



中华人民共和国国家知识产权局  
STATE INTELLECTUAL PROPERTY OFFICE OF THE PEOPLE'S REPUBLIC OF CHINA

## 郑重声明

- 一、经授课教师同意，本课件仅作为交流学习使用，并欢迎广泛传播，但禁止作为商业用途。
- 二、在交流使用过程中，请尊重版权。
- 三、课件中涉及的观点仅代表授课教师本人立场。
- 四、使用课件中的数据、图表时请注明来源，保证完整性，避免断章取义。
- 五、课件中涉及的政策法规或其它信息的有效性，请以相关主管部门(单位)公布为准。



欢迎关注微信公众号  
“专利文献众享”或扫  
描左侧二维码，获取  
最新公益讲座信息及  
专利文献服务。



欢迎添加微信  
331546945，  
加入微信交流  
群，获取更多  
多公益讲座资  
讯和帮助。

**公益讲座**  
[www.sipo.gov.cn/wxfw](http://www.sipo.gov.cn/wxfw)



中华人民共和国国家知识产权局  
STATE INTELLECTUAL PROPERTY OFFICE OF THE PEOPLE'S REPUBLIC OF CHINA

# 医药生物领域专利检索策略

国家知识产权局专利局  
医药生物发明审查部  
杨振宇



欢迎关注微信公众号“专利文献众享(patdoc-sipo)”或扫描左侧二维码，  
获取最新公益讲座信息及专利文献服务。



欢迎添加微信  
“331546945”，  
加入微信交流群，  
获取最多公益讲座  
资讯和帮助。

**公益讲座**  
[www.sipo.gov.cn/wxfw](http://www.sipo.gov.cn/wxfw)

# 内容概要



- 专利检索概述
- 检索资源
- 检索要素及特点
- 检索策略
- 案例

# 专利检索概述——概念与特点



- 概念
  - 获取与专利相关的技术信息
- 特点
  - 技术导向性
  - 时效性
  - 地域性

# 专利检索概述——类型



- 法律状态检索
- 专利性检索
- 侵权检索
- 专利分析

# 检索资源



- 专利检索资源
- 非专利检索资源
- 生物序列检索资源

# 检索资源——专利



- 各国专利局官方检索资源
  - 中国专利检索及分析系统（PSS系统）
    - [www.pss-system.gov.cn](http://www.pss-system.gov.cn)
  - USPTO PatFT&AppFT数据库
    - [patft.uspto.gov](http://patft.uspto.gov)
    - 检索指导和资源链接 ([www.uspto.gov/patents-application-process/search-patents](http://www.uspto.gov/patents-application-process/search-patents))
  - EPO Espacenet检索系统
    - [worldwid.espacenet.com](http://worldwid.espacenet.com)

# 检索资源——专利



- 其他官方资源入口
  - SIPO网站
  - 文献服务
  - 互联网专利检索系统链接

# 检索资源——专利



- 商业性专利检索资源
  - Patentis
    - [www.patentics.com](http://www.patentics.com)
  - FPO
    - [www.freepatentsonline.com](http://www.freepatentsonline.com)

# 检索资源——非专利



- 知识平台
  - 中国知网 (CNKI)
    - [www.cnki.net](http://www.cnki.net)
  - 万方数据知识服务平台
    - [www.wanfangdata.com.cn](http://www.wanfangdata.com.cn)
  - Pubmed
    - [www.ncbi.nlm.nih.gov/pubmed](http://www.ncbi.nlm.nih.gov/pubmed)
  - STN
    - [stnweb.cas.org](http://stnweb.cas.org)

# 检索资源——非专利



公益讲座  
www.sipo.gov.cn/wxfw

- 互联网搜索引擎

- 谷歌学术

- 百度学术

- 必应学术

# 检索资源——生物序列



## Genbank和EMBL

数据库	GenBank	EMBL
数据覆盖	非专利数据覆盖较全	非专利数据覆盖较全, 专利数据比GenBank全
检索工具	BLAST	FASTA/BLAST
检索速度	快	慢
检索库	nr, pat	核酸: EMBL-Bank, Patents 蛋白质: UniProt Knowledgebase, UniProtKB/Swiss-Prot, UniProtKB/Swiss-Prot isoforms, UniProtKB/TrEMBL, Patents

