

# 连锁不平衡 (linkage disequilibrium, LD)

指群体内不同位点等位基因间的非随机关联，重点就三个字：**非随机**。如果不好理解，先从反面来看，以此相对的另一个概念就是**连锁平衡** (linkage equilibrium, LE)，也就是不同位点等位基因之间是随机遗传的，这个我们熟悉，不就是两个独立事情么，今天下雨和后天下雨之间没有联系，是独立的；抛两次硬币都正面向上的概率是 $0.5 \times 0.5 = 0.25$ 。所以我一直不喜欢这个连锁不平衡的概念，非得这么绕把人搞懵，简单来说，连锁不平衡就是**连锁，不独立**。【知乎】

**QTL**全称是Quantitative Trait Loci。顾名思义就是基因组上的一些位点，对一些特定性状具有某种量化的影响。这里的重点是，我们关心的性状是一个有多个不同水平的**可定量** (Quantitative) 的性状，而不是定性 (Qualitative) 的性状。**QTL定位的基本原理**，就是测定一群个体的某个数量性状 (表型)，以及它们的基因型 (就是基因组上的一些遗传标记，例如SNP/RFLP等等，但不一定是全基因组)，然后寻找基因型和表型的对应关系。【知乎】

QTL 定位**不能准确地定位出复杂性状** (受多个基因调控以及环境影响) 的相关基因在基因组上的准确位置，但是定位精确度可达 2 cM (厘摩尔) 以下甚至更精确的距离。

GWAS依赖连锁不平衡性

提高研究群体样本量以及结合多组学的方法，是打破固有群体高连锁不平衡性。

提高样本的大小有利于提高GWAS的统计力

## 后GWAS多组学

## GWAS和QTL区别

GWAS用的是关联分析，QTL用的是连锁分析。GWAS是全基因组关联分析，用的群体一般是自然群体，标记用的一般是高密度的SNP标记。定位到的基因一般都能具体到**某个或者某几个SNP**上。QTL用的是连锁分析，群体一般是遗传背景相近的遗传连锁群，标记密度一般没必要特别高，但也可以用snp标记，只是说有点大材小用了。定位到的一般是某一个**区域**，精确度不如GWAS。当然上诉都是数学计算上的理论结果，两个方法谁精确还是要具体问题具体分析。

作者：大烧饼

链接：<https://www.zhihu.com/question/382223018/answer/1220150347>

来源：知乎

【chatAI】GWAS（全基因组关联研究）和QTL（数量性状位点）是两种常见的遗传研究方法，它们用于理解基因与**性状或疾病之间的关系**，但在研究设计和应用方面有所不同。

1. 定义和目标：GWAS旨在鉴定基因组范围内与性状或疾病相关的遗传变异，而QTL研究主要关注鉴定已知家系的**数量性状位点**。
2. 研究对象：GWAS是一种无偏的研究设计，可以对**大规模人群**进行分析，探索基因突变和复杂性状之间的关联；而QTL研究通常使用**特定模型有限的家系群体**，并关注数量性状的遗传变异。
3. 方法和数据：GWAS使用基因芯片或基因测序等高通量技术，分析数十万个基因位点和成千上万的个体，通过统计学方法来寻找遗传变异与性状之间的关联。QTL研究通常使用连锁分析或关联分析等方法来确定特定数量性状位点与家系中性状变异之间的关联。
4. 验证和功能解释：GWAS通常基于**关联性**，需要进一步的验证和功能解释来确定是否存在因果关系。QTL研究则更关注已知家系中数量性状位点的遗传效应和作用机制。

综上所述，GWAS和QTL是两种不同的遗传研究方法，分别用于鉴定基因组范围内的遗传变异与性状或疾病的关联，以及特定家系中的数量性状位点。它们在研究设计、数据分析和结果解释上有所区别。

## 性状变异

性状变异是指同一物种中个体之间在**某一性状方面表现出差异**的现象。这种差异可以是由于**遗传因素**所引起的，也可以是由于**环境因素**所引起的。性状变异是自然选择的基础，因为它为适应环境变化提供了不同的可能性。在进化过程中，性状的变异可以通过**基因突变或基因重组**等方式产生，并在后代中传递下去。因此，**性状变异是生物适应环境变化和进化的一种重要机制**。

当谈到性状变异时，我们可以将其分为两大类：**离散性状变异和连续性状变异**。

离散性状变异指的是物种中某一性状的表现可以分为有限的几个离散类别，例如眼睛颜色、花瓣数目、血型等。在这种情况下，每个个体只能表现出这几个预定的离散状态，没有中间过渡状态。

而连续性状变异则是指某一性状的表现存在连续的变化范围，没有明确的离散状态。例如身高、体重、种子大小等都是连续性状变异的例子。在这种情况下，个体可以在性状的连续范围内表现出各种不同的状态，没有明确的分界点。

