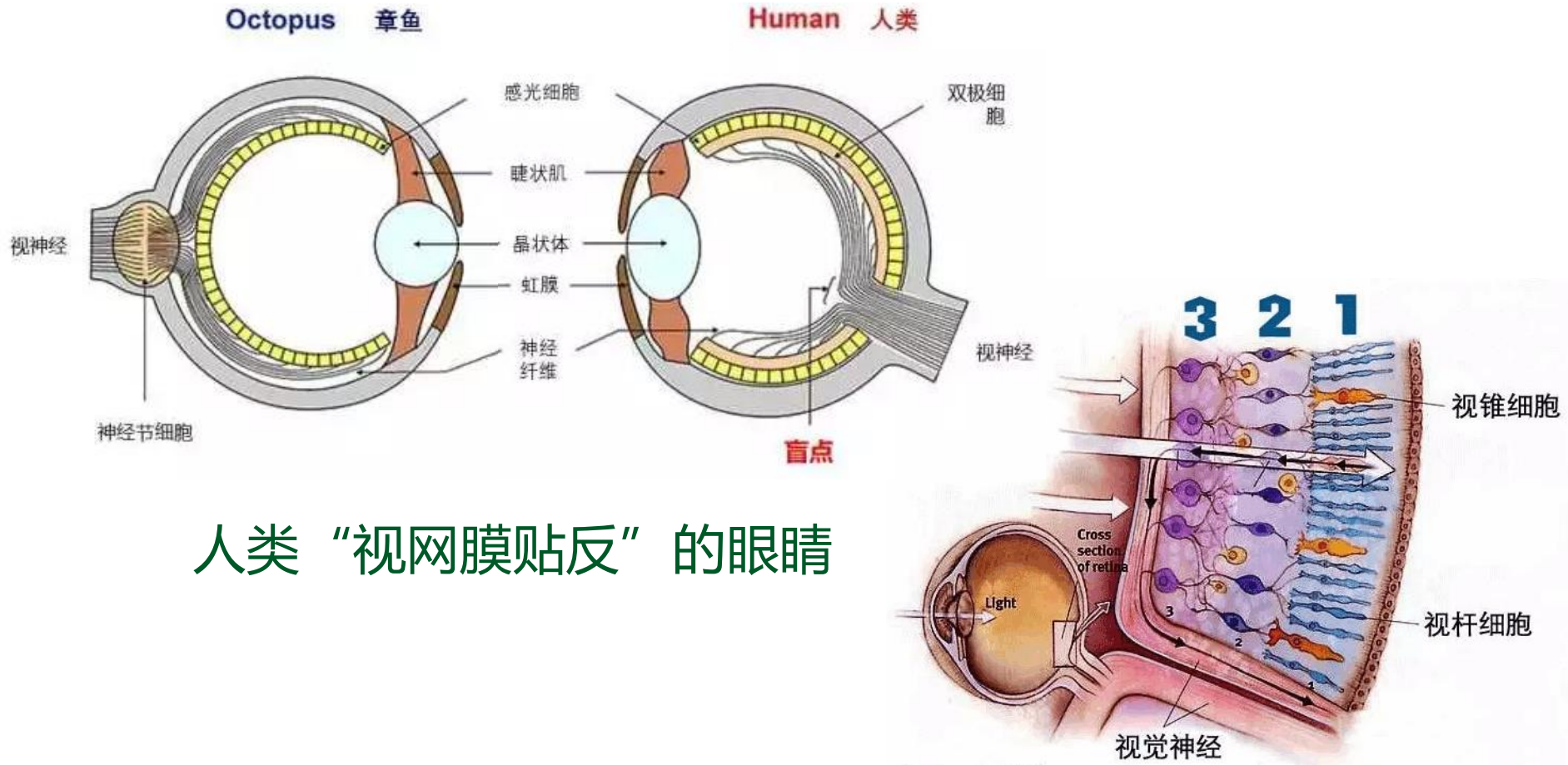




第十七章 生物演化与群体遗传

- 生物演化与物种形成
- 基因频率与哈代-温伯格平衡
- 突变、自然选择与遗传漂变

一、生物演化与物种形成



人类“视网膜贴反”的眼睛

生物演化学：研究生物物种的起源、演变的根本原因、机制和历史过程。

物种 (species): 具有一定形态和生理特征、分布在一定区域内的生物类群, 是生物分类的基本单元, 也是生物繁殖和演化的基本单元。界定物种的主要标准是: 是否存在生殖隔离。

隔离机制的分类

1. 合子前机制阻止杂种合子形成。

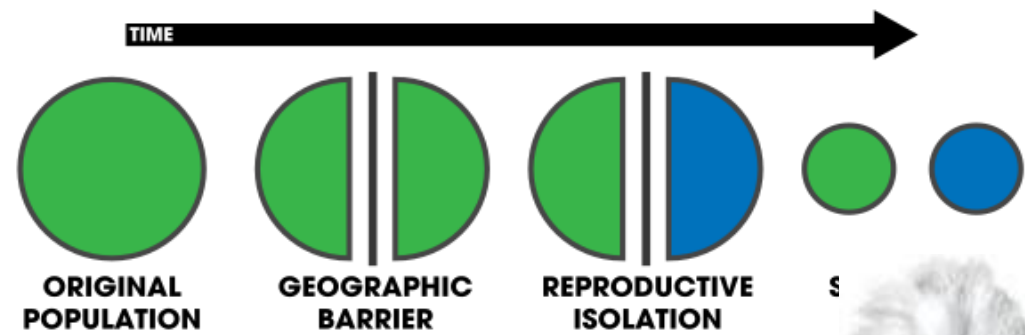
- (a) 生态隔离。在同一个大的地区, 群体占领不同的栖息地。
- (b) 时间隔离。交配或开花期发生在不同的季节。
- (c) 行为隔离。不同物种间两性间的互相吸引很弱。
- (d) 机械隔离。生殖器或花的部分物理不相容性阻止了交配或花粉传递。
- (e) 通过不同的传粉者隔离。近缘的开花植物种可能专门性地吸引不同的昆虫作为传粉者。
- (f) 配子隔离。配子可能不相互吸引, 或配子可能在另一性的生殖道中或花的柱头上无生活力。

