International Institute of Health Management Research (IIHMR)
New Delhi
Batch: 2020 - 2022

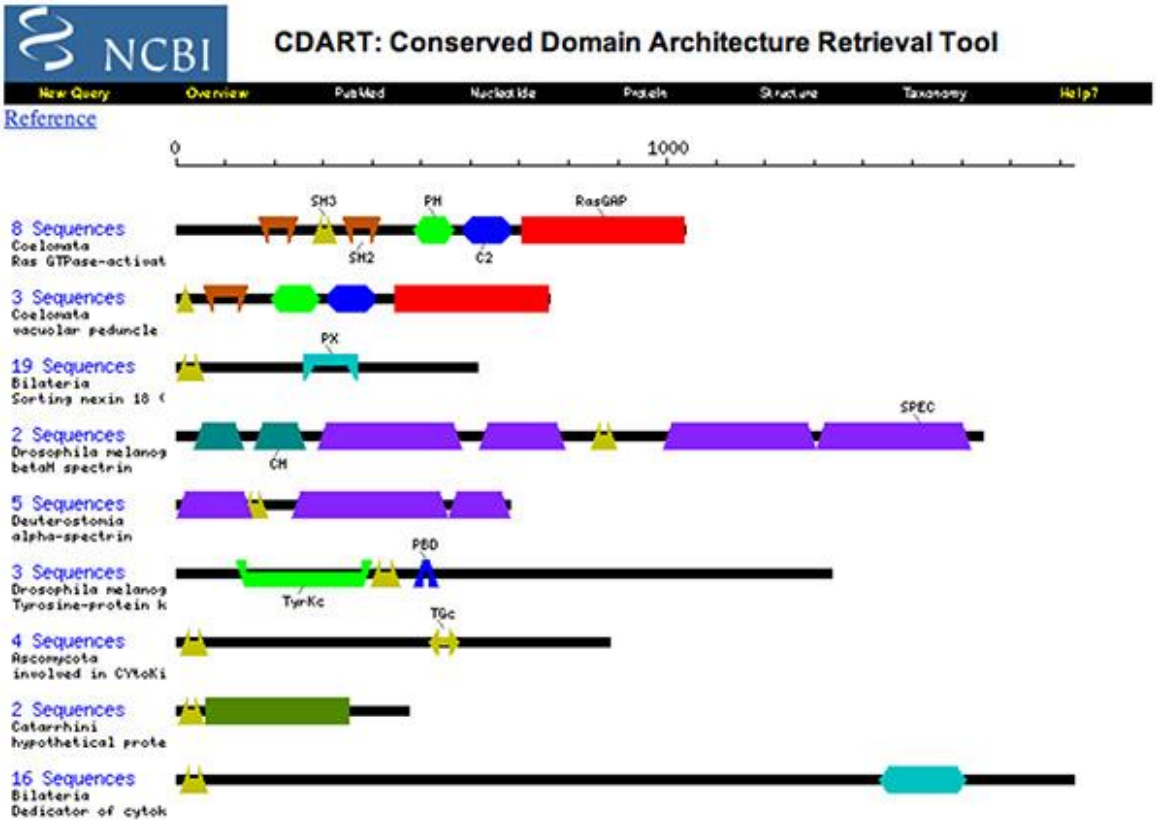
Total marks: 70 **Bioinformatics (HIT – 711)** **Time: 2:00 HR**

Note: Q.No 6 needs to be uploaded

Section A

Answer all the following Questions **5 x 8 = 40**

1.



This is a Conserved Domain Architecture retrieval Tool (CDART) entry. This screenshot shows the first of ten pages of results for the eukaryotic SH3 domains. Answer the following questions based on these: a) briefly define homologs, orthologs and paralogs. b) Are the proteins depicted above homologs?

2. What is drug repurposing? Can you explain with an example? What is the role of bioinformatics in this?

3. How has Next generation sequencing helped in fighting the pandemic?

4. Write short note on BRENDA

5. How is bioinformatics linked to health explain with an example (No theory)?

Section B

Answer all the following Questions

6. Let S1 = AATTCGCGTA and S2 = TATCGCTACA Obtain the optimal global alignment using dynamic programming method. Use any scoring scheme of your choice 10 marks

7. Disease Cause Is Pinpointed With Genome, New York Times, **By Nicholas Wade** March 10, 2010 - Case Study

Two research teams have independently decoded the entire genome of patients to find the exact genetic cause of their diseases. The approach may offer a new start in the so far disappointing effort to identify the genetic roots of major killers like heart disease, diabetes and Alzheimer's. In the decade since the first full genetic code of a human was sequenced for some $500 million, less than a dozen genomes had been decoded, all of healthy people.

Geneticists said the new research showed it was now possible to sequence the entire genome of a patient at reasonable cost and with sufficient accuracy to be of practical use to medical researchers. One subject's genome cost just $50,000 to decode.

"We are finally about to turn the corner, and I suspect that in the next few years human genetics will finally begin to systematically deliver clinically meaningful findings," said David B. Goldstein, a Duke University geneticist who has criticized the current approach to identifying genetic causes of common diseases.

