



**CHANDIGARH  
UNIVERSITY**

Discover. Learn. Empower.

# INSTITUTE-UNIVERSITY INSTITUTE OF ENGINEERING

## ACADEMIC UNIT-II

Computer Science Engineering

Subject Name-Biology For Engineers

Subject Code- 20SZT148


**BIOINFORMATICS AND BLAST**

DISCOVER . **LEARN** . EMPOWER

# TYPES OF TRANSDUCERS

## Course Outcome

CO Number	Title	Level
<b>CO1</b>	It gives an idea about the about the basic cell biology.	<b>Understanding</b>
<b>CO2</b>	It deals with the idea of uses of biology in engineering.	<b>Understanding</b>
<b>CO3</b>	It provide knowledge about the uses of softwares in biology field.	<b>Remembering</b>



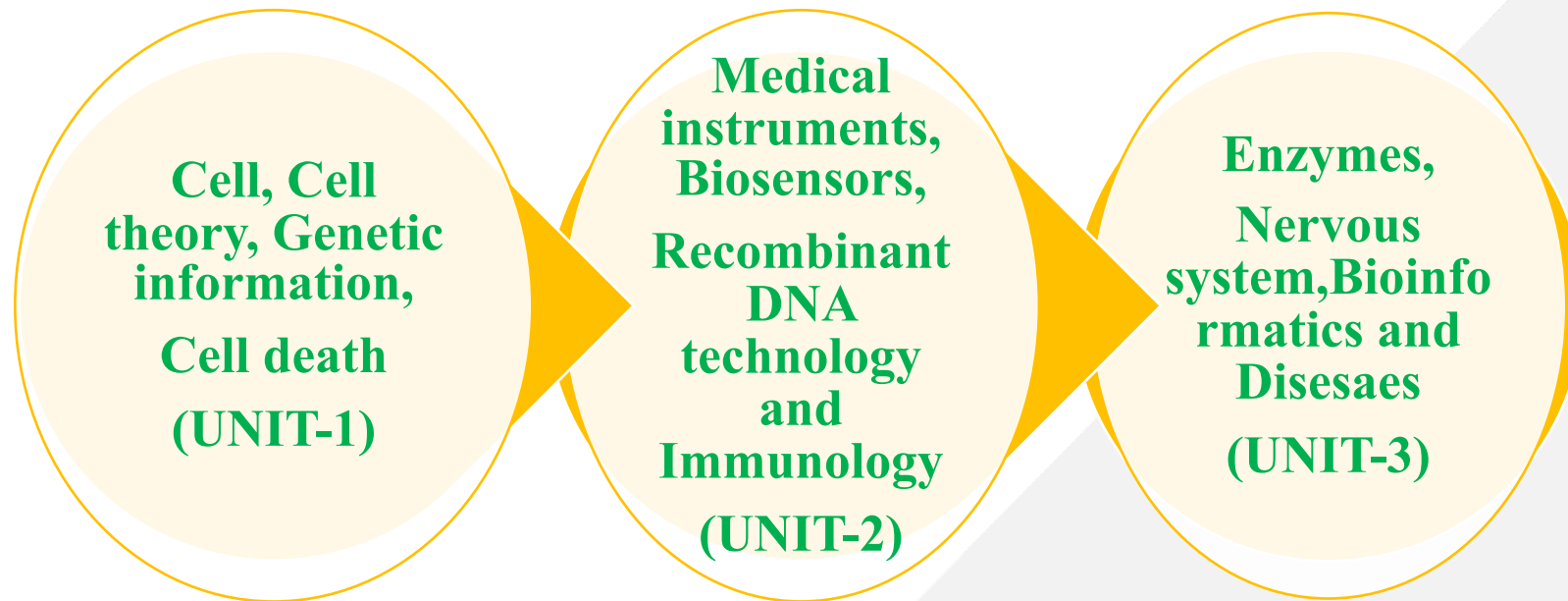
2015-03-23

47

Will be covered in this lecture

<https://www.electronicshub.org/types-of-transducers/>

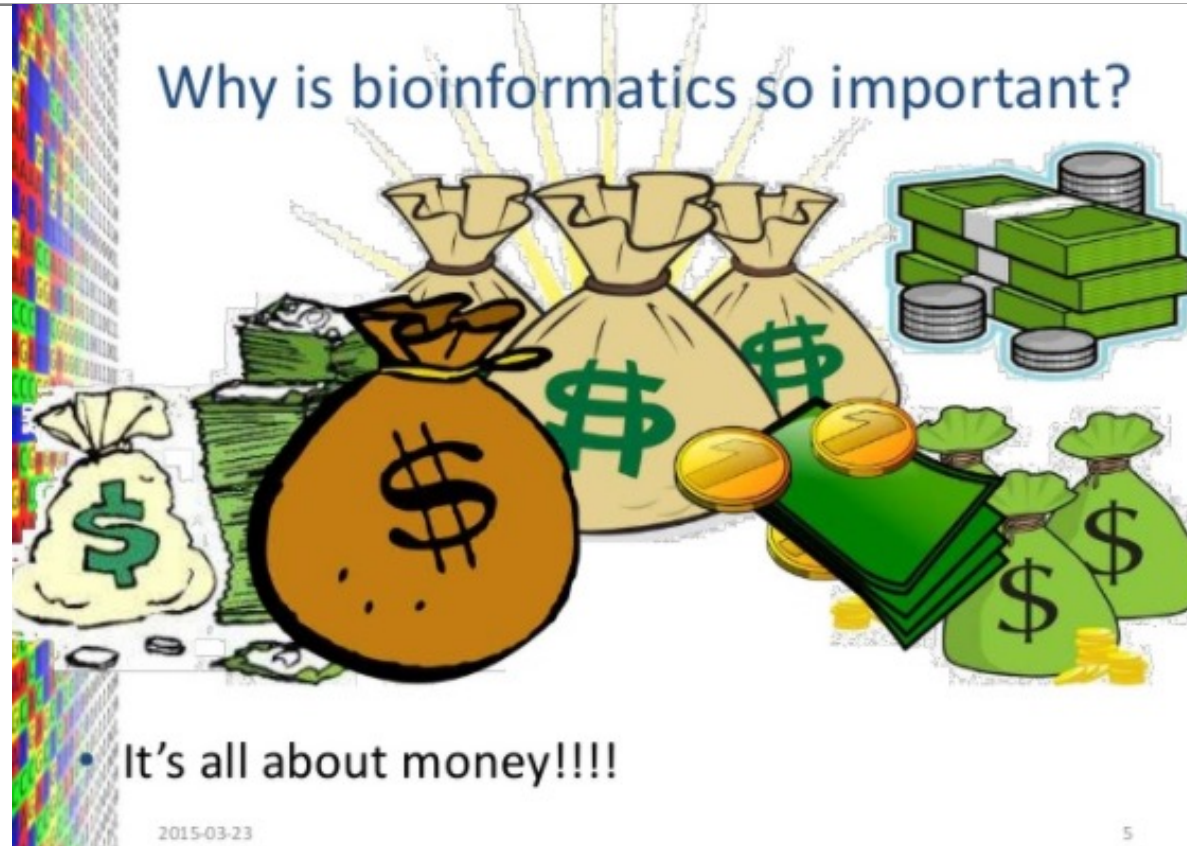
# BIOLOGY FOR ENGINEERS



# BIOINFORMATICS

- Bioinformatics is an interdisciplinary field that develops methods and software tools for understanding biological data.
- As an interdisciplinary field of science, bioinformatics combines computer science, statistics, mathematics, and engineering to study and process biological data.

# BIOINFORMATICS



<https://image.slidesharecdn.com/blast-170204042954/95/blast-basic-local-alignment-search-tool-1-638.jpg?cb=1486182905>

# NEED FOR BIOINFORMATICS

- The need for bioinformatics has arisen from the recent explosion of publicly available genomic information, such as resulting from the Human Genome Project.
- Gain a better understanding of gene analysis, taxonomy, & evolution.
- To work efficiently on the rational drug designs and reduce the time taken for the development of drug manually.