2. 合子后机制减小杂交合子的生存能力或生育能力。

- (g) 杂种无生活力。杂种合子是无生活力的或是生活力降低的。
- (h) 杂种不育。其中一种性别或两种性别的杂种不能产生有功能的配子。
- (i) 杂种衰败。子二代或回交杂种生殖力或生活力降低。

1. 隔离的类型

地理隔离 (geographically isolation)：指的是两个群体占据着不连续的分布区间，空间上的隔离阻止了两个群体间个体的交配，从而阻止了基因交流。



异域物种形成模式

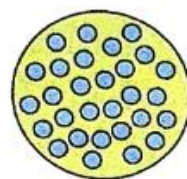


生殖隔离 (reproductive isolation) : 自然条件下, 群体之间不能通过两性生殖得到生活力和生育力正常的后代的现象。

1) 受精前生殖隔离: 不能形成受精卵的隔离现象。包括季节隔离、行为隔离、受精隔离。

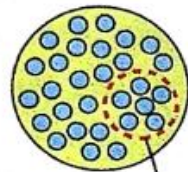


同域物种形成



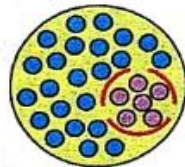
步骤

- 1 一个个体间相互交配的种群栖息在一个地域内



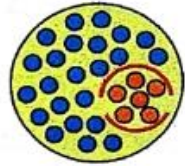
步骤

- 2 两种群间的部分生态隔离限制了基因流动



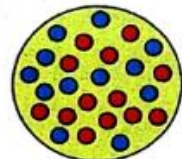
步骤

- 3 两种群基因流动更多的限制增加了遗传分歧



4

- 部分生殖隔离产生,引起遗传分歧的增加



步骤

- 5 生殖隔离完成,产生两个不同的物种

2) 受精后生殖隔离：杂合子的生活力或生殖力很弱。包括杂种不活，杂种不育。

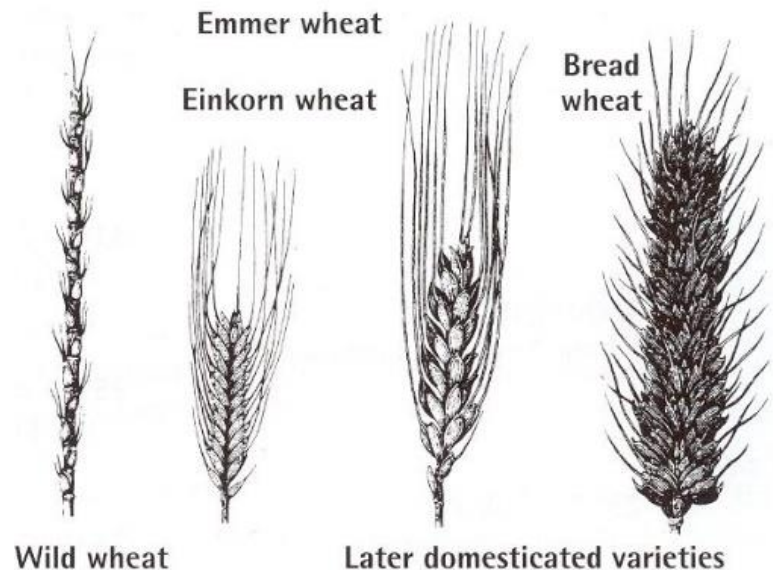
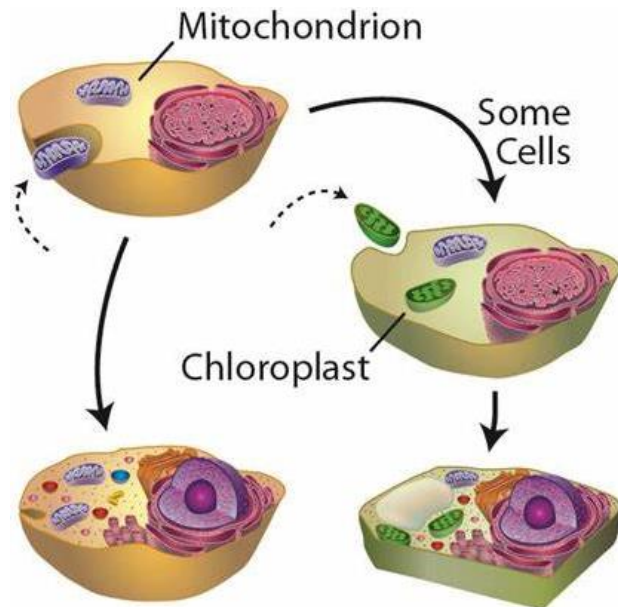


2. 物种形成的方式

渐变式：在一个长时间内旧的物种逐渐演变形成新的物种，是物种形成的主要形式。也是自然选择学说所描述的新物种形成方式。

“基因突变 + 环境选择”