# 检索资源——Genbank



- Genbank中的专利序列数据
  - Pat数据库主要来自US专利公开文件，更新及时
  - 目前(18年1月) 约有214万蛋白质序列，3279万核酸序列

# 检索资源——EMBL



- EMBL中的专利序列数据
  - 来自EPO、JPO、KIPO和USPTO
  - 按季度更新
  - Level-1非冗余专利序列
    - 相同长度上100%相同
  - Level-2非冗余专利序列
    - 相同并属于同一专利族
    - 丰富了来自专利文件的生物信息和附加数据

# 检索要素



- 概念
  - 表述技术方案的信息
  - 向检索系统提交的信息

# 检索要素表达



- 表达方式
  - 关键词
  - 专利分类号
  - 化学结构
  - 生物序列
- 表达方式本身的特征
- 检索系统相应反馈的特征

# 检索要素表达——关键词



- 获取方法
  - 阅读技术方案
  - 扩展
    - 检索前
    - 检索过程中
- 本领域的特点
  - 精准性
  - 列举不全易造成漏检
  - 信息量较低（难以表述复杂的概念）

# 检索要素表达——分类号



- 常用类型
  - IPC分类号
  - CPC分类号(更细化)
- 获取方法
  - 查分类表
  - 利用关键词检索进行分类号统计
- 本领域的特点
  - 精度低
  - 信息量高 (可表述复杂的概念)

# 检索要素表达——举例



- 关键词：曲妥珠单抗、Trastuzumab、赫赛汀、Herceptin（一种抗HER2单克隆抗体）
  - 精度高
  - 多种表述
- 分类号：C07K 16/28（抗受体，细胞表面抗原或细胞表面决定因子的免疫球蛋白），C07K 16/30（来自肿瘤细胞）
  - 精度低
  - 表述复杂的概念

# 检索要素表达——生物序列



公益讲座  
www.sipo.gov.cn/wxfw

- 表达方式
  - 核酸序列: 单字母代码
  - 氨基酸序列: 单字母、三字母代码
- 特点
  - 精确度高
  - 可进行同源性检索
  - 不同检索系统需要采用不同的形式, 非序列数据库中有漏检可能
- 例: 短肽**YGFGG, TYR GLY PHE GLY GLY**
  - PSS系统: YGFGG 或 TYR w GLY w PHE w GLY w GLY (w为邻近算符)
  - Google学术: YGFGG或“TYR GLY PHE GLY GLY”
  - 若目标文献中“YGFGG”被空格打断, 容易造成漏检。

# 中国专利检索及分析系统



- 简介
  - 专利检索功能
  - 专利分析功能
  - 支持多语种的检索（英、法、德、日、韩、俄等）
  - 简单、快捷、功能丰富
- 收录范围
  - 103个国家、地区和组织的专利数据
  - 引文、同族、法律状态信息

# PSS系统——药物检索



- 药物检索
  - 深度加工标引的药物专利数据库
  - 药物检索
    - 高级检索、方剂检索
  - 词典
    - 中药词典、西药辞典

- 检索式编辑
  - 格式：“字段” = “检索表达式”
  - 例：1. 关键词=嵌合抗原受体 AND (VEGF OR 血管内皮生长因子)
- 检索式运算
  - 检索历史
  - 运算符 AND、OR、NOT、()
  - 例：2. IPC分类号=(C07K19/00)
  - 3. 1 AND 2
- 检索式管理
  - 个人中心—我的检索式

- 注意事项
  - ()为半角符号
  - 输入中有空格时需加半角双引号，否则空格视为 OR 运算，例如：“regulatory T cells”
  - 需要检索保留关键字时加半角双引号，例如：“and”
  - 分类号中有“/”时，需使用半角双引号或括号，如：IPC分类号=(C07K19/00)

# PSS系统——截词符



- 截词符
  - # 表示任意一个存在的字符
  - ? 表示任意一个字符或没有字符
  - +表示包含任何长度任何字符的字符串
- 用于检索英文字符串
- 例：
  - cell# 匹配 cells
  - cell? 匹配 cell, cells
  - isolat+ 匹配 isolate, isolating, isolated, isolation