# FIELD OF BIOINFORMATICS

- Molecular Medicine
- Gene Therapy
- Drug Development
- Microbial genome applications
- Crop Improvement
- Forensic Analysis of Microbes
- Biotechnology
- Evolutionary Studies
- Bio-Weapon Creation

# WHERE BIOINFORMATICS HELP

- Experimental Molecular Biology
- In Genetics and Genomics
- In generating Biological Data
- Analysis of gene and protein expression
- Comparison of genomic data
- In Simulation & Modeling of DNA, RNA & Protein



# APPLICATION OF BIOINFORMATICS

- Molecular medicine
- Personalized medicine
- Preventative medicine
- Gene therapy
- Drug development
- Microbial genome applications
- Waste cleanup

# SOFTWARE AND TOOL

- Software tools for bioinformatics range from simple command-line tools
- To more complex graphical programs and standalone web-services available from various bioinformatics companies or public institutions.
- The computational biology tool best-known among biologists is probably BLAST, an algorithm for determining the similarity of arbitrary sequences against other sequences, possibly from curated databases of protein or DNA sequences.

# SOFTWARE AND TOOL

- BLAST is one of a number of generally available programs for doing sequence alignment.
- The NCBI provides a popular web-based implementation that searches their databases.

# SOFTWARE AND TOOL

## High throughput sequencing

- 454 Roche  ~10 hours | 400bp | 500 Mbp
- Illumina  ~250 hours | 2x100bp | >150 Gbp
- PacBio  14 hours | 1..10kbp | ~50 Mbp
- Ion Torrent  4 hours | 400bp | <10 Gbp
- Nanopore  1..? hours | 1..?bp | 1..?bp