爆发式：短期内以飞跃形式形成新的物种，往往没有复杂的中间亚种阶段。主要在高等植物普遍存在。

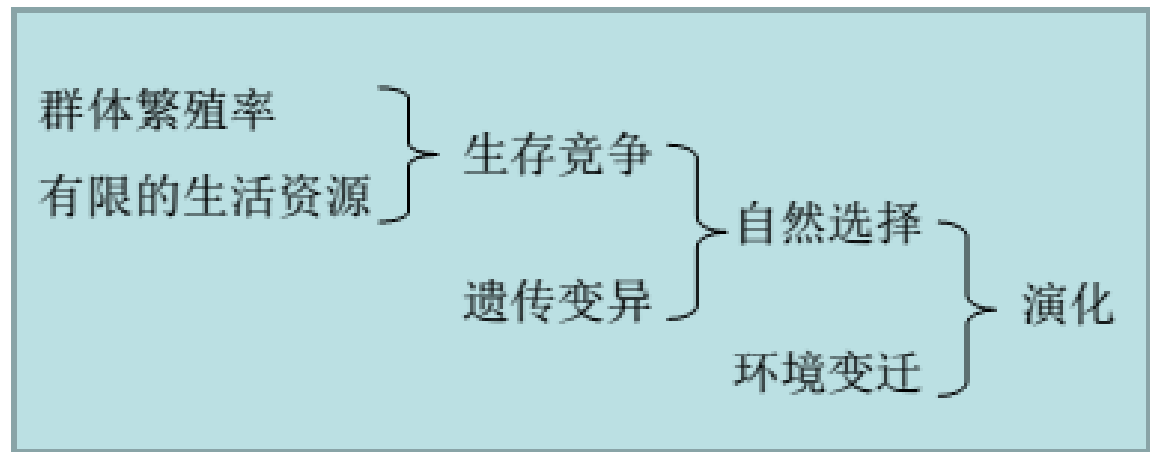




现代达尔文主义

又称为现代综合演化论，其基本思想是：

- 1) 共享一个基因库的群体是生物演化的基本单位，而生物演化机制的研究应属于群体遗传学的研究范畴；
- 2) 突变是演化的原材料，它通过自然选择被保留并累积那些适应性的变异，再通过隔离机制阻止群体间的基因交流，从而形成物种的演化。





最小可存活种群 (MVP)

- 最小存活种群 (minimum viable population, MVP) 在种群统计随机性、环境随机性、自然灾害和遗传随机性影响下, 在给定时间内有较高存活概率的最小隔离种群大小。
- 许多研究认为, 种群数量级在 10^3 以上才能满足一般物种较高概率的中期或长期存活目标。MVP给出了满足保护要求的种群数量下限, 对自然保护区设计、物种的受威胁等级划分和保护措施的制定等有重要意义。



二、群体的遗传平衡

1. 群体 (population) :

- (生态学)群体 —— 某一空间内生物个体的总和。包括**全部物种**的生物个体。
- (遗传学)群体、种群、孟德尔群体 —— **有相互交配关系、能自由进行基因交流的同种生物个体的总和。一个群体内全部个体共有的全部基因称为基因库(gene pool)。**
- **最大的孟德尔群体就是整个物种(不存在生殖隔离)。**



- **群体遗传学**：研究群体遗传结构变化的规律、原因以及在生物进化与新物种形成中作用的学科
- 以孟德尔群体为基本研究单位——基因库问题；
- 以**基因频率**和**基因型频率**描述群体遗传结构；
- 采用数学和统计方法进行研究；

2. 基因型频率: 一个群体内某种特定基因型所占的比例。

在一个个体数为N的二倍体生物群体中，一对等位基因(A, a)的三种基因型的频率如下表所示：

基因型	个体数	基因型频率
AA	D'	$D=D'/N$
Aa	H'	$H=H'/N$
aa	R'	$R=R'/N$
	N	1

基因频率: 一个群体内某特定基因座(locus)上某种等位基因占该座位等位基因总数的比例，也称为等位基因频率。

在一个个体数为N的二倍体生物群体中，一对等位基因(A, a)的共有**2N**个基因座位，两种基因的频率如下表所示：

等位基因	基因座数	基因频率	
A	$2D'+H'$	$p=(2D'+H')/2N$	$D+\frac{1}{2}H$
a	$2R'+H'$	$q=(2R'+H')/2N$	$R+\frac{1}{2}H$
	2N	1	1

Hardy-Weinberg 定律 (1908)



- 在**随机交配**的大群体内，基因频率和基因型频率在没有**迁移、突变和选择**的理想条件下，世代相传保持不变。也称为遗传平衡定律。
- 平衡条件：群体很大；随机婚配；没有突变；没有选择；没有迁移。
- 该定律揭示：**尽管显性等位基因的作用可以遮盖隐性等位基因的作用，但是在世代传递过程中，群体的基因频率及基因型频率都维持不变，隐性等位基因不会消亡。**
- 平衡群体中，基因型频率和基因频率的关系为：

$$D=p^2 \quad H=2pq \quad R=q^2$$

$$p+q=1; \quad D+H+R=1$$

Generation I



Genotypes

AA Aa aa

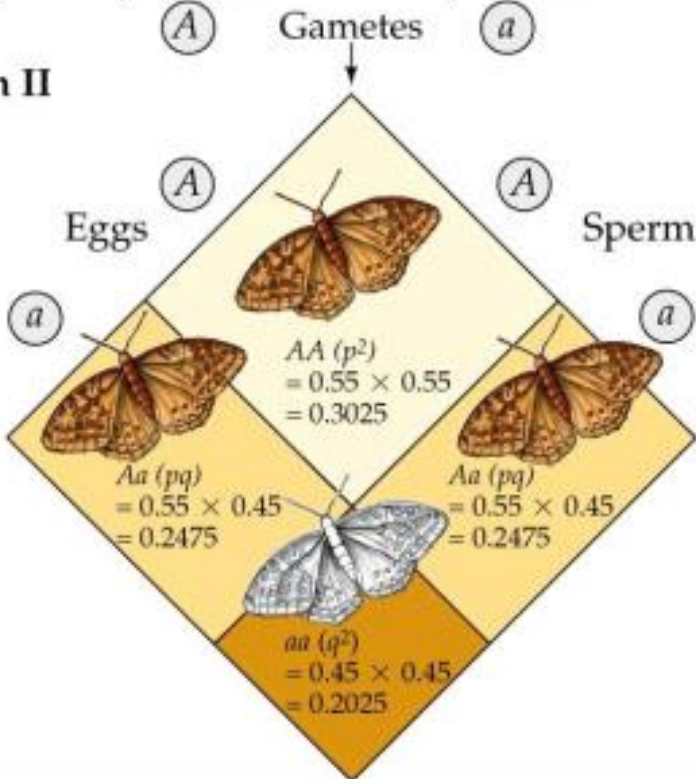
Frequency of
genotypes in
population

0.45 0.20 0.35

Frequency of
alleles in
population

$0.45 + 0.10$ $0.10 + 0.35$
 $p = 0.55$ $q = 0.45$

Generation II



$$p = D + \frac{1}{2} H$$

$$q = R + \frac{1}{2} H$$

任何一个不平衡的大群体，无论其基因频率如何，只要经过一个世代的随机交配，一对常染色体基因型频率就接近平衡，形成平衡群体；若没有其它因素的影响，一直进行随机交配，这种平衡状态始终不变。



