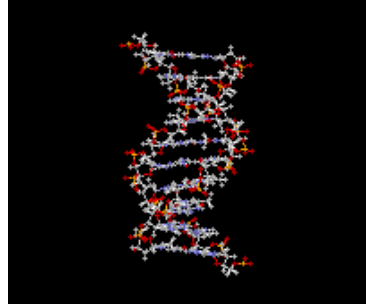




FACULTY OF ENGINEERING & TECHNOLOGY  
DEPARTMENT OF BIOTECHNOLOGY

Dr. Simranjit Singh  
Assistant Professor  
Department of Biotechnology  
Rama University, Kanpur

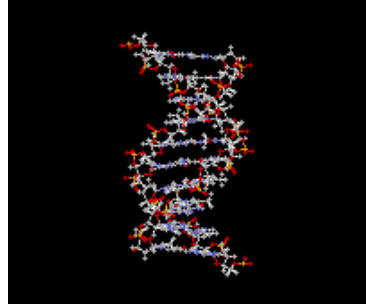
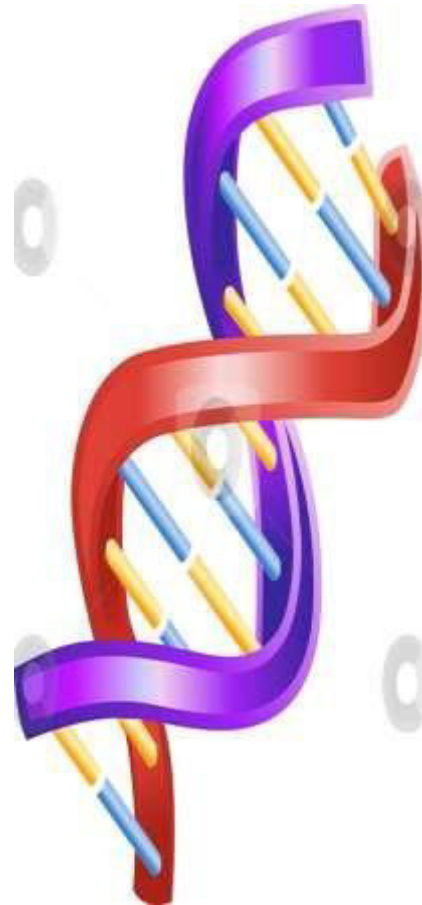


# **EXPRESS SEQUENCE TAG A**

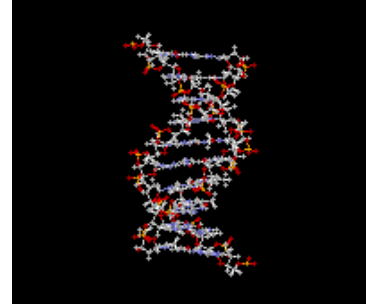
## **Tool In Molecular Biology**

# Content

- ✓ Expressed sequence tag ??
- ✓ Use of EST.
- ✓ Tag preparation.
- ✓ cDNA to EST.
- ✓ EST and NCBI.
- ✓ Gene mapping tool.
- ✓ EST format on NCBI site.
- ✓ Limitation.
- ✓ Collection.

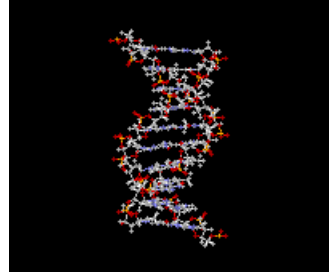


## Express Sequence Tag (EST) ???



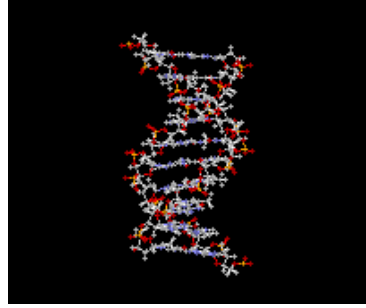
- ESTs are **small pieces of DNA sequence**.
- Short, single pass **sequence read from mRNA**.
- Bits of DNA sequence , **Represent gene express** in single cells, tissues or organ from different organisms.

# Use of EST



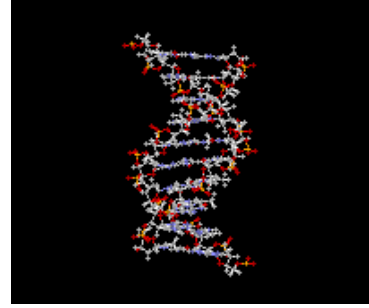
- ✓ Identify unknown gene and map their position in a genome.
- ✓ Provide simple and inexpensive path for discovering new gene.
- ✓ Genome map construction.
- ✓ Characterization of expressed gene.

