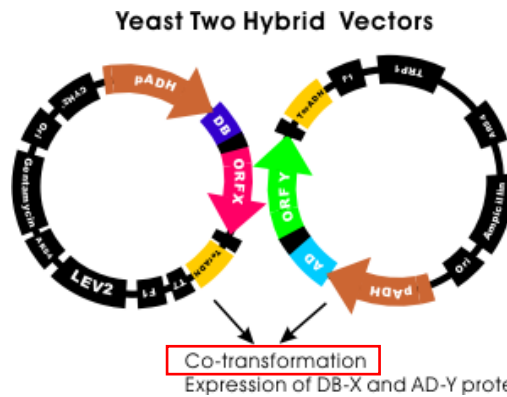
2.2.2 pBD-A 与 pAD-I 两种质粒共转化酵母细胞。

分别设置如下对照组：

pBD / pAD；

pBD-A / pAD；

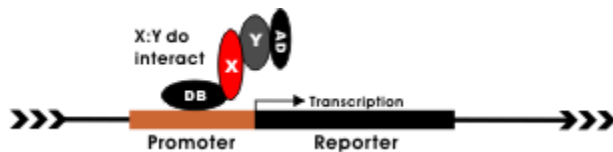
pBD / pAD-I；



2.2.3 蓝白斑法验证 *LacZ* 基因的表

**说明**：如果蛋白 A 和蛋白 I 能够相互作用，那么 GAL4 DNA 结合结构域和 GAL4 激活结构域就会相互作用，从而激活 *lacZ* 报道基因的表达。

如果设置的实验组长出蓝色克隆而对照组没有的话基本上可证明蛋白 A 和蛋白 I 有相互作用。



	His <sup>+</sup> (3AT <sup>r</sup> )	β-Gal	URA <sup>+</sup>	5FOA
X:Y do not interact	-	White	-	+
X:Y do interact	+	Blue	+	-

Phenotype of interaction

### 概述

1) 研究蛋白质相互作用的意义：细胞生命活动的基础

{ 稳定相互作用--protein complex  
瞬时相互作用--enzyme-substrate

蛋白质相互作用如无相应功能是没有研究意义的！

蛋白质相互作用  $\longleftrightarrow$  生物学功能

2) 蛋白质相互作用的亲和力---一切蛋白质相互作用的基础

$AB = A+B$   $K_d = [A][B]/[AB]$  (  $K_d$ : dissociation constant )

3) 蛋白质相互作用的应用

1、利用抗原和抗体的相互作用

Western blot: 变性抗原，主要用于检测蛋白质的存在与否。



免疫共沉淀：非变性抗原，用于鉴定不同蛋白质间的相互作用。

Ab—Pr1—Pr2：免疫共沉淀 Pr1—Ab—Pr2：非特异性结合

染色质沉淀：用抗体沉淀抗原得到与抗原结合的 DNA，用 PCR 鉴定 DNA，从而证明蛋白质和特定 DNA 间的相互作用

## 2、利用已知的相互作用建立 tag

GST pull down：已知蛋白质和 GST 融合，与细胞或组织的“蛋白质汤”混合，用谷胱甘肽亲和层析分离和已知蛋白质结合的新蛋白质，即 GST tag。

Biotin—Avidin 结合：Biotin 标记已知蛋白质，通过 Avidin 结合 biotin 标记的蛋白质，从而得到与 biotin 标记蛋白结合的新蛋白质。优点是 biotin—avidin 的结合是最牢靠的；缺点是细胞内结合 biotin 的蛋白质很多。

## 3、直接利用蛋白质的相互作用

**蛋白质亲和层析**：用蛋白质当亲和层析的亲和基团。要有大量的蛋白质。

**酵母双杂交**：优点是简单；缺点是不能得到修饰调控型的相互作用蛋白质。还有衍生的酵母三杂交以及单杂交。

**phage display**：Physically 固定蛋白质，结合 phage，洗脱，扩增 phage，再结合；这样做好几轮。

**Bait 蛋白质筛表达库**：bait 蛋白质用同位素、GST、等标记，再筛库。

蛋白质组：这是检测和鉴定蛋白质的方法。不管用什么方法，只要你能得到结合的蛋白质，理论上蛋白质组都可以分析。但由于检测和鉴定方法上的提高，使得许多蛋白质间的相互作用都可以被用于发现新的蛋白质和新的相互作用。缺点是发现的相互作用蛋白质太多常常超出人们的辨别能力和研究能力。

### 1) 发现相互作用的生物学意义

{ 相互作用的证实：不同的发现相互作用的技术都可以被用来验证相互作用的蛋白质。  
相互作用的生物学功能：有重要生物学功能的相互作用、一般功能的相互作用、还是没有功能的相互作用。