# PSS系统——截词符



- 布尔算符
  - AND      OR      NOT
- 检索要素之间的运算关系
- 例：
  - 1. 关键词=(免疫球蛋白 OR 抗体) AND (分离 OR 纯化)
  - 2. 发明名称=试剂盒
  - 3. 1 NOT 2

# PSS系统——邻近算符



- $W, nW, =nW$ 
  - 两个元素按顺序出现在同一句中
  - $=n$ , 正好间隔n个词
  - $n$ , 至多可间隔n个词
- $D, nD, =nD$ 
  - 元素间无顺序要求, 其余同上
- 例:
  - 白 2W 介素, 白介素、白细胞介素
  - 口腔 2D 崩解, 口腔崩解片、口腔内崩解片剂、口腔快速崩解片、可崩解口腔膜剂

# PSS系统——同在算符



- S, 同一句
- P或L, 同一段落
- F, 同一字段
  - 注: “关键词”检索项内有三个字段“发明名称”、“摘要”、“权利要求”

# PSS系统——同在算符案例



- 案例：一种用于组培苗生根的培养基，配方中含有IBA和MET。

- 检索要素

生根 培养基 IBA MET

- 构建检索式

权利要求=生根 S 培养基 S IBA S MET

检索结果：3篇

# 检索策略——检索流程



- 头脑风暴
  - 接收并加工巨量的信息
  - 完整确定的检索流程、预先制定检索策略、及时反馈和调整、良好的记录习惯
- 一般流程
  - 理解发明→确定检索要素→检索要素表达→检索操作→对比文件阅览与选取
  - 全程的反馈与调整

# 检索策略——理解发明



- 检索的基础
- 思路
  - 背景技术
  - 解决的技术问题
  - 解决技术问题的手段
  - 提炼发明构思
- 发明构思与技术特征
  - 发明构思是更高一层级的概念

# 检索策略——理解发明



- 发明构思为导向 vs. 技术特征为导向
  - 发明构思为导向
    - 仍是以具体的技术特征为入口
      - Why? How?
      - 思路更深远
  - 技术特征为导向
    - 假阳性结果
    - 思路受限

# 检索策略——确定检索要素



- 以理解发明为基础
- 一般检索要素
- 核心检索要素
- 本领域常见检索要素
  - 物质名称、关联物质
  - 生物序列
  - 功能、效果
  - 操作过程、参数

# 检索策略——检索要素表达



- 分类号
  - 限定检索范围
  - 一般与关键词结合使用
- 关键词
  - 异名、同义词、近义词
  - 上下位概念
- 结构式
  - 核心结构
  - 取代基

# 检索策略——检索要素表达



- 生物序列
  - 氨基酸序列，三字母代码/单字母代码
  - 注意文献排版对检索的影响（空格，换行）
- 例1：一种 $\alpha$ -促黑素细胞激素( $\alpha$ -MSH)，是包含序列 Ser Tyr Ser Met Glu His Phe Arg Trp Gly Lys Pro Val (SEQ ID NO: 8)(SYSMEHFRWGKPV)的多肽。
  - SYSMEHFRWGKPV
  - +EHFRW+
  - ...Tyr W Ser W Met W Glu...

# 检索策略——检索过程



- 构建检索式
  - 构建检索块
    - 以同一组概念为单位  
(曲妥珠单抗 OR Trastuzumab OR 赫赛汀 OR Herceptin)
    - 避免概念混杂  
(抗HER2抗体 OR 曲妥珠单抗 OR Trastuzumab OR 赫赛汀 OR Herceptin)
  - 记录检索操作和思路
  - 反馈与调整
    - 关键词、分类号扩展
    - 检索要素取舍

# 检索策略——对比文件阅览选取



公益讲座  
www.sipo.gov.cn/wxfw

- 关键词高亮显示
- 记录备选文献
  - 文献编号
  - 公开内容
  - 引用位置
  - 技术特征表
- 文献追踪

# 检索策略——反馈和调整



- 结果过多
  - 原因分析
    - 噪音（噪音来自哪些检索要素）
    - 检索要素的限定效果
  - 调整
    - 调整关键词
    - 运用分类号
    - 使用邻近算符、同在算符

