



# 分子流行病学

## -概述与生物标志

主讲：邓晓蓓 讲师

系科：流行病学与生物统计学系

Email：dengxiaobei@126.com





# 目录 CONTENTS



概述

生物标志物

主要研究方法

应用与展望



# 目录 CONTENTS



● 概述

● 生物标志物

● 主要研究方法

● 应用与展望



# 分子流行病学定义

**(一) 分子流行病学**：是阐明**人群和相关生物群体中生物标志的分布**及其与疾病和健康的关系和影响因素，并研究防制疾病，促进健康的策略与措施的科学。

- **生物标志** ( biological markers,简称M或BM )：指能代表**生物结构和功能的可识别物质**。
- **分子生物标志** ( molecular biomarker)：指代表生物结构和功能的生物大分子特征，如**DNA、RNA、蛋白质**等，是生物标志的主要部分。
- **相关生物群体**：是指与人类健康和疾病具有较密切关系的生物群体。



# 血清流行病学定义

**(二) 血清流行病学(seroepidemiology)**：是应用血清学的方法对**血清**中各种成分(包括**抗原、抗体、代谢产物、生化物质、营养成分及遗传因子**等)的出现和分布进行研究，以阐明疾病及健康状态在人群中的分布及其影响因素，并在采取预防控制措施后应用血清学方法考核其效果。



# 血清学检测

2017-10-04 11:58:35: 白细胞, 0.00 个/ul,; 红细胞, 0.00 个/ul,; 颜色, 浅黄,; 浑浊度, 清澈,; 比重, 1.005,; PH, 6.00,; 尿糖, 阴性(-),; 尿蛋白, 阴性(-),; 尿胆原, 阴性(-),; 胆红素, 阴性(-),; 酮体, 阴性(-),; 亚硝酸盐, 阴性(-),; 隐血, 阴性(-),; 白细胞酯酶, 阴性(-),; 维生素 C, 阴性(-),; 鳞状上皮, 0.00 个/ul,; 非鳞状上皮, 0.00 个/ul,; 黏液, 0.00 个/LPF,;  
2017-10-04 11:50:54: 颜色, 黄色,; 形状, 稀糊状,; 白细胞, 0,; 红细胞, 0,;  
2017-10-05 09:00:53: 红细胞沉降率, 2.00 mm/h,;

2017-10-05 08:31:38: Cpn 肺炎衣原体(IgM), 阴性(-),; Mp 肺炎支原体(IgG), 阴性(-),; Mp 肺炎支原体(IgM), 阴性(-),;  
2017-10-05 08:47:44: A 族乙型溶血性链球菌, 阴性(-),;

2017-10-05 09:35:10: 降钙素原, 0.10 ng/ml,;

2017-10-05 11:46:22: 尿素氮, 1.6 mmol/L, ↓; 肌酐(酶法), 21 umol/L, ↓; GFR(MDRD 计算值), 732.7,; 丙氨酸氨基转移酶, 34 U/L,; 乳酸脱氢酶, 253 U/L, ↑; 肌酸激酶, 72 U/L,; 肌酸激酶同工酶 MB 活性, 24 U/L,; α 羟基丁酸脱氢酶, 171 U/L,;

2017-10-05 10:05:54: 白细胞,  $13.75 \times 10^9/L$ , ↑; 中性粒细胞比值, 21.2 %N, ↓; 淋巴细胞比值, 69.0 %L, ↑; 单核细胞比值, 8.2 %M,; 嗜酸性粒细胞比值, 1.2 %E,; 嗜碱性粒细胞比值, 0.4 %B,; 中性粒细胞绝对值,  $2.91 \times 10^9/L$ ,; 淋巴细胞绝对值,  $9.49 \times 10^9/L$ , ↑; 单核细胞绝对值,  $1.13 \times 10^9/L$ , ↑; 嗜酸性细胞绝对值,  $0.17 \times 10^9/L$ ,; 嗜碱性粒细胞绝对值,



# 人类基因组流行病学定义

(三) **人类基因组流行病学**：应用**流行病学与基因组信息**相结合的研究方法，开展以**人群**为基础的研究，评价基因组信息(基因或基因变异，及其相应编码的产物)，对人群健康和疾病的流行病学意义，是遗传流行病学与分子流行病学交叉的前沿领域。



# 分子流行病学的发展

- **流行病学**是研究人群中**疾病和健康的分布**，探索影响分布的**因素**，并在此基础上**制定预防策略**的措施的一门学科。
- 一直以来，流行病学在**疾病的预防和治疗**，**疾病危险因素的研究和控制**方面都起着十分重要的作用。
- 近二三十年来，生物学技术尤其是分子生物学技术取得了突破性进展，并逐渐渗透到各个医学领域。
- **分子生物学的理论和技术**不断应用于传统的流行病学，并在此基础上形成了一门新的分支学科——**分子流行病学**。





# 分子流行病学的产生和发展

## (一) 分子流行病学产生的背景

### 1. 疾病防制中的新问题

- 一些传统的表型 ( phenotype)检测方法已不能完全满足传染性疾病的传染来源与传播途径精确辨别的需要；
- 不能完全适应非传染性疾病的病因及其致病机制的探讨。



# 分子流行病学的产生和发展

## 2.分子生物学的发展

- **理论上**：DNA双螺旋模型、基因表达调控理论、遗传中心法则、遗传密码、RNA逆转录等
- **技术上**：凝胶电泳、体外核酸扩增、DNA测序、分子杂交、基因克隆、色谱分析等

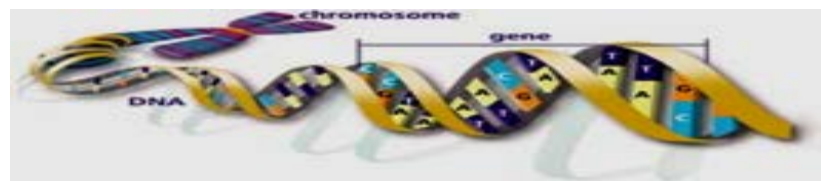




# 分子流行病学的发展历程

## 1. 分子流行病学概念的演变

- 最早使用的是Kilbourne博士（1972），他在一篇关于流感病毒抗原的变异与流感大流行的关系的论文中首次应用了分子流行病学这一术语。
- Higginson（1977）给分子流行病学下定义为：应用复杂精细的技术进行生物材料的流行病学研究。
- 1997年俞顺章教授提出一个简明的解释。
- 1993年出现了世界上第一本专著，由美国的Schulte所著。
- 国内首部分子流行病学专著是由徐德忠教授主编完成的，出版于1998年。





# 分子流行病学的应用现状

在疾病的预防控制中发挥着越来越重要的作用，主要表现在：

- (1) **研究内容更加丰富**：从最初的研究传染病开始，到疾病和健康状态相关生物标志的分布等；
- (2) **研究手段越来越多**：从最初的质粒图谱、核酸分子杂交等到现在的基因芯片技术；
- (3) **应用范围不断扩大**：不仅限于预防医学，基础医学、环境科学等研究。



# 分子流行病学的应用范围

**1.探索病因**

**2.危险因素致病机理的研究**

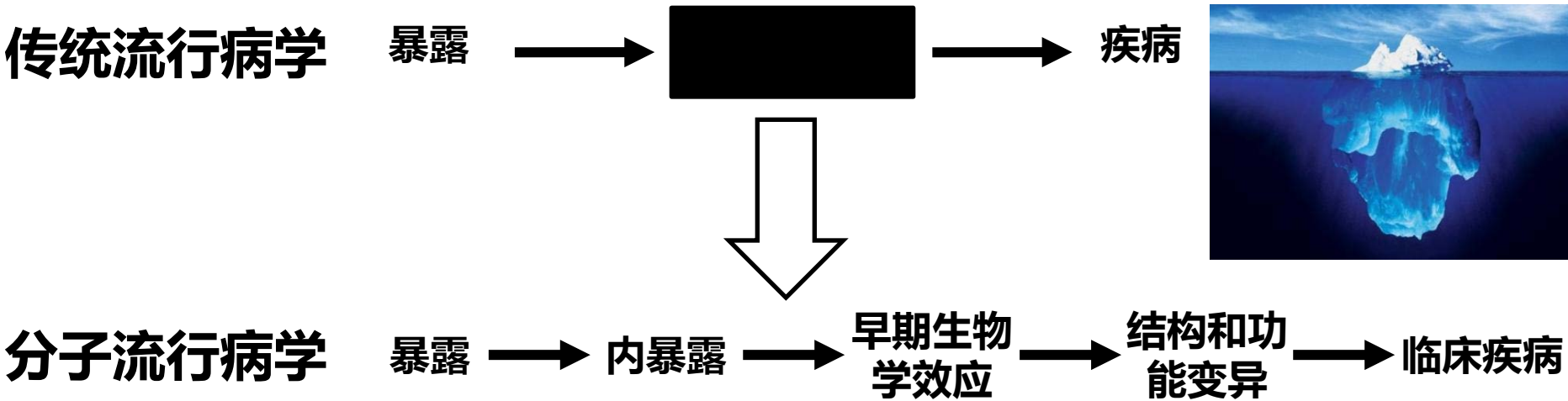
**3.疾病易感性的测定**

**4.疾病流行规律的研究**

**5.疾病防制措施的研究**



# 传统流行病学与分子流行病学的关系

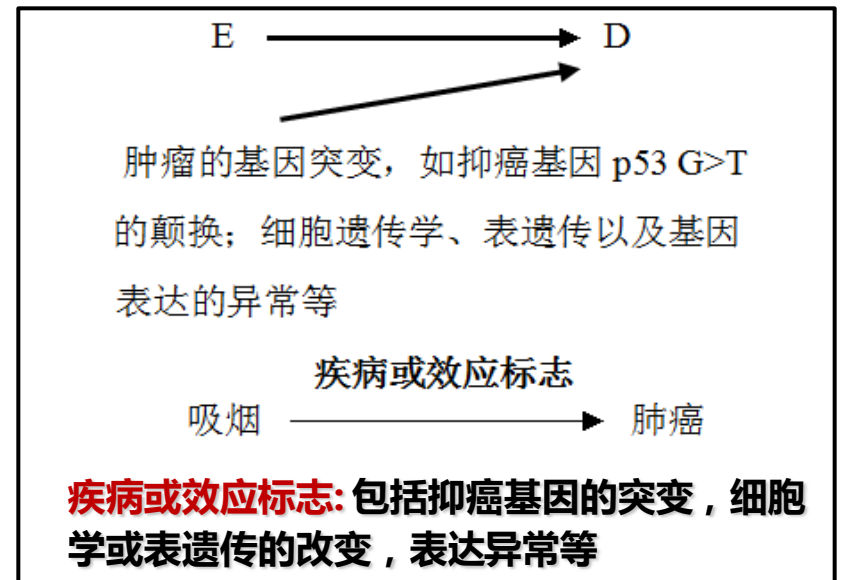
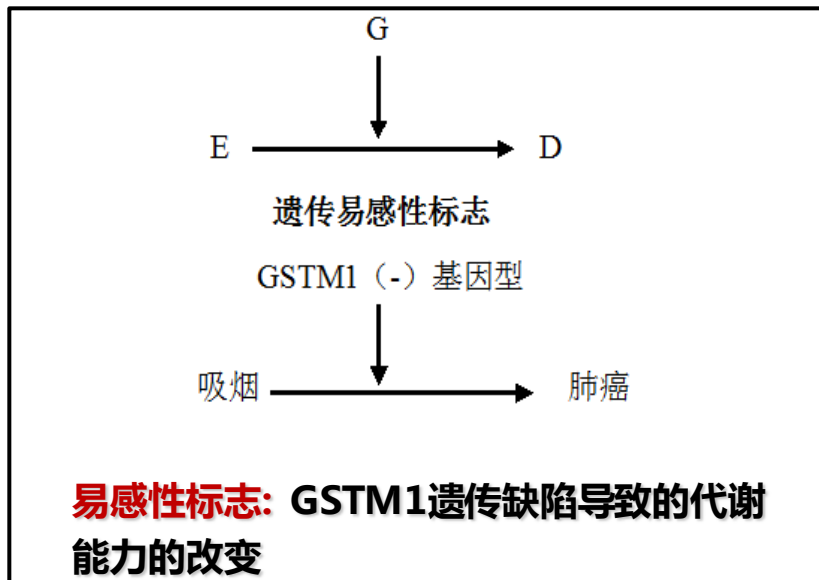
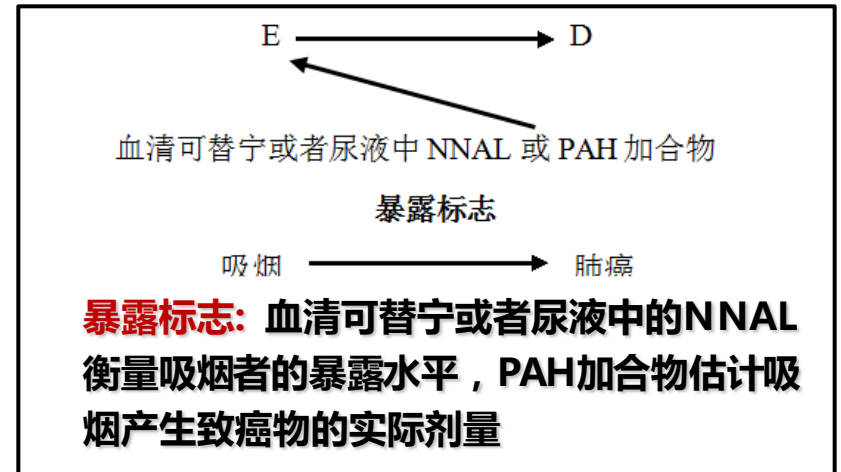
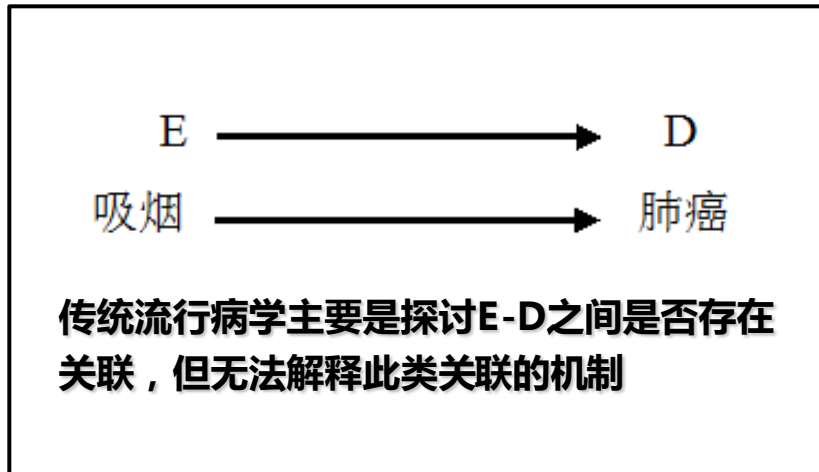


- **传统流病学**在探讨疾病的病因、研究疾病的危险因素以及对疾病进行预防和控制等方面起着重要作用。但是，难以得到**暴露和疾病**之间关系的直接证据。
- **分子流病学**应用分子生物学技术从分子和基因水平,阐明**生物标志**在人群中的**分布及其疾病/健康的关系和影响因素**，为病因学研究和预防措施评价开辟新的途径。



# 传统流行病学与分子流行病学的关系

(Elston RC, 2002)





# 传统和分子流行病学的比较

	传统流行病学	分子流行病学
推动力	公共卫生	现代科学
研究水平	人群	个体/器官组织/细胞/分子
研究样式	人口统计学/社会科学	临床/实验
认识论策略	自上而下	自下而上
预测水平	人群	个体