# Tag Preparation



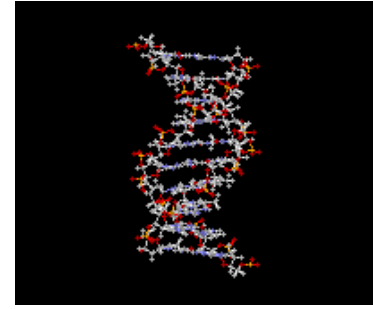
- It is a conversion of **m RNA** in to **cDNA**.
- **Stability** of cDNA.
- **Expressed DNA** sequences only.

# cDNA -----ESTs



- ✓ **5' EST-** sequencing only beginning portion of cDNA.
- ✓ Code for protein. These region tend to **conserved** and not change within a gene family.
- ✓ **3' EST-** sequencing only ending portion of cDNA.
- ✓ less conservation because it is **un translated region**.

# Gene mapping tool



## ✓ ETS – Genome marker.

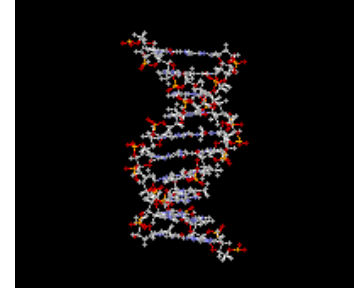
- Unique to a particular.
- Easily recognizable and occur only once in genome or chromosome.

## ✓ Resource of gene discovery.

- Tool for detection of hereditary disease involving gene.
- Generation of ETS is rapid and inexpensive.

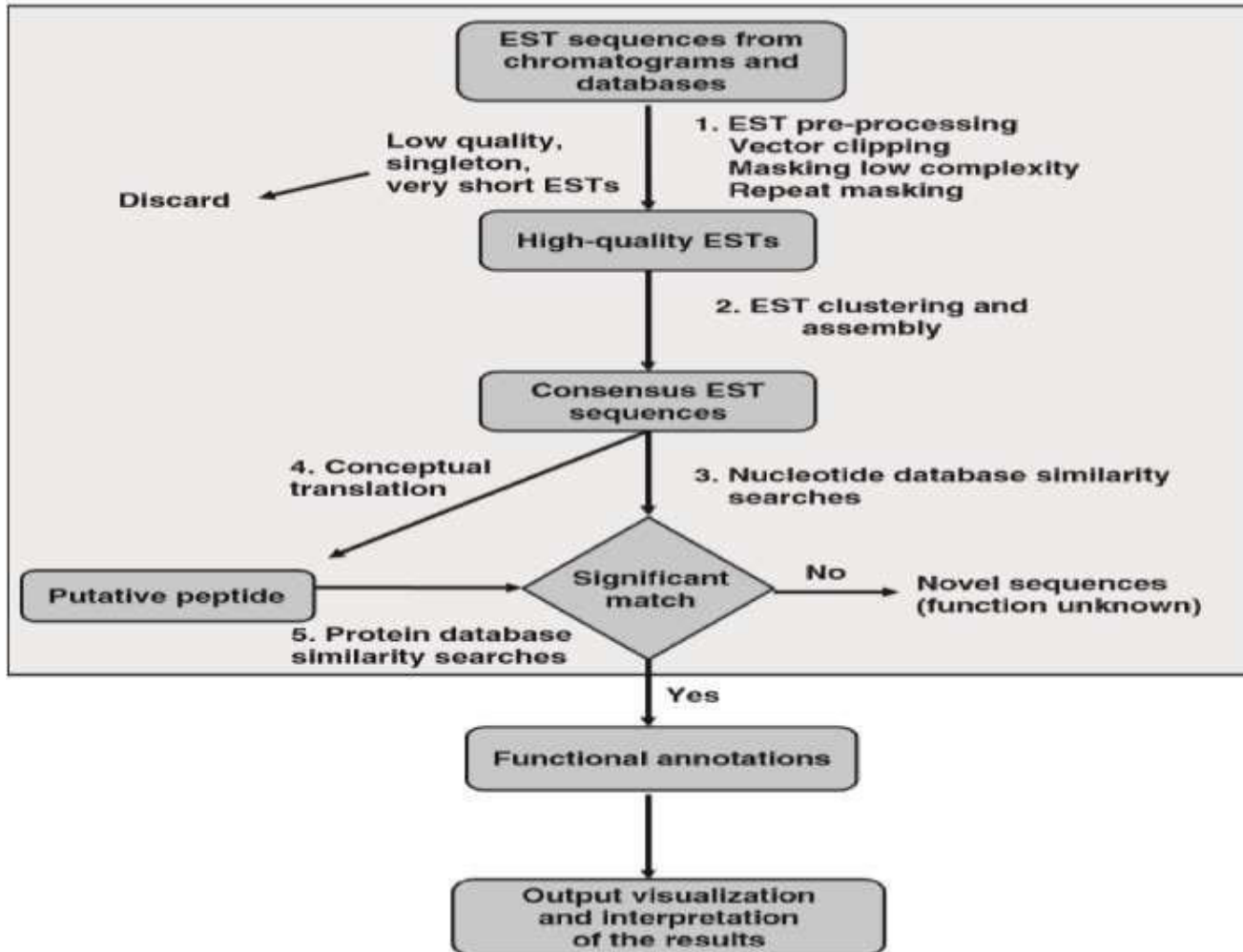
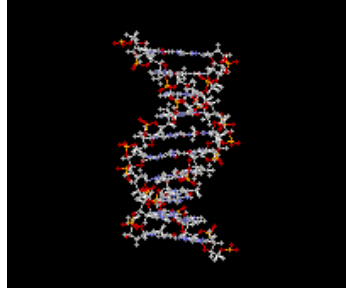