# 检索策略——反馈和调整



- 结果过少
  - 原因分析
    - 拼写错误
    - 要素限定过于具体
  - 调整
    - 检查检索式
    - 检索要素扩展（同近义词、上位概念）
    - 检索要素舍弃

# 检索策略——反馈和调整



- 检索方向调整
  - 基于技术方案
  - 基于已获得的文献
- 检索全过程的反馈和调整

# 常见检索策略——生物序列



- 涉及的主题
  - 基因、引物、蛋白质（肽、抗体、酶）等
- 序列检索
  - 序列数据库Genbank、EMBL、STN等
  - 专利数据库
- 关键词、分类号检索
  - 名称、来源、功能、发明人

# 常见检索策略——抗体



- 常用检索要素
  - 分类号、名称、抗原、杂交瘤、生产方法
  - 序列
    - 表位序列
    - CDR序列
- 序列检索
  - Genbank中检索组装的CDR1-3
  - 单个CDR序列直接作为关键词

# 常见检索策略——基因检测方法



公益讲座  
www.sipo.gov.cn/wxfw

- 常用检索要素
  - 基因或靶点名称
  - 疾病/致病菌的名称
  - 检测手段（PCR、探针、荧光定量等）
  - 探针、引物、分子标记等的具体核酸序列
- 注意异名和上位概念的扩展

# 常见检索策略——生物制备方法



公益讲座  
www.sipo.gov.cn/wxfw

- 常用检索要素
  - 反应物、催化剂、产物
  - 反应原理
  - 反应过程、参数

# 检索案例1

- 一种人源抗huIFN- $\alpha$  抗体，其包含重链CDR1和重链CDR2，其中所述的重链CDR1的氨基酸序列为SYAMS，重链CDR2的氨基酸序列为AISGSGGSTYYADSVKG。
- 使用CDR序列构建检索式  
说明书= SYAMS AND AISGSGGSTYYADSVKG
- PSS系统中可获得多篇对比文件

# 检索案例2——技术方案

1. 一种用于诱导对象产生针对多种HIV病毒的广谱中和抗体的疫苗，所述疫苗包含两种或更多种来自不同种慢病毒的膜抗原或编码所述膜抗原的多核苷酸。
2. 权利要求1的疫苗，其包含2种、3种、4种、或5种来自不同种慢病毒的膜抗原或编码所述膜抗原的多核苷酸。
3. 权利要求1的疫苗，其中所述慢病毒选自人免疫缺陷病毒（HIV）、猴免疫缺陷病毒（SIV）、马传染性贫血病毒（EIAV）和猫科免疫缺陷病毒（FIV）中的一种或多种。

# 检索案例2——理解发明

- 说明书中的记载
- 之前的研究中使用HIV不同亚型进行交叉免疫，以诱导机体免疫反应从而对抗HIV。
- 本发明开发出基于来自不同种慢病毒抗原的交叉免疫方案以及相应的疫苗和抗体。
- ...不同慢病毒的膜抗原顺序免疫或混合免疫...
- 慢病毒选自人免疫缺陷病毒（HIV）、猴免疫缺陷病毒（SIV）、马传染性贫血病毒（EIAV）、和猫科免疫缺陷病毒（FIV）中的一种或多种

# 检索案例2——确定发明构思



- 除HIV以外的其他慢病毒（SIV、EIAV、FIV）可以提供对抗HIV的交叉免疫
- 免疫方案包括不同种慢病毒的膜抗原的同时或依次施用

# 检索案例2——确定检索要素

1. 一种用于诱导对象产生针对多种HIV病毒的广谱中和抗体的疫苗，所述疫苗包含两种或更多多种来自不同种慢病毒的膜抗原或编码所述膜抗原的多核苷酸。
3. 权利要求1的疫苗，其中所述慢病毒选自人免疫缺陷病毒（HIV）、猴免疫缺陷病毒（SIV）、马传染性贫血病毒（EIAV）、和猫科免疫缺陷病毒（FIV）中的一种或多种。