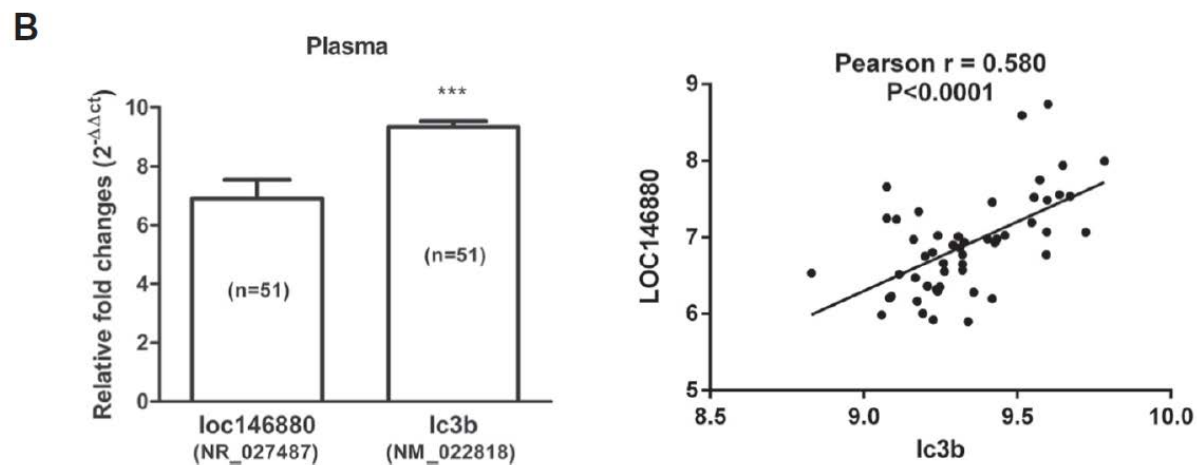
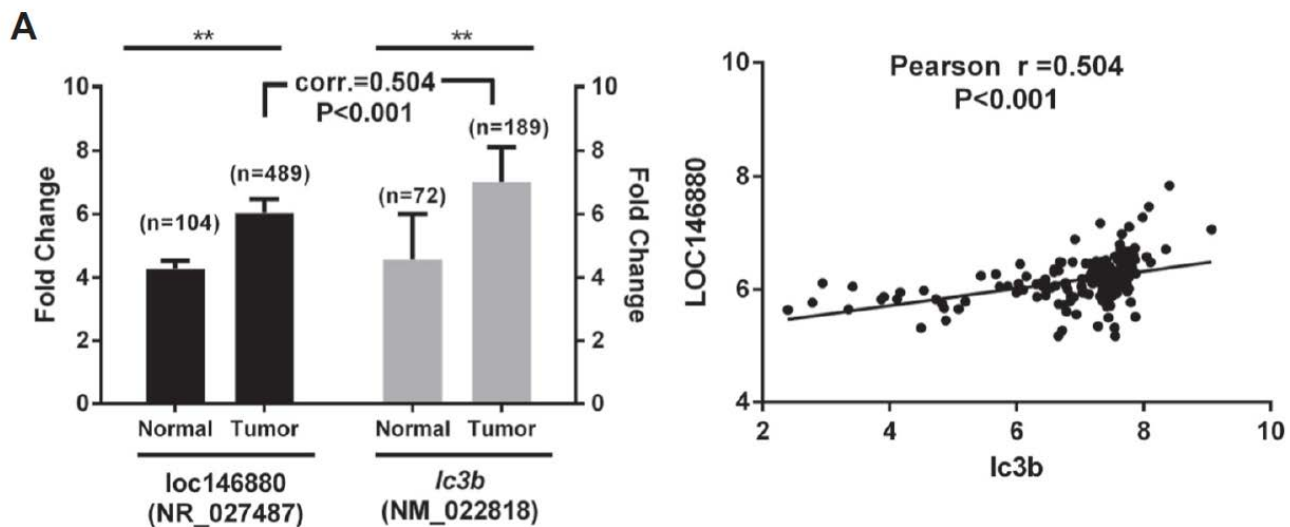


细胞内部的结构 (7分钟)





# 进一步寻找肺癌样本中的生物标志物





# 分子流行病学研究内容和范围（一）

## 1. 疾病病因的探讨

- 疾病的发生，大多是**环境因素与遗传因素**同时起作用。只是有些疾病以环境因素为主，有些则正好相反。
- 遗传因素起重要作用的疾病往往与基因有关；
- 很多与环境因素有关的疾病，也常常需从基因方面去研究，因为**环境因素**可以通过**宿主的基因突变或异常表达**引起疾病。



# 疾病易感性举例

- 代谢酶尿苷二磷酸葡萄糖醛酸转移酶 1A7(UDP-glucuronosyltransferase, UGT1A7)可催化香烟中的致癌物质苯并芘、亚硝胺NNK和杂环胺PhIP与葡萄糖醛酸结合使之失活，在解毒机制中起重要作用。
- 与携带UGT1A7\*1/\*1基因型个体比较，携带UGT1A7\*3/\*1基因型的个体患肺腺癌的风险增高 1.80 倍，携带UGT1A7\*3基因型的个体患肺腺癌的风险增高1.59倍。
- UGT1A7基因多态可能是中国人肺癌遗传易感性因素。



# 分子流行病学研究内容和范围（二）

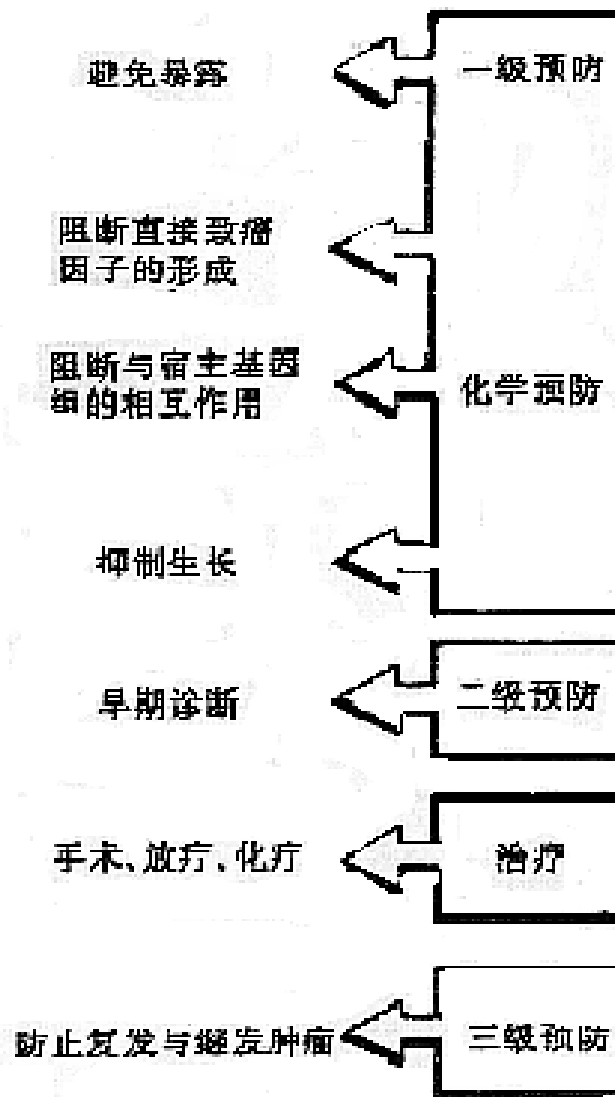
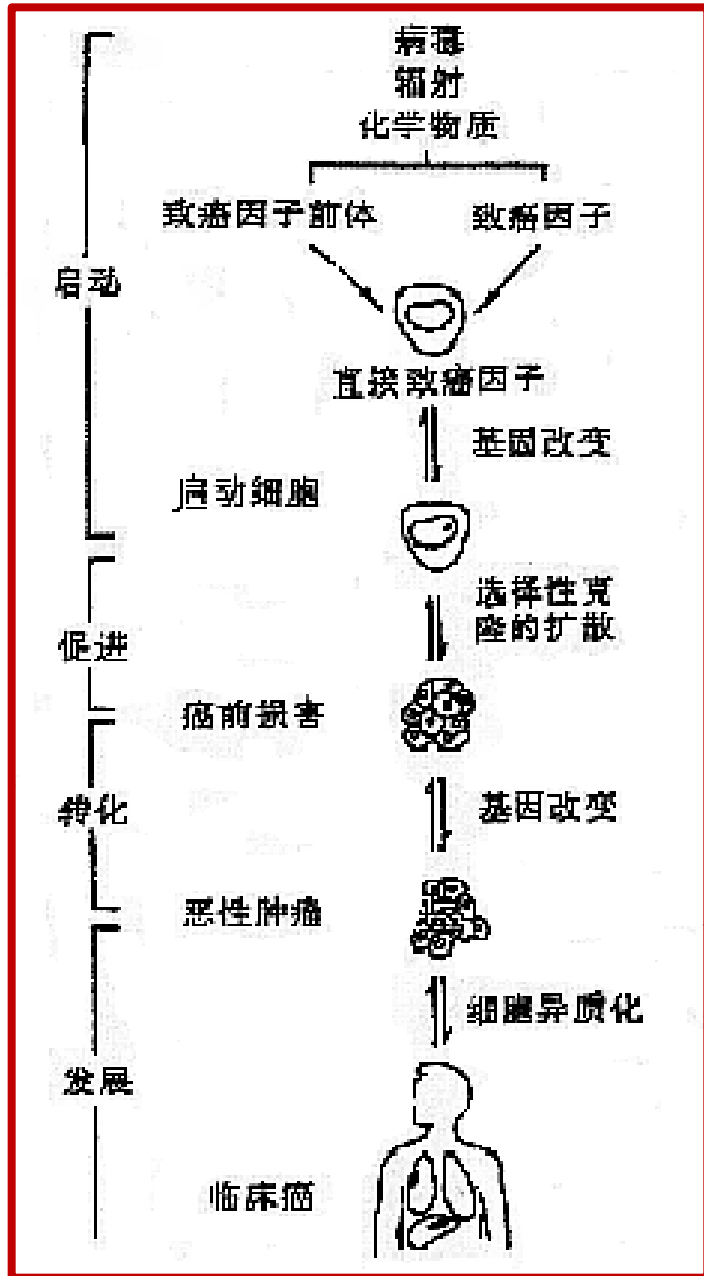
## 2. 危险因子致病机制的研究

- 分子流行病学不仅研究环境与宿主体内的**致病因子**，而且着重研究这些**致病的危险因素**，特别是化学物质与生物活性物质从暴露开始至疾病发生全过程的各个环节。



# 癌症发病的分子流行病学研究

- 20多年来，对与癌相关的各种生物学标志，致癌物的种种生物效应标志，如DNA加成物（ adduct ）等等进行了十分广泛的研究，有关报道相当丰富，积累了很多资料，提出了一些很有意义的假设。
- 正常细胞转变为癌细胞经历了致癌作用累积而又连续的几个阶段：**启动（ initiation ）**、**促进（ promotion ）**，**转化（ conversion ）**和**发展（ progression ）**。



分子流行病学和癌的发展过程及其三级预防策略



# 分子流行病学研究内容和范围（三）

## 3. 疾病易感性的测定

- 近来分子生物学与免疫学研究发现，机体对很多疾病也同样存在着易感性，且往往与**个体的某些基因型别或序列或其表达产物有关**，如果与遗传有联系，往往称之为遗传易感性（尤其慢性病）。



# 疾病易感性举例

- 目前，与遗传易感性相关研究较多的一个基因区域是人类第6号染色体短臂上的**人类白细胞抗原**（HLA），其原因可能有二，首先**HLA基因区有高度的多态性**，其次为HLA的功能是介导对**特异性抗原的免疫反应**。
- Erlich的研究发现，HLA II类单倍体（DRB1 \* 1501-DQB1 \* 0602）的比例在受到人乳头瘤状病毒（HPV）16型感染的宫颈癌与重度异常增生的病人中明显升高。





# 分子流行病学研究内容和范围（四）

## 4. 流行规律的研究

- 这方面研究最多的是传染病。在**传播范围的确定，传播途径的判断与传染源的追索方面**，分子流行病学可以发挥无与伦比的作用。
- 因为以往常常是根据**宏观流行病学调查分析**，再辅**以病原体分离与血清学分型**来分析流行规律。
- 然而近年的很多分子流行病学的研究表明，上述的方法有时不够精确。例如，尽管传染源与接触者之间的病原体的血清学型别一致，但如果基因型不一致，就不能确切判定两者的传播关系。



# 流行规律的研究（一）

## （1）传播范围的确定

- 按照传染病流行病学基本理论，在一次爆发或流行中，其受染范围的确定应根据如下原则：在爆发或流行期间，有共同或有效暴露史，经过一个**平均潜伏期发病**或出现**急性感染的生物学标志**可判定为受染者，然后**总计受染人员**，**结合这些人员活动的区域**，以确定**受染或流行涉及范围**。
- 可是，过去测定的生物学标志主要是属于血清学、免疫学、微生物学范围，缺乏基因型标志，因而可能出现错误的判断。
- 分子流行病学研究**缩小了常规流行病学调查所推测受染范围**，从基因水平上证实爆发的受染人数与传染来源。



# 流行规律的研究（二）

## （2）传播途径的判定

- 近来分子流行病学研究引入一些新的技术，如**分子进化诊断**（molecular evolutionary diagnosis）。

**案例：**1990年7月，美国佛罗里达州一名AIDS妇女，怀疑在当地一牙科诊所手术时感染了HIV，因为该诊所的牙科医生也是一名AIDS患者。随后当地卫生部门对该牙医诊治过的一千余名患者进行HIV检测，共发现7名HIV感染者有在该诊所的就诊史。

**那么这7名HIV感染者是否与该牙医有关呢？**



针对这一问题可采取的分析方法：

- 1、传统流行病学分析方法：接触者调查，分析感染者可能的危险行为如输血、静脉吸毒、不洁性交史等。
- 2、分子流行病学方法：比较HIV病毒基因序列的差异，分析不同病例感染株遗传关系的远近。



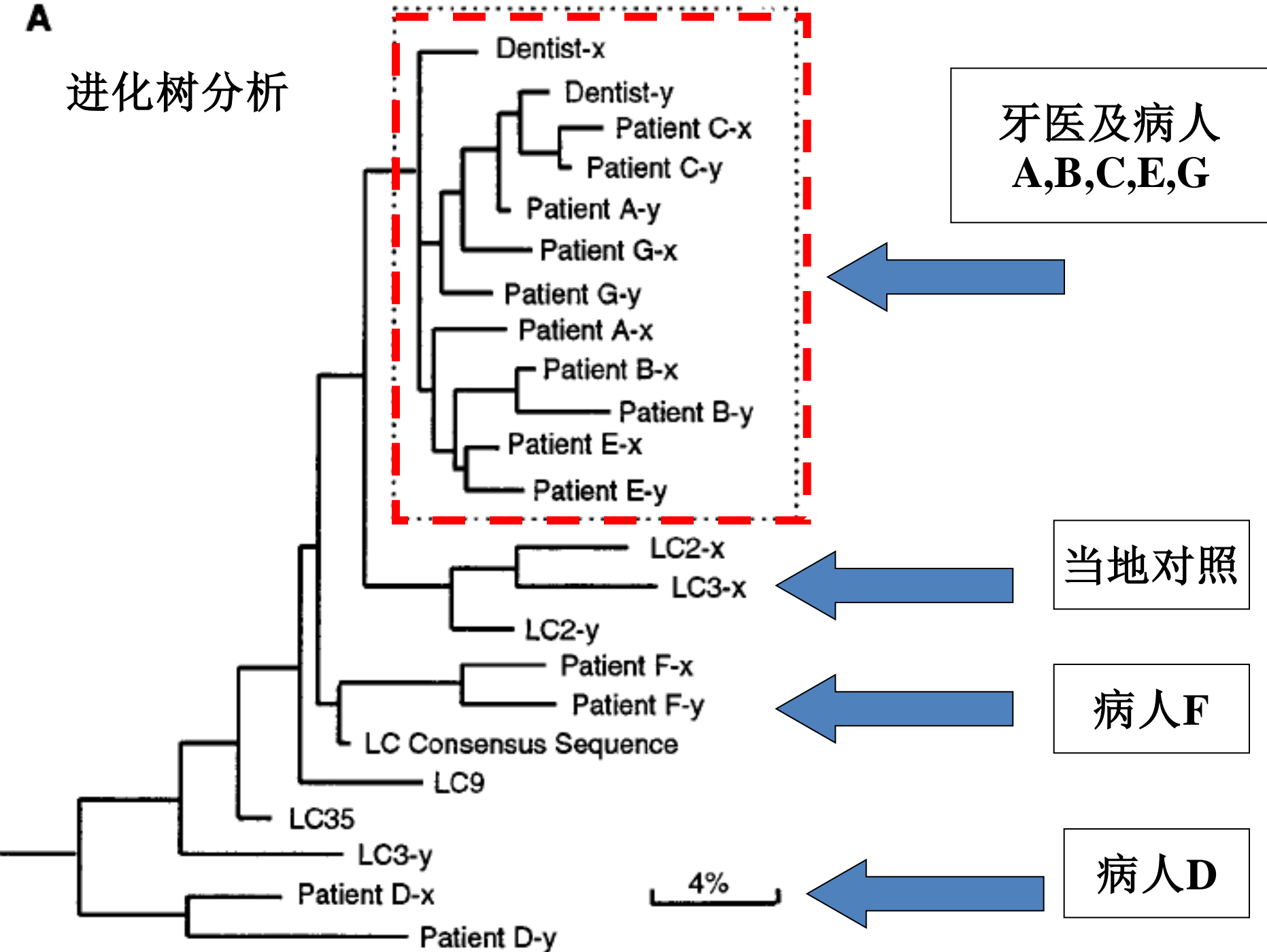
## 与牙医感染株的遗传变异性比较

## 与对照感染株的遗传变异性比较

Person	Sex	Known risk factor	Clinical status*	M13 clones (no.)	Intraperson variation† (%)	Interperson variation† (%)	
						To dentist	To 30 LCs‡
Dentist	M	Yes	AIDS	6	3.3 (0.8–5.4)		11.0 (5.8–16.0)
Patient A	F	No	AIDS	6	2.0 (0.0–4.5)	3.4 (0.8–6.2)	10.9 (5.4–14.8)
Patient B	F	No	Asymptomatic (CD4 = 222/μl)	12	1.9 (0.4–3.7)	4.4 (2.1–7.0)	11.2 (6.2–16.5)
Patient C	M	No§	Asymptomatic (CD4 = <50/μl)	5	1.2 (0.4–1.6)	3.4 (2.1–4.9)	11.1 (7.0–15.6)
Patient E	F	No	Asymptomatic (CD4 = 567/μl)	6	2.1 (0.4–3.7)	3.4 (1.2–6.6)	10.8 (5.8–14.8)
Patient G	M	No	Asymptomatic (CD4 = 400/μl)	5	2.8 (1.6–3.7)	4.9 (2.9–7.0)	11.8 (6.2–16.9)
Patient D	M	Yes	AIDS	5	7.5 (0.0–9.9)	13.6 (11.5–15.6)	13.1 (7.8–17.3)
Patient F	M	Yes	Asymptomatic (CD4 = 253/μl)	6	3.0 (0.8–5.8)	10.7 (8.2–13.6)	11.9 (7.0–17.3)