2015-03-23

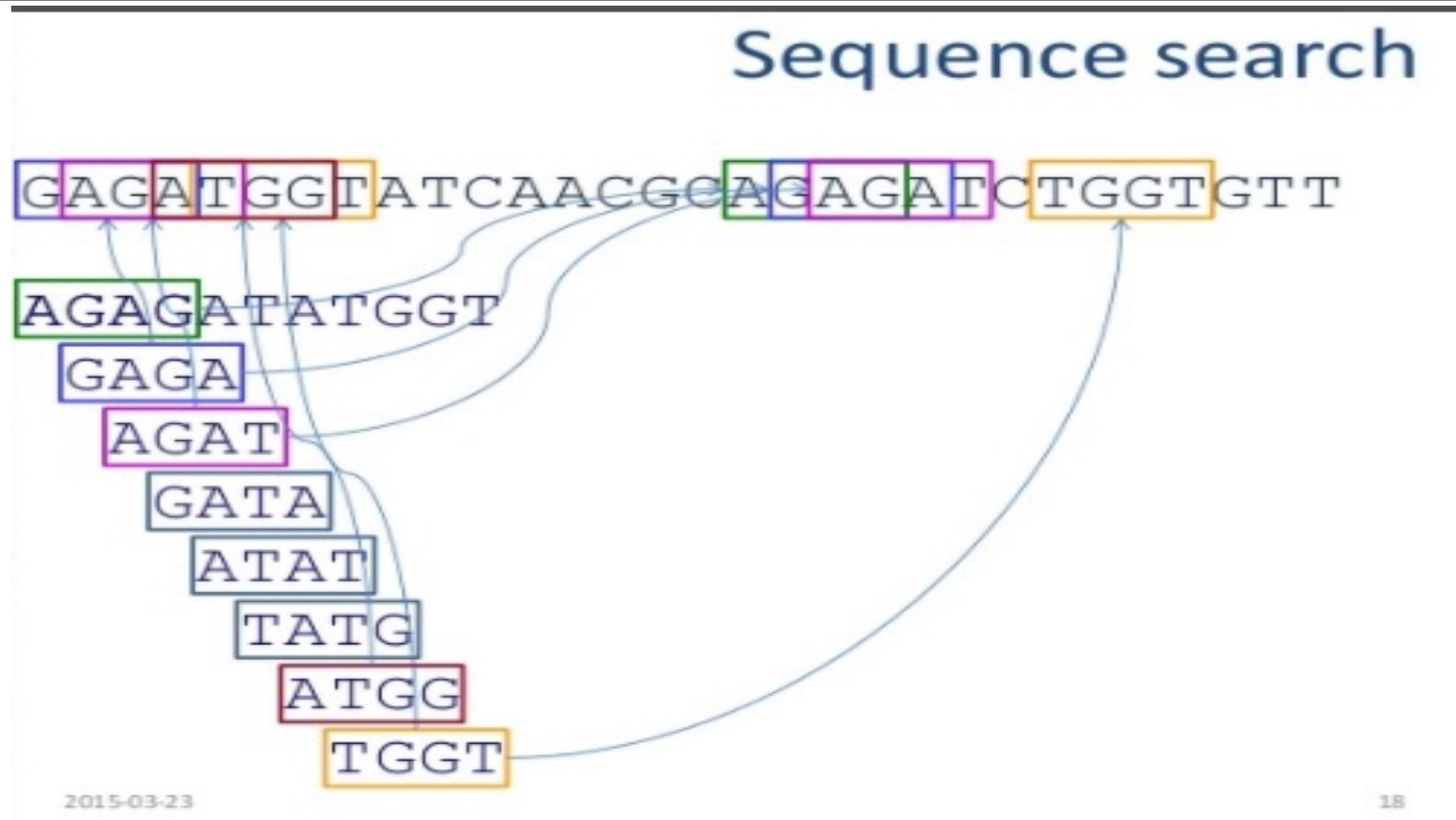
<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-13-638.jpg?cb=1427042135>

# SEQUENCE SEARCH (BLAST)

BLAST is one of the most commonly used bioinformatics software

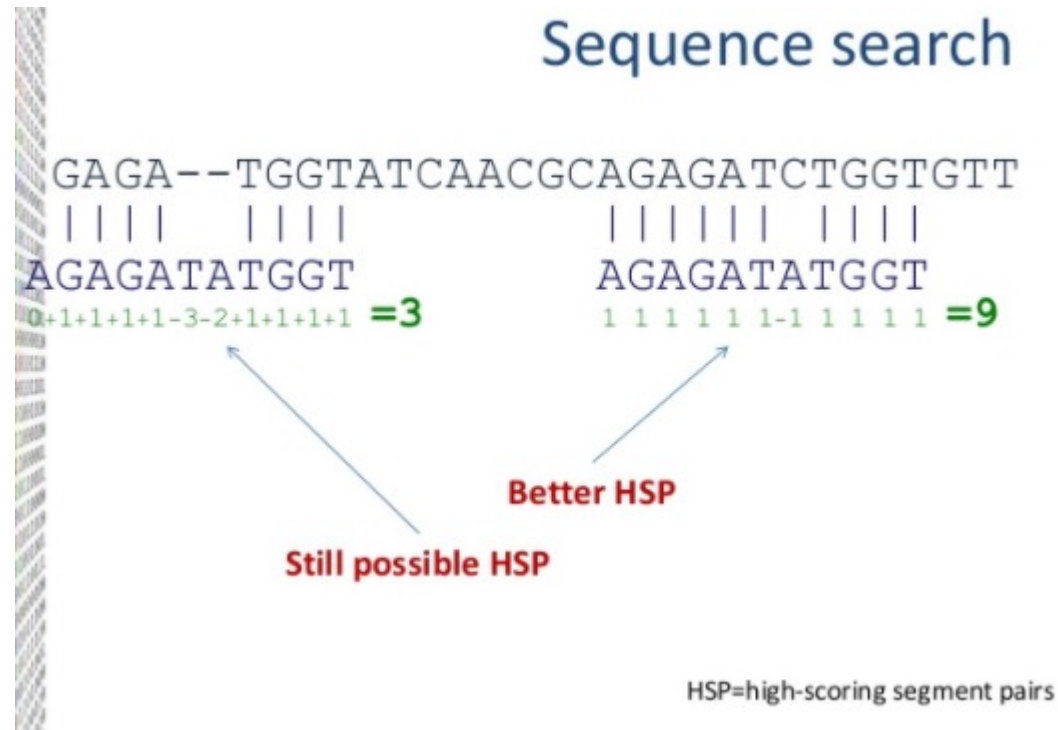
- It finds small sub-sequences of your query in the subject sequence
- Uses word to match with the database of subject and then uses heuristics to verify and extend match

# SEQUENCE SEARCH (BLAST)



<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-18-638.jpg?cb=1427042135>

# SEQUENCE SEARCH (BLAST)



# SEQUENCE SEARCH (BLAST)

- Global alignment – global optimization that "forces" the alignment to span the entire sequences (Needleman–Wunsch algorithm or *Clustal style*)
  
- Local alignment – identify short regions of similarity within long divergent sequences (Smith–Waterman algorithm or BLAST style)