3. 基因频率的计算

1) 共显性或不完全显性下基因频率的计算

- 例题：某一人群中MN血型的分布是M型250人，MN型500人，N型250人，请问等位基因 M 和 N 的基因频率。

由表型直接可以识别、计算基因频率：

$$p = (2D' + H')/2N$$

$$q = (2R' + H')/2N$$

$$M \text{ 的频率} : (250 \times 2 + 500) / (1000 \times 2) = 0.5$$

$$N \text{ 的频率} : (250 \times 2 + 500) / (1000 \times 2) = 0.5$$



2) 等位基因完全显性的隐性遗传病，突变基因频率的计算

例题：一个平衡群体中，PKU的发病率在1/10000左右，那么突变基因频率为？携带者频率为？

$$R = q^2$$

$$p = 1 - q$$

$$q = (1/10000)^{1/2} = 0.01$$

$$p = 1 - 0.01 = 0.99$$

$$H = 2pq \approx 2q = 0.02$$



3) 完全显性的显性遗传病，突变基因频率的计算

例题：在一个平衡群体中，软骨不全发病率是 $1/40000$ ，那么突变基因频率为？正常基因频率为？

在平衡群体中，群体发病率 = $AA + Aa = p^2 + 2pq$

实际条件中，由于突变基因频率 p 通常很低， AA 纯合个体少， p^2 可以忽略；

$$p^2 \approx 0, q \approx 1,$$

$$\text{发病率} = p^2 + 2pq \approx 2pq \approx 2p$$

$$\text{显性遗传病基因频率: } p = \frac{1}{2} \times \text{发病率}$$

$$\text{致病基因 } A \text{ 频率 } p = \frac{1}{2} \times 1/40000 = 0.0000125$$

$$\text{正常基因 } a \text{ 频率 } q = 1 - p = 0.9999875$$

复等位基因的遗传平衡——ABO血型



例如：如果随机婚配的平衡人群中表型频率：A型=0.45、B型=0.13、O型=0.36、AB型=0.06，计算复等位基因 I^A 、 I^B 、 i 的频率？

设： $I^A = p$ 、 $I^B = q$ 、 $I = r$ ，基因型及其频率如下：

表型	基因型	基因型频率	表型频率
A	$I^A I^A$	p^2	$p^2 + 2pr$
	$I^A i$	$2pr$	
B	$I^B I^B$	q^2	$q^2 + 2qr$
	$I^B i$	$2qr$	
AB	$I^A I^B$	$2pq$	$2pq$
O	ii	r^2	r^2

因为O型为 ii ，表型=基因型， i 基因的频率 $r = \sqrt{O} = \sqrt{0.36} = 0.6$ ；A型+O型（表型频率） $= p^2 + 2pr + r^2 = (p+r)^2 = 0.45 + 0.36 = 0.81$ ，所以 $p+r = \sqrt{0.81} = 0.9$ ，而 $r = 0.6$ ，故 $p = 0.9 - 0.6 = 0.3$ ；又因 $p+q+r=1$ ，所以 $q = 1 - p - r = 0.1$ 。



4. 伴性基因的遗传平衡

女 性		男 性	
基因型	基因型频率	基因型	基因型频率
$X^A X^A$	p_f^2	$X^A Y$	p_m
$X^A X^a$	$2p_f q_f$		
$X^a X^a$	q_f^2	$X^a Y$	q_m