# ESTs and NCBI

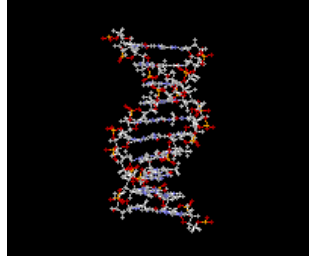


- ✓ In 1992, scientist at NCBI developed a new database designed to serve as a **collection point for ESTs**.
- ✓ Once EST that was submitted to **GenBank** had been screen and **annotated**, it was then deposited in this new database, called **dbEST**.

# ESTs and NCBI



# EST Format



NCBI Resources How To Sign in to NCBI

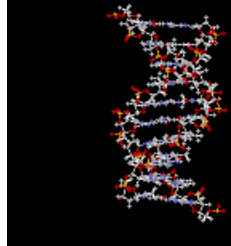
Search NCBI databases Help

est Search

Results found in 35 databases for "est"

Literature			Genes		
Books	1,157	books and reports	EST	76,032,739	expressed sequence tag sequences
MeSH	17	ontology used for PubMed indexing	Gene	27,714	collected information about gene loci
NLM Catalog	2,193	books, journals and more in the NLM Collections	GEO DataSets	2,363	functional genomics studies
PubMed	13,104	scientific & medical abstracts/citations	GEO Profiles	142,293	gene expression and molecular abundance profiles
PubMed Central	82,881	full-text journal articles	HomoloGene	268	homologous gene sets for selected organisms
Health			PopSet	554	sequence sets from phylogenetic and population studies
ClinVar	2	human variations of clinical significance	UniGene	4,615	clusters of expressed transcripts
dbGaP	175	genotype/phenotype interaction studies	Proteins		
GTR	2	genetic testing registry	Conserved Domains	3	conserved protein domains
MedGen	5	medical genetics literature and links	Protein	538,932	protein sequences
OMIM	2,362	online mendelian inheritance in man	Protein Clusters	1	sequence similarity-based protein clusters
PubMed Health	169	clinical effectiveness, disease and drug reports	Structure	127	experimentally-determined biomolecular structures
Genomes			Chemicals		
Assembly	0	genome assembly information	BioSystems	2,030	molecular pathways with links to genes, proteins and chemicals
BioProject	535	biological projects providing data to NCBI	PubChem BioAssay	1,447	bioactivity screening studies
BioSample	28,935	descriptions of biological source materials	PubChem Compound	44	chemical information with structures, information and links
Clone	0	genomic and cDNA clones	PubChem Substance	193	deposited substance and chemical information
dbVar	529	genome structural variation studies			
Epigenomics	0	epigenomic studies and display tools			
Genome	25	genome sequencing projects by organism			
GSS	81,512	genome survey sequences			
Nucleotide	3,600,333	DNA and RNA sequences			
Probe	46,262	sequence-based probes and primers			
SNP	6,549	short genetic variations			
SRA	3,432	high-throughput DNA and RNA sequence read archive			
Taxonomy	0	taxonomic classification and nomenclature catalog			