**A**

## 进化树分析





# 流行规律的研究（三）

## （3）传染源的追索

- 判定传染来源或两人之间的传播关系应符合如下4条原则：
  1. **A（传染源）与B（受染者）的接触是有效接触（暴露）；**
  2. **接触发生在A的传染期内；**
  3. **B的发病时间是在暴露后该病最短与最长潜伏期之间；**
  4. **A与B受染致病的血清学型别一致。**
- 如今看来，第④条应该用基因型别来代替，而基因序列的测定与型别差异的鉴定虽然比较复杂与精细，但可靠得多。



- 美国1990~1991年报道佛罗里达一个口腔诊所有1名患有艾滋病的牙医，被他看过病的患者中有7名后来感染了HIV，然而经过分子流行病学研究发现，其中5名患者的HIV与牙医的HIV有共同的氨基酸特异模式（signature pattern），而且序列的差异率 $<5.0\%$ ，为同一病毒株；而另2名患者的HIV则与牙医的不同。因此，通过序列分析，从分子水平确定了前5名病人的HIV感染是由牙医引起，而后2名病人的传染来源则不是牙医。





# 分子流行病学研究内容和范围（五）

## 5. 防控措施的研究

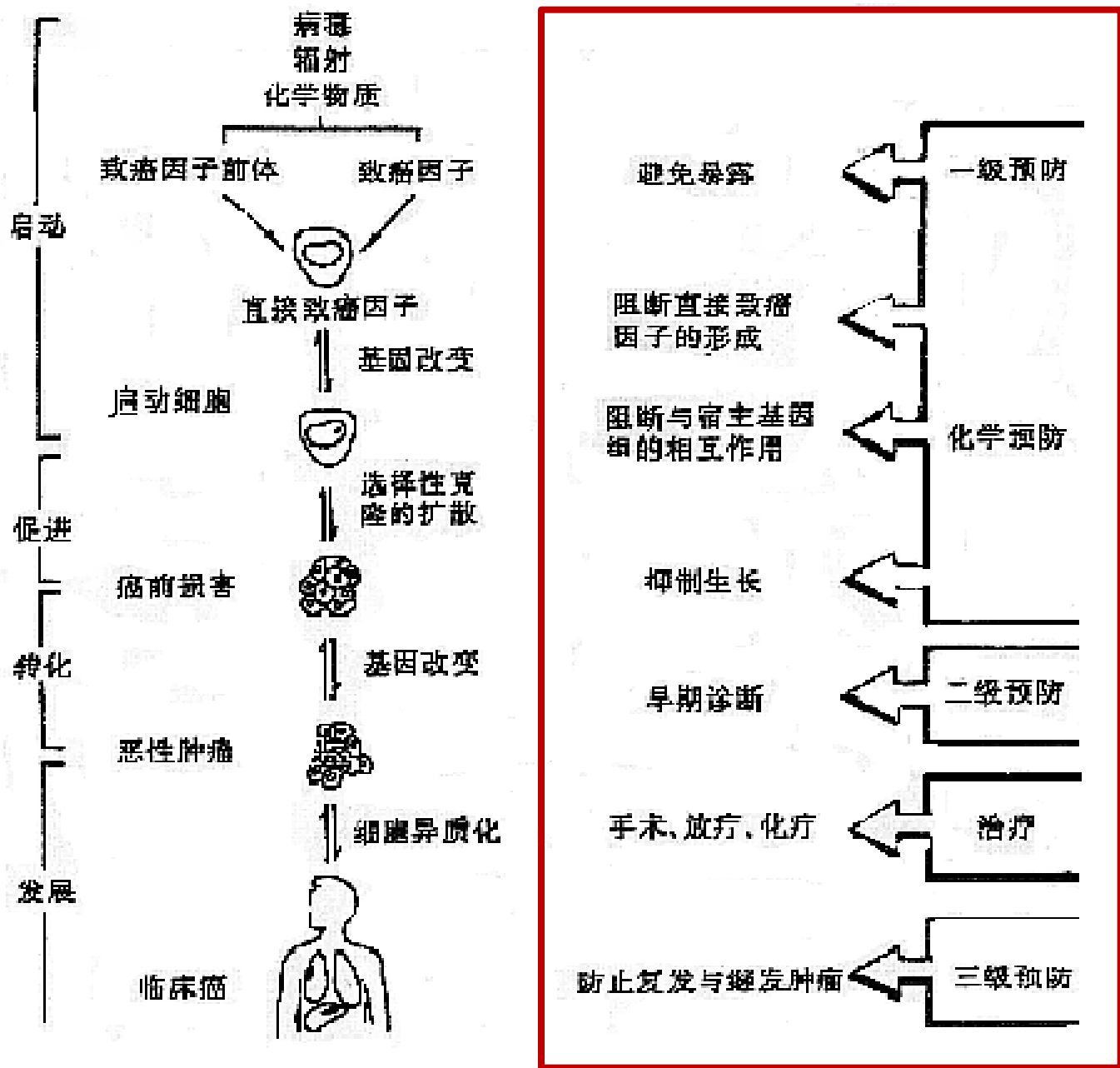
### （1）传染病预防

- 因为在传染病流行的研究中应用了分子生物学等现代技术，对**传播与流行范围的确定**，**传播途径**、**传播媒介与因子**以及**传染来源**的推断更为精确，所以提出的控制措施更有针对性，更有成效。



## (2) 非传染病的预防

- 由于分子流行病学不仅研究疾病的危险因素，而且研究其引起疾病整个过程中的所有环节，也即研究**这些致病因子的致病机制**。因此，这就为疾病的三级预防特别是第I、II级预防采取针对性措施奠定了科学的基础。



分子流行病学和癌的发展过程及其三级预防策略



# 要点

- 分子流行病学是流行病学的一个分支(分子生物学？群体遗传学？)
- 分子流行病学的主要研究内容是各种生物标志物及相关过程
- 分子流行病学的研究方法是随着分子生物学的进展不断更新的。



# 目录 CONTENTS



概述

**生物标志物**

主要研究方法

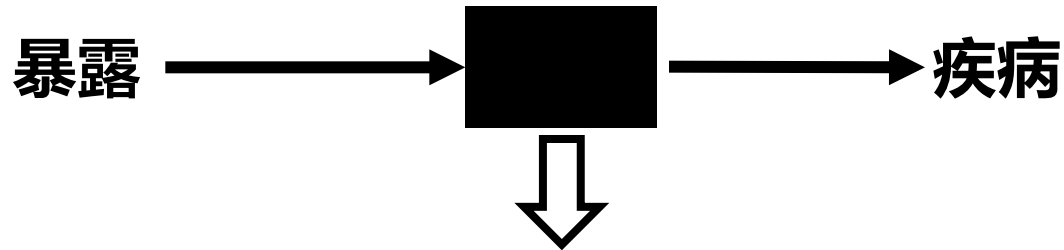
应用与展望



# 生物标志物的分类

( National Research Council, 1987 )

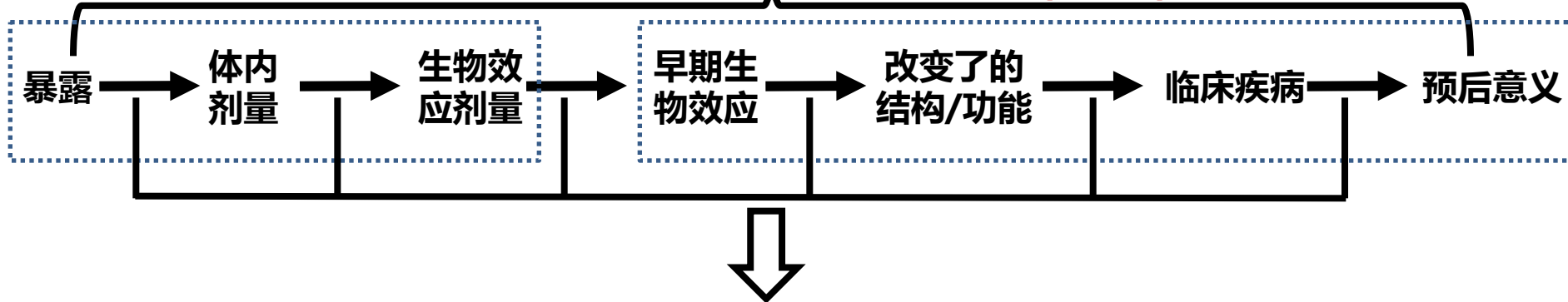
## 传统分子流病学



## 分子流行病学

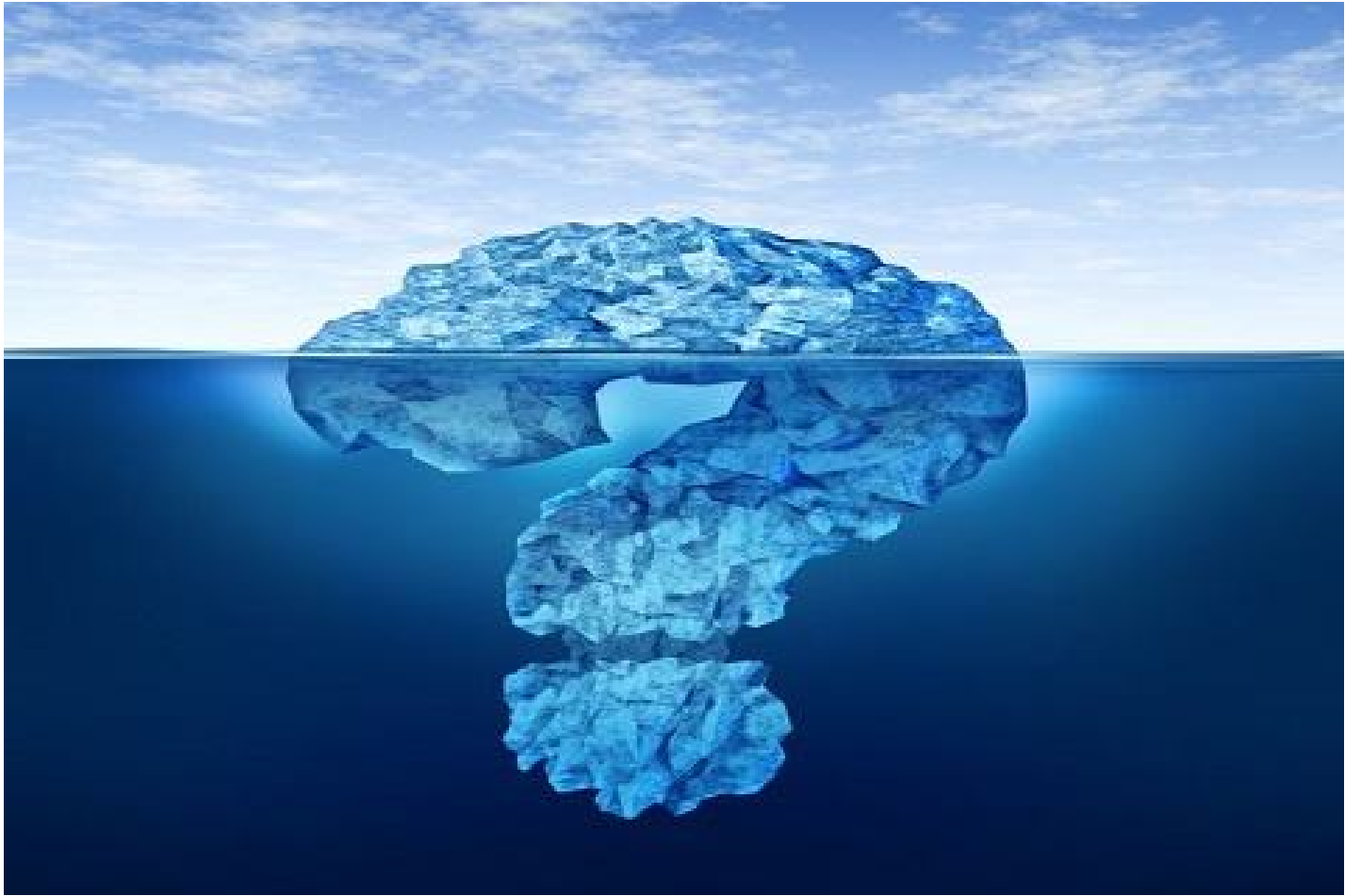
### 暴露标志

### 效应 ( 疾病 ) 标志



## 易感性标志







# 生物标志的分类（一）

## （一）生物标志的种类

1. **暴露标志**（exposure marker,  $M_E$ ）：与疾病或健康状态有关的暴露因素的生物标志。其包括外暴露标志和内暴露标志。

- **外暴露标志**：暴露因素进入机体之前的标志和剂量，如病毒、细菌、吸烟烟雾等，如个人所处直接环境（空气、水、土壤或食物）中的某物质浓度或强度，是通过环境监测来估计。
- **体内剂量**：暴露因素进入机体之后的标志，即外源性物质于暴露后在体内可测量到的剂量标志，是外源性物质进入人体的可靠证据。
- **生物学效应剂量**：表示已与靶组织细胞内DNA或蛋白质相互作用的外源性物质或其反应产物的含量。





# 体内剂量标志

- **体内剂量标志 (ID)** : 外源性物质于暴露后在体内可测量到的剂量标志, 表示被人体吸收的外源性物质 (性质未改变的或其代谢产物) 的数量。
- **吸烟** — 血或尿中的烟草代谢产物可的宁浓度 ;
- **多环芳烃** — 尿中的1- 羟基芘水平 ;
- **饮食暴露** — 黄曲霉毒素B1(AFB1)和尿中N-亚硝基化合物的浓度 ;
- **工作暴露和环境污染** — 血清或脂肪组织中的杀虫剂DDT、多氯联苯PCBs、二氧杂芘环氧化物含量。



# 生物有效剂量标志

- **生物有效剂量标志 (BED)** : 表示已与靶组织细胞内DNA或蛋白质相互作用的外源性物质或其反应产物的含量。
- 它是反映靶细胞分子内接触剂量的生物标志物, 主要包括蛋白质加合物 (Protein adduct)、DNA 加合物 (DNA adduct)。