伴性隐性遗传病，男患者:女患者 = $q_m : q_f^2$

伴性显性遗传病，男患者:女患者 = $p_m : (p_f^2 + 2p_f q_f)$

性染色体基因频率群体平衡的重要标志是：

$$p = p_m = p_f \quad q = q_f = q_m$$



例如，红绿色盲是伴X隐性遗传病，男性8%发病率。假设突变基因在男女中以同样的频率出现，则：

男性隐性基因型率（发病）= 基因频率= 8%

女性红绿色盲的基因型频率 $q^2 = 8\% \times 8\% = 0.0064 = 0.64\%$

男性隐性基因 a 的表型频率，要比女性高得多

例：调查知：某一群体内，男性色盲患者为7%，女性患者为0.25%，请问：



1) 该群体是否平衡？

设： $q = q_f = q_m = 0.07$ （假设群体平衡）

预计女性中色盲患者为 (0.07) 的平方 = 0.49%

实际上，女性中患者为0.25%，群体不平衡。

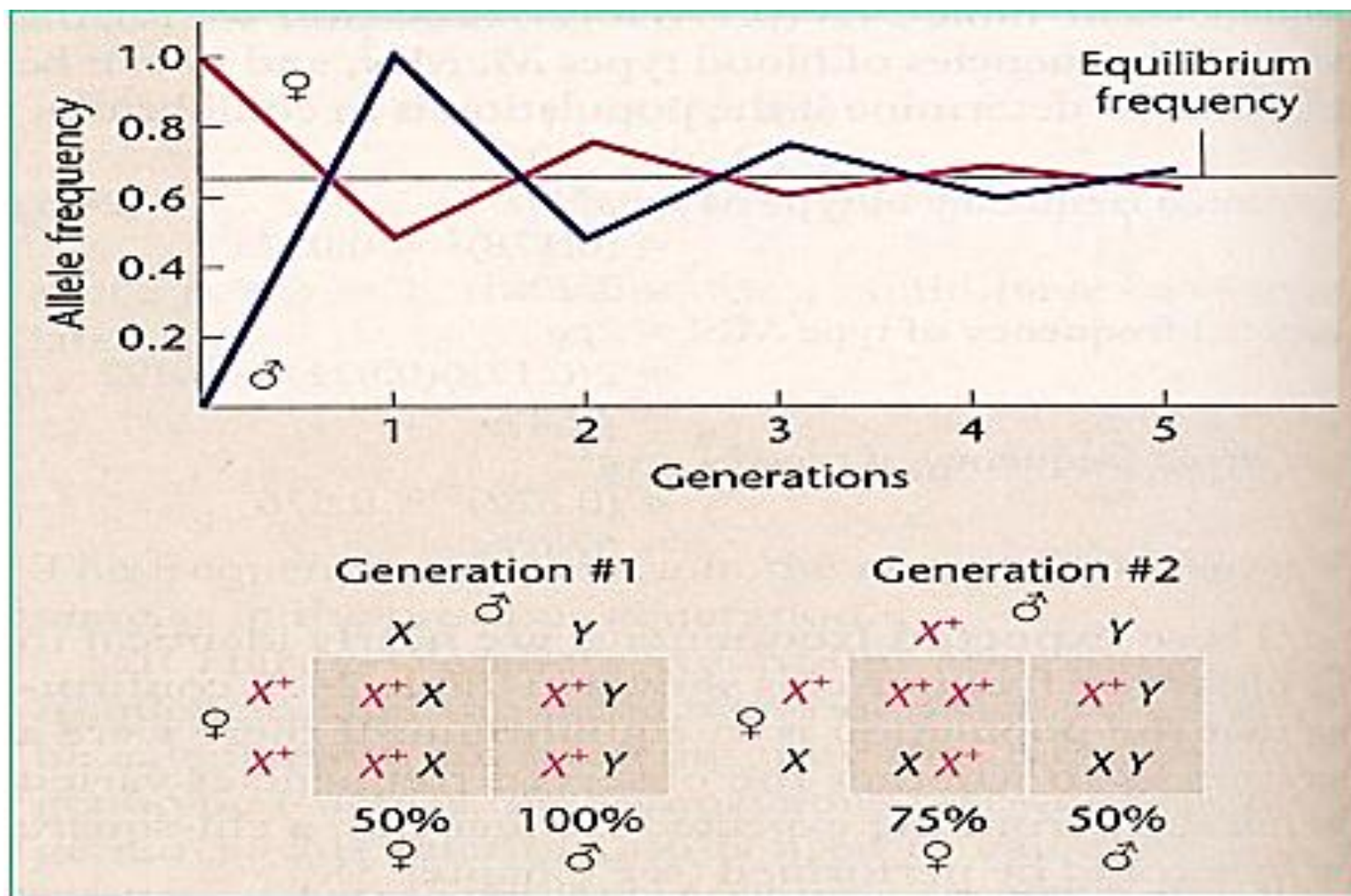
2) 该群体随机婚配一代后，色盲基因携带者的频率？

X隐性遗传病携带者只可能是女性，其基因型为 X^+X^- 。已知

$q_f = 0.05$, $q_m = 0.07$ ，所以，下代中携带者的基因型频率为：

$$\begin{aligned} H &= p_m \times q_f + p_f \times q_m = (1 - q_m) \times q_f + (1 - q_f) \times q_m \\ &= (1 - 0.07) \times 0.05 + (1 - 0.05) \times 0.07 = 0.05 \end{aligned}$$

对于**性染色体基因**，如果原有群体中雌体和雄体中基因频率不同：随交一代不可能达到平衡，事实上永远不会达到平衡，只会无限接近平衡。以 p_f 为例：





三、突变、自然选择与遗传漂变

从群体遗传学角度，推动演化（即改变等位基因频率）的因素包括：

- 增加或减少等位基因类型：

 - 突变（遗传变异的材料来源）

 - 迁移（不同群体的基因交流）

- 改变等位基因频率：

 - 自然选择（与个体生存力、生殖力有关）

 - 遗传漂变（基因库的随机取样）



1. 突变情况下群体中基因频率的改变

设初始频率为: $a = q$; $A = 1-q$,

突变率: $A \rightarrow$ 正向突变 $\rightarrow a(u)$; $a \rightarrow$ 回复突变 $\rightarrow A(v)$

每代中有 $(1-q)u$ 的 $A \rightarrow a$, qv 的 $a \rightarrow A$

当 $(1-q)u > qv$, a 的频率增加;

$(1-q)u < qv$, A 的频率增加;

处于平衡时:

$$(1-q)u = qv$$

$$q = \frac{u}{u+v}$$

由于多数基因突变频率很低($10^{-4} \sim 10^{-7}$), 因此突变对基因频率的改变要经过很多世代。时间的长短则与世代周期长短密切相关。

给定一对等位基因的正反突变频率, 就可以计算平衡状态的基因频率。

例: $u = 1 \times 10^{-6}$, $v = 5 \times 10^{-7}$, 则: $q = 67\%$, $p = 33\%$;

$u = v = 1 \times 10^{-6}$, 则: $q = 50\% = p$

2. 选择与群体基因频率的改变

群体内遗传上不同的个体或基因型有差别的个体，其生活力和繁殖力不同，导致基因频率发生改变。

自然选择(nature selection): 自然界中逐渐淘汰适合度低的个体，选择生活力和繁殖力高的个体作为下一代亲本的过程。

选择有正、负、中性选择之分。

从十九世纪初开始至今，英国曼彻斯特城市的黑色椒花蛾的基因型频率从1%上升到99%。



同一种基因型在不同条件下的适合度不同

- **适合度 (fitness, w)**: 群体中一个个体相对于其他个体存活并传递其基因到下一代的能力。适合度具有两个基本成分: 存活力 (viability) 和生殖成功 (reproductive success); 一般用相对生育力 (relative fertility) 来衡量 w 。
- **选择系数 (selection coefficient, s)**: 在选择的作用下降低的适合度, $s=1-w$ 。
- **遗传负荷 (genetic load)**: 如果一个群体的突变不断积累, 并且这些突变是有害的, 就会出现适合度下降, 这种现象称之为遗传负荷。用被选择淘汰的基因型频率占选择保存的基因型频率之比度量遗传负荷, 即 $(1-w)/w$ 。

