

BIOINFORMATICS ASSIGNMENT 1 (Day 6 - 10)

Rasitha Arafa R

rasithaarafarafiudeen@gmail.com

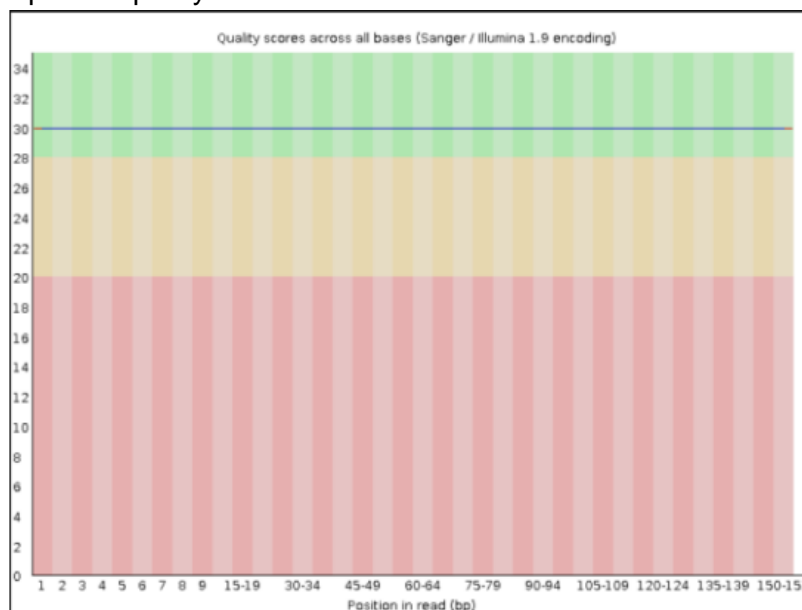
NGS DATA QUALITY CHECK (DAY 6)

1. SRA accession number: **SRR24518778**
2. NGS platform: **ILLUMINA sequencing**
3. Layout: **Pair-end sequencing**
4. Basic statistics:

Measure	Value
Filename	SRR24518778_fastq_gz.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	3516206
Sequences flagged as poor quality	0
Sequence length	35-151
%GC	38

- This basic statistic report gives the overall outline of the sequencing strategy and inferences from the input.
- The given input data was sequenced through illumine sequencing which has a total sequence of 3516206 but the average sequence length is referred as 35-151 with a GC content of 38%

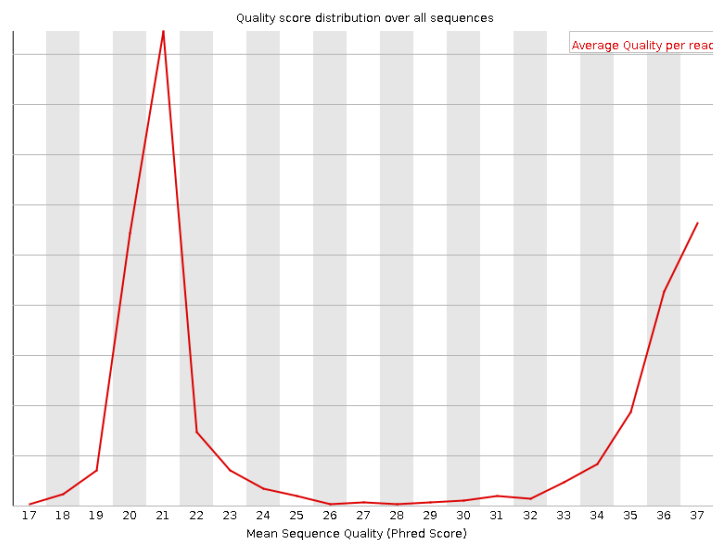
5. Per Base sequence quality:



- This defines the quality of each base read.
- The mean quality, represented by the blue line
- The y-axis shows the quality scores.

- The x-axis is the base position.
- This sample contains reads up to 151 bp long.
- The higher the score, the better the base call.
- The background of the graph divides the y-axis into very good quality scores (green), scores of reasonable qualities (orange), and reads of poor quality (red).
- In this all the reads are come out in the green region showing greater quality.

