

3.5 GENE SEQUENCING

Databases have been developed since the 1960's to help store complex information.

a) **GenBank** ® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences. GenBank is part of the International Nucleotide Sequence Database Collaboration, which comprises the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.

An application:

If a species splits to form two separate species, differences between base sequences of the genes of those two species will gradually accumulate. The number of differences can give an indication of how long ago species diverged from the common ancestral species.

Find the base sequence of human hemoglobin:

1. Go to: <http://www.ncbi.nlm.nih.gov>
2. Press "Gene" on the right side of the page
3. Add "HBB human" (haemoglobin protein) and press "search"
4. Click on the first option Name/Gene ID: HBB ID: 3043
5. Search and click on the word "FASTA"
6. Observe the number of chromosome and the base sequence of human hemoglobin.

b) Loci of human genes

The loci of genes can be found using the OMIM website (Online Mendelian Inheritance in Man).

1. Search for OMIM and open the home page
2. Choose OMIM Gene Map
3. Enter HBB

4. Check the locus: 11p15.4
11 is the number of chromosome
p indicates the short arm of the chromosome (q indicates the long)
15.4 the specific region in the short arm
5. Enter 11p15.4 to observe the locus on the chromosome