例如:

- 丹麦一地区的统计发现: 该地区软骨发育不全遗传病侏儒 108 人生 27 个子女, 正常对照 457 人生 582 个子女;
- 侏儒的相对生育率是: $w = 27/108 \div 582/457 = 0.20$
- 因此: 软骨发育不全性侏儒的选择系数为

$$s = 1-w = 1-0.20 = 0.80$$



选择对隐性纯合体不利时，基因频率的改变

表 17 - 10 显性完全,选择对隐性纯合体不利时,基因频率的改变

	AA	Aa	aa	合计	基因 a 频率
初始频率	p^2	$2pq$	q^2	1	q
适合度	1	1	$1 - s$		
选择后频率	p^2	$2pq$	$q^2(1 - s)$	$p^2 + 2pq + q^2 - sq^2 = 1 - sq^2$	
相对频率	$\frac{p^2}{1 - sq^2}$	$\frac{2pq}{1 - sq^2}$	$\frac{q^2(1 - s)}{1 - sq^2}$	1	$\frac{pq + q^2(1 - s)}{1 - sq^2} = \frac{q(p + q - sq)}{1 - sq^2}$ $= \frac{q(1 - sq)}{1 - sq^2}$
基因 a 频率的改变 $\Delta q = \frac{q(1 - sq)}{1 - sq^2} - q = \frac{q - sq^2 - q + sq^3}{1 - sq^2} = \frac{-sq^2 + sq^3}{1 - sq^2} = \frac{-sq^2(1 - q)}{1 - sq^2}$					



$$\Delta q = q_1 - q_0 = \frac{-sq_0^2(1-q_0)}{1-sq_0^2}$$

当 s 很小时, 分母 $1-sq_0^2 \rightarrow 1$, 则

$$\Delta q = -sq_0^2(1-q_0)$$

当 q_0 值很小时, 每代基因频率的改变量是很小的;

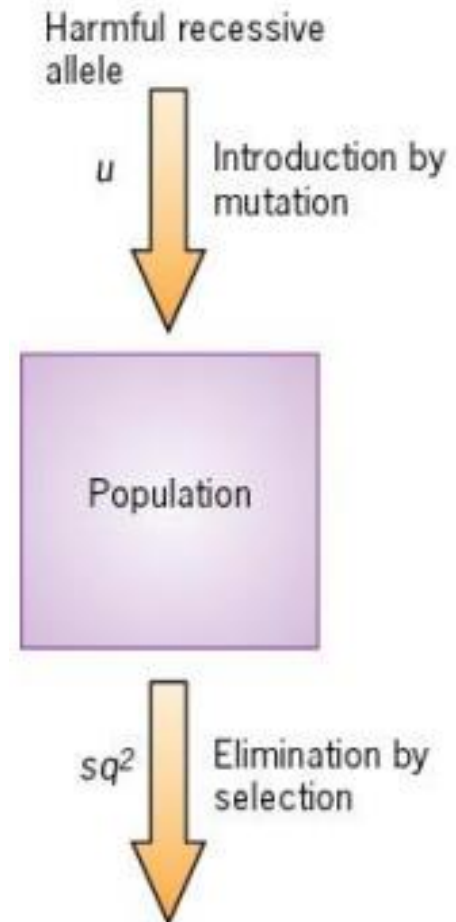
当 q_0 值较大时, q 的改变量也较大;



突变与选择的共同作用下群体基因频率的改变

- 当选择对纯合隐性个体不利时，a 频率的改变：
- $q_1 - q_0 = -sq^2(1-q) / 1-sq^2$ ，
即a基因的频率 q 每代减少 $sq^2(1-q)$
- 新产生的隐性突变基因 (A→a) 的频率 $pu = u(1-q)$
- 平衡时： $sq^2(1-q) = u(1-q)$ ； $q^2 = u/s$