---

**生物学标志**

**环境暴露物质**

**生物材料**

---

**DNA加合物**

**苯并 ( a ) 芘**

**白细胞**

**蛋白质加合物  
( 血红蛋白 )**

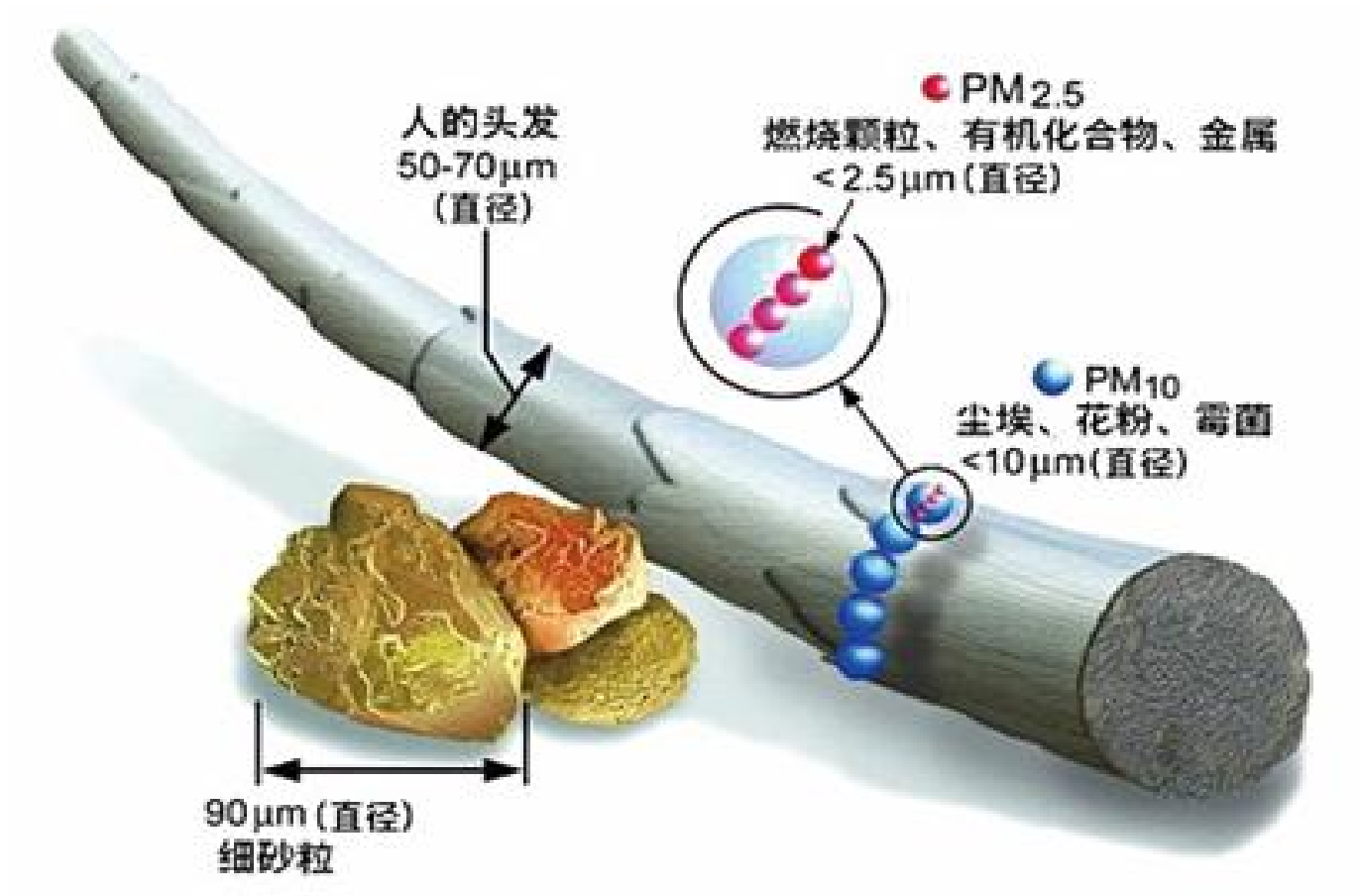
**环氧乙烯**

**红细胞**

---



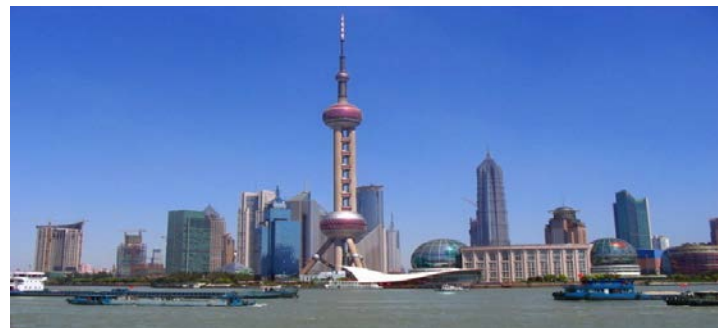
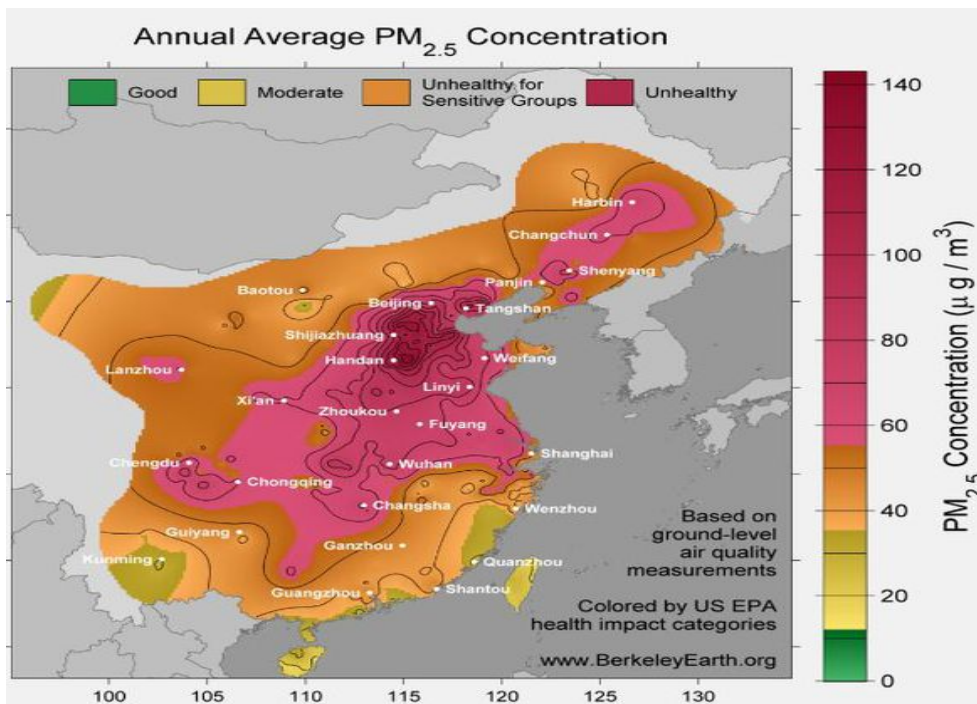
# 常见的空气污染物-PM<sub>2.5</sub>



PM<sub>2.5</sub>的介绍 (2分27秒视频)



# 大气污染现状



25%



雾霾的介绍 (1分42秒视频)



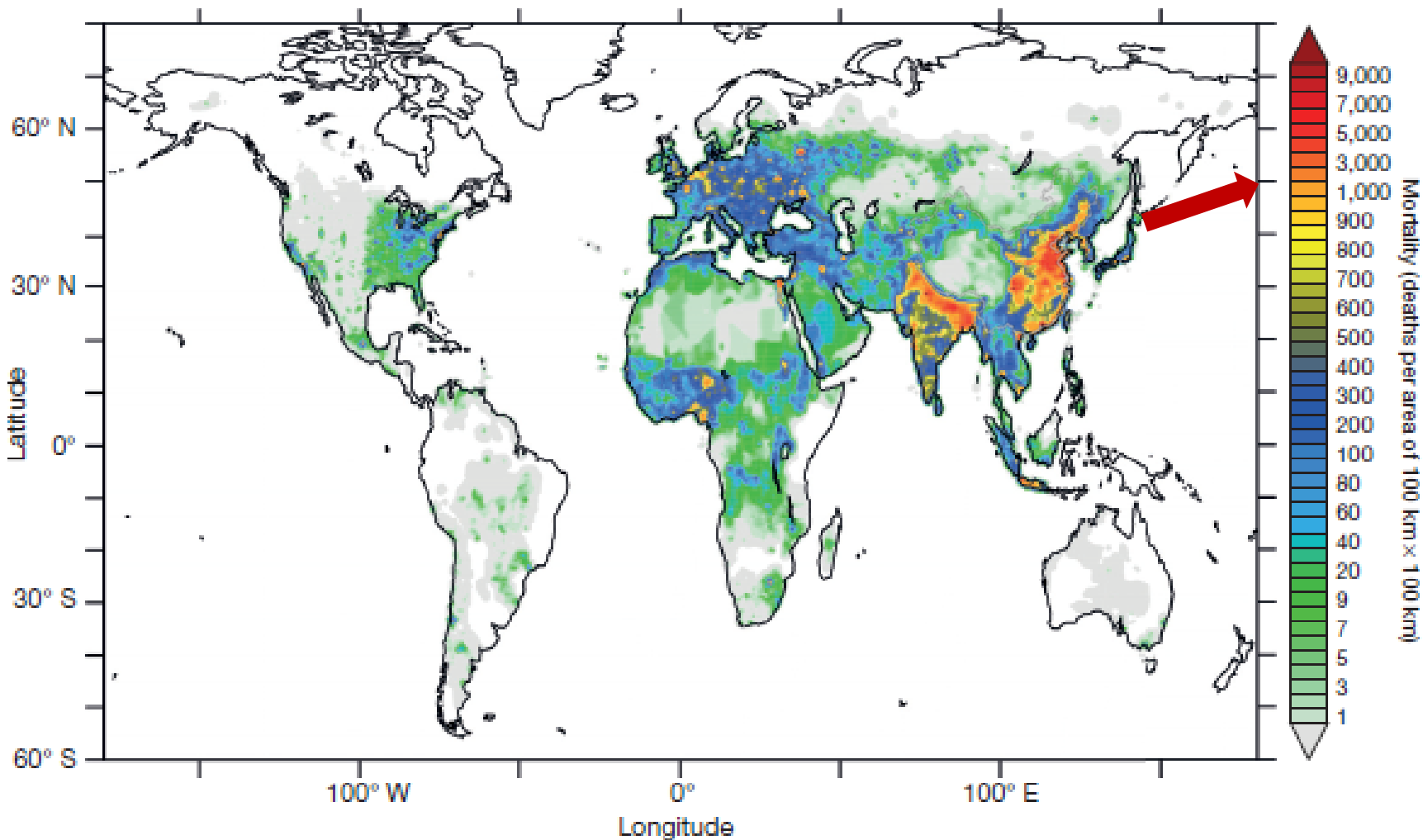
**PM2.5**

**PM2.5**  
直径小于等于2.5微米的  
颗粒物,是造成灰霾天气  
的“元凶”之一



# 大气颗粒物与健康

## 我国空气污染导致较高死亡率



### 2010年全球空气污染与死亡率之间的关系

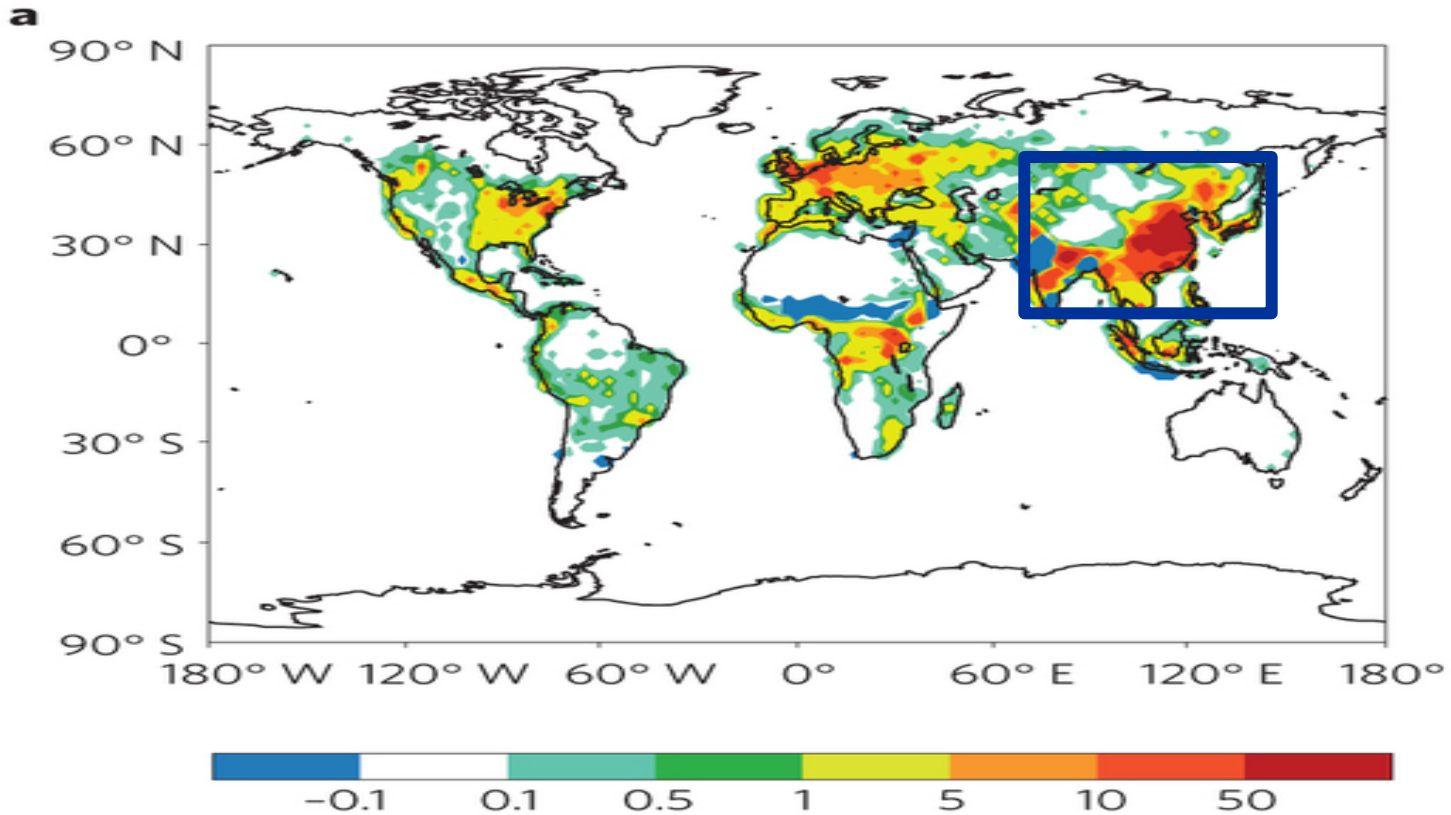
PM2.5导致死亡率和新生儿肺炎的介绍 (2分10秒视频)

J. Lelieveld et al, Nature, 2015.09. Doi:10.1038/nature15371



# 大气颗粒物对健康的影响

## 我国空气污染物导致肺癌的高发病率和死亡率



2010年全球空气污染与肺癌死亡率之间的关系

PM<sub>2.5</sub> 10  $\mu\text{g}/\text{m}^3$   $\uparrow$  肺癌的死亡率 8%  $\uparrow$



# PM<sub>2.5</sub> 化学组分含量

PAHs	Concentration (ng/mg)	OC/EC	Concentration (µg/mg)
萸烯 Acenaphthylene	22.99 ± 0.57	OC	4538 ± 411
二氢萸 Acenaphthene	13.70 ± 0.55	EC	90 ± 2
蒽 Anthracene	5.67 ± 1.16		
苯并[a]蒽 Benz[a]anthracene	157.40 ± 1.00		
<b>苯并[b]荧蒽 Benzo[b]fluoranthene</b>	<b>234.16 ± 21.83</b>		
苯并[k]荧蒽 Benzo[k]fluoranthene	84.93 ± 5.02		
苯并[a]芘 Benzo[a]pyrene	102.97 ± 8.11		
苯并[e]芘 Benzo[e]pyrene	92.40 ± 0.56		
苯并[g,h,i]苝 Benzo[ghi]perylene	90.94 ± 21.93		
屈 Chrysene	222.49 ± 7.09		
晕苯 Coronene	16.47 ± 5.58		
二苯并[a,h]蒽 Dibenz[a,h]anthracene	13.21 ± 2.64		
荧蒽 Fluoranthene	228.44 ± 45.70		
芴 Fluorene	9.81 ± 0.74		
茚并[1,2,3-cd]芘 Indeno[123-cd]pyrene	72.36 ± 8.11		
萘 Naphthalene	3.47 ± 0.52		
菲 Phenanthrene	196.63 ± 26.37		
芘 Pyrene	214.27 ± 41.32		

Metals	Concentration (ng/mg)
Mg	4602 ± 24
Pb	285 ± 1
Ca	2126 ± 144
Ti	351 ± 1
V	38 ± 0.1
Cr	78 ± 1
Mn	641 ± 2
Fe	1736 ± 15
Ni	55 ± 1
Cu	342 ± 2
<b>Zn</b>	<b>4659 ± 732</b>
As	335 ± 0.4





# Air pollution and cancer: biomarker studies in human populations

- **Markers of DNA or protein adducts:** 1-hydroxypyrene, benzene metabolites (phenylmercapturic acid (PMA) and trans-trans-muconic acid (ttMA)), 8-oxo-7,8-dihydro-2-deoxyguanosine (8-oxodG) in **urine**, DNA strand breaks, base oxidation, 8-oxodG and PAH bulky adducts in **lymphocytes**;
- **Markers of oxidative stress-related gene:** genotypes of glutathione transferases (GSTs) and NADPH: quinone reductase (NQO1) in plasma;
- **Markers of Methylation:** SAT $\alpha$  and NBL2 in blood leukocytes, APC, p16, p53 and RASSF1A in peripheral blood leukocyte;
- **Markers of macrophages:** Carbon loading in airway macrophages
- Three biomarkers of oxidative stress were found to be significantly associated with the personal PM<sub>2.5</sub> exposure, namely **8-oxodG, PLAAS and MDA**.