Besides identifying disease genes, one team, in Seattle, was able to make the first direct estimate of the number of mutations, or changes in DNA, that are passed on from parent to child. They calculate that of the three billion units in the human genome, 60 per generation are changed by random mutation — considerably less than previously thought.

The three diseases analyzed in the two reports, published online Wednesday, are caused by single, rare mutations in a gene.

In one case, Richard A. Gibbs of the Baylor College of Medicine sequenced the whole genome of his colleague Dr. James R. Lupski, a prominent medical geneticist who has a nerve disease, Charcot-Marie-Tooth neuropathy.

In the second, Leroy Hood and David J. Galas of the Institute for Systems Biology in Seattle have decoded the genomes of two children with two rare genetic diseases, and their parents.

More common diseases, like cancer, are thought to be caused by mutations in several genes, and finding the causes was the principal goal of the $3 billion human genome project. To that end, medical geneticists have invested heavily over the last eight years in an alluring shortcut.

But the shortcut was based on a premise that is turning out to be incorrect. Scientists thought the mutations that caused common diseases would themselves be common. So they first identified the common mutations in the human population in a $100 million project called the HapMap. Then they compared patients' genomes with those of healthy genomes. The comparisons relied on ingenious devices called SNP chips, which scan just a tiny portion of

the genome. (SNP, pronounced “snip,” stands for single nucleotide polymorphism.) These projects, called genome-wide association studies, each cost around $10 million or more.

The results of this costly international exercise have been disappointing. About 2,000 sites on the human genome have been statistically linked with various diseases, but in many cases the sites are not inside working genes, suggesting there may be some conceptual flaw in the statistics. And in most diseases the culprit DNA was linked to only a small portion of all the cases of the disease. It seemed that natural selection has weeded out any disease-causing mutation before it becomes common.

Dr. James R. Lupski, a medical geneticist with a nerve disease, had his whole genome decoded. Credit...Michael Stravato for The New York Times

The finding implies that common diseases, surprisingly, are caused by rare, not common, mutations. In the last few months, researchers have begun to conclude that a new approach is needed, one based on decoding the entire genome of patients.

The new reports, though involving only single-gene diseases, suggest that the whole-genome approach can be developed into a way of exploring the roots of the common multigene diseases.

“We need a way of assessing rare variants better than the genomewide association studies can do, and whole-genome sequencing is the only way to do that,” Dr. Lupski said.

With 10 genomes of healthy humans sequenced, Dr. Gibbs, a specialist in DNA sequencing, decided it was time to decode the genome of someone with a genetic disease and asked his colleague Dr. Lupski to volunteer. Mutations in any of 39 genes can cause Charcot-Marie-Tooth, a disease that impairs nerves to the hands and feet and causes muscle weakness.

Fifty thousand dollars later, Dr. Lupski turned out to have mutations in an obscure gene called SH3TC2. The copy of the gene he inherited from his father is mutated in one place, and the copy from his mother in a second.

Both his parents had one good copy of the gene in addition to the mutated one. A single good copy can generate enough, or nearly enough, of the gene’s product for the nerves to work properly. Dr. Lupski’s mother was free of the disease and his father had only mild symptoms.

In the genetic lottery that is human procreation, two of their eight children inherited good copies of SH3TC2 from each parent; two inherited the mother’s mutation but the father’s good copy and are free of the disease; and four siblings including Dr. Lupski inherited mutated copies from both parents. These four all have Charcot-Marie-Tooth disease. The results are reported in The New England Journal of Medicine.

In Seattle, Dr. Hood and Dr. Galas have also applied whole-genome sequencing to disease. They analyzed the genome of a family of four, in which the two children each have two single-gene diseases, called Miller syndrome and ciliary dyskinesia. With four related genomes available, the researchers could identify the causative genes. They also improved the accuracy of the sequencing because DNA changes that did not obey Mendel’s rules of inheritance could be classed as errors in the decoding process.

The Seattle team believes whole-genome sequencing can be applied to the study of the common multigene diseases and plans to sequence more than 100 genomes next year, starting with multigenerational families.

The family whose genomes they report in Science were sequenced by a company with a new DNA sequencing method, Complete Genomics of Mountain View, Calif., at a cost of $25,000 each. Clifford Reid, the chief executive, said that the company was scaling up to sequence 500 genomes a month and that for large projects the price per genome would soon drop below $10,000. “We are on our way to the $5,000 genome,” he said.

Dr. Reid said the HapMap and genomewide association studies were not a mistake but “the best we could do at the time.” But they have not yet revolutionized medicine, “which we are on the verge of doing,” he said.

Dr. Goldstein, of Duke University, said the whole-genome sequencing approach that was now possible should allow rapid progress. “I think we are finally headed where we have long wanted to go,” he said.

Answer the following questions: (20 marks)

a) Describe what Professor Lupski found in his genome.

b) What questions do you still have about his genetic disease?

c) What information in the Times article would you use to find the original publication of the scientific findings about Professor Lupski?

d). How many people were involved in the publication of Professor Lupski’s genome?