**存在亲和不等于存在相互作用，存在相互作用不等于存在相应功能**

### 2) 生物学功能的研究

**功能的获得**：转基因表达，使得原本不表达该蛋白质的细胞表达了蛋白质，从而和已有的蛋白质发生相互作用，细胞有了新的功能。

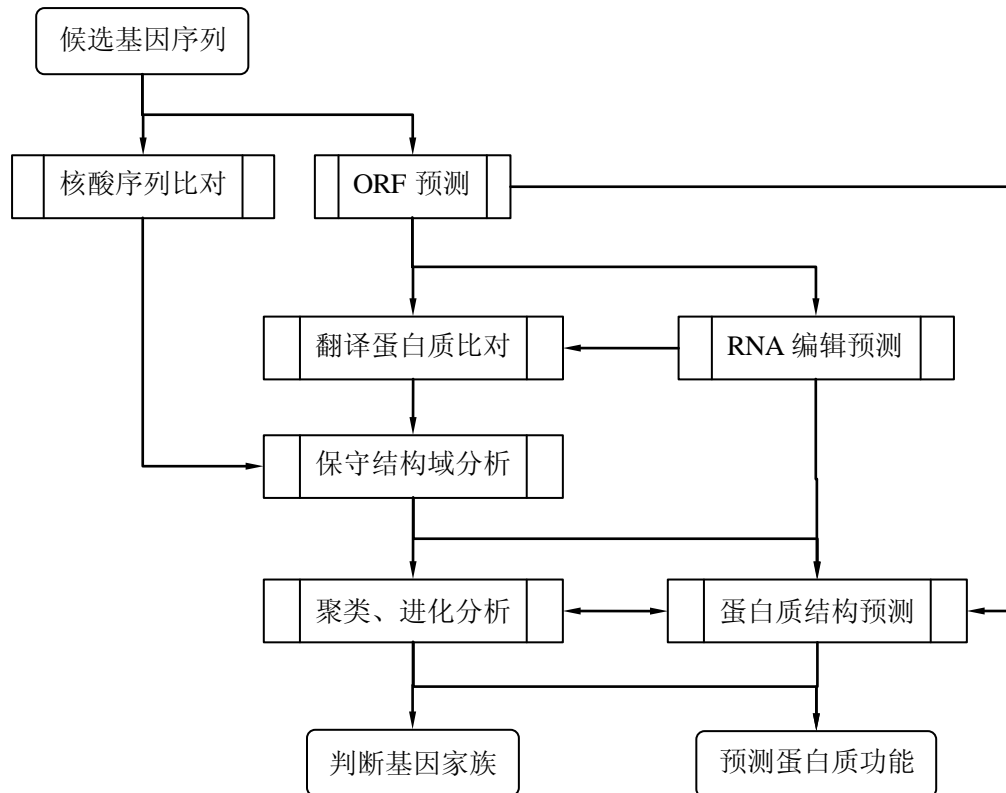
常用技术：细胞内转基因过表达蛋白质、转基因老鼠

**功能的失去**：抑制蛋白质的表达或用 dominant negative 突变体，使原有的蛋白质缺失或失活，细胞功能的丧失。

常用技术：RNA 干扰使蛋白质不表达、蛋白质部分功能片断的影响、基因敲除的细胞、基因敲除的老鼠

## 系统进化树

1. 从某个新近测序的基因组中发现了一个候选基因，需要进一步预测其功能并鉴定其是否为某个基因家族的成员。试用框图给出相关的生物信息学方法及步骤。



**核酸序列比对:** BLASTn, tBLASTx

**ORF 预测:** ORF Finder, 对找到的 ORF 进行评价 (TATA 盒, 启动子)

**RNA 编辑预测:** intro 和 exon 预测, 选择性剪切, RNA 编辑 (多框阅读)

**翻译蛋白质比对:** BLASTx, tBLASTn, tBLASTx

**保守结构域分析:** 保守区和可变区预测, 特征 Motif, 保守域序列预测

**蛋白质结构预测:** 二级结构预测, 三级结构预测 (从头预测, 同源预测)

**聚类、进化分析:** 基因家族的判定

**蛋白质功能预测:** 同源类比, 虚拟筛选或模拟生物过程

## 2. BLAST

Blastn: 核酸序列→核酸序列

Blastp: 蛋白质序列→蛋白质序列

Blastx: 核酸序列 (翻译) →蛋白质序列

Tblastn: 蛋白质序列→核酸序列 (翻译)

Tblastx: 核酸序列 (翻译) →核酸序列 (翻译)

**Blast 流程 Blast**——Basic Local Alignment Search Tool

在浏览器中输入 NCBI 网址即可进入 BLAST 的主页, 主页面按不同的搜索任务做了分类, 选择所需的类别进行搜索。进入搜索页面后, 把查询序列粘贴到“search”编辑框中。在 database 选择项中, 有若干数据库可以选择, BLAST 的缺省数据

库是 nr 数据库，如果有特别需要可以选别的数据库。点击 BLAST 之后，程序就会把序列提交给 BLAST 服务器，服务器会返回给用户一个 requestID，用于搜索 BLAST 结果。输入 ID，点击相应按钮，浏览器就会弹出新的窗口显示 BLAST 结果。

BLAST 结果通常包括三个部分：

1、序列在数据库中搜索得到的 Hits 分布图，把用户序列在数据库中能找到的匹配序列从高到底用不同颜色的线条图形表示出来。

2、在数据库中检索到的匹配序列的从高到低的排序，BLAST 显示所有 E 值小于设定的 E 值的匹配序列。这个部分包括三个数据：

-----第一列：序列的名称

-----第二列：Score（分值），用户序列和数据库序列中命中序列比对之后的相似性打分，打分越高则序列相似性越高。

-----第三列：E 值，是随机产生一个比所得分值高的对位排列的概率。是衡量分值可靠性的测度，E 值越小，所得相似序列的可靠性越高。

如果用户有大量的数据，则需要将 BLAST 程序下载到用户本地的计算机中，进行本地化分析。

### 直系同源 / 并系同源

直系同源基因(orthologous gene)是指在不同物种之间同源相似的基因（不同的物种中来源于共同的祖先），而并系同源基因(paralogous gene)是指一个物种内的同源基因（同一物种内由于基因复制而分离的同源基因）。

**双向 BLAST:** 用物种 A 的基因序列搜索(BLAST)物种 B 的基因序列，搜到的最高相似的序列反过来再 BLAST 物种 A 的基因组(阈值 E 小于  $1 \times 10^{-20}$ )。如果两条序列在这两个 BLAST 中均是最相似的，称为“双向 BLAST 最高相似蛋白”，则认为**是直系同源关系。**

**Mega-Blast:** 可用于搜索近似完全的匹配，可以处理一批核苷酸查询，比标准 BLAST 查询速度快，NCBI 进行基因组 BLAST 查询时的默认程序。

**PSI-BLAST:** 位置特定的迭代 BLAST(Position Specific Iterated BLAST)，搜索数据库以找出与查询序列同一蛋白质家族的成员，揭示亲缘关系较远的蛋白质间的关系。

### 两个序列间的查询

- 比较两个序列的相似性，不需要传统 BLAST 的数据库查询
- BLAST2
- 局部对位排列，获得结构域或序列内重复信息
- 建议不超过 150 kb

从 GenBank 等分子数据库中检索收录号为 AF166093 的 DNA 序列，用 BLAST 等工具对该序列进行同源性搜索，用所获得的序列构建进化树。

Flow:

1. 登录 NCBI 主页-点击 BLAST-点击 TBLASTX-在 Search 文本框中粘贴检测序列-点击 BLAST!-点击 Format-得到 result of BLAST。

2 . **Mega** 建 树  
构建进化树的主要步骤是比对，建立取代模型，建立进化树以及进化树评估。

(1) 首先用 Clustal X 对上述检测到的序列进行比对, 并将开始和末尾处长短不同的序列剪切整齐, 得到目标同源序列。

(2) 采用 mega 软件, 建立系统树。

关于用 Mega 建立系统树的具体步骤, 有兴趣的同学可参考  
<http://liucheng.name/603/>

2005 年考题:

背景知识 1: 干扰素 (Interferon, IFN) 是一个具有多种生物学功能的蛋白质家族, 具有抗病毒、细胞生长抑制和免疫调节的作用。目前, 将 IFN 分为 2 个亚型: I 和 II, 其中 I 型基因分为  $\alpha$ 、 $\beta$ 、 $\omega$ 、 $\kappa$ 、 $\delta$  和  $\tau$ 。对所有哺乳动物研究发现,  $\alpha$  干扰素家族成员均没有内含子, 而且序列相似性高。人类 IFN- $\alpha$  2 的序列号是 NM\_000605。

背景知识 2: 直系同源基因是指在物种形成过程中从祖先物种“继承”的基因, 其在结构和功能等方面都具有保守性; 并系同源基因是存在于同一个基因组中的同源基因。可以根据并系同源基因起源时间与物种形成时间的关系, 进一步划分并系同源基因, 其中“inparalog”基因特指那些物种形成之后形成的同源基因。

题目:

根据上述背景知识, 用 BLAST 工具搜索  $\alpha$ -干扰素基因相关序列, 并据此构建哺乳动物  $\alpha$ -干扰素基因的进化树。此外, 根据并系同源基因的概念讨论  $\alpha$ -干扰素基因的进化模式。

(提示: 物种分类可以参考在线 BLAST 结果中的 Taxonomy report)