检索要素：“疫苗”、“广谱”、“中和”、“抗体”、“HIV”、“两种或更多多种”、“不同种”、“慢病毒”、“膜抗原”、“多核苷酸”、“SIV”、“EIAV”、“FIV”、“交叉免疫”

# 检索案例2——确定检索要素

现有要素：“疫苗”、“广谱”、“中和”、“抗体”、“HIV”、“两种或更多种”、“不同种”、“慢病毒”、“膜抗原”、“多核苷酸”、“SIV”、“EIAV”、“FIV”、“交叉免疫”

- 排除“广谱”、“中和”、“多核苷酸”
- 确定上下位关系：“慢病毒”——“HIV”、“SIV”、“EIAV”、“FIV”
- “两种或更多种”、“不同种”、“交叉免疫”是体现发明构思的检索要素

# 检索案例2——检索要素分析

- “慢病毒” or “HIV”/“SIV”/“FIV”/“EIAV”？
- “两种或更多种”、“不同种”、“交叉”，怎么表达？
- 其他检索要素的重要性排序？  
“膜抗原”、“疫苗”、“抗体”、“抗原”

# 检索案例2——核心检索要素



- 病毒名称  
HIV / SIV / FIV / EAIV
- 免疫方案  
不同物种、两种或多种（抗原）
- 抗原名称  
膜抗原
- 产品类型  
疫苗

# 检索案例2——检索要素表达

- 病毒名称

HIV OR SIV OR FIV OR EAIV

- 抗原名称

膜抗原 OR envelope+

- 产品类型

疫苗 OR vaccine OR 免疫原性组合物?

- 免疫方案

两种 OR 三种 OR 多种 OR 不同种 OR 异种 OR 异源?

交叉免疫?

# 检索案例2——检索要素表达



- 产品类型  
疫苗 OR vaccine OR 免疫原性组合物?  
→更适合用分类号表达  
A61K39/00 含有抗原或抗体的医药配制品
- 免疫方案  
→ A61K39/295 多价病毒抗原
- 一个分类号表达两个检索要素

# 检索案例2——检索过程



- 检索式：
  1. 关键词=HIV or SIV or EIAV or FIV
  2. 关键词=膜抗原 or 膜蛋白 or envelope+
  3. IPC分类号=(A61K39/295)
  4. 1 and 2 and 3
- PSS系统中73篇结果

# 检索案例2——检索调整



- 关键词=HIV or SIV or EIAV or FIV

修改为

- 关键词=HIV and (SIV or EIAV or FIV )
- 调整后检索结果为3篇， 获得对比文件

# 检索案例2——小结



- 理解发明
- 提取一般检索要素
- 基于对发明构思的理解确定核心检索要素
- 选用合适的方式表达检索要素
- 检索式构建与调整

# 检索案例3

- 一种酶法合成  $\gamma$ -D-谷氨酰-L-色氨酸的方法，其特征在于将  $\gamma$ -谷氨酰转肽酶加入到由L-色氨酸和  $\gamma$ -D-谷氨酰胺组成的底物溶液中，通过  $\gamma$ -谷氨酰转肽酶的转肽反应制备  $\gamma$ -D-谷氨酰-L-色氨酸。

# 检索案例3

- 检索要素分析
  - 产物:  $\gamma$ -D-谷氨酰-L-色氨酸
  - 酶:  $\gamma$ -谷氨酰转肽酶/ $\gamma$ -GT
  - 底物:  $\gamma$ -D-谷氨酰胺, L-色氨酸
  - 反应机理: 转肽反应
- 分析确定核心检索要素
  - 产物名称中包含两种底物的特征
  - 酶名称中包含底物和反应机理

# 检索案例3

- 检索过程
  - 使用产物和酶的英文关键词在百度学术中检索
- 检索结果
  - Development of an efficient enzymatic production of  **$\gamma$ -D-glutamyl-L-tryptophan** (SCV207) , a prospective medicine for tuberculosis , with bacterial  **$\gamma$ -glutamyltranspeptidase**.

# 谢谢聆听！



公益讲座  
[www.sipo.gov.cn/wxfw](http://www.sipo.gov.cn/wxfw)