# Biomarkers of DNA or protein adducts

**Carcinogenesis vol.26 no.11 pp.1846–1855, 2005.**

**Table I.** Studies on DNA or protein adducts in human populations exposed to different air pollution levels

Reference	Country	Population	Measure	Levels <sup>a</sup>	P-value	Notes
DNA adducts						
Perera <i>et al.</i> (10)	Poland	Highly-industrialized area	PAH–DNA adducts	30.4/10 <sup>8</sup> versus 11 (rural area)	<i>P</i> < 0.05	Winter levels
Peluso <i>et al.</i> (8)	Italy	Police officers	Bulky DNA adducts	1.3/10 <sup>8</sup> versus 0.9	<0.05	In summer: 2.8 versus 0.8 ( <i>P</i> < 0.001)
Nielsen <i>et al.</i> (11)	Denmark	Bus drivers	PAH–DNA adducts	1.2 fmol/microg versus 0.585	0.04	Rural controls: 0.074, <i>P</i> < 0.001)
Palli <i>et al.</i> (12)	Italy	Traffic workers	Bulky DNA adducts	13.7/10 <sup>9</sup> versus 11.0	0.10	Among never-smokers <i>P</i> = 0.03
Georgiadis <i>et al.</i> (13)	Greece	Students with different air pollution exposures	Bulky DNA adducts	1.25/10 <sup>8</sup>	<0.001 versus 1.54 <sup>b</sup>	
Ruchirawa <i>et al.</i> (9)	Thailand	Police officers	Bulky DNA adducts	1.6/10 <sup>8</sup> versus 1.2	0.03	
Sorensen <i>et al.</i> (14)	Denmark	Students with different pollution exposures	8-oxo-dG	0.01 <sup>c</sup>	0.007	
Peluso <i>et al.</i> (15,123)	10 European countries	Residents with different pollution levels	Bulky DNA adducts	0.066 <sup>d</sup>	0.0095	Never or ex-smokers
Protein adducts						
Pastorelli <i>et al.</i> (20)	Italy	Newspaper vendors	Benzopyrene–hemoglobin adducts	0.3 fmol/mg versus <0.1	0.09	
Richter <i>et al.</i> (21)	Germany	Children in towns with different pollution levels	4-ABP–hemoglobin adducts	30.7 pg/g Hb 20.7	<0.001	

<sup>a</sup>More exposed vs less exposed.

<sup>b</sup>Higher adducts levels in the least polluted area.

<sup>c</sup>Regression coefficient between unit increments of PM<sub>2.5</sub> and adduct levels.

<sup>d</sup>Regression coefficient between unit increments of O<sub>3</sub> and adduct levels.



# 效应标志

- **效应标志 ( effect marker , MEF)** : 宿主暴露后产生功能性或结构性变化的生物标志。
- **生物学反应标志 ( BR )** : 反映了由于结合到靶细胞的外源性物质的持续作用 , 进一步引起细胞与组织的生物学或生物化学的变化。这些变化主要发生在细胞特定部位 , 尤其在基因的某些特定序列 ; 这些变化可引起机体某些不可逆转的生物学效应。
- **疾病标志 ( MD )**
- **健康状态标志 ( MH )**



# 生物学反应标志

生物学标志	环境暴露物质	生物材料
染色体	诱变化学物	
畸变	苯并 ( a ) 芘	白细胞
姊妹染色体单体交换	苯并 ( a ) 芘	白细胞
微核 ( micronuci )	苯并 ( a ) 芘	白细胞
癌基因激活	苯并 ( a ) 芘	组织
原嘌呤比例升高	铅	红细胞
乙酰胆碱脂酶降低	有机磷杀虫剂	血浆



# 亚临床（早期）疾病的生物学标志

---

**生物学标志**

**疾病**

---

**血清甲胎蛋白**

**肝癌、肠道疾病、胎儿神经管缺失**

**癌胚抗原**

**胃肠道癌、其它胃肠道疾病**

**肿瘤特异性抗原**

**各种癌症**

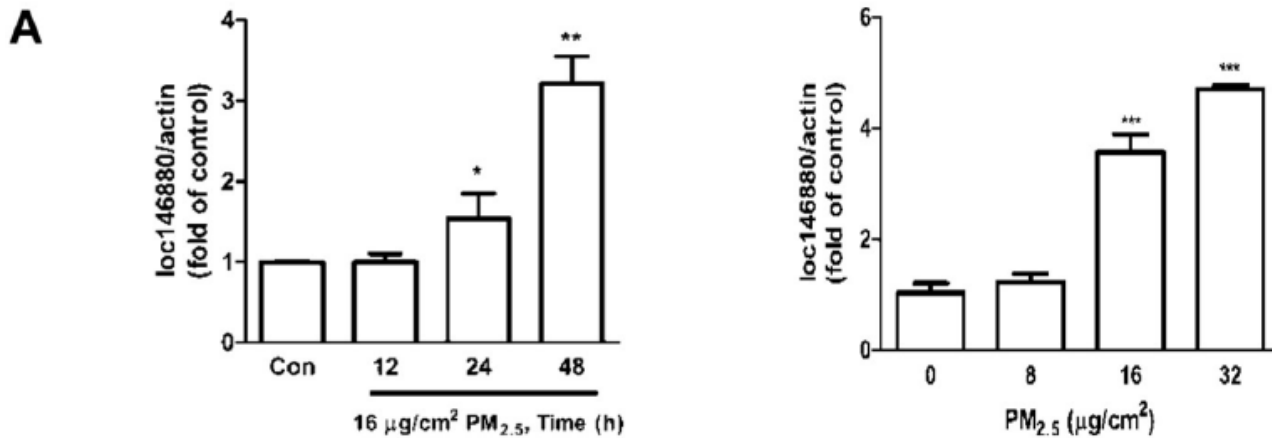
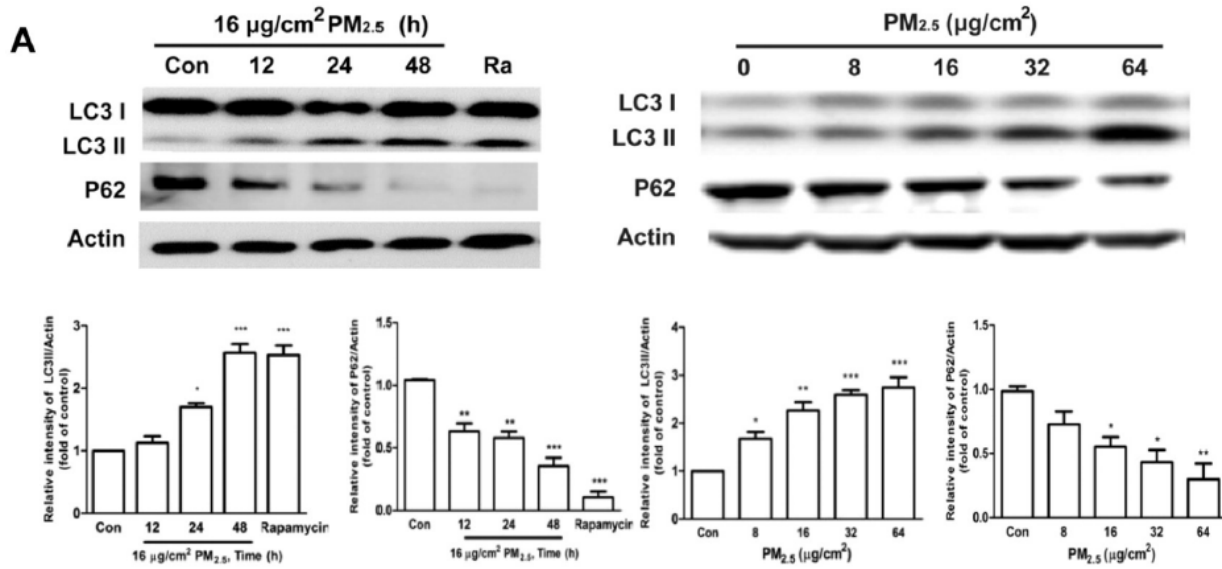
**血清谷草转氨酶**

**心肌梗塞**

---



# 举例-效应标志物





# Air pollution and cancer: biomarker studies in human populations

- **Markers of DNA or protein adducts:** 1-hydroxypyrene, benzene metabolites (phenylmercapturic acid (PMA) and trans-trans-muconic acid (ttMA)), 8-oxo-7,8-dihydro-2-deoxyguanosine (8-oxodG) in **urine**, DNA strand breaks, base oxidation, 8-oxodG and PAH bulky adducts in **lymphocytes**;
- **Markers of oxidative stress-related gene:** genotypes of glutathione transferases (GSTs) and NADPH: quinone reductase (NQO1) in plasma;
- **Markers of Methylation:** SAT $\alpha$  and NBL2 in blood leukocytes, APC, p16, p53 and RASSF1A in peripheral blood leukocyte;
- **Markers of macrophages:** Carbon loading in airway macrophages
- Three biomarkers of oxidative stress were found to be significantly associated with the personal PM<sub>2.5</sub> exposure, namely **8-oxodG, PLAAS and MDA**.



# Biomarkers of gene mutations and gene promoter

Table II. Summary of findings from human biomarker studies and animal *in vivo* studies on gene mutations and gene promoter hypermethylation associated with exposure to indoor or outdoor air pollution, some of their major components, or cigarette smoke (as a model exposure for PAHs)

Alteration/molecular biomarker studied	Cell type/tissue studied	Type of exposure associated	Comments	Reference
<b>Gene mutations</b>				
<b>Human studies</b>				
<i>HPRT</i> gene mutations in healthy newborns	Cord blood lymphocytes	Polluted outdoor air (urban)	Mutation frequency correlated with presence of aromatic DNA adducts. Negative findings in the mothers	(73)
			Negative findings in peripheral blood lymphocytes in adults	(71,72)
<i>TP53</i> gene and/or <i>RAS</i> gene mutations in non-lung cancer patients who are smokers, or non-smokers without evidence of lung cancer	Lung tumour tissue; non-malignant epithelial cells from sputum	PAH-rich emissions from smoky coal combustion in unvented fireplaces or stoves (indoor exposure)	Frequent in lung tumour tissue. In non-malignant cells, <i>TP53</i> mutations present with a low frequency	(88,90,91)
<b>Experimental studies</b>				
Mutations of the <i>lambda/lacI</i> transgene in rat transgenic assay	Lung tissue	Diesel exhaust (inhalation exposure, 4 weeks)	Also <sup>32</sup> P-labelled aromatic DNA adducts and 8-oxo-dG increased	(37)
			Negative findings on mutations in other rodent studies	(43,52)
<i>p<sup>um</sup></i> allele deletions in mouse embryos	Retinal pigment epithelium	Diesel exhaust particles (transplacental exposure for embryonic days 10.5–15.5 following oral exposure of the pregnant dams)	70 kb deletions spanning exons 6–18 of the <i>p<sup>um</sup></i> allele in <i>p<sup>um</sup>/p<sup>um</sup></i> offspring mice. <sup>32</sup> P-post-labelling adducts and 8-oxo-dG levels not significantly increased	(53)
<i>p53</i> gene and <i>K-ras</i> gene mutations in rats	Lung tumour tissue	Diesel exhaust; carbon black (inhalation exposure for 24 months)	Infrequent	(54,55)
<i>K-ras</i> gene mutations in rats	Lung adenomas and adenocarcinomas	Diesel exhaust particles (intracheal instillation for 10 weeks, tumours studied after 30 months)	Frequent after intratracheal instillation but not increased after inhalation exposure.	(56)
<b>Promoter methylation<sup>d</sup></b>				
<b>Human studies</b>				
<i>p16<sup>INK4A</sup></i> gene methylation in smoking lung cancer patients	Lung tumour tissue, precursor lesions to lung carcinoma, non-malignant bronchial epithelial cells from brush and sputum samples, and serum DNA	Cigarette smoking	Frequent in NSCLC. Promoter methylation of various other genes also frequently detected	(Reviewed in 100)
<i>p16<sup>INK4A</sup></i> gene methylation in cancer-free smokers	Non-malignant bronchial epithelial cells from brush and sputum samples	Cigarette smoking	Present with a lower frequency as compared to the tumour tissue	(103,104,108, 111–113)
<b>Experimental studies</b>				
<i>DAPK</i> gene, and <i>RAR-β</i> gene methylation in mice	Lung tumour tissue	Cigarette smoke (whole body exposure for 30 months)	Frequent	(116,117)
<i>p16<sup>INK4A</sup></i> gene methylation in rats	Lung tumour tissue	Diesel exhaust; carbon black (inhalation exposure for 24 months)	Frequent	(115)

<sup>d</sup>Data available mainly originates from exposure to cigarette smoke (smokers and experimental data). Abbreviations: NSCLC, non-small cell lung cancer; 8-oxo-dG, 8-oxo-7,8-dihydro-2'-deoxyguanosine.





# 易感性标志

- **易感性标志** ( susceptibility markers, MS ) 个体对疾病发生、发展易感程度的生物标志。
- 易感性主要与宿主的遗传特征，以及生长发育、营养、免疫、机体活动状态等有关。**不同疾病的不同阶段，都有可能具有不同的易感性标志。**
- **易感性影响着其它生物学标志在机体内的水平；**不同个体的易感性不同，与之相关的其它生物学标志的水平也可以不同。

**Table 1.** Characteristics of the Health 2006 study population

	<i>n</i> /total (%)
Sex	
Men	1553/3471 (44.7%)
Women	1918/3471 (55.3%)
Age in years	
18–28	269/3471 (7.8%)
≥28–38	444/3471 (12.8%)
≥38–48	872/3471 (25.1%)
≥48–58	854/3471 (24.6%)
≥58–69	1032/3471 (29.7%)
GSTM1 genotype*	
GSTM1 null	1770/3364 (52.6%)
GSTM1 one copy	1322/3364 (39.3%)
GSTM1 two or more copies	272/3364 (8.1%)
GSTT1 genotype*	
GSTT1 null	493/3344 (14.7%)
GSTT1 one copy	1578/3344 (47.2%)
GSTT1 two or more copies	1273/3344 (38.1%)
Smoking status	
Current	773/3437 (22.5%)
Occasionally	113/3437 (3.3%)
Previously	1116/3437 (32.5%)
Never	1435/3437 (41.8%)
Type of housing	
Apartment	963/3424 (28.1%)
House	1629/3424 (47.6%)
Row house	782/3424 (22.8%)
Other	50/3424 (1.5%)
Size of residence; square meters (m <sup>2</sup> )	
Less than 70 m <sup>2</sup>	364/3438 (10.6%)
70–110 m <sup>2</sup>	1380/3438 (40.1%)
More than 110 m <sup>2</sup>	1694/3438 (49.3%)



美国当地时间5月14日，好莱坞知名女星、联合国难民事务高级办事处亲善大使安吉丽娜·朱莉（Angelina Jolie）在《纽约时报》（the New York Times）发表了一篇名为《My Medical Choice》的专栏文章，文章披露她已于4月27日接受了预防性的双侧乳腺手术，以降低罹患风险。

朱莉写道，自己之所以进行这项手术，是因为她有基因缺陷（携带BRCA1基因突变），罹患乳癌和卵巢癌风险较高。原因是她担心自己有遗传自母亲的基因缺陷、罹患乳腺癌的风险高，此法以降低其患癌的风险。文章透露，朱莉的母亲Marcheline Bertrand与癌症作斗争了近十年，并在56岁去世。她的医生估计，她有87%的患乳腺癌的风险，50%的卵巢癌的风险。她要主动将风险降到最低。手术后，她患乳腺癌的几率从87%下降到5%以下。在文章中，她表示，“我的这个决定，丝毫不会削弱我的女人味”。



## **易感性基因标志物：**

**1. 人类基因组多态性（SNPs）**

**2. 基因组表遗传修饰（甲基化，乙酰化）**



# 全基因组关联分析

- GWAS ( Genome-wide association study ) ，即全基因组关联分析，是指在人类全基因组范围内找出存在的序列变异，即单核苷酸多态性 ( SNP ) ，从中筛选出与疾病相关的SNPs。
- 有疾病的人 ( 病例 ) 和相同条件的无该疾病的人 ( 对照 ) 。每个人都提供些作为样本的细胞，如从口腔内侧擦下的表皮细胞，可以从这些细胞中提取DNA，并涂布在基因芯片上，该芯片上可以读取上百万个DNA序列。这些芯片被读入计算机，在那里通过生物信息学技术对其进行分析，而不是阅读的全部DNA序列，这些系统通常读的是各个SNP，这些SNP便是成组的DNA变异 ( 单倍型 ) 的标记。