系统树的建立过程同上。

并系同源基因概念: 并系同源基因 (paralogous gene) 是指一个物种内的同源基因。一个生物物种的基因组中, 两个基因或开放读码在各自全长的 60% 以上范围内, 同一性不少于 30% 时, 称为同源体。

进化模式的讨论: 根据所建立的系统进化树, 讨论哺乳动物  $\alpha$ 、 $\beta$ 、 $\omega$ 、 $\kappa$ 、 $\delta$  和  $\tau$  编码基因, 通过比较  $\alpha$  等亚型建立系统进化树确定其同源关系。

补充资料: 系统进化树构建及数据分析的简介:  
<http://liucheng.name/577/>

## 好好思考一下第 2 题

### 1. 肖华胜

请简述基因芯片的原理及在系统生物学中的应用，并设计一个用基因芯片研究肿瘤分子分型的实验方案。

原理：碱基互补配对

基因芯片(Gene chip)技术是指通过 (Microarray) 技术将高密度 DNA 片段通过高速机器人或原位合成方式以一定的顺序或排列方式使其附着在如膜、玻璃片等固相表面，以同位素或荧光标记的 DNA 探针，借助碱基互补杂交原理，进行大量的基因表达及监测等方面研究的最新革命性技术。

应用：基因组比较，基于芯片的比较基因组杂交：应用基因芯片研究 DNA 结构变化、拷贝数变化。待检 DNA 和参照 DNA 在被标记后，点样于完全基因芯片上进行杂交，与其比较可得出被删除或被扩增的区域。

(病理：SNP)；ChIP on Chip (蛋白质与 DNA 相互作用)；表达谱分析 (病理)

Biomarker 搜寻，药物筛选与药理研究，基因组分析等

实验设计：

探针：肿瘤分型标记，PCR 制备，overlapped

模板：标记的肿瘤 cDNA

结论：肿瘤分子分型

实验设计：

1) 目的；2) 对象；3) 所用库、材料；

4) 根据 1、2、3 选择方法，并确定方法相应的检测手段与标准

5) 结果与方法学验证

2 请结合自己的研究课题设计一个用基因芯片研究基因表达和功能的实验方案并简述主要的实验过程

3 简述生物芯片的种类和各自的原理，用高覆盖率基因表达谱芯片研究癌和癌旁组织的基因表达为例，设计一套实验方案

没做过相关实验的话，你只能理解着记吧

## 1 promoter analysis (启动子分析)

其主要原理是将不同长度的 DNA 序列通过基因工程的方法连入有报告基因的质粒载体中，然后将质粒转入受体细胞（如大肠杆菌），然后可以通过检测报告基因的表达情况（mRNA 或蛋白表达水平或酶活）即可确定启动子的核心序列。报告基因表达越高说明启动子活性越强。

## 2 Northern Hybridization

此技术主要用来检测 RNA 的表达水平及大小，利用的原理是碱基互补配对形成双链（RNA-RNA RNA-DNA DNA-DNA），即用放射性标记的探针检测目标序列的技术。其主要步骤及注意事项课件上已经说的很清楚，此不再赘述。

## 3 Ribonuclease Protection

核糖核酸酶保护实验(Ribonuclease protection assay, RPA)是近十年发展起来的一种全新的 mRNA 定量分析方法。

其基本原理是将标记的特异 RNA 探针(32P 或生物素)与待测的 RNA 样品液相杂交，

标记的特异 RNA 探针按碱基互补的原则与目的基因特异性结合，形成双链 RNA；未结合的单链 RNA 经 RNA 酶 A 或 RNA 酶 T1 消化形成寡核糖核酸，而待测目的基因与特异 RNA 探针结合后形成双链 RNA，免受 RNA 酶的消化，故该方法命名为 RNA 酶保护实验。

对于 32P 标记的探针，杂交双链进行变性聚丙烯酰胺凝胶电泳，用放射自显影或磷屏成像系统检测被保护的探针的信号；对于生物素标记的探针，杂交双链经过变性聚丙烯酰胺凝胶电泳后电转移至尼龙膜，采用链霉亲和素-辣根过氧化物酶（Streptavidin-HRP）和化学发光底物与膜上 biotin 标记的探针结合，X 射线胶片或化学发光图像分析仪检测杂交信号。

## 4, Real time PCR

实时定量 PCR (real-time quantitative PCR)是指在 PCR 指数扩增期间通过连续监测荧光信号强弱的变化来即时测定特异性产物的量，并据此推断目的基因的初始量，不需要取出 PCR 产物进行分离。在 PCR 反应体系中加入荧光集团，利用荧光信号的变化实时监测 PCR 扩增反映中每一个循环扩增产物量的变化，通过 Ct 值和标准曲线的分析对起始模板进行定量分析。

与常规 PCR 相比，它具有以下优点：

1，封闭反应，无需 PCR 后处理

2，特异性强，灵敏度高，有效解决 PCR 污染问题、自动化程度高

3，采用对数期分析，摒弃终点数据，定量准确

简单的说：  
探针保护目的片段；检测目的片段目的：定量检测——其定量灵敏



- 4, **定量范围宽**, 可达到10个数量级
- 5, **仪器在线式实时检测**, 结果直观, 避免人为判断
- 6, 可实现一管双检或多检
- 7, 操作安全, 缩短时间, 提高效率。

高效、在线、范围广、灵敏、封闭、准确

主要实验步骤:

**1, 反转录:** 依赖反转录酶将 RNA 反转录成 cDNA

**2, 扩增:** 用 PCR 的方法扩增 cDNA

**3, 检测:** 实时检测和定量扩增的产物

实时荧光定量 PCR 技术的主要应用:

1. DNA 或 RNA 的绝对定量分析: 包括病原微生物或病毒含量的检测, 转基因动植物转基因拷贝数的检测, RNAi 基因失活率的检测等
2. **基因表达差异分析:** 例如比较经过不同处理样本之间特定基因的表达差异(如药物处理、物理处理、化学处理等), 特定基因在不同时相的表达差异以及 cDNA 芯片或差显结果的确证
3. 基因分型: 例如 SNP 检测, 甲基化检测等

**Realtime PCR 常用的两种方法分别为: Sybr green (荧光染料掺入法) 和 Taqman probe (探针法)**

1, SYBR green 荧光染料

在 PCR 反应体系中, 加入过量 SYBR 荧光染料, SYBR 荧光染料特异性地掺入 DNA 双链后, 发射荧光信号, 而不掺入链中的 SYBR 染料分子不会发射任何荧光信号, 从而保证荧光信号的增加与 PCR 产物的增加完全同步。

此方法适用:

- 1、**灵敏度高:** 使用 SYBR 可使荧光效果增强到1000倍以上
- 2、**通用性好, 不需要设计探针, 方法简便, 省时, 价格低廉。**
- 3、**通用型方法**, 在国内外科科研中普遍使用。
- 4、高通量大规模的定量 PCR 检测
- 5、**专一性要求不高的定量 PCR 检测。**