# SEQUENCE SEARCH (BLAST)

## Sequence alignment

- Global alignment

```
FTFTALILLAVAV
|  |||  |||  ||
F--TAL-LLA-AV
```

- Local alignment

```
FTFTALILL-AVAV
|||  ||  ||
--FTAL-LLAAV--
```

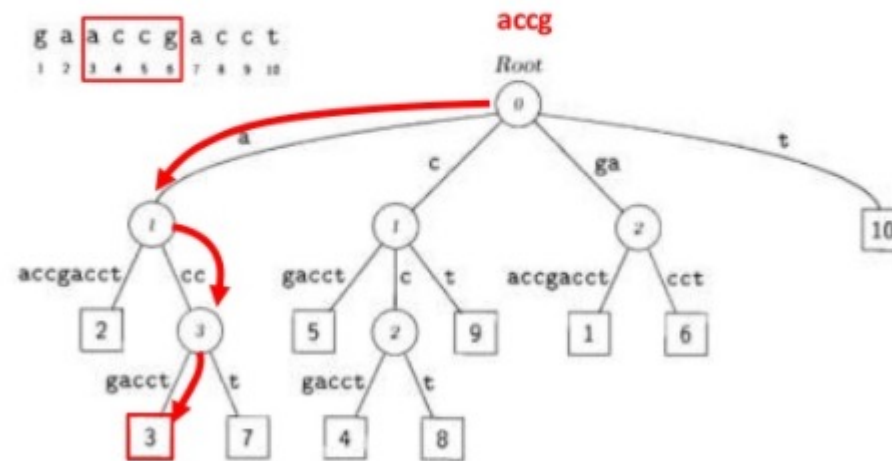
<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-21-638.jpg?cb=1427042135>

# GENOME ALIGNMENT

- Glocal alignment
- Uses a word matching method
- Creates suffix tree for faster search
- Searches suffix tree for exact matches of words clusters them and then uses local alignment methods to extend match

# GENOME ALIGNMENT

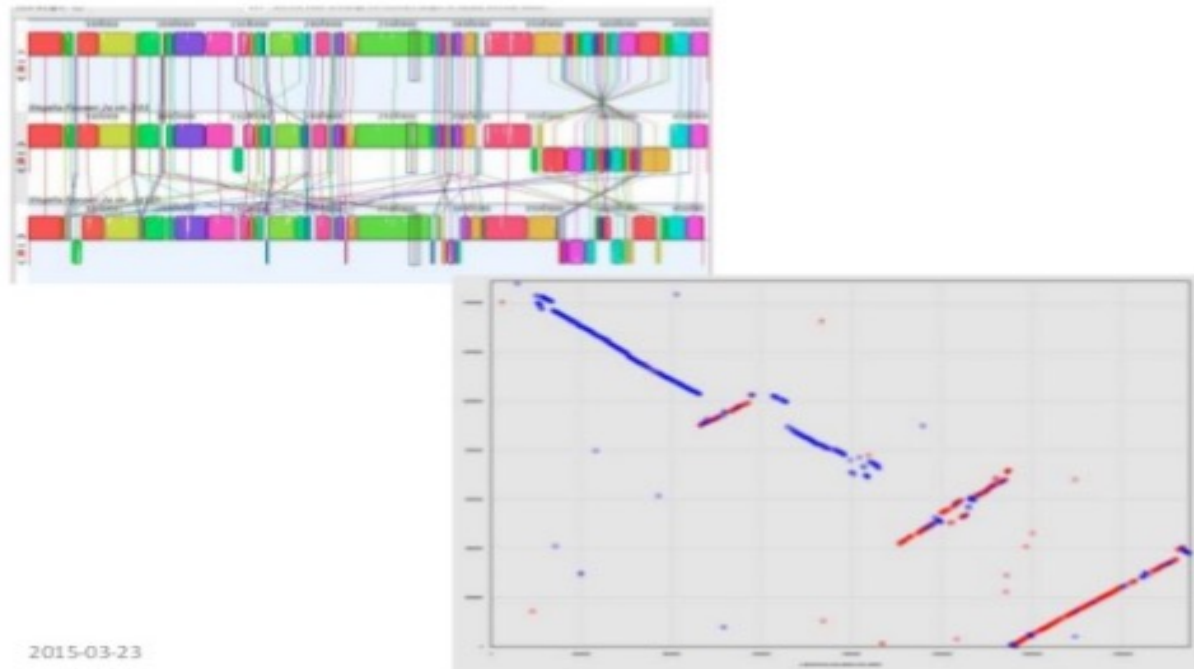
## Suffix tree



<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

# GENOME ALIGNMENT

## Genome alignment

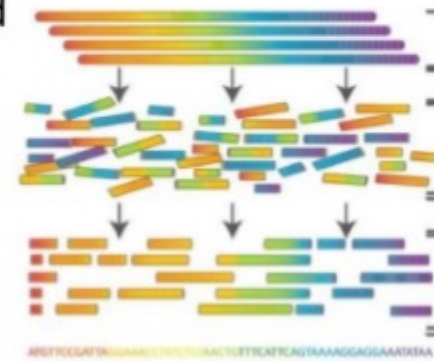


<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

# ASSEMBLY

## Assembly

- Short read assembly is extremely difficult and computationally intensive!
- For longer reads an Overlap Consensus (OLC) assemblers are used
- For shorter reads (and in high numbers) De Bruijn Graph assemblers are better