# What is SNP?

**A mutation in the DNA sequence with a frequency of >1%**

## Spelling Option

L I V E  
L O V E

## Genetic Alphabetic Order 4 Letters

A T C G

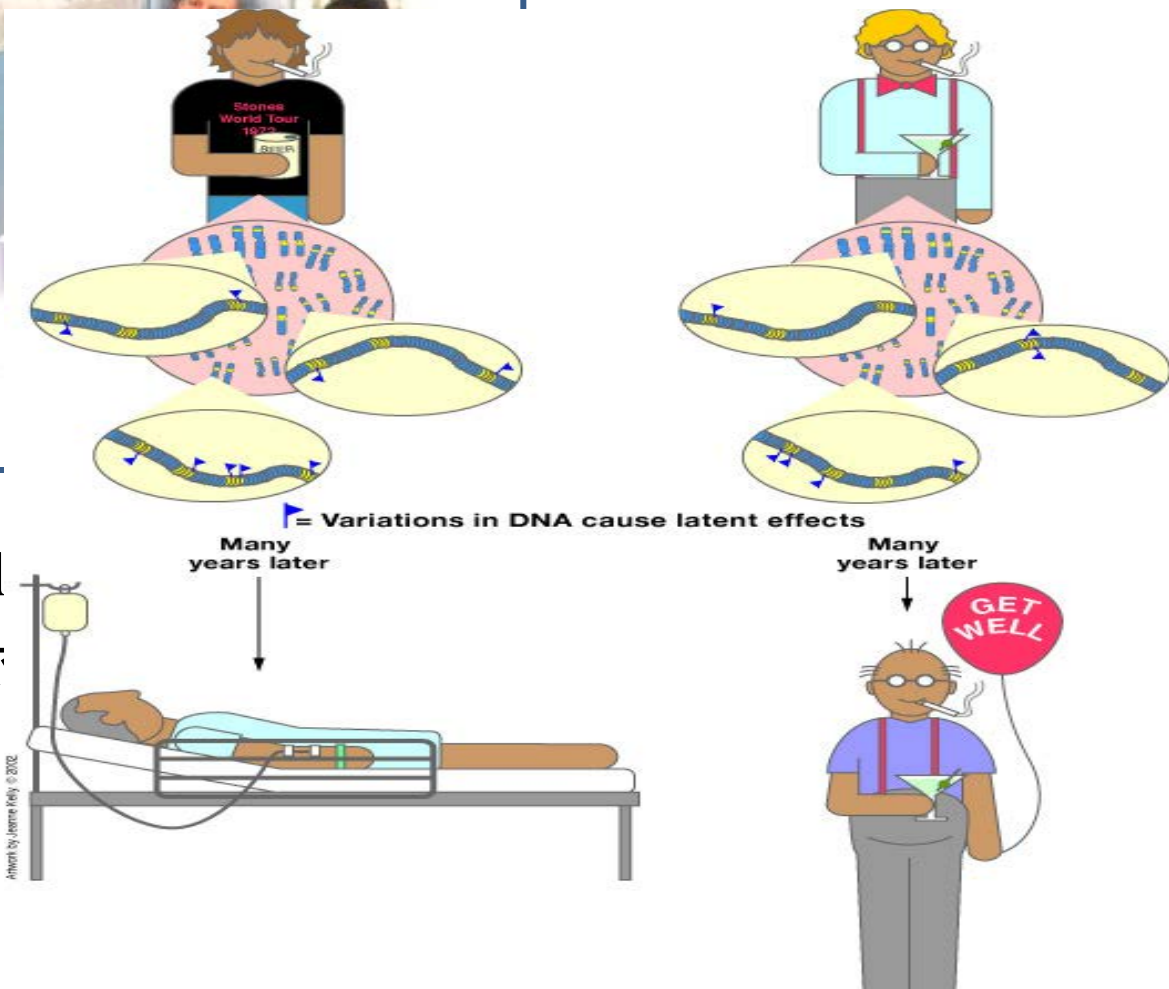
....A T T C C G **G** G T A C T A C T....  
....A T T C C G **A** G T A C T A C T....

▲  
**SNP**



(ism, SNP)

突变。SNP存在于整  
在一个SNP，估计总



- 研究表明，正是这0.1小小的差异还与个体对应性差别有密切关系。





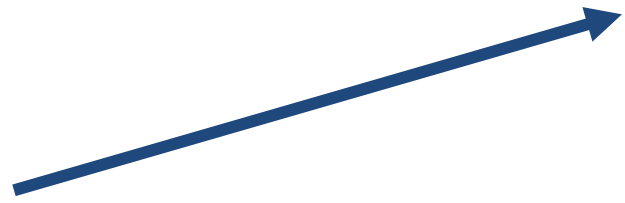
# 研究问题 1:

## Why only < 20% smokers develop lung cancer?

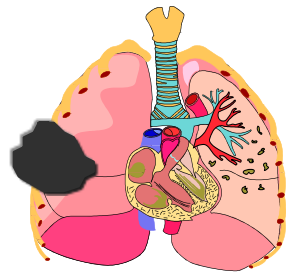
**Exposure**



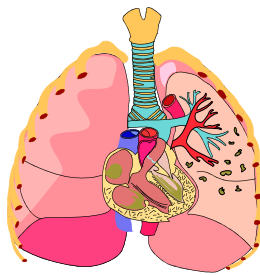
**Genetic susceptibility?**



**SNPs?**



**Cancer**



**Cancer-free**

**CIR by 70 yrs**

**M: 15.9%**

**F: 9.5%**

(Peto, BMJ, 2000)

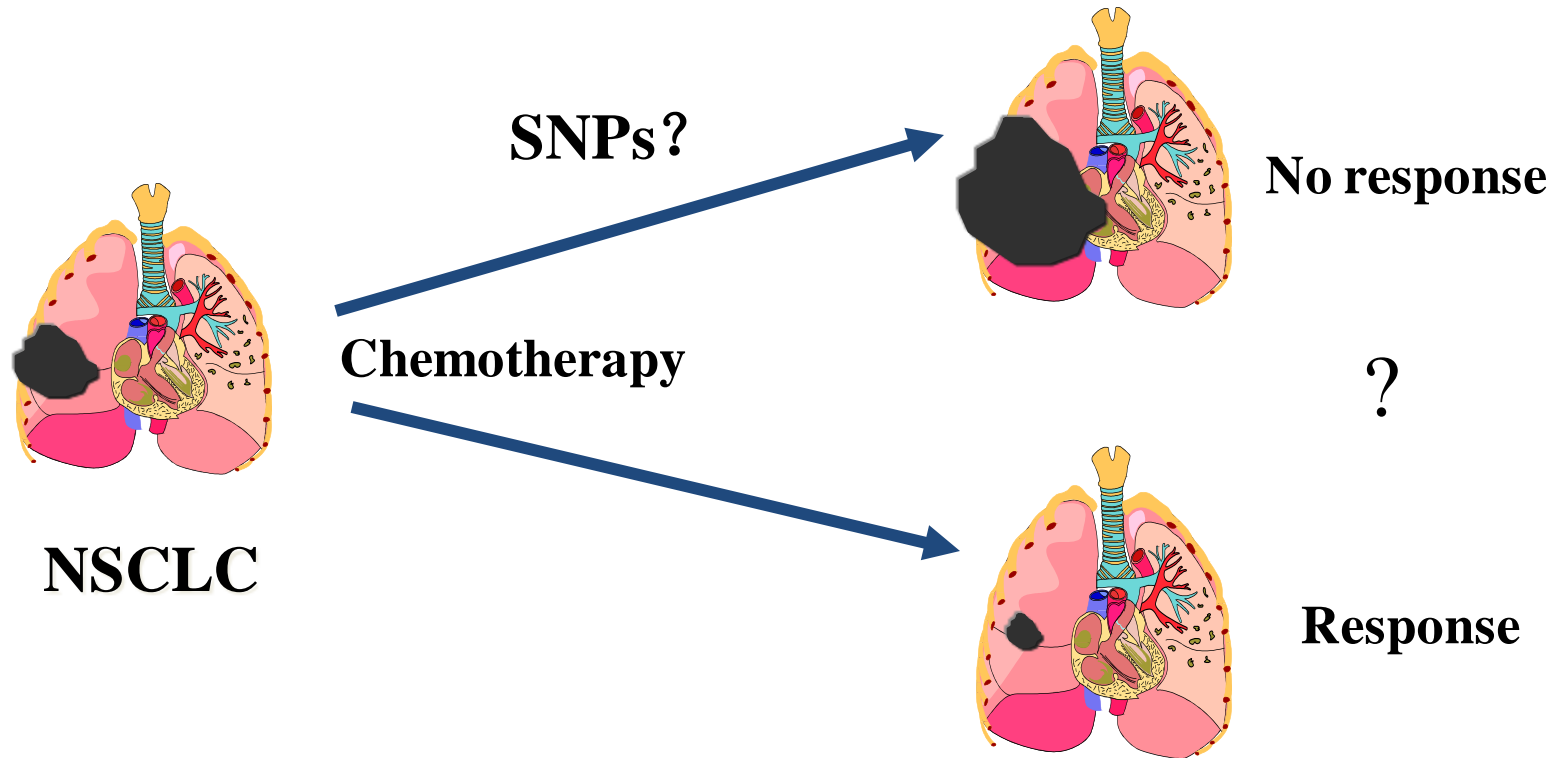
**Screening, prevention**





## 研究问题 2:

Why some lung cancer patients have no response to chemotherapy?



SNPs & Chemosensitivity



# GWAS of Never-smoking Female Lung Cancer

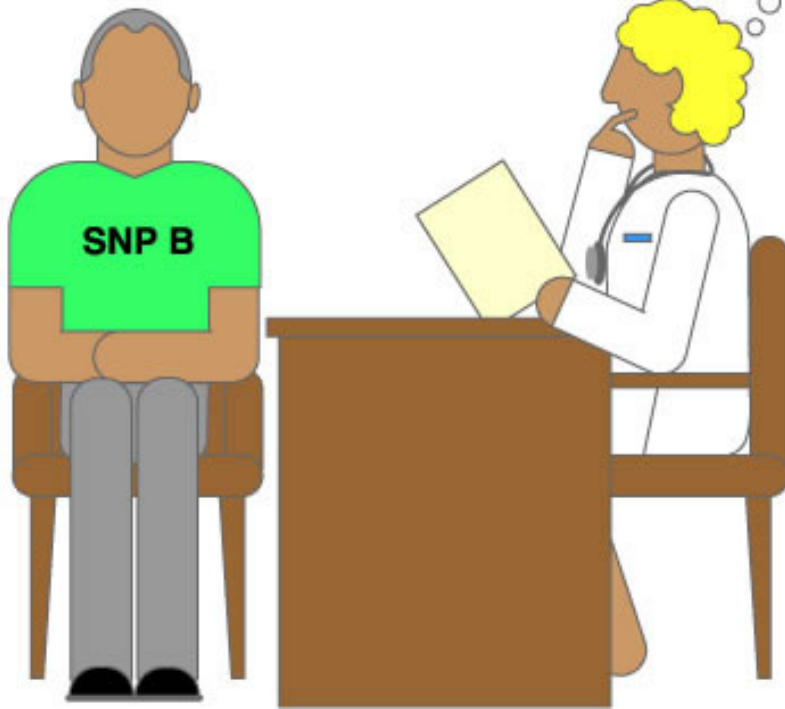
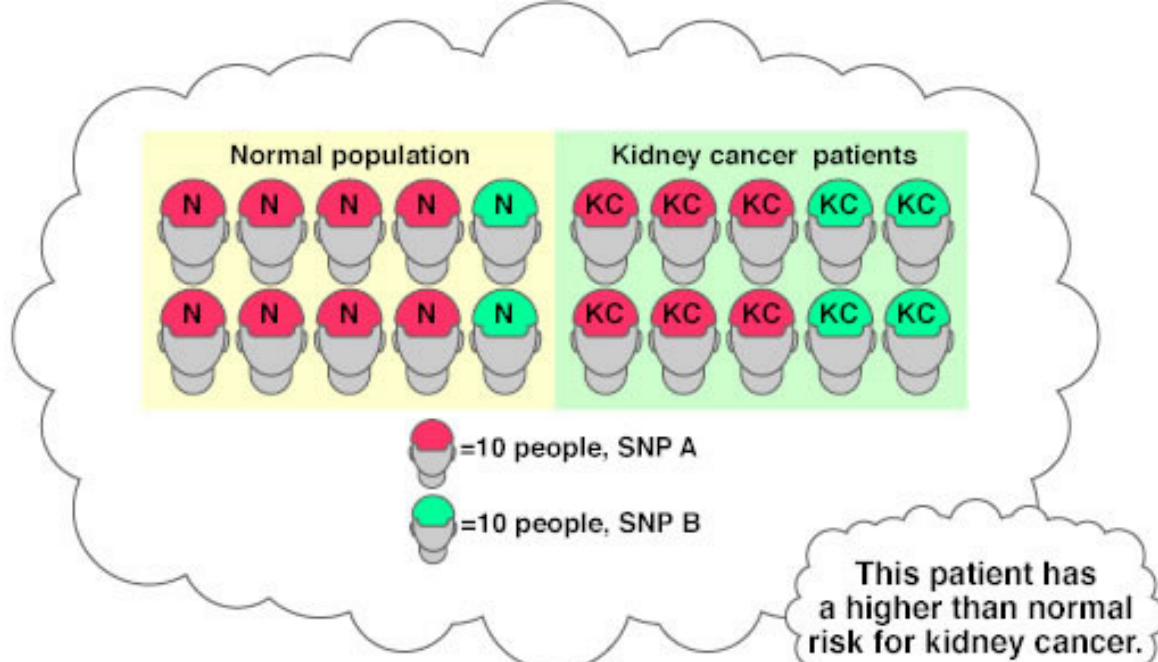
nature  
genetics

## Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia

**Table 1** New loci associated with lung cancer in a GWAS of never-smoking Asian females

SNP	Plausible candidate gene(s)	Chromosome position	Subset	Allele <sup>a</sup>	MAF <sup>b</sup>		Subjects		OR (95% CI)	<i>P</i> <sub>trend</sub>
					Control	Case	Control	Case		
rs7086803	<i>VTG1A</i>	10q25.2	GWAS	G/A	0.26	0.32	4,492	5,457	1.32 (1.24–1.41)	5.04 × 10 <sup>-17</sup>
			Replication	G/A	0.27	0.31	2,887	1,085	1.23 (1.10–1.37)	3.36 × 10 <sup>-4</sup>
			Combined	G/A	0.27	0.31	7,379	6,542	1.28 (1.21–1.35)	3.54 × 10 <sup>-18</sup>
rs9387478	<i>ROS1, DCBLD1</i>	6q22.2	GWAS	C/A	0.50	0.46	4,542	5,510	0.85 (0.81–0.90)	7.79 × 10 <sup>-8</sup>
			Replication	C/A	0.49	0.47	2,891	1,091	0.92 (0.83–1.01)	0.088
			Combined	C/A	0.50	0.46	7,433	6,601	0.85 (0.81–0.90)	4.14 × 10 <sup>-10</sup>
rs2395185 <sup>c</sup> (rs28366298)	HLA class II region	6p21.32	GWAS	G/T	0.35	0.38	4,541	5,504	1.16 (1.09–1.23)	2.60 × 10 <sup>-6</sup>
			Replication	A/C	0.37	0.42	2,880	1,008	1.20 (1.08–1.33)	7.93 × 10 <sup>-4</sup>
			Combined	Meta			7,421	6,512	1.17 (1.11–1.23)	9.51 × 10 <sup>-9</sup>

<sup>a</sup>Minor allele listed second. <sup>b</sup>Minor allele frequency. <sup>c</sup>For the HLA class II region, because a TaqMan assay could not be designed for rs2395185, we instead genotyped rs28366298, its perfect surrogate ( $r^2 = 1.0$ ), by TaqMan. The reported *P* value is based on meta-analysis of the rs2395185 results in the GWAS and the rs28366298 results in the TaqMan set.