SYRB法要排除非特异性扩增的干扰, 而 TaqMan探针则可排除此问题

2, Taqman Probe 荧光探针

PCR 扩增时在加入一对引物的同时加入一个特异性的荧光探针, 该探针为一寡核苷酸, 两端分别标记一个报告荧光基团和一个淬灭荧光基团。探针完整时, 报告基团发射的荧光信号被淬灭基团吸收; PCR 扩增时, Taq 酶的 5' -3' 外切酶活性将探针酶切降解, 使报告荧光基团和淬灭荧光基团分离, 从而荧光监测系统可接收到荧光信号, 即每扩增一条 DNA 链, 就有一个荧光分子形成, 实现了荧光信号的累积与 PCR 产物形成完全同步

此方法适用:

- 1、**具有高适应性和可靠性, 实验结果稳定重复性好, 特异性更高。**
- 2、**适用于扩增序列专一的体系的检测。**

### 3、样品中靶基因含量过低的定量 PCR 检测。

- 4、靶基因的特异序列较短，无论怎样优化引物设计条件都不能解决。
- 5、存在与靶基因同源的序列，在 PCR 中容易出现非特异性扩增，对特异性要求较高的定量。
- 6、广泛用于人类传染病的诊断和病原定量，在动物病原体基因的检测，畜禽产品的检验检疫，生物制品的鉴定。

## 5, siRNA knockdown of a gene

### RNAi 实验方法：

#### 一、RNAi 的分子机制

通过生化和遗传学研究表明，RNA 干扰包括起始阶段和效应阶段 (initiation and effector steps)。在起始阶段，加入的小分子 RNA 被切割为 21-23 核苷酸长的小分子干扰 RNA 片段 (small interfering RNAs, siRNAs)。证据表明：一个称为 Dicer 的酶，是 RNase III 家族中特异识别双链 RNA 的一员，它能以一种 ATP 依赖的方式逐步切割由外源导入或者由转基因，病毒感染等各种方式引入的双链 RNA，切割将 RNA 降解为 19-21bp 的 **双链 RNAs (siRNAs)**，每个片段的 3' 端都有 2 个碱基突出。

在 RNAi 效应阶段，**siRNA 双链结合一个核酶复合物从而形成所谓 RNA 诱导沉默复合物 (RNA-induced silencing complex, RISC)**。激活 RISC 需要一个 ATP 依赖的将小分子 RNA 解双链的过程。**激活的 RISC 通过碱基配对定位到同源 mRNA 转录本上，并在距离 siRNA 3' 端 12 个碱基的位置切割 mRNA**。尽管切割的确切机制尚不明了，但每个 RISC 都包含一个 siRNA 和一个不同于 Dicer 的 RNA 酶。另外，还有研究证明含有启动子区的 dsRNA 在植物体内同样被切割成 21-23nt 长的片段，这种 dsRNA 可使内源相应的 DNA 序列甲基化，从而使启动子失去功能，使其下游基因沉默。

#### 二、进行 RNAi 试验

##### (一) siRNA 的设计

1. 在设计 RNAi 实验时，可以先在以下网站进行目标序列的筛选：

<http://www.genesil.com/business/products/order2.htm>

[http://www.ambion.com/techlib/misc/siRNA\\_finder.html](http://www.ambion.com/techlib/misc/siRNA_finder.html)

<http://www.ic.sunysb.edu/Stu/shilin/rnai.html>

<http://design.dharmacon.com/rnadesign/default.aspx?SID=45358710>

##### 2. **RNAi 目标序列的选取原则：**

(1) 从转录本 (mRNA) 的 AUG 起始密码开始，寻找“AA”二连序列，并记下其 3 端的 19 个碱基序列，作为潜在的 siRNA 靶位点。有研究结果显示 GC 含量在 45%—55% 左右的 siRNA 要比那些 GC 含量偏高的更为有效。Tuschl 等建议在设计 siRNA 时 **不要针对 5 和 3 端的非编码区 (untranslated regions, UTRs)**，原因是这些地方有丰富的调控蛋白结合区域，而这些 UTR 结合蛋白或者翻译起始复合物可能会影响 siRNP 核酸内切酶复合物结合 mRNA 从而影响 siRNA 的效果。

与 primer 设计类似，应避免连续的 AAAA 序列，以及自身互补等

(2)将潜在的序列和相应的基因组数据库（人，或者小鼠，大鼠等等）进行比较，排除那些和其他编码序列/EST 同源的序列。例如使用 BLAST

（[www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)）

(3)选出合适的目标序列进行合成。通常一个基因需要设计多个靶序列的 siRNA，以找到最有效的 siRNA 序列。

### 3. 阴性对照

一个完整的 siRNA 实验应该有阴性对照，作为阴性对照的 siRNA 应该和选中的 siRNA 序列有相同的组成，但是和 mRNA 没有明显的同源性。通常的做法是将选中的 siRNA 序列打乱，同样要检查结果以保证它和目的靶细胞中其他基因没有同源性。

### 4. 目前已证实的 siRNA 可以在下面的网页找到：

<http://design.dharmacon.com/catalog/category.aspx?key=49>

[http://www.ambion.com/techlib/tb/tb\\_502.html](http://www.ambion.com/techlib/tb/tb_502.html)

<http://web.mit.edu/mmcmanus/www/siRNADB.html>

[http://python.penguindreams.net/Order\\_Entry/jsp/BrowseCatalog.jsp?Category=Published](http://python.penguindreams.net/Order_Entry/jsp/BrowseCatalog.jsp?Category=Published)

## (二) siRNA 的制备

目前为止较为常用的方法有通过化学合成，体外转录，长片断 dsRNAs 经 RNase III 类降解（e.g. Dicer, E. coli, RNase III）体外制备 siRNA，以及通过 siRNA 表达载体或者病毒载体，PCR 制备的 siRNA 表达框在细胞中表达产生 siRNA。

### 1. 化学合成生产 siRNA

优点：方便，研究人员几乎不需要做什么工作。

缺点：价格较贵，效率只有转录合成的 shRNA 的 1/10-1/40，基因抑制持续时间短，对细胞毒性大，转染效率低，此外由于合成工艺上存在不可弥补的缺陷，此方法不能合成 shRNA，不能纠正合成中产生的 20%左右的碱基错误。

适用于：建议不再采用此种合成方法。

不适用于：基本上对目前所有的实验室来说均不适用。

### 2. 生物合成转录生产编码 siRNA

优点：价格较低，抑制效率高，低浓度的 siRNA 即可达到抑制效果。体外转录合成，比较接近生理状态。

缺点：无法保证有效性，实验规模受到限制，不能进行大量的生产，不能维持长时间抑制效应，而且需要用户参与操作，难以保证实验的一次成功性。

适用于：筛选 siRNAs 特别是需要制备多种 siRNAs，考虑合成价格时。

不适用于：实验需要大量的一个特定的 siRNA。长期研究。

尽管有些公司推出了 shRNA 和 siRNA 试剂盒，但这些试剂盒大都突出操作简单，毕竟涉及 RNA 操作，用户自己难免会产生很多预料不到的困难，上海迪奥生物科技有限公司可以完全为您解决后顾之忧，您可以直接拿到有效的 siRNA 或 shRNA 序列。