# EST Format



NCBI Resources (2) How To (2) Sign In to NCBI

EST [EST] [est] Search

Create alert Limits Advanced Help

See [est.alpha/beta.hydrobase](#) in the Gene database  
[est](#) reference sequences (Epubs,LU)

Display Settings: Summary, 20 per page, Sorted by Default order Send to: -

Items: 1 to 20 of 76032739

Page [1] of 380167 Next - Last -

Found 70714664 nucleotide sequences. Nucleotide (360933) EST (76032739) 655 (81812)

- [Z1\\_EST1569 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
672 bp linear mRNA  
Accession: J2478143.1 GI: 512289528  
EST GenBank FASTA
- [Z1\\_EST1568 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
530 bp linear mRNA  
Accession: J2478142.1 GI: 512289531  
EST GenBank FASTA
- [Z1\\_EST1567 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
517 bp linear mRNA  
Accession: J2478141.1 GI: 512289530  
EST GenBank FASTA
- [Z1\\_EST1566 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
746 bp linear mRNA  
Accession: J2478140.1 GI: 512289529  
EST GenBank FASTA
- [Z1\\_EST1565 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
529 bp linear mRNA  
Accession: J2478139.1 GI: 512289528  
EST GenBank FASTA
- [Z1\\_EST1564 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
511 bp linear mRNA  
Accession: J2478138.1 GI: 512289527  
EST GenBank FASTA
- [Z1\\_EST1563 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
750 bp linear mRNA  
Accession: J2478137.1 GI: 512289526  
EST GenBank FASTA
- [Z1\\_EST1562 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
482 bp linear mRNA  
Accession: J2478136.1 GI: 512289525  
EST GenBank FASTA
- [Z1\\_EST1561 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
421 bp linear mRNA  
Accession: J2478135.1 GI: 512289524  
EST GenBank FASTA
- [Z1\\_EST1560 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
730 bp linear mRNA  
Accession: J2478134.1 GI: 512289523  
EST GenBank FASTA
- [Z1\\_EST1579 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
297 bp linear mRNA  
Accession: J2478133.1 GI: 512289522  
EST GenBank FASTA
- [Z1\\_EST1578 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
800 bp linear mRNA  
Accession: J2478132.1 GI: 512289521  
EST GenBank FASTA
- [Z1\\_EST1577 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence](#)  
309 bp linear mRNA  
Accession: J2478131.1 GI: 512289520  
EST GenBank FASTA

Filter your results:

- All (76032739)
- Bacteria (2714)
- mRNA (76032739)
- Manage Filters

Results by taxon

Top Organisms [Taxa]

- Homo sapiens (2709028)
- Mus musculus (4953604)
- Caenorhabditis elegans (2049602)
- Gus scrofa (1676489)
- Bos taurus (7522417)
- All other taxa (37784399)
- More...

Find related data

Database: [Select]

Search details

est [All] [Fields]

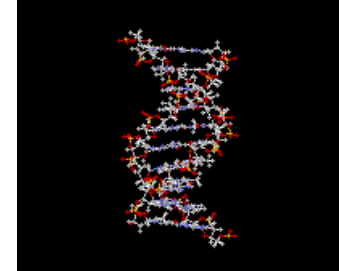
Search See more...

Recent activity

est (76032739) Turn Off Clear

- The detection of DNA-binding proteins by protein blotting
- Dual Positive And Negative Control: the Arabidopsis Operon - Moders
- See more...

# EST Format



NCBI Resources How To Sign In to NCBI

EST EST [Search] Limits Advanced Help

Display Settings: - EST Send to: -

### ZJ\_EST588 Ziziphus jujuba cv. Dongzao fruit EST library Ziziphus jujuba cDNA, mRNA sequence

GenBank: JZ475142.1  
SraBank: EASIA

**IDENTIFIERS**

dbEST id: 78679602  
EST name: ZJ\_EST588  
GenBank Acc: JZ475142  
GenBank gi: 515289591