# 基因检测- 危险度评价



# Human Genome Project

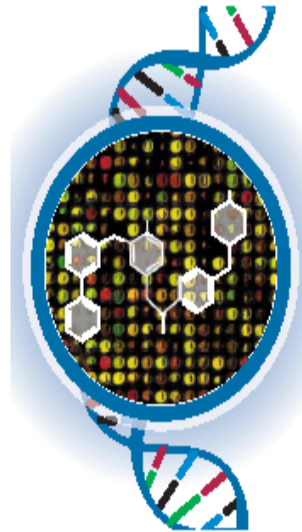
**Biology**



**SNPs**



**Response**



**Disease**



**Prevention**

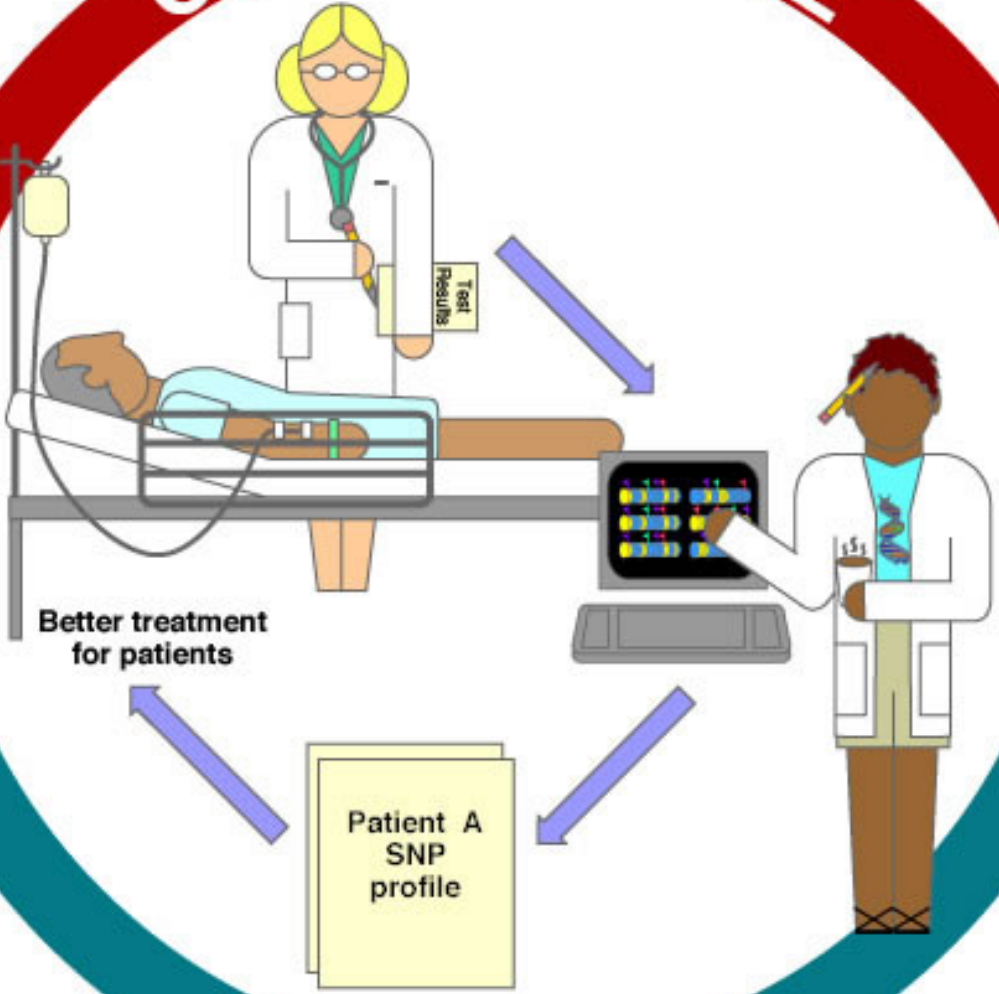


**SNPS=Single Nucleotide Polymorphisms**

**人类基因组计划**

Collins et al., Nature, 2003

CLINIC/BEDSIDE



Better treatment for patients

Patient A  
SNP  
profile

RESEARCH

从科学研究  
→ 临床实践

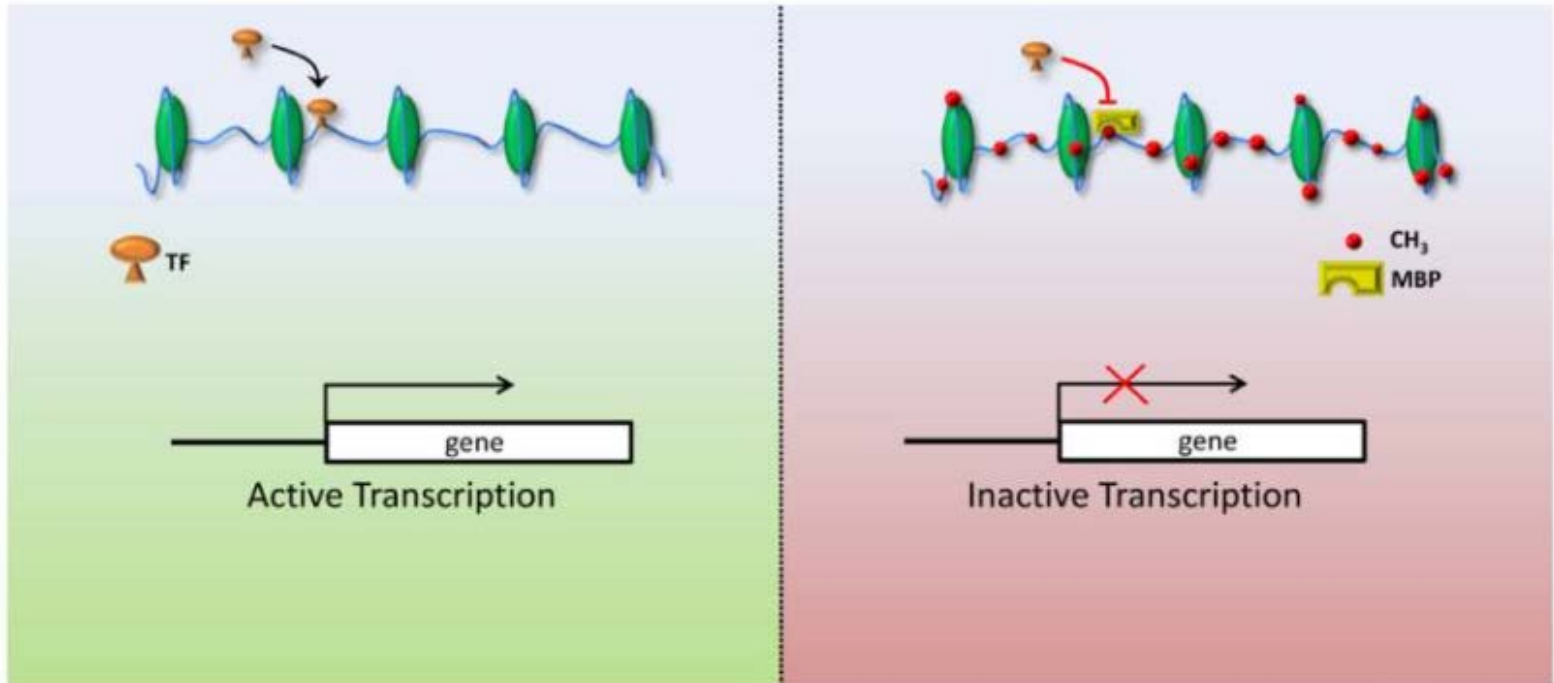
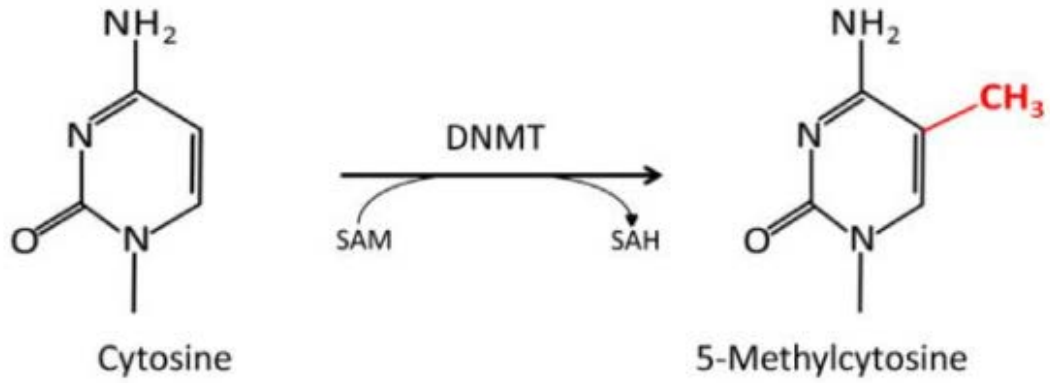




# DNA甲基化

- DNA甲基化是在DNA甲基化转移酶的作用下将甲基选择性地添加到胞嘧啶上形成5-胞嘧啶的过程；
- 刚被发现时被定义为第五种碱基，实际上它是一种重要的表观遗传学标记；
- 在调控基因表达、维持染色质结构、基因印记、X染色体失活以及胚胎发育等生物学过程中发挥着重大的作用。
- 它就像DNA的一顶最神奇的“帽子”。





DNA甲基化修饰和对基因表达的转录调控作用



# 生物标志的选定（一）

## 1.生物标志的筛选：

被研究的生物标志必须具备如下条件：

- 1) 较强的特异性和稳定性；
- 2) 检测的方法具有较高的灵敏度和特异度；
- 3) 检测方法快速、简便。





# 生物标志的选定（二）

## 2. 生物标志的特性

- 分子特性、
- 时相特性、
- 个体内变异、
- 个体间变异、
- 群体间变异、
- 储存变异。



# 生物标志的选定（三）

## 3.检测方法的实用性和可信性

生物标志检测的“最佳”方法：成熟稳定、操作简便、标本容易采集。



# 生物标志的选定（四）

## 4. 检测方法的有效性

生物标志检测方法的灵敏度和特异度

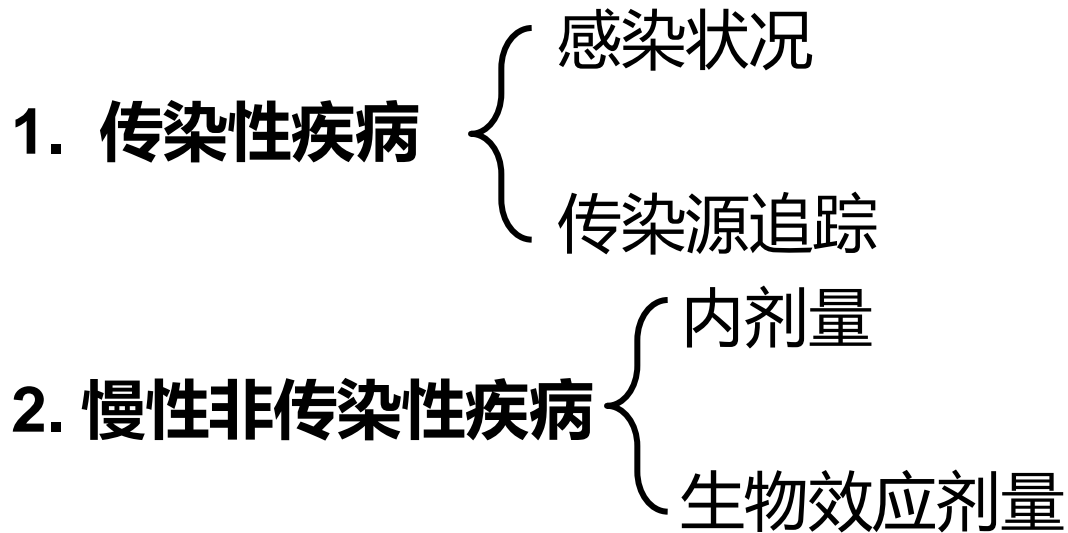


# 暴露测量

## (一) 外暴露研究



## (二) 内暴露研究





# 效应测量

## (一) 传染病

免疫效应、病理性效应

## (二) 慢性非传染性疾病(NCD)

- 基因表达异常和代谢异常；
- 基因突变或染色体畸变

## (三) 健康状况



# 易感性测量

## (一) 遗传性疾病

- Huntington病(HD)： IT15基因第1外显子内部不稳定CAG三核苷酸重复序列异常扩展而发病；

## (二) NCD

- 载脂蛋白E与III型高脂蛋白血症有密切关系：ApoE多态性是动脉粥样硬化早期及发展过程中个体差异的主要原因；

## (三) 传染病

- 利用基因标志对病原体致病的遗传易感性进行评判。



# 疾病防治效果评价

## (一) 传染性疾病预防效果评价

- 预防接种效果评价和预防接种相关发病研究

## (二) NCD防治效果评价

- 分子流行病学中最早的生物标志作为结局进行测量，缩短了效果评价的时间，对于不同措施的效果评价更加客观和准确。



# 目录 CONTENTS



● 概述

● 生物标志物

● **主要研究方法**

● 应用与展望





# 主要研究方法-选择测量指标

## 一般原则：

- 1.生物标志应特异、稳定；
- 2.标本采集、储存方便；
- 3.检测方法比较简单、实用，而且操作规范，便于与同类研究结果比较；
- 4.灵敏度和特异度高。



# 主要研究方法-选择测量指标（二）

## 测量指标：

- 1.暴露指标：(内暴露和外暴露)
- 2.效应指标：最早期的生物效应指标
- 3.易感性指标：

传染性疾病--抗体水平

NCD----易感基因及其表达产物



# 标本的采集和储存

## 一般要求：

- 1.在采集和储存过程中不能受到“污染”，包括外界生物的、化学的和其他标本的“交叉污染”。
- 2.储存的生物标本在任何时候进行检测都可以获得一致的结果。
- 3.所有生物标本都应有详细的背景资料和鉴别标识。
- 4.储存时应注意储存条件，一般为冷藏保存。



# 现场调查研究方法

分子流行病学研究中的指标检测需要一定的实验条件，  
样本宜小不宜大。

常用的流行病学的方法：描述性研究、分析性研究、  
实验性研究等都在分子流行病学研究中有所应用，描述  
性研究多为小样本的抽样调查，分析性研究中以病例对  
照研究较为常见。



# 常用分子生物学相关技术简介

**(一) 核酸技术**

**(二) 蛋白质技术**

**(三) 酶学技术**

**(四) 生物芯片技术**

**(五) 其他**



# 常用分子生物学相关技术简介

## (一) 核酸技术

### 1. 核酸电泳图谱

### 2. 核酸酶切电泳图谱

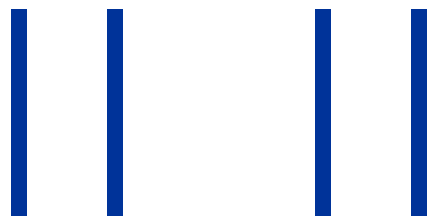
- 限制性核酸内切酶
- 限制性片段长度多态性 ( RFLP )

例：HaeIII识别DNA特异序列是

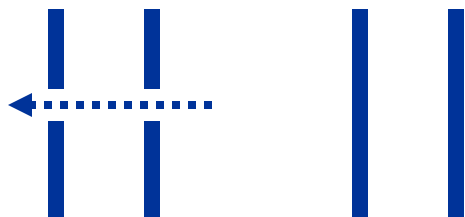
GG CC

CC GG

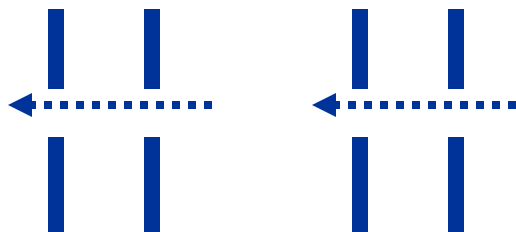
# 限制性内切酶酶切结果示意图



纯合未切开型



杂合型

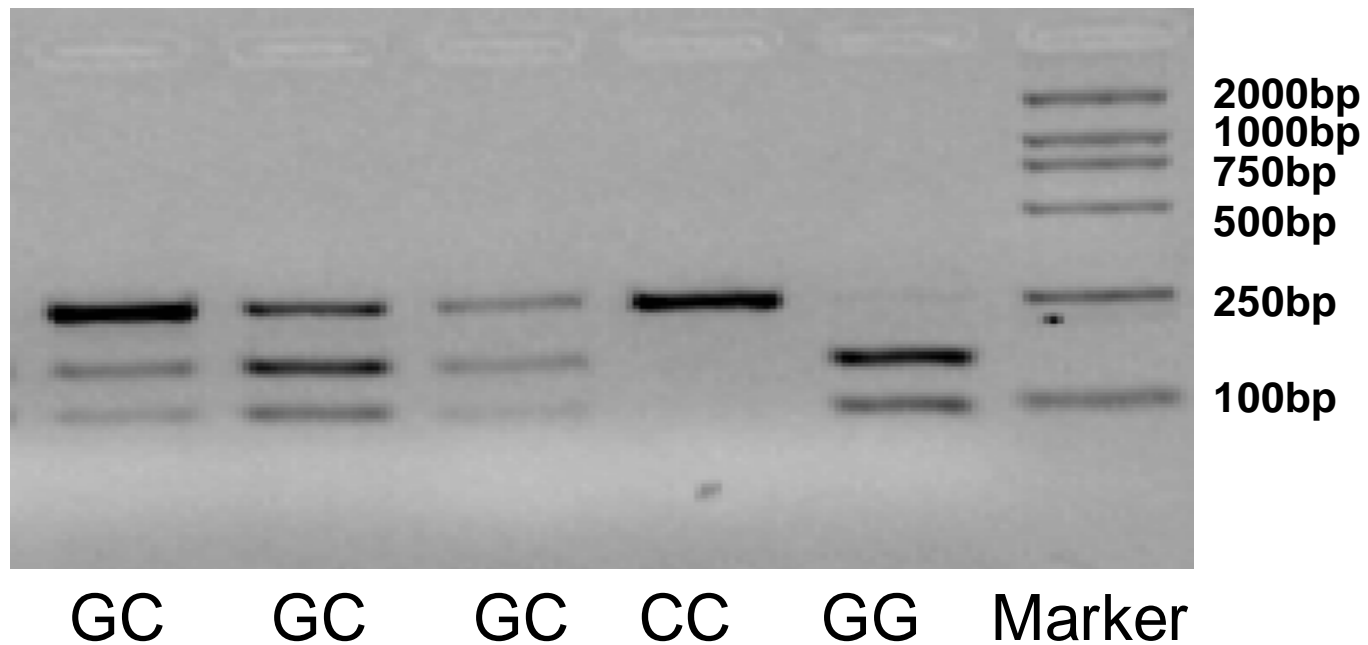


纯合切开型



# 限制性酶切结果

- HLA-DQB1基因片段经SacI限制性内切酶消化后，可分出三种基因型：未突变纯合子（基因型为CC），基因序列中无SacI 限制性酶切基因座，所以仅见一条电泳带；突变纯合子（基因型为GG），可见两条电泳带；突变杂合子（基因型为GC），可见三条电泳带。