### 3. Dicer 酶法生产 siRNA

优点：可以跳过筛选与检测有效 siRNA 序列的步骤，节省时间和开支。

缺点：有可能引发非特异性的基因沉默，尤其是同源或者密切相关的基因。

适用于：快速而经济地研究某个基因功能缺失的表型

不适用于：长期研究项目，或者是需要一个特定的 siRNA 进行研究

#### 4. 编码 shRNA 的载体生产 siRNA

优点：制备质量可控性好，带有抗生素标记的载体可以在细胞中持续抑制靶基因的表达，持续数星期甚至更久，可以大量扩增。

缺点：转录后的 shRNA 的正确折叠率不能保证，有可能因此造成非特异性抑制，质粒载体转染效率不稳定，与细胞类型有关。

适用于：已知一个有效的 siRNA 序列，需要维持较长时间的基因沉默，或者需要用抗生素筛选能表达 siRNA 的细胞。长期研究。

不适用于：筛选 siRNA 序列（其实主要是指需要多个克隆和测序等较为费时、繁琐的工作

#### 5. 病毒粒子生产 shRNA

优点：易于制备、纯化、浓缩，宿主范围广，感染率高，理化性质稳定，不整合宿主基因组，遗传毒性和免疫毒性低，是目前最好的 RNA 干扰技术之一。

缺点：耗费时间比较长，即使是用最快的也得60天左右，对实验条件要求较高，容量偏小，有可能伴随部分炎症反应。

适用于：转染效率低无法解决，需要维持较短时间的基因沉默的 siRNA

不适用于：大量筛选 siRNA 或长时间的研究，主要原因是价格因素

#### 6. PCR 制备 shRNA 表达框生产 siRNA

优点：成本较低，方便快捷，可以用于有效序列的筛选，可以用于质粒载体构建。

缺点：转染效率低，不适合大规模制备。

适用于：筛选 siRNA 序列，在克隆到载体前筛选最佳启动子

不适用于：长期抑制研究。

### 开展 RNAi 的实验过程：

#### 1. 确立需要干扰的目标基因

#### 2. 检查使用细胞的转染情况

可以选择一个4kb 左右大小的质粒和可实施的转染方法，检测转染效率。

转染效率超过40%：可以选择任何一种 RNA 干扰方法。

转染效率介于10%-40%：可以选用体外合成的 si RNA 或者带有筛选标记的 shRNA 质粒表达载体。

转染效率低于10%：带有筛选标记的 shRNA 质粒表达载体或病毒载体。

#### 3. si RNA 序列选择设计

制备20-30个 si RNA 序列，选择有效序列或委托专业化公司设计（3个序列保证1个有效）

#### 4. 开始实验

近年来的研究表明，将与 mRNA 对应的正义 RNA 和反义 RNA 组成的双链 RNA(dsRNA)导入细胞，可以使 mRNA 发生特异性的降解，导致其相应的基因沉默。这种转录后基因沉默机制(post-transcriptional gene silencing, PTGS)被称为 RNA 干扰 (RNAi)。

## 凝胶迁移或电泳迁移率实验 (EMSA)

凝胶迁移或电泳迁移率实验 (EMSA-electrophoresis mobility shift assay) 是一种研究蛋白与其相互结合的 DNA 序列之间相互作用的技术, 可用于定性和定量分析。这一技术最初用于研究 DNA 结合蛋白, 目前已用于研究 RNA 结合蛋白和特定的 RNA 序列的相互作用。

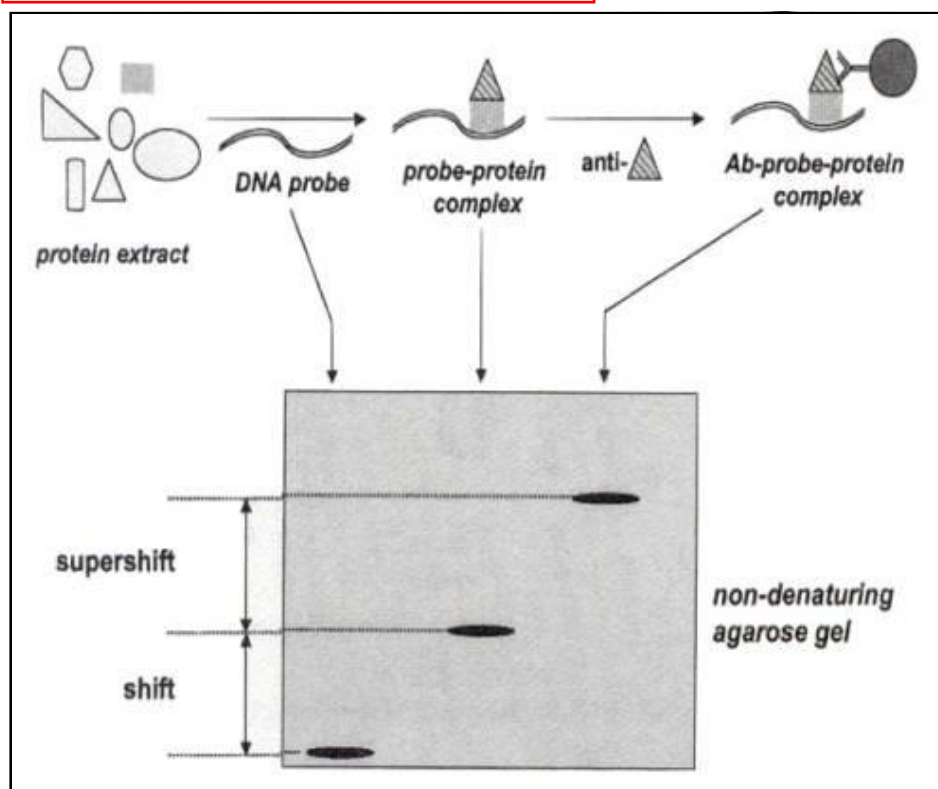
**原理:** 通常将纯化的蛋白和细胞粗提液和  $^{32}\text{P}$  同位素标记的 DNA 或 RNA 探针一同保温, 在非变性的聚丙烯酰胺凝胶电泳上, DNA-复合物或 RNA-复合物比非结合的探针移动得慢, 根据目的条带的位置分离出复合物, 随后将凝胶干燥并进行放射自显影。

**探针的标记:** 同位素标记的探针依研究的结合蛋白的不同, 可是双链或者是单链。

- 1、DNA 核苷酸探针用  $^{32}\text{P}$  和 T4 多核苷酸激酶来作末端标记;
- 2、同位素标记的 RNA 用噬菌体 RNA 聚合酶和同位素标记的核苷酸在体外合成

