

# ANNOTAZIONE DI GENOMI \*

**Definizione:** conoscere la localizzazione, la struttura e la funzionalità di ogni elemento che compone il genome

base: modello genico

formato: GFF3

## annotazione funzionale

caratterizzare ogni singolo gene assegnando una funzione biologica ad ogni proteina codificata

## annotazione genica

definire all'interno del genome: localizzazione e struttura dei geni e trascritti alternativi

### 1. allineamento di evidenze sperimentali

elaborate ed allineate al genome (cDNA, EST, proteine omologhe, ...)

## metodi

3. metodi basati sulla predizione genica ab initio, guidati da evidenze sperimentali mix

### 2. metodi ab initio

NO EVIDENZE SPERIMENTALI  
uso algoritmi e modelli:

**sensori segnale**

giunzioni esone-introne

**sensori contenuto**

regioni codificanti di lunghezza variabile

## Predittori

→ dati di esempio-training  
→ dati di valutazione - testing  
**augustus**

**Accuratezza** =  $(SN + SP) / 2$

Sensibilità → quanto sono in grado di predire

Specificità → quanto sono corretto

### 3 lvl d'indagine

- locus genico
- regioni esoniche
- giunzioni esone-introne

## Risultato

Creazione di un CONSENSUS: sequenza che rappresenta sequenze relative e allineate

## \* GENOME BROWSERS

Sono web server con tool di annotazione automatica

### → ENSEMBL

genome browser per vertebrati  
Annota i geni, calcola gli allineam.  
BLAST, VEP, e fa validazione

### → GENSCAN

esegue una predizione genica tramite HMM, e poi raffina per dati sperimentali

→ EST seq. veloci di tessuti interi