# 常用分子生物学相关技术-核酸技术

## 3.核酸分子杂交：

- 原位杂交、斑点杂交、转印杂交(Southern blot和Northern blot)

## 4.核酸体外扩增

## 5.核酸序列分析



**Genetic  
Technologies**



## PCR扩增仪

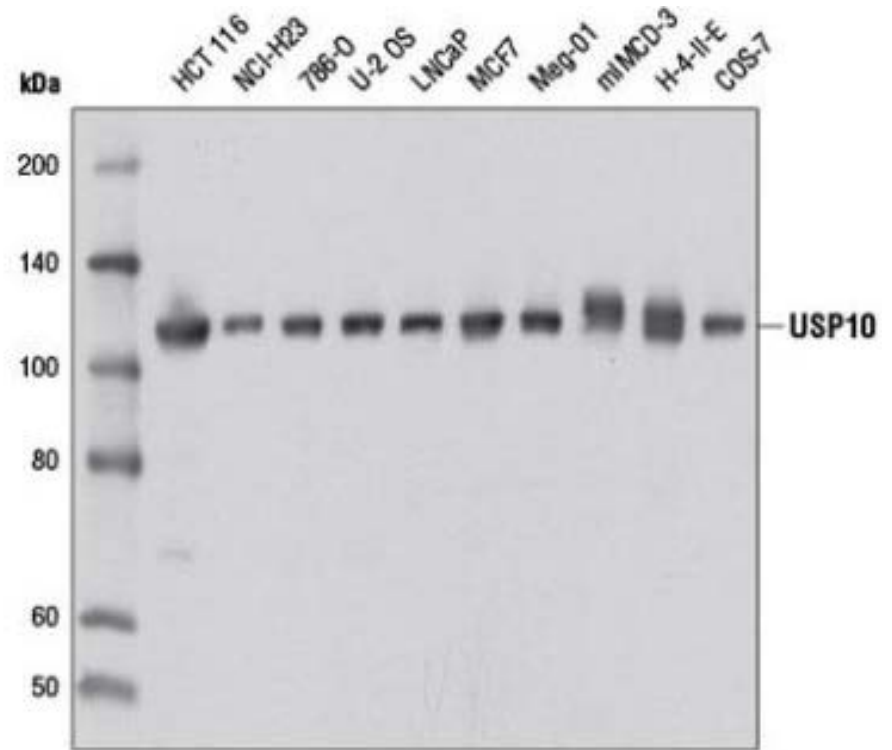
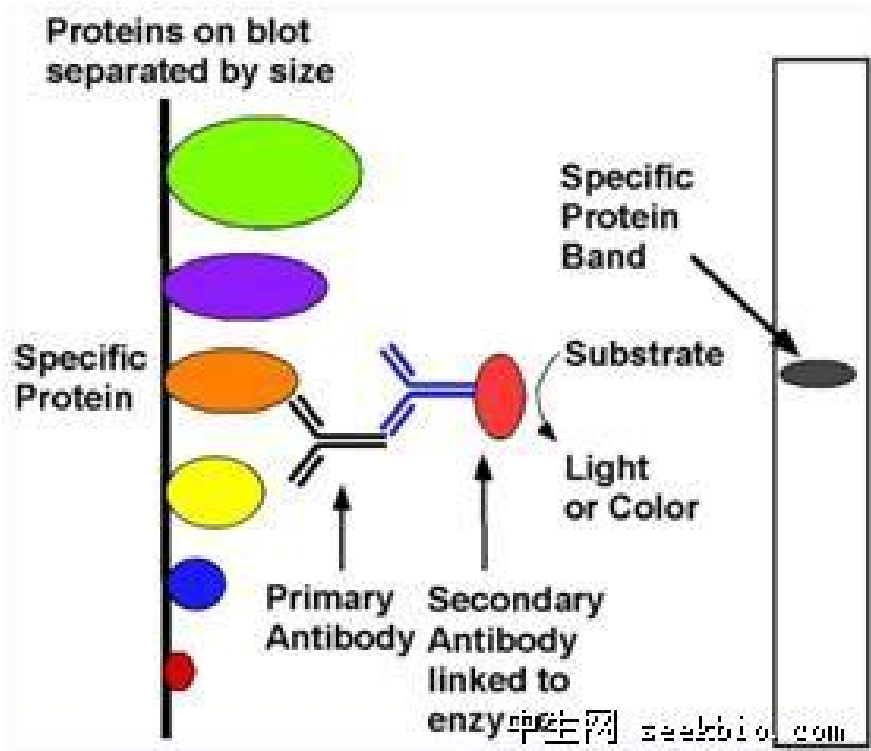


## 实时定量荧光PCR仪



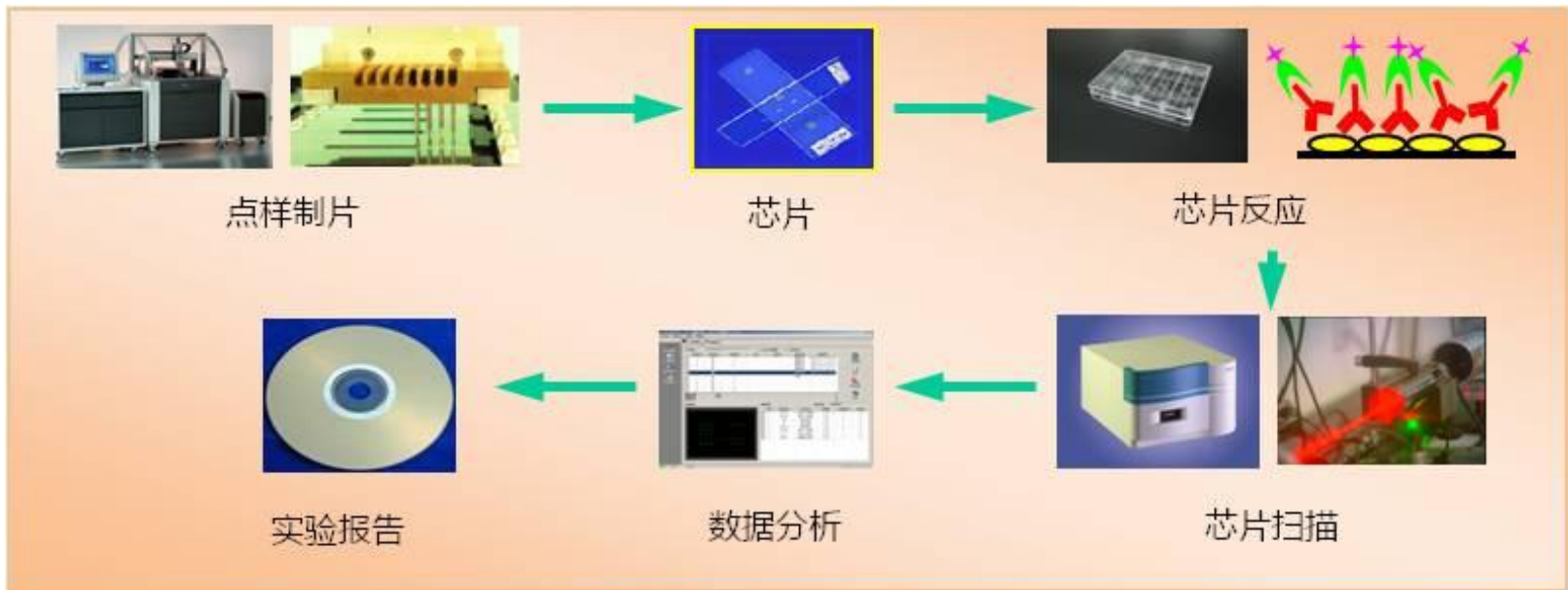


# 蛋白技术-western blotting

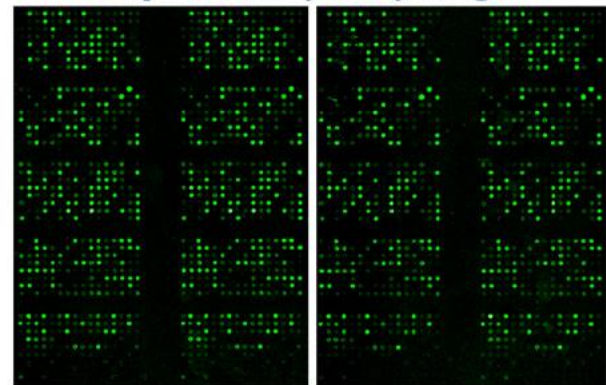




# 蛋白技术-蛋白芯片



Explorer Array Assay Images



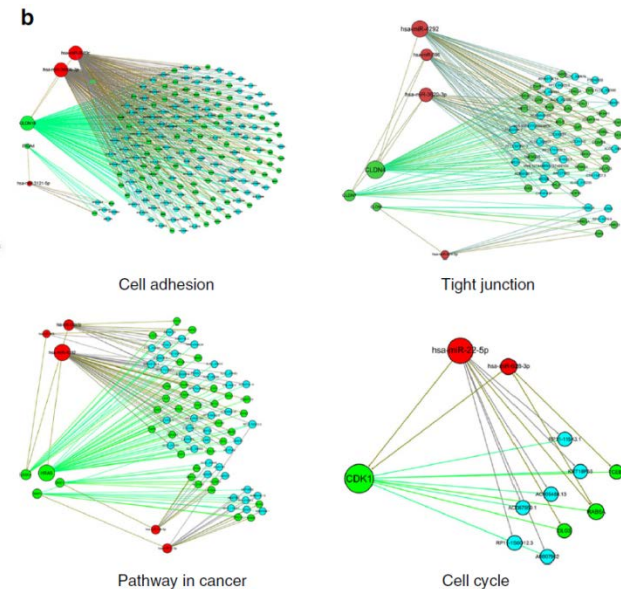
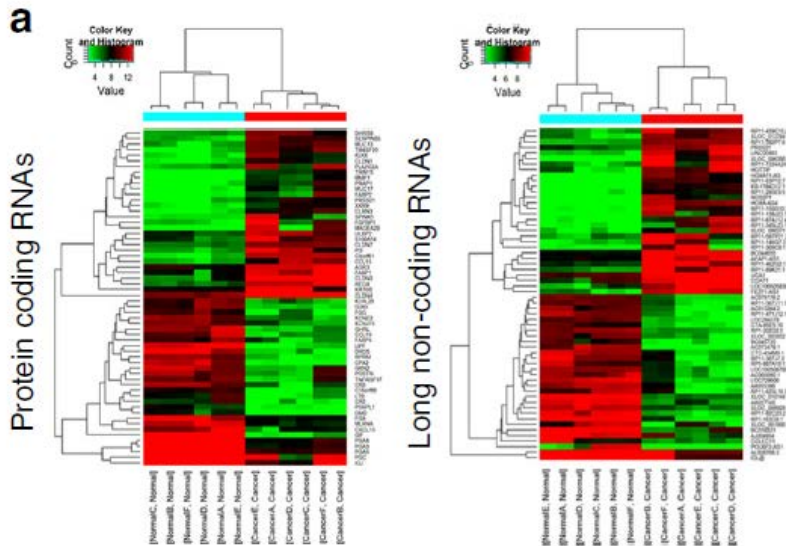
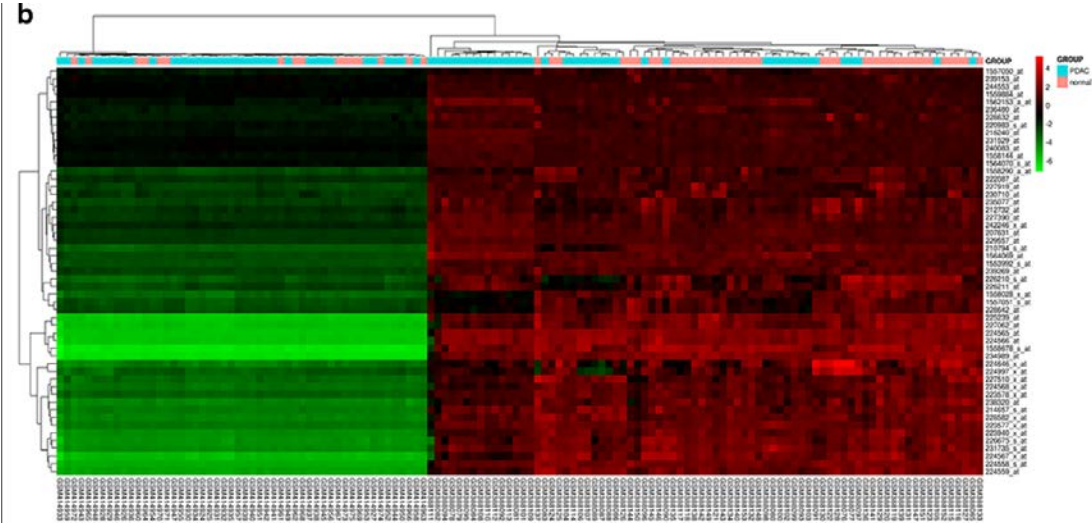
Control Sample

Treated Sample

656个点，14个对照；片内2次重复



# 生物芯片





# 研究设计与资料分析

1. 确定研究目的
2. 确定测量指标
3. 确定适宜的测量方法
4. 调查方法与样本量
5. 结果与分析
6. 质量控制
7. 注意事项



# 研究设计与资料分析

## 1.传统流行病学指标

如：生物标志发生率、检出率、危险度估计等。





# 研究设计与资料分析（二）

## 2.血清流行病学指标

如：抗体几何平均滴度、阳转率、感染率。





# 研究设计与资料分析（三）

## --遗传关系分析

- **基因型** (genotype) : 某 ( 些 ) 基因相同的生物 , 称为同一基因型。
- **克隆** (clone): 许多生物特征 ( 表型或基因型 ) 相同的生物体 , 虽然分离来源、地区、时间可能不同 , 但它们的这些特征几乎只能解释为共同来源 , 则这些生物体称为一个克隆。
- **克隆群** (clone cluster): 遗传关系非常紧密的多个生物克隆 , 可以划分为一个克隆群。
- **克隆关系** : 一个克隆 ( 或遗传关系非常紧密 , 只能解释为一个克隆 ) 的生物 , 称它们之间具有克隆关系。



# 研究设计与资料分析（四）

## 4. 遗传多态性分析

- 遗传多态度
- 等位基因频率
- 平均遗传多态度
- 遗传杂合度



# 研究实例

**题目：**6号染色体MHC II 区精神分裂症易感基因筛查研究

**测量指标：**SNPs(DQB1、DRB1、DRB3等)

**测量方法：**PCR，限制性酶切，电泳等

**研究对象：**精神分裂症核心家系

**结果分析：**DQB1位点与精神分裂症相关联 ( $X^2=4.112$  ,  $P=0.043$  )

**结论：**DQB1位点与精神分裂症相关联



# 目录 CONTENTS



概述

生物标志物

主要研究方法

**应用与展望**

# 分子流行病学的应用前景

- 分子流行病学是一门新兴的边缘学科，近几十年内随着分子生物学技术的发展得到了长足的进步，目前在传染病、NCD及健康状态的研究中显示出了巨大的作用;
- 随着先进的生物标志检测技术、计算机技术、信息技术和统计学方法的不断深入，分子流行病学会为维护人类健康发挥更大，更重要的作用。



**谢谢大家！**

**邓晓蓓**