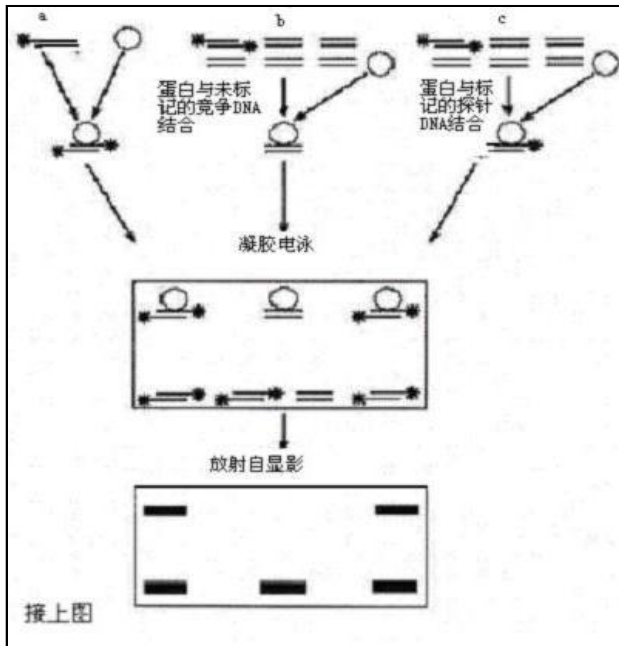
下面两幅图分别描述了 EMSA 的流程。

第一张图是用已知的 DNA 去拉未知的蛋白  
第二张图是用已知的蛋白拉未知的 DNA



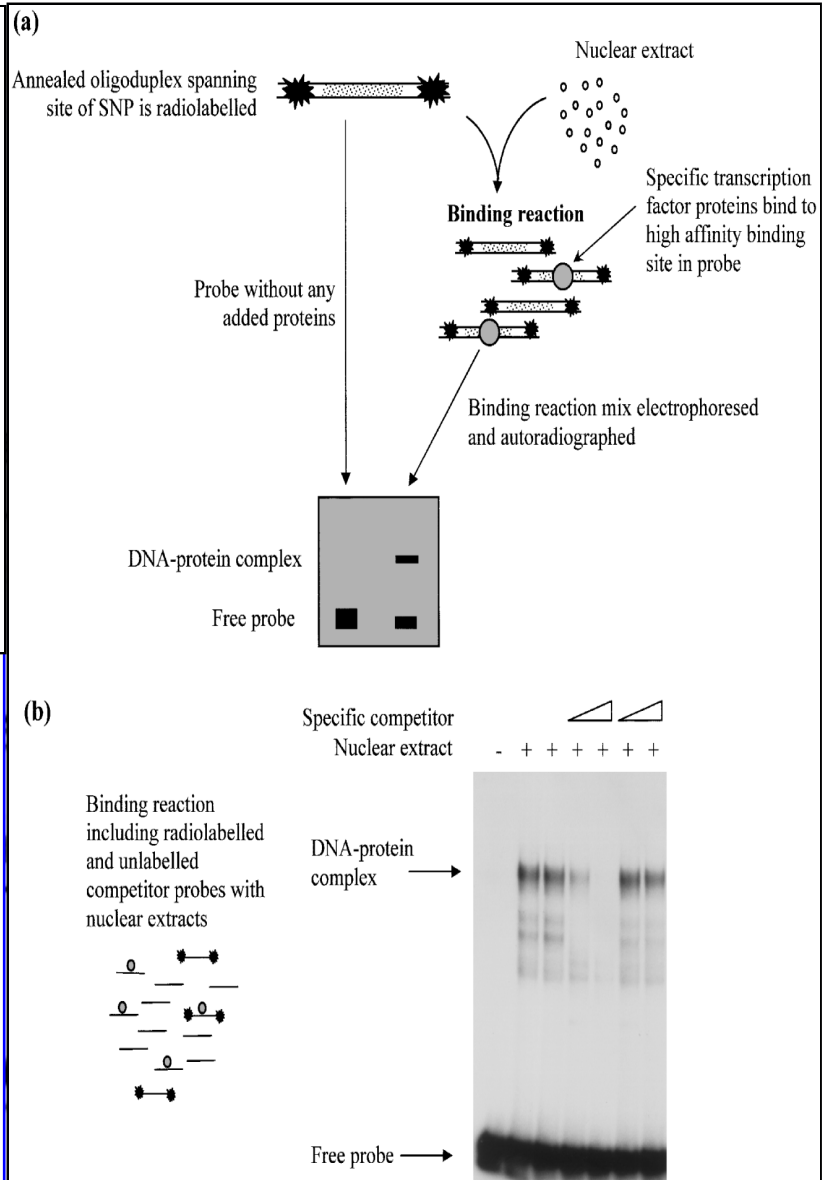
凝胶迁移实验基本原理图

标记了的DNA由于与同一种细胞蛋白结合, 于是在凝胶电泳中移动速度变慢, 放射自显影中呈滞后的带条。



如图所示，已知DNA，则合成标记的DNA probe，Free probe迁移最快，与蛋白结合后变慢，而加入相关蛋白的抗体，则更慢。而已知蛋白来探测未知DNA：

- (1) 蛋白与其相结合DNA的标记probe结合
- (2) 蛋白与竞争性DNA（未标记）结合，标记量减少，以此说明DNA确实与蛋白结合；此外，若再用非相关的DNA片段来竞争，则不会引起减少（排除假阳性）



### EMSA 的应用：

- 1、可用于测定转录因子结合的最短的 DNA 序列
- 2、基于 EMSA 的方法之上，我们可以用 DNA 亲和色谱来纯化特异性的 DNA 结合蛋白

### EMSA 的优缺点：

- 优点：1、可以分离各种不同类型的复合物，单体或者二聚体
- 2、DNA-蛋白质复合物的位置很容易识别并分离来进行下一步实验
- 不足：1、是一种体外的实验方法，不能完全反应体内的 DNA 和蛋白质相互作用的情况。
- 2、不能知道结合的 DNA 序列，需要进一步借助其他方法