不论是离散性状变异还是连续性状变异，它们都为自然选择提供了基础。自然选择是指环境选择某些性状的个体具有更好的生存和繁殖优势，从而使这些性状在种群中逐渐增多。当环境发生改变时，性状变异可以为物种提供多样的适应策略，增加其生存的机会。

总之，性状变异是自然界中普遍存在的现象，通过遗传和环境的相互作用，它为物种的适应和演化提供了重要的基础。

## 罕见的多态性

当我们说林木存在罕见的多态性时，意味着在同一个物种的林木群体中存在着**非常显著的多样性或变异性**。这意味着不同个体之间在某些重要特征上表现出明显的差异。

多态性可以表现在各种不同的方面，如树高、树干直径、形状、叶子颜色、叶子形态、花朵形状等。这些多态性特征可能是由于**遗传因素**、**环境因素**或两者的相互作用所导致的。

罕见的多态性使得同一物种的林木群体在外观和表现上具有很大的变化，可能会是自然景观中的一个独特和吸引人的特征。多态性也可以为物种在不同环境中找到更多的适应性。

然而，我们需要注意，多态性并不适用于所有物种。有些物种具有较少的变异性，而某些物种则具有更显著的多态性。多态性的存在取决于物种的遗传背景、环境条件和其他因素的相互作用。

在生态学和植物学中，研究和理解罕见的多态性对于保护生物多样性、了解物种适应性和进化过程都非常重要。

榿树（科学名称：Cunninghamia lanceolata）是一种古老的常绿针叶树，主要分布在中国南部和东南亚地区。榿树在形态上具有一定的变异性，但并不被认为是典型的多态性树种。

在榿树的种群中，可能存在一定程度的形态上的变异，例如树高、叶子大小和形状等方面的差异。这些变异在一定程度上受到遗传和环境因素的影响。然而，相比于一些其他树种，**榿树的变异程度**并不算很高。

尽管如此，在其分布范围内，某些地理区域或特定生境下的榿树种群可能会呈现出更为显著的形态上的变异。这种变异可能在一定程度上为该物种的适应性和生存策略提供了多样性。

总的来说，榿树可以在一定程度上呈现出形态上的变异，但并不典型于多态性树种。对于植物的变异和多态性，实际情况往往受到多种因素的影响，包括遗传背景、环境条件、地理分布等。

## 低频单核苷酸多态性

指的是基因组中的单个核苷酸位置上存在多种不同等位基因（alleles），并且这些不同等位基因的**频率很低**，即它们在种群中的频率较低。

在基因组中，DNA的编码序列是由四种不同的核苷酸（腺嘌呤A、胞嘧啶C、鸟嘌呤G、胸腺嘧啶T）组成的，因此在同一个位置上可能会存在不同的核苷酸取代。当这种核苷酸取代在种群中的频率很低时，就被称为低频单核苷酸多态性。

低频单核苷酸多态性在生物学和遗传学中具有一定的重要性。尽管它们在整个种群中的频率很低，但它们可能对个体的表型和遗传特征产生影响，并且可能在一些特定条件下发挥**重要的生物学功能**。因此，研究低频单核苷酸多态性对于理解种群遗传结构、生物适应性和进化过程具有一定的意义。

此外，在人类疾病研究和遗传疾病诊断方面，低频单核苷酸多态性也可能扮演重要角色，因为它们可能与一些疾病的易感性以及药物代谢途径等方面相关联。

**频率很低的**现象可能暗示着某种情况，具体取决于所涉及的领域和具体情境。以下是一些可能的情况：

1. **遗传多样性**：在遗传学和生物学领域，低频率的遗传变异可能暗示着特定等位基因在整个种群中的出现频率较低。这可能意味着**该等位基因相对罕见**，或者是最近才在种群中出现**的新变异**。对于物种的遗传多样性和群体遗传结构来说，低频率的等位基因可能具有一定的重要性。
2. **稀有事件或异常情况**：在某些情况下，低频率的事件可能暗示着可能性较低的或相对罕见的事件发生。这可能适用于各种领域，如自然灾害、疾病爆发，或者是极不寻常的现象。

3. 小概率事件：在数学和统计学中，低频率可能意味着事件发生的概率较低。这可能意味着在给定的情况下，该事件发生的可能性相对较小，或者是随机事件的偶然发生。
4. 少见的现象或特征：在其他领域，如社会科学或经济学，低频率可能意味着某些现象或特征相对较为罕见，或者是在整体群体中不太典型的情况。

总的来说，频率很低可能提示着一种罕见或不太寻常的情况，而具体含义则取决于所讨论的具体情境和领域。

## 多效应基因

指的是**对多个性状或表型特征产生影响的基因**。这些基因编码的蛋白质或调控元件可能在细胞或生物体中参与多个生物学过程，从而对多个性状产生影响。

多效性基因是遗传研究中的一个重要概念，它们可以**解释为什么在某些情况下，一个基因突变可能影响多个相关的性状或表型特征**。这种情况可能是因为这些基因编码的蛋白质或调控元件在细胞中担任多种功能或参与多个途径。因此，单个基因的突变可能导致多种表型变化。