平衡时基因频率由选择系数和突变率决定，与原来的基因频率无关。

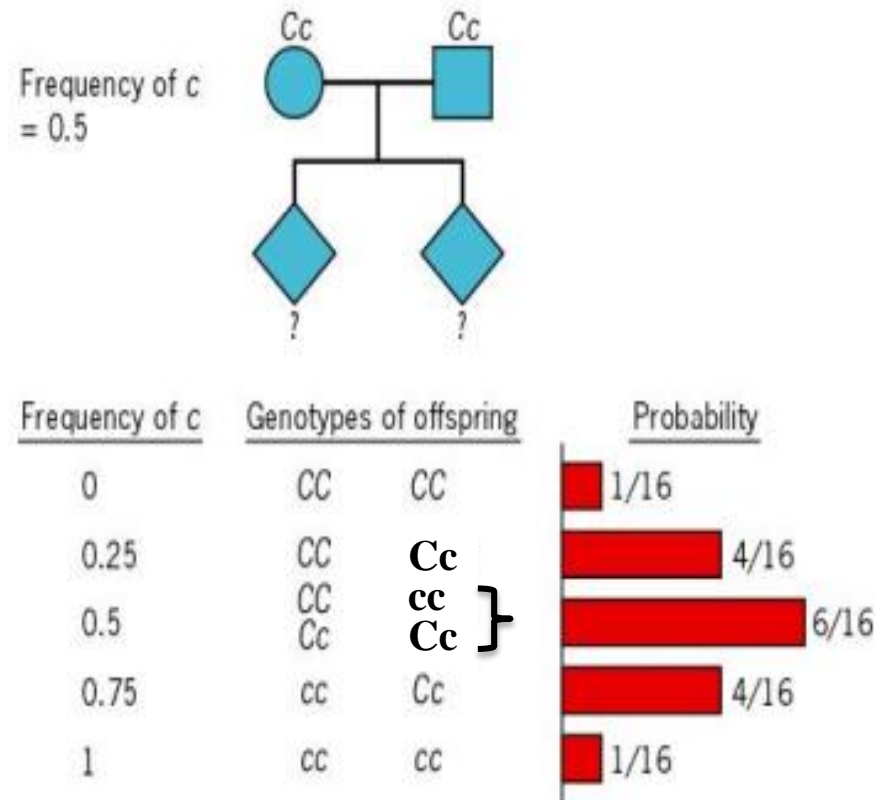




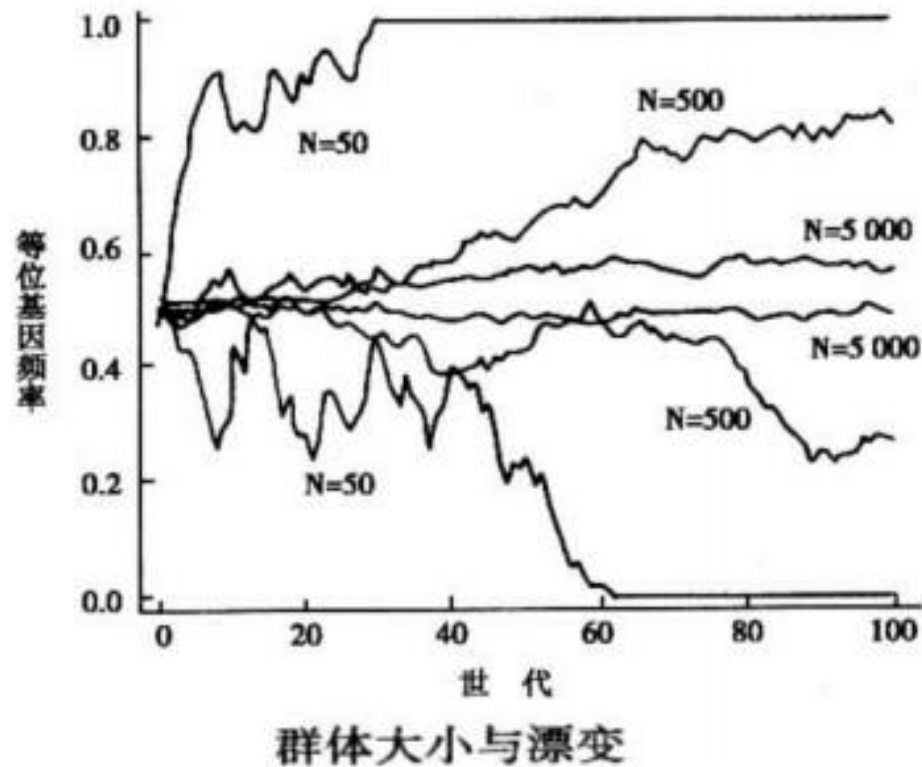
3. 遗传漂变与群体基因频率

■ **遗传漂变 (genetic drift):** 群体内由于抽样误差造成的等位基因频率的随机波动。包括两个方面：1) 配子形成过程中等位基因发生分离，导致后代所携带的等位基因类型具有随机性；2) 各种亲本组合所产生的子代数不同，不同等位基因类型的后代个体数目亦有不不确定性。

- ① 假设一对夫妻在某一基因座上的基因型均为Cc。那由他们建立的这个群体中，C和c基因频率均为0.5。假设他们生育了两个孩子，那么由这两个孩子建立的新群体的基因频率有10/16的概率会发生变化。
- ② 这个变化不是由突变或自然选择造成的，而是单纯地源于孟德尔遗传机制。
- ③ 随机漂变是发生在生物繁殖时基因取样过程中的必然结果。

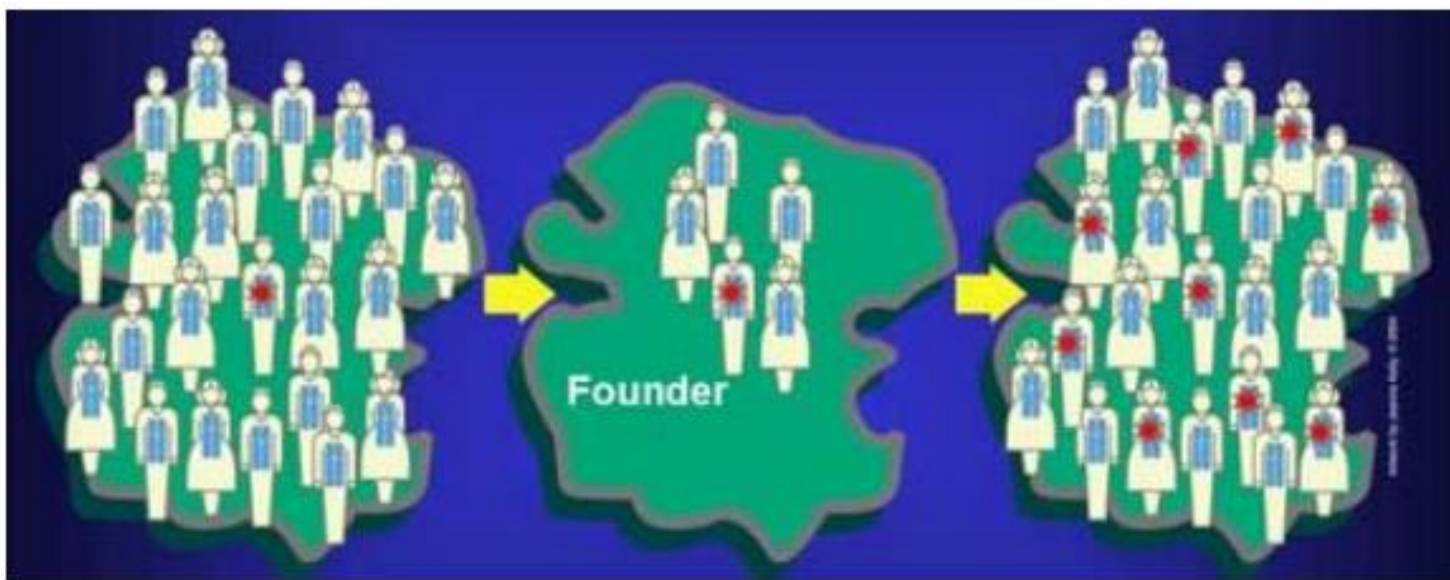


- 遗传漂变的结果是**等位基因频率的随机变化**，群体中低频率的基因可以因为遗传漂变直接变成高频率基因，反之亦如此。
- 遗传漂变对基因的弃留**不同于选择**，它对有利/不利基因或即无利又无害的“中性基因”都是一视同仁的，**只取决于群体的大小和生殖时的机遇。**



