

# Objects in DNA analysis

Nasser M. Abbasi

sometime in 1997

Compiled on May 23, 2020 at 5:15am

This is a description of objects used in DNA/Genomics software development. I need to first come up with such a list, and the relationship between them to base any further software development of a sw library.

Each object is described and listing of what other objects it contains and what attributes the objects has.

Each object will be represented as a class. The object attributes will help to define the access methods of the object.

## 1. Chromosome.

contains: contig objects. one DNA Molecule

Attributes:

- (a) Enochromatic size, in kb.
- (b) number of contigs it contains.
- (c) number of bases.
- (d) Name or ID of chromosom (1, 2, ... Y, X).

## 2. contig.

contains: physical position object, cytogenetic position object.

Attributes:

- (a) Name or ID. such as NT\_00039
- (b) size (in kb in bases)

## 3. physical position of contig in the chromosom.

Attributes (these are estimates from RH map positions):

- (a) start position. number
- (b) end position. number

## 4. Cytogenetic position of contig in the chromosom.

Attributes (based on annotation from GenBank entry)

- (a) start id. such as 1p34.3
- (b) end id. such as p36.13

## 5. DNA molecule (or DNA sequence)

contains number of Gene objects, contains one sequence of DNA base pairs objects.

- (a) number of genes.
- (b) number of base pairs.

## 6. Gene

contains number of exons and number of introns.

- (a) number of exons.
- (b) number of introns.

- |  |   |
|--|---|
| <p>7. exon. a region in a gene that is a coding region for protein. Contains number of codons, and a DNA sequence.</p> <ul style="list-style-type: none"><li>(a) size in bases.</li><li>(b) number of codons.</li><li>(c) start position in gene</li><li>(d) end position in gene.</li></ul> | <p>8. intron. a region in a gene that is a not a coding region for protein. Contains a DNA sequence</p> <ul style="list-style-type: none"><li>(a) size in bases.</li><li>(b) start position in gene</li><li>(c) end position in gene.</li></ul> |
|--|---|