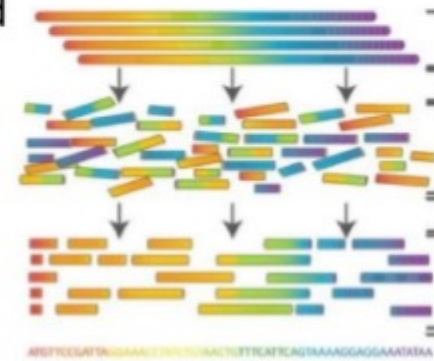


<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

# ASSEMBLY

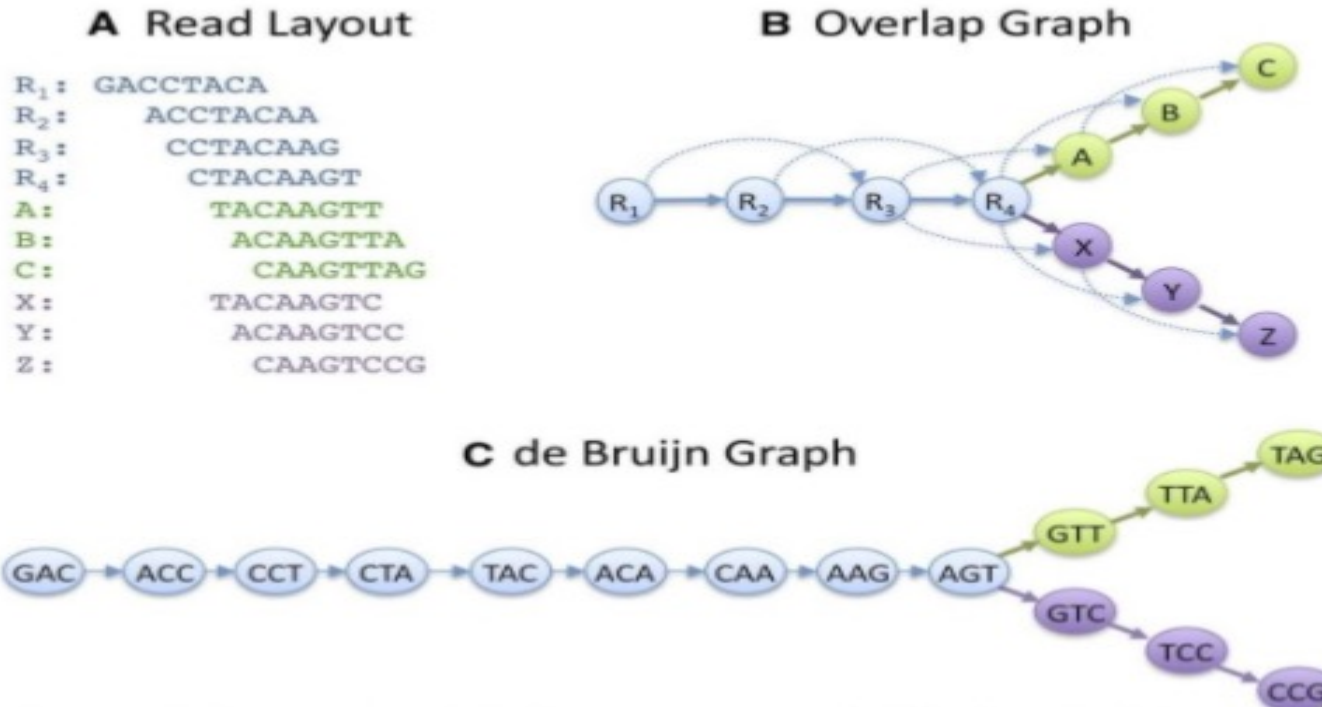
## Assembly

- Short read assembly is extremely difficult and computationally intensive!
- For longer reads an Overlap Consensus (OLC) assemblers are used
- For shorter reads (and in high numbers) De Bruijn Graph assemblers are better



<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

# ASSEMBLY



Source: Schatz, Delcher, Salzberg, <http://www.genome.org/cgi/doi/10.1101/gr.101360.109>.  
2015-03-23 26

<https://image.slidesharecdn.com/introductiontobioinformatics-1503191/5041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

# Genome annotation

- Prediction of:
  - Genes
  - Repeats
  - Non coding RNAs (rRNAs, tRNAs, miRNAs, snRNAs, siRNAs, ta-siRNAs)
  - Promoters
  - Enhancers
  - Protein binding sites
  - ...