多效性基因对于理解遗传疾病的发病机制以及表型多样性的形成具有重要的影响。它们可以解释为什么一些遗传疾病表现出不同的症状和临床表现，即使它们是由同一个基因突变引起的。此外，多效性基因还可以解释为什么一些基因相关研究在不同个体或种群中产生不一致的结果，因为**这些基因可能对多个性状产生不同程度的影响**。

虽然多效性基因在遗传学中起着重要的作用，但研究和理解它们的具体机制仍然是一个复杂的课题，需要进一步的实验和研究来揭示。

## 遗传效应

遗传效应指的是**遗传因素**对特定性状或表型特征的影响程度。遗传效应可以通过各种方法来确定，并被广泛研究和讨论。

遗传效应通常分为以下几种类型：

1. 基因型效应：基因型效应是指基因型对特定性状或表型特征的直接影响。不同基因型的个体可能展现出不同的表型差异。例如，单核苷酸多态性（SNP）位点上的不同等位基因可能会直接影响某一性状的表现。
2. 加性效应：加性效应是指多个基因对一个性状的叠加效应。这意味着每个对该性状有贡献的基因的效应是可加的，它们的组合决定了最终的表型。加性效应是广泛应用于遗传研究和遗传评估的常见假设。
3. 互作效应：互作效应是指不同基因之间相互作用对特定性状的影响。这些基因之间的相互作用可能会增强或减弱彼此的效应，导致复杂的表型变化。互作效应的研究对于理解遗传背后的更深层次的机制至关重要。
4. 环境交互效应：环境交互效应是指基因型与环境之间的相互作用对特定性状的影响。环境条件可以改变基因表达和基因效应，从而显著影响表型特征。环境交互效应的研究通常涉及对不同环境条件下基因型-表型关系的比较和分析。

遗传效应的研究和确定往往需要大规模的遗传数据、统计分析方法以及模型假设的建立。这些研究对于理解遗传疾病、性状遗传的可变性以及遗传改良等方面具有重要意义。它们也为我们提供了更深入地了解遗传与环境之间的复杂相互作用的机会。

确定位点的遗传效应是遗传学和基因组学研究中的关键问题之一。下面是一些常见的方法和策略来确定位点的遗传效应：

1. **关联分析**：关联分析是通过比较**基因型和表型数据**之间的相关性来确定位点的遗传效应。常见的关联分析方法包括单核苷酸多态性 (SNP) 关联分析和基因组关联分析。这些方法可以帮助确定某个位点的基因型与特定表型特征之间的相关性。
2. **基因型和表型相关性**：通过研究不同基因型和特定表型特征之间的关系，可以确定位点的遗传效应。这可以通过对基因型和表型数据进行统计分析来实现，例如卡方检验、Fisher 精确检验、线性回归等。
3. **基因表达和功能研究**：通过研究位点的基因表达和功能变化，可以推断位点的遗传效应。这可以通过转录组学研究、蛋白质组学研究、功能实验等方法来实现。这些研究可以揭示位点突变对基因表达或蛋白功能的影响，从而推测其对表型特征的遗传效应。
4. **家系和亲缘关系研究**：家系和亲缘关系研究可以通过分析在家族中共同出现的基因型和表型特征来确定位点的遗传效应。家系研究可以通过纵向或横向的家系数据进行分析，而亲缘关系研究可以利用家族间的遗传相似性来确定位点的遗传效应。

以上是一些常见的方法和策略来确定位点的遗传效应。在实际研究中，通常会使用多种方法的组合，并结合大规模的样本数据和统计分析来确保结果的可靠性和有效性。

## 遗传因素

是指在遗传学中对个体特征或性状产生影响的因素。这些因素包括基因、染色体、基因组以及遗传变异等。

1. **基因**：基因是遗传信息的基本单位，它们编码着生物体内部的蛋白质或RNA。基因携带着细胞中各种生物过程的指令，对于个体特征的表达和传递起着至关重要的作用。
2. **染色体**：染色体是细胞核中的结构，它们携带了基因，并在细胞分裂时传递给子代。人类细胞中有46条染色体，其中包含了几乎所有的基因。
3. **基因组**：基因组是特定生物体细胞中所有基因的集合。基因组包括物种特异性的基因序列，它们对于决定物种的特征和适应性具有重要作用。
4. **遗传变异**：遗传变异指的是基因和基因组中存在的个体间遗传差异。这些差异可以是单核苷酸多态性 (SNP)、插入/缺失、拷贝数变异等。

遗传因素通过决定基因型来影响个体的表型特征。不同的基因型可能导致个体之间在某些特征上的差异。遗传因素与环境因素之间的相互作用也会对个体表型产生影响。

理解遗传因素对个体特征的影响对于遗传学、基因组学、遗传疾病的研究以及农业改良等领域具有重要意义。通过对遗传因素的研究，我们可以更好地理解基因对于个体差异和性状变异的贡献，促进健康和疾病的管理和预防，以及提高农作物品质和产量。