1. Gene: a segment of DNA on a chromosome that codes for a specific protein and thus determines a trait

Modern definition: a locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions, and or other functional sequence regions

1. Genome: the entirety of an organism’s hereditary information.(includes both the genes and the non-coding sequences of the DNA/RNA)
2. Genomice: the molecular characterization of whole genomes

Gene doesn’t code for a protein——RNA gene(rRNA,tRNA),non-coding RNAs(small RNAs, microRNA)

Gene families: related genes may be organized in several clusters at different locations

Point mutation: change of a single nucleotide. Includes the substitution, deletion, insertion of one nucleotide in a gene

Gene polymorphisms/single nucleotide polymorphism(SNP): a single-letter change in DNA, part of the natural genetic variation within a population that creates diversity

Ribonucleic acid(RNA)

1. Single-stranded
2. Ribose sugar, rather than deoxyribose
3. Uracil(U) instead of thymine(T)
4. RNA can move out of the nucleus to the cytoplasm
5. In enkaryotes, the primary transcript(RNA) is edited before it moves to the cytoplasm

Translation

Carried out by the ribosome

Makes a chain if amino-acids from mRNA

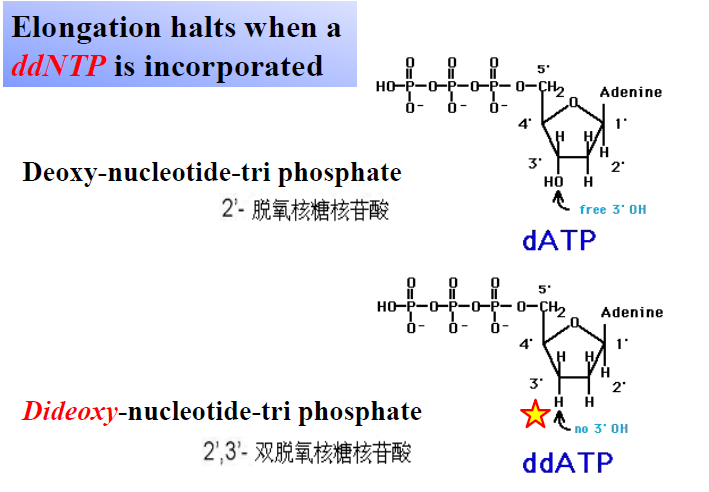
3 bases→1 amino-acid

Start with methionine(AUG), ends with stop codon(UAA,UAG,UGA)

Protein is often modified after translation, initial methionine loss

Important supporting techniques for large-scale sequencing

1. Sanger dideoxynucleotide

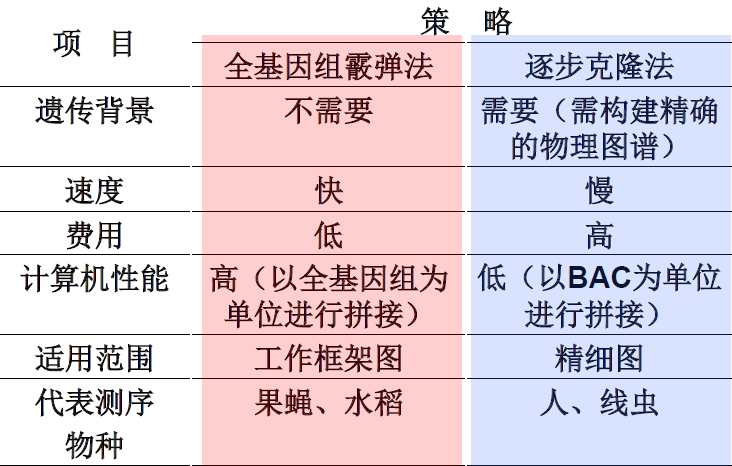


1. Automatic DNA sequencing

四种荧光染料的测序 pcr产物可在一根毛细管内电泳，从而避免了泳道间迁移率差异的影响，大大提高了测序的精确度。由于分子大小不同，在毛细管电泳中的迁移率也不同，当其通过毛细管读数窗口段时，激光检测器窗口中的ccd(charge-coupled device)摄影机检测器就可对荧光分子逐个进行检测，激发的荧光经光栅分光，以区分代表不同碱基信息的不同颜色的荧光，并在ccd摄影机上同步成像，分析软件可自动将不同荧光转变为dna序列

Two strategies large-scale sequencing

1. Clone by clone: 使用BAC（细菌染色体克隆载体），插入待测序的多个基因组片段，然后通过这个载体对基因组片段进行克隆。对每个克隆出来的片段进行霰弹法切割成很多个小片段，对小片段进行克隆之后进行双脱氧链终止反应进行测序，之后进行组装。相当于每个亚单位基因组测序组装后成一个单位，每个单位测序组装成整个基因组
2. Whole genome shotgun:他们首先在整个水稻基因组上生成许多已知长度的ＤＮＡ（脱氧核糖核酸）切片，然后使它们按ＤＮＡ序列的重合区域进行排列。这些切片数量足以覆盖水稻基因组４次。科学家们接着确定每个切片的碱基对序列，并用计算机程序将其组装成更长的片段，然后将这些片段排序、装配成１０万多个被称为支架的更大组件。



1. PCR(polymerase chain reaction)
2. Bioinformatics

Genomics:

structural genomics: characterizes the physical nature of whole genomes. Includes genetic mapping, physical mapping and sequencing of entire genomes

functional genomics: make use of the vast wealth of data produced by genomic projects to describe gene and protein functions and interactions

comparative genomics: a field where genomic features of different organism (diverse/related) are compared. The genomic features may include DNA sequence, genes ,gene order, regulatory sequences, and other genomic structural landmarks

Genome size: the length of DNA associated with one haploid complement of chromosomes

Gene number: the number of genes included in a genome

Gene density: the average number of genes per Mb of genomic DNA

Important concepts

Gene/Genome

Genomics

DNA sequencing principle

Shortgun sequencing

Structural genomics

Gene density

Comparative genomics

Functional genomics