<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

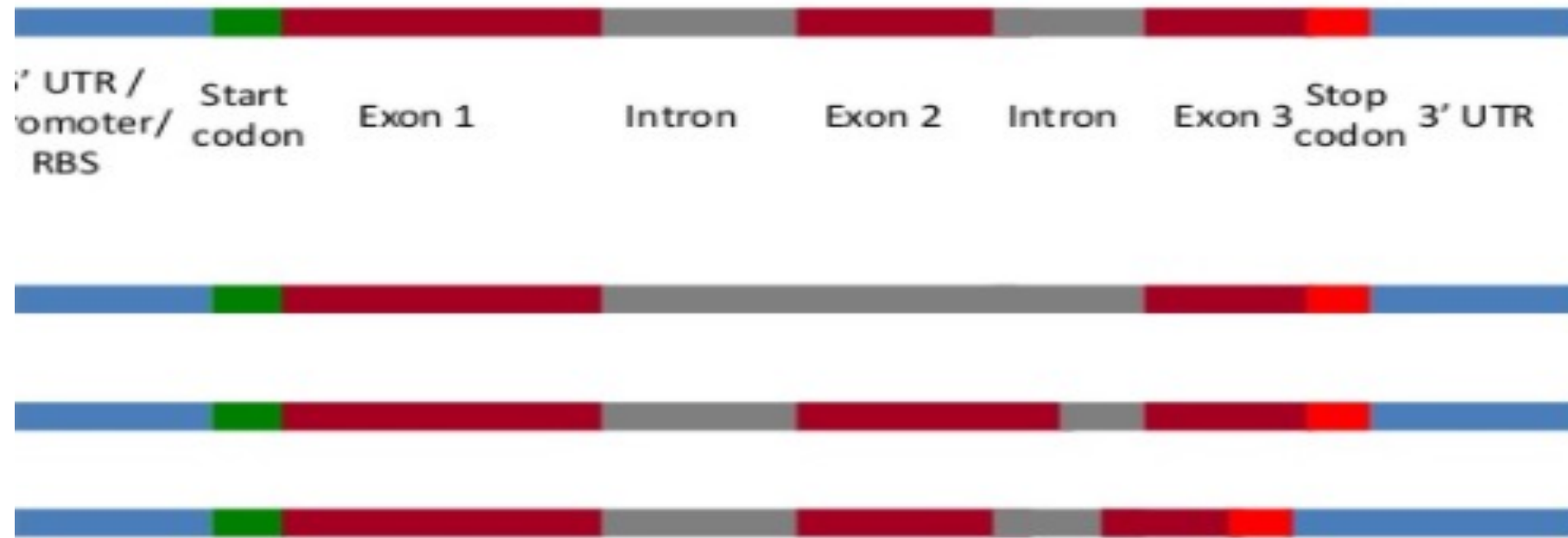


# Genome annotation

- Prediction of:
  - Genes
  - Repeats
  - Non coding RNAs (rRNAs, tRNAs, miRNAs, snRNAs, siRNAs, ta-siRNAs)
  - Promoters
  - Enhancers
  - Protein binding sites
  - ...

<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

# Genome annotation



<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

# Genome annotation

ACCTCTCACTCTTTCTTCTCATCTTCTTCAATTATAACAACCTAACCATGTCTTTAGAACACAAGAAT  
TTCTCTTCCTAATTCCTTCAACAATGGCCAATAATCTCAATCTCTTCCTTTGTTTCCTTCTCTTTATT  
TGTTTCTTCACTCTCTGCCTTAGCCCTGGTGGTCTAGCTTGGGCTCTAATTTCAAAGCCAAAAAA  
CCAATCCATCATTCTAGAAATCAGCTTATGAGCTTTTGTTTCATAGAGCAATGGGGTTTGCTCCAT  
TGGAGAGTGTGTTTGTGTTGGTGATGAAGAAAATAAGAACAGCCCTAATGATAATGAAGATGAA  
GATTTTGTTGATGTGTTGCTTGATTGGGAAAAGGAAAACAACTCACCGACTCTGATATGATTGC  
TGTGTTAGGTACGTATGTGTATTATAATTTCTTGTTTCATTACTATTTTGATATTTTCTACTGCACT  
TCAATTTTAATCGGTTTGAAATGATTTTTTAATATGCTCTTACAAGATTATGACTTGGGGAAAGATT  
TTACATCTTTAAATATTTCATTTTTTGTGTGATACATGAAATGCATGACTGTTTTTTACTTGCGATT  
TACATGTTGAAATTTTCTTTACTTTGATATTCTATGTTTTTTTAAACAAATTTTCTCTTAAATAAATGA  
CATGTAGGAAATGACCTCCCAAATCTTCCTTATCTCCATGCAGTCGTCAAAGAGACTCTTAGAAT  
GCACCTCCCGGCCCACTTCTCTTGGGGCACGCCTTGCCATCCATGACACCCATGTGCGCAGGC  
CACTACATTCCTGCTGGCACCCTGCGATGGTCAACATGTGGGCCATAACCCACGACGACCAAA  
CTGTGGCTCGCTCAGTTAGTTCATAAGTTCGAATGGGTTCAAGCTGATGAATCGAAAATCAAAG  
TGGATTGTCTGAGTGTTTGAAGCTATCTCTGGAAATGAAACACCCTTTGATTGTAGGGCTATC  
CCGAGGAATGTAGGGTTCGAGTCTCACCTGATCATGCATGACAGATTAAAAAAAAGAA  
AAGGCACATCTAGGGGAGCTTATTATGATATTATCATATGTTGAAAATTAAATGTGTTTGTGCTTT  
CTTTCTTTTTTTCTTTTCTTTCTTTCTTTCTTTAATCAATTGATATTATATCTTGTGTGGAAACA  
ATAGTATCGGATTGAGATTAAATGTTGGGATAATCCTTAAATGTAATTCCGTTATTAAGTGTA