EMSA 和 SDS-PAGE 的主要区别：前者用的是非变性胶，后者用的是变性胶

### Foot Printing（足迹法）

一般用 DNA-Foot printing

是一种用来测定特异性的 DNA 序列结合蛋白的方法，它可以用来研究蛋白质



-DNA 相互作用，体外和细胞内均可。

该方法的流程图如下：

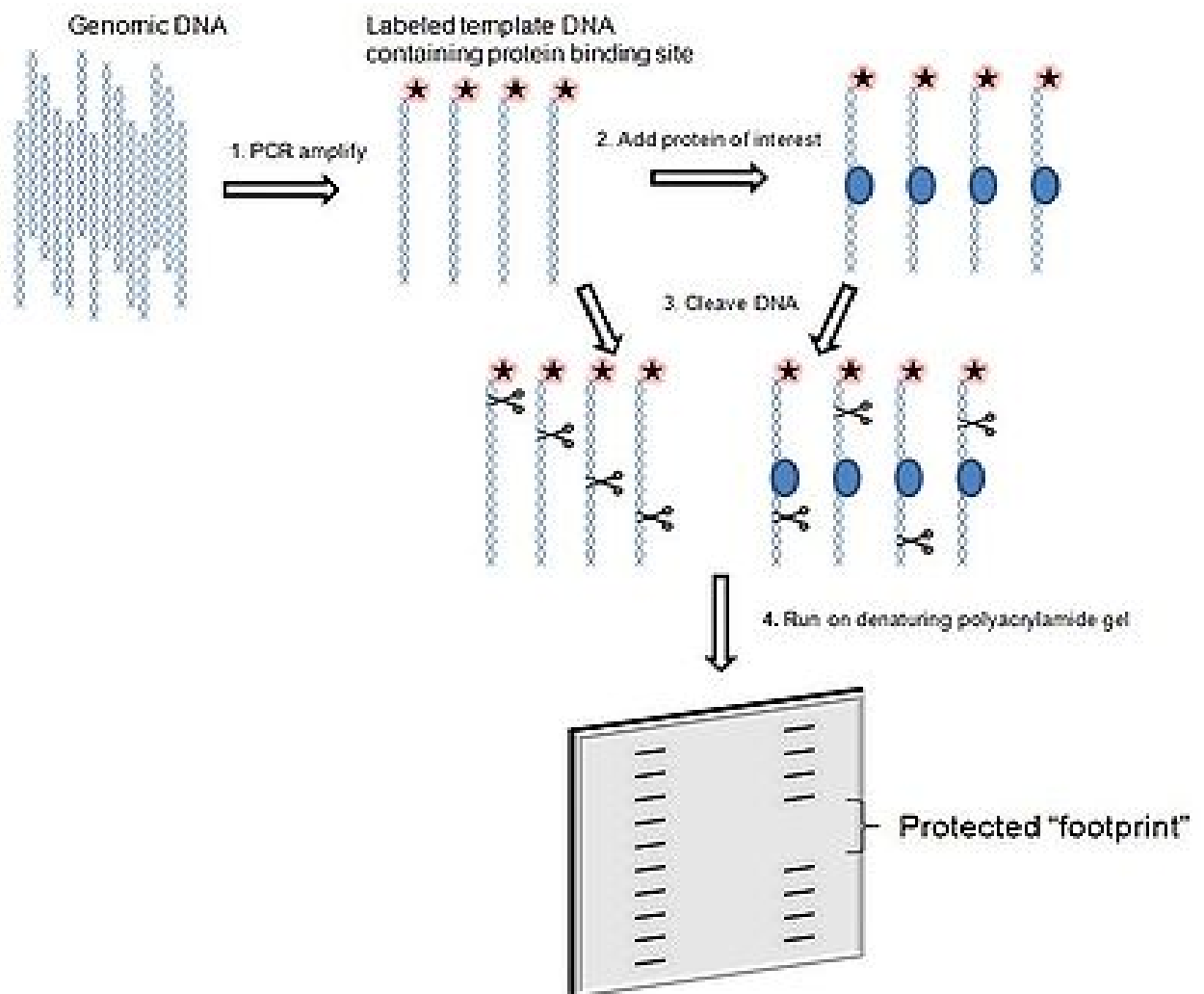
步骤为：1、PCR 扩增模板 DNA，理想的扩增子的大小在 50-200bp 之间，并进行标记

2、在模板 DNA 中加入待研究的蛋白，蛋白会和相应的 DNA 序列结合。

（注意留下部分模板 DNA 不加蛋白质作为对照）

3、加入核酸外切酶或化学剪切试剂，可以随意酶切序列，被蛋白质结合的 DNA 序列由于受到了保护而免于遭到酶切

4、聚丙烯酰胺凝胶电泳进行跑胶，对照的模板 DNA 组由于没有蛋白保护形成了梯度图谱，而有蛋白结合的实验组会有一段没有条带，类似脚印，所以叫足迹法。



标记法有两种：放射性标记和荧光标记

使用的剪切试剂和方法有：

1、**Dnase I**：是一种双链核酸内切酶，可加入 EDTA 随时终止它的的剪切反应；缺点是它的剪切位点不是随机的，会受到局部 DNA 结构的影响，因此得到的 DNA 图谱可能是不均一的梯度，而限制了结合位点分析的准确性

2、**羟自由基**：羟基可以和 DNA 结合使其断裂，因为羟基很小，所以得到的 DNA footprinting 有很高的分辨率，缺点是反应进行的时间非常长

3、**紫外照射**：可用于 in vivo foot printing，紫外线可穿透细胞膜进行 DNA 剪切而不损害细胞状态，然后将细胞裂解，分离出 DNA-蛋白质复合物进行分析，可反应 in vivo 的状态

### **Chromosome IP (染色质免疫沉淀 CHIP)**

这是目前唯一研究体内蛋白质与 DNA 相互作用的方法。

**原理**：在活细胞状态下固定 DNA-蛋白质复合物，并将其随机切断为一定长度范围内的染色质小片段，然后通过免疫学方法沉淀此复合物，特异性的富集目的蛋白结合的 DNA 片段，通过对目的片段的纯化与检测，从而获得蛋白质 DNA 相互作用的信息。

**应用**：1、可以检测体内反式因子与 DNA 的动态作用

2、可以用来研究组蛋白的各种共价修饰与基因表达的关系

3、CHIP 与基因芯片相结合建立的 CHIP-on-chip 方法已广泛用于特定反式因靶基因的高通量筛选

4、CHIP 与体内足迹法相结合，用于寻找反式因子的体内结合位点

5、RNA-CHIP 用于研究 RNA 在基因表达调控中的作用

ChIP 的一般流程：

甲醛处理细胞---收集细胞，超声破碎---加入目的蛋白的抗体，与靶蛋白-DNA 复合物相互结合---加入 ProteinA，结合抗体-靶蛋白-DNA 复合物，并沉淀---对沉淀下来的复合物进行清洗，除去一些非特异性结合---洗脱，得到富集的靶蛋白-DNA 复合物---解交联，纯化富集的 DNA 片段---PCR 分析

**ChIP 的不足**：1、需要一个特异性蛋白质抗体，有时难于获得；

2、为了获得高丰度的结合片段，必须实验演示胞内条件下靶蛋白质的表达情况；

3、调控蛋白质的基因的获取可能需要限制在组织来源中。

### **RNase Protective Assay (RNA 酶保护实验 RPA)**

**原理**：标记的特异 RNA 探针按碱基互补的原则与目的基因特异性结合，形成双链 RNA；未结合的单链 RNA 经 RNA 酶 A 或 RNA 酶 T1 消化形成寡核糖核酸，而待测目的基因与特异 RNA 探针结合后形成双链 RNA，免受 RNA 酶的消化，故该方法命名为 RNA 酶保护实验。

对于 <sup>32</sup>P 标记的探针，杂交双链进行变性聚丙烯酰胺凝胶电泳，用放射自显影或磷屏成像系统检测被保护的探针的信号；对于生物素标记的探针，杂交双链经过变性聚丙烯酰胺凝胶电泳后电转移至尼龙膜，采用链霉亲和素-辣根过氧化物酶 (Streptavidin-HRP) 和化学发光底物与膜上 biotin 标记的探针结合，X 射线胶片或化学发光图像分析仪检测杂交信号。

**关于探针的制备**：

探针的制备是将要研究的目的基因的部分片段克隆到一个载体上，并且使用 SP6, T3 或 T7 启动子来控制它的表达，这些启动子被 DNA 依赖的 RNA 聚合酶所识别，选择正确的体外转录试剂盒，转录出要研究的 RNA 的互补链，再纯化得到，制备过程相对比较麻烦。