奠基者效应 (founder effect)：指由带有亲代群体中部分等位基因的少数个体重新建立新的群体，是遗传漂变的一种形式。

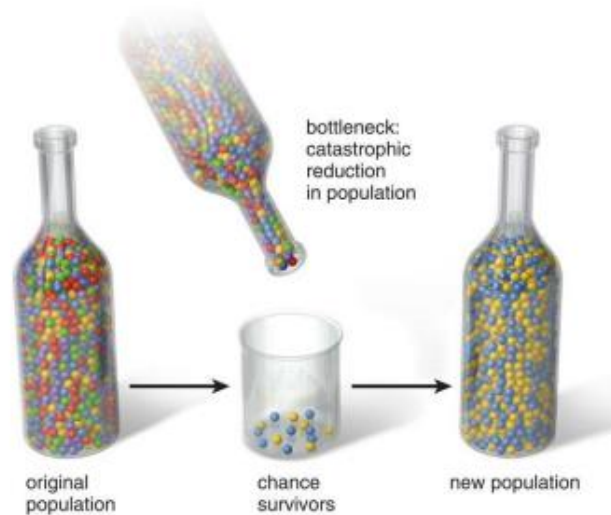
尽管该群体数量会逐渐增加，但因隔离关系无法与其他生物群体交配繁殖，导致群体内部等位基因频率显著不同于外部群体。



- **案例**：在委内瑞拉的偏僻山村中具有非常高的亨廷顿舞蹈病的突变基因频率，这是一位始祖妇女携带了该基因，随着世代繁衍，突变基因在村庄中扩散开来。

种群瓶颈 (population bottlenecks)：指的是由于外界环境的剧烈变化使群体的个体数急剧减少，此时群体的等位基因频率亦发生急剧改变。

遭受瓶颈的种群将**经历遗传漂变**，重新建立的群体的等位基因频率显著不同于初始群体。

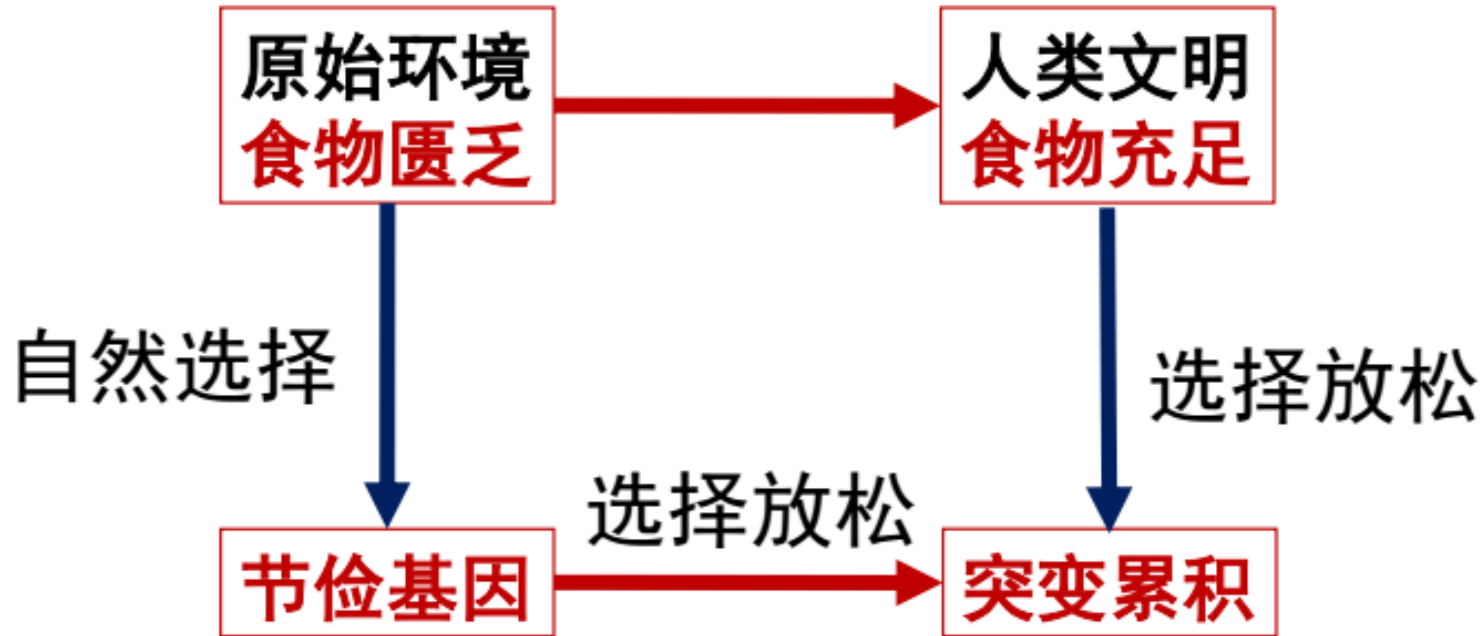


- **案例**：一种AR遗传的先天性失明，在东卡罗林群岛的Pingelap人中，有相当高的发病率。1970年调查该岛的1 500余人中，盲人的比例竟高达4%~10%。据记载在1780~1790年间，一次飓风袭击了该岛，造成大量死亡，岛上只留下9个男人和数目不多的女人。最初留下的这些人中“先天性失明”基因频率较高，虽然该基因在选择上处于劣势，但在以后的数个世代内，遗传漂变导致了这种高频率。



选择放松假说

Selection Relaxation Hypothesis



节约基因就是能让肌体代谢机制处于节约状态的基因，这是多年以来人们适应恶劣环境的产物。