<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

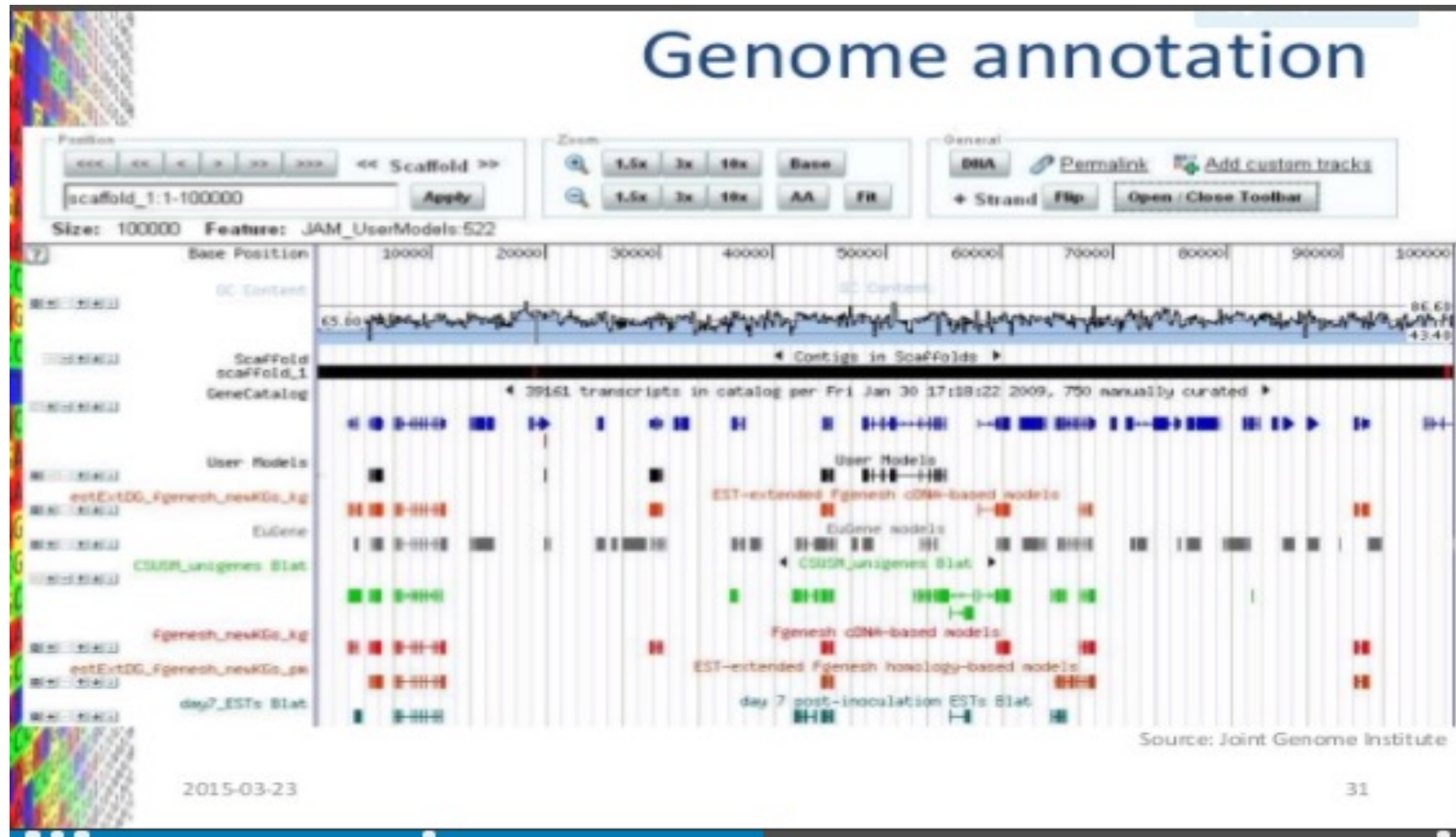


# Genome annotation



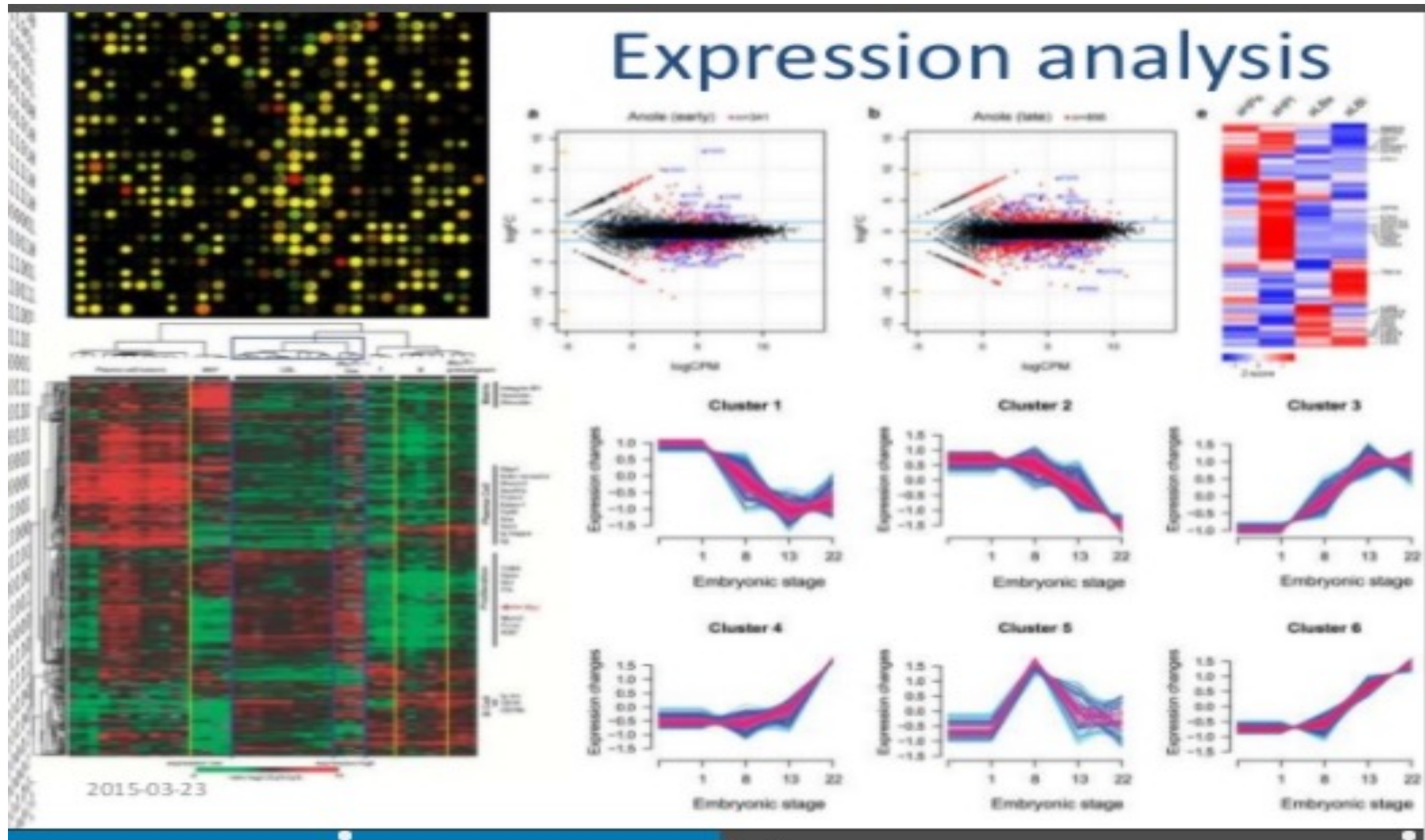
ACCTCTCACTCTTTCTTCTCATCTTCTTCAATTATAACAACCTAACC[ATG]TCTTTAGAACAGAAT  
 TTCTCTTCCTAATTCCTTCAACAATGGCCAATAATCTCAATCTCTTCCTTTGTTTCCTTCTCTTTATT  
 TGTTTCTTCACTCTCTGCCTTAGCCCTGGTGGTCTAGCTTGGGCTCTAATTTCAAAGCCAAAAA  
 CCAATCCATCATTCTAGAAATCAGCTTATGAGCTTTTGTTCATAGAGCAATGGGGTTTGCTCCAT  
 TGGAGAGTGTGTTTGTGTTGGTGATGAAGAAAATAAGAACAGCCCTAATGATAATGAAGATGAA  
 GATTTTGTGATGTGTTGCTTGATTGGAAAAGGAAAACAAACCTCACCGACTCTGATATGATTGC  
 TGTGTTAG | GTACGTATGTGTATTATAATTTCTTGTTTCATTACTATTTTGATATTTTCTACTGCACT  
 TCAATTTTAATCGGTTTGAAATGATTTTTTAATATGCTCTTACAAGATTATGACTTGGGAAAGATTC  
 TTACATCTTTAAATATTCAATTTTTGTGTGATACATGAAATGCATGACTGTTTTTTACTTGCGATT  
 TACATGTTGAAATTTCTTTACTTTGATATTCTATGTTTTTTTAACAAATTTCTCTTAAATAAATGA  
 CATGTAG | GAAATGACCTCCCAAATCTTCCTTATCTCCATGCAGTCGTCAAAGAGACTCTTAGAAT  
 GCACCTCCCGGCCCACTTCTCTTGGGACAGCCTTGCCATCCATGACACCCATGTCGCAGGC  
 CACTACATTCTGCTGGCACCCTGCGATGGTCAACATGTGGGCCATAACCCACGACGACCAAA  
 CTGTGGCTCGCTCAGTTAGTTCATAAGTTCGAATGGGTTCAAGCTGATGAATCGAAAATCAAAG  
 TGGATTTGTCTGAGTGTGTTGAAGCTATCTCTGGAAATGAAACACCCTTGATTGTAGGGCTATC  
 CCGAGGAATGTAGGGTTCGAGTCTCACC[TGA]TCATGCATGACAGATTAAAAAAGAA  
 AAGGCACATCTAGGGGAGCTTATTATGATATTATCATATGTTGAAAATTAAATGTGTTTGTGCTTT  
 CTTTTCTTTTTCTTTTCCTTCTTCTTCTTCTTAATCAATTGATATTATCTTGTGTGGAAACA  
 ATAGTATCGGATTGAGATTAATGTTGGGATAATCCTTAAATGTAATTCCGTTATTAAGTGTGAA

<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>



<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>





<https://image.slidesharecdn.com/introductiontobioinformatics-150319175041-conversion-gate01/95/introduction-to-bioinformatics-23-638.jpg?cb=1427042135>

# USES OF BLAST

- Uses of blast BLAST can be used for several purposes.
- These include:
  - the sequence of interest. Identifying Species: With the use of BLAST, you can possibly correctly identify a species and/or find homologous species.
  - This can be useful, for example, when you are working with a DNA sequence from an unknown species.
- Locating Domains: When working with a protein sequence you can input it into BLAST, to locate known domains within

# USES OF BLAST

- Establishing Phylogeny: Using the results received through BLAST, you can create a phylogenetic tree using the BLAST web-page.
- Comparison: When working with genes,
- BLAST can locate common genes in two related species,
- and can be used to map annotations from one organism to another.



# CONCLUSION

- Bioinformatics is an interdisciplinary field that develops methods and software tools for understanding biological data.
- Software tools for bioinformatics range from simple command-line tools
- To more complex graphical programs and standalone web-services available from various bioinformatics companies or public institutions.

# REFERENCES

- C.B.Powar, 2010.Cell Biology.5<sup>th</sup> Ed,Himalyan Publishing House.
- Leshie Cromwell, Fred.J. Weibell and Erich.A.Pfeiffer. 2003. Biomedical instrumentation and measurements. 2<sup>nd</sup> edition, PHI.
- John G. Webster 1998. Medical Instrumentation: Applications and Design, 3<sup>rd</sup> edition, Jon Wiley and Sons, New York.
- Jeremy M. Berg, John L. Tymoczko and Lubert Stryer. 2006. “Biochemistry,” 6th Ed. W.H. Freeman and Co. Ltd.
- Robert Weaver. 2012 “Molecular Biology,” 5th Edition, MCGraw-Hill.
- Jon Cooper, , 2004. “Biosensors A Practical Approach” Bellwether Books.
- Martin Alexander, 1994 “Biodegradation and Bioremediation,” Academic Press.



# THANK YOU

For queries  
Email: [subject\\_code\\_2020@gmail.com](mailto:subject_code_2020@gmail.com)