**PRA 的优点**：1、检测灵敏度高，步骤少，数据真实性较高

2、RNA-RNA 杂交体系的稳定性较高，无探针自身复性的现象。

不足：1、核酸酶消化时，酶量的控制困难，容易消化过头，而没有条带。  
2、如上所说，探针的制备比较麻烦。

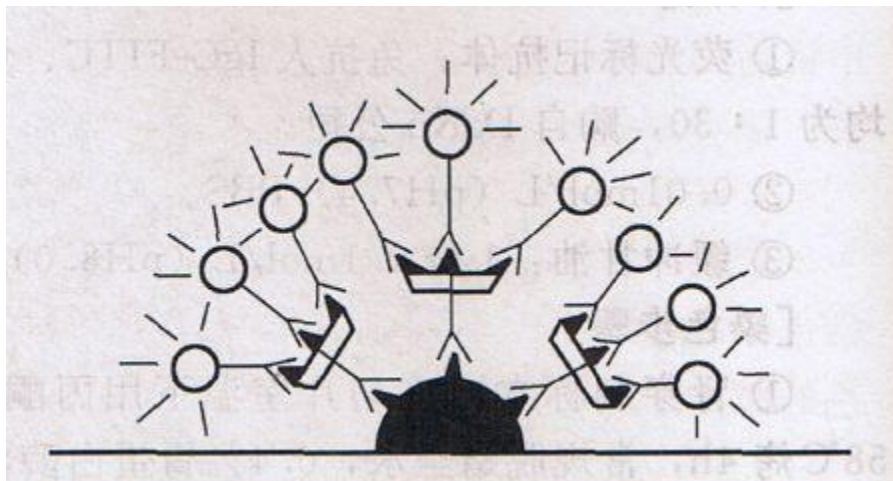
### 免疫荧光染色 (Immunofluorescent staining)

原理：其原理类似于 western blot，一抗与待检测的蛋白质特异性结合，连接有不同荧光基团的二抗和一抗相结合，在荧光显微镜下可以观察到不同的荧光，从而显示目的基因的表达情况。免疫荧光由于有较高的敏感性可以显示出基因表达的亚细胞的结构，所以通常用来作为基因定位的方法。



#### 荧光抗体染色方法：

- 1、**直接法**：用特异荧光抗体直接滴加于待检测抗原标本上，抗原抗体发生特异性结合，产生荧光，根据荧光分布和形态确定抗原性和部位。
- 2、**间接法**：用一抗和抗原结合，再用荧光标记的二抗与一抗结合，间接地显示出组织和细胞中抗原的位置。
- 3、补体法：抗原抗体反应时加入补体，补体结合于抗原抗体复合物上，用抗补体 C3 等荧光抗体直接作用于组织切片，形成抗原-抗体-补体-抗补体荧光抗体复合物，在荧光显微镜下显示阳性荧光的部位就是免疫复合物上补体存在处，常用于肾穿刺组织活检诊断。（如图所示）



- 应用：
- 1、血液中 T 及 B 细胞及其亚群的鉴定
  - 2、血清中自身抗体的检测
  - 3、组织中免疫球蛋白及其补体组分的检测
  - 4、恶变组织中肿瘤特异性抗原的鉴定
  - 5、特殊染色体的鉴定等。

1. Please describe the differences and similarities of the two basic initial phase determination methods, MAD (multi-wavelength anomalous dispersion) and MIR (multiple isomorphous replacement) in 1-2 pages.

**MAD (多波长反常散射):** 晶体衍射中有一条弗里德耳定律, 就是说不论晶体中是否存在对称中心, 在晶体衍射中总存在着对称中心, 也即有  $F_{HKL} = F_{\bar{H}\bar{K}\bar{L}}$ 。但是当使用的 X 射线波长与待测样品中某一元素的吸收边靠近时, 就不遵从上述定律, 也即  $F_{HKL} \neq F_{\bar{H}\bar{K}\bar{L}}$ 。这是由电子的反常散射造成的, 利用这一现象可以解决待测物的相角问题。一般, 这一方法常与重原子同晶置换法结合使用。在收得同晶置换物的衍射数据后, 改变入射线波长至靠近重原子的吸收边处, 再次收集数据, 这套数据是存在反常散射的, 可利用这两套数据来求位相。有如多同晶置换法, 如采用几个不同波长的 X 射线, 对所含不同元素收集几套反常散射数据, 则可得更正确、更完整的相位信息, 是为多波长反常衍射法(MAD)。

**MIR (多同晶置换):** 把对 X 射线散射能力大的重金属原子作为标识原子。这种置换入重原子的大分子应与无重原子时的原晶体有相同的晶胞参数和空间群, 且绝大多数原子的位置相同, 故称同晶置换。从这些含重原子晶体的衍射数据, 利用基于派特逊法的方法可解出重原子的位置, 据此算出其结构因子和相角, 进而利用相角关系计算出没有重原子的原晶体的相角, 解出结构。经常使用不只一种重原子进行置换, 以得几种同晶置换衍生物, 称多对同晶置换法。

两者的相同点: 都是利用重原子的特性来解决相角问题。

两者的差别: MAD 是基于 MIR 的基础之上的, 采用多种波长完备所需的信息。

2. 运用蛋白质结晶学解析蛋白质结构时遇到的最大难题是什么? 列举两种常用的解决此问题的实验方法, 并简单说明其方法的基本原理. 2004

在大分子结晶学中, 相角问题, 也就是相角的确定是最困难的难题之一。在进行 X 射线晶体结构分析时, 数据收集只能纪录各个衍射线的强度, 而与振幅相对应的 X 射线散射的相位不能被直接测定, 因此衍射线的相角信息必须通过其他方法来进行测定, 这就是相角问题。一般采用多重同晶置换 (MIR)、多波长反常散射 (MAD) 等技术确定初始相位。

3. 举例说明几种结晶蛋白质的方法

预测: 研究蛋白质 3D 结构的重要性。

测定生物大分子三维结构所使用的三种主要实验技术方法

1. 晶体 X 射线衍射 (X-ray diffraction) 技术

所能测定的生物大分子的分子量范围是很宽的, 可以从 1kDa 以下到 400kDa 甚至更大, 而且技术的成熟度比较高、应用成本比较低廉。是必须预先制备出适合于 X 射线衍射的单晶体。

2. 溶液多维核磁共振 (NMR) 技术

该技术不需要预先制备晶体。该方法技术正在酝酿着革命性的突破, NMR 技术的应用成本远高于晶体 X 射线衍射技术。适用于分子量 < 30kDa 的可溶性蛋白, 须同位素标记; 优点是动态的溶液结构, 更接近于生理状态。

节省环节: 数据收集和处理 (几个星期), 12 NMR 数据/年 (600MHz 仪器)。

3. 低温电子显微镜三维电子衍射图象重构技术

非常适合测定分子量非常巨大的生物大分子的复合体, 例如病毒、膜蛋白的复合体等等。但分辨率低, 方法还有待发展。

即X光衍射 (成熟、但单晶要求高); 多维 NMR (分子量有限、须标记; 但保证了生理状况); 电子显微镜三维电子衍射 (分析分子量大、可保持天然构象、但目前分辨率低)