**CLONE INFO**

DNA type: cDNA  
PRIMERs  
5-prime Tail: none

**SEQUENCE**

```
CTTCAAGAAAGCTGCAGCTTCAGAGCCAAAGCTCTCTTTTATTACAGAGCCAAAGAGCTCG  
AAGTTATAAACACAGCACTGTCGCTTCGGGTTTGATCGTAAACAAATCCCGATGACCGG  
GCAATCTGGAAATGGATGATTCGGTGGCCACCTTTGGATCAGGCTTCACGCTGATC  
CAGCTCCGATTCCTCGCTTACTCTTCGCTCTTGGAAATCGGTCGCGAAATTCGAGGT  
TCAGCTTAATTAATTTGGCTGTGGGTTCTGATATGGGAAAGGTTTTCGGATGGCTCTCGGG  
GTGGGCTTTGATGATTTTTCCTCTTTGGGTTGATGCTTTTTCATGGCGGCTTTGATGGGTT  
TATGGATATGGAAATCGAATGGCTCGCTCATACCCAAATCATTCCTCTCTATATACCT  
GGAGTTTACTCTTTCGACAGGACAGGATGGTTAACCGTTGCACTGTGCCACCCCAATTG  
CCTCTCTTACTGCTGCATCAGCCAGTGTGTATGCACTGCATACAGCCACTGCATTAA  
TCGGTTGACCTCCGCTTCATATTCCTTCGCGCTCTCAATTAATTCAGAGCTCTTGA  
GACCAAGCAAGTAGGTTGCAACACACATTTTCAGCACAAACATCCCATAGGTTCC  
TTGCTATGGGCTCTCTGCTGCAATAGTTTATGATTCATATTAAGTTCAAGGCTTAGGA  
GTCTGCTCAGATGCCATTTGCATGGGAAAGGACTGCTATTATTAGACATTGTGTGTGT  
GGGATGCATGCTCTTCTAGGCTTAGTTTCAAGCTGTCTCTATTCTTA
```

Entry Created: Jun 27 2013  
Last Updated: Feb 4 2014

**COMMENTS**

Identified from expressed sequence tag (EST) library of the fruit of *Ziziphus jujuba*

**LIBRARY**

Lib Name: LIBEST\_028238 *Ziziphus jujuba* cv. Dongzao fruit EST library  
Organism: *Ziziphus jujuba*  
Database: Genbank  
Organ: Fruit  
Develop. stage: 30, 60 and 110 days after flowering (DAF)  
Vector: pCD18 vector  
Description: Fruit samples of Dongzao were harvested at 30, 60, 110 days after flowering (DAF), immediately frozen in liquid nitrogen and stored at -80 C. Then the fruit employed for EST libraries construction were collected at jujube fruits of three different periods.

**SUBMITTED**

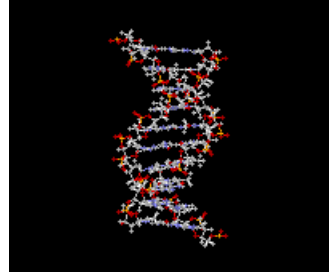
Name: Xiaoming Fang  
Lab: College of Biological Sciences and Technology  
Institution: Beijing Forestry University  
Address: No. 35, Qinghua Eastern Road, Haidian District, Beijing 100083, P.R. China  
Tel: +86-10-62337061  
Fax: +86-10-6233164  
E-mail: xmfang@bjfu.edu.cn

**CITATIONS**

Title: A Chinese jujube (*Ziziphus jujuba* Mill.) fruit expressed sequence tag (EST) library: annotation and EST-SSR characterization  
Authors: Liu, J., Li, H., Ma, L., Wang, S., Q., Sun, J., Li, Y., Wu, R. L., Fang, X. M.  
Citation: Sci. Hortae. 166: 90-106 2014

**MAP DATA**

# EST Limitation

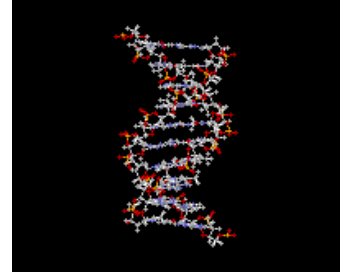


- Very **difficult to isolate mRNA** from some tissue and cell types.
- **No indication** as to the gene from which it is derived.

# How to organize EST collection?

- ✓ **Clustering** :associate individual EST sequence with unique transcripts or gene.
- ✓ **Assembling**: derive consensus sequence from overlapping ESTs belonging to same cluster.
- ✓ **Mapping**: Associate ESTs with exons in genomic sequences.
- ✓ **Interpreting**: Find correct coding region.

# Private EST database



- ✓ **SANBI**- South Africa
- ✓ **MIPS**- Munich and SIB
- ✓ **TIGEM**- Italy
- ✓ **DOTS**- Pennsylvania



Thank  
you