6. Per sequence quality score:




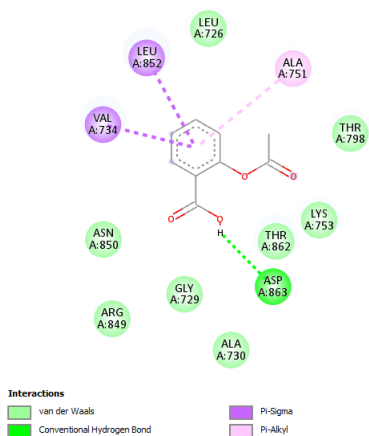

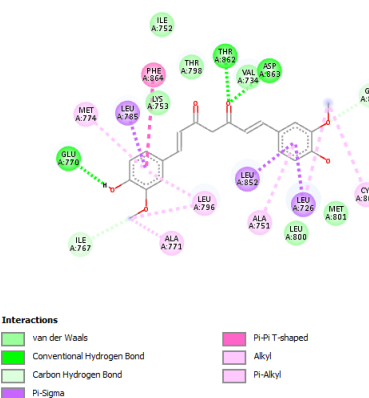

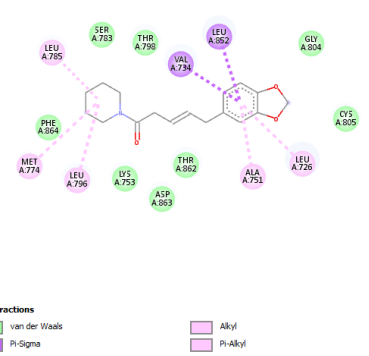
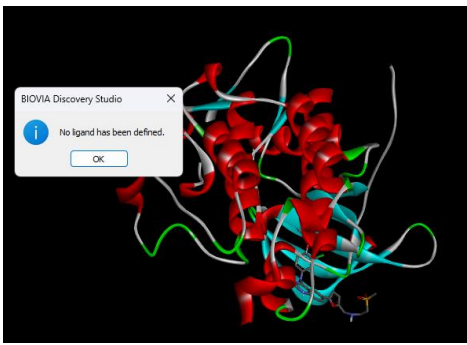
- This defines the quality of the sequence reads.
- The phred score is calculated on the basis of errors per 1000 sequence reads.
- It is fine to have 1 error with 1000 base reads to be considered as good quality.
- The phred score should be less than 30 always.
- The x-axis is the average quality score over the full length of all reads.
- The y-axis gives the total number of reads with this score.
- The read quality is peak in the upper range of the plot.
- Here the graph shows higher peak at 151 range on y axis around phred score of 21, which signifies good overall quality but in the range of 70 – 100 there is a phred score increase but that is not significant.

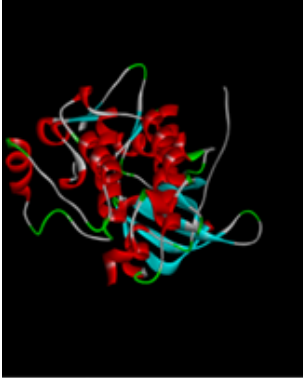
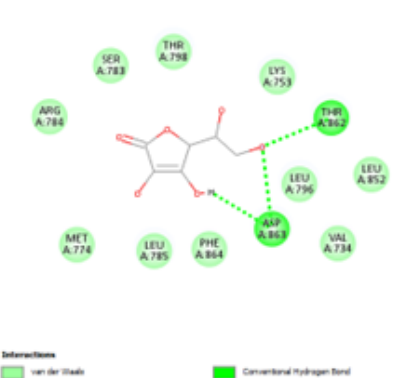
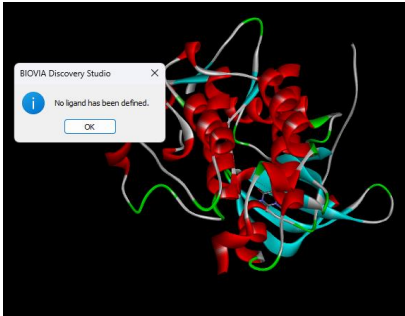

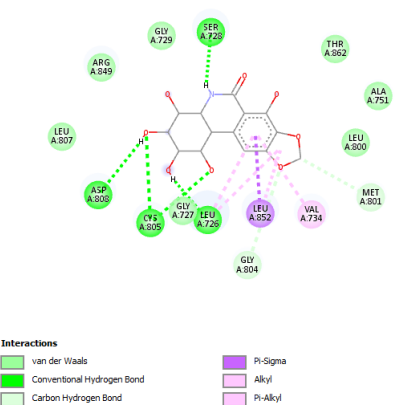

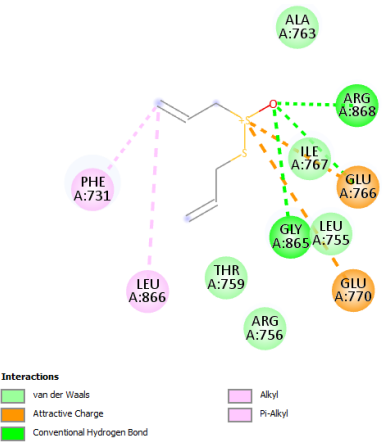
GitHub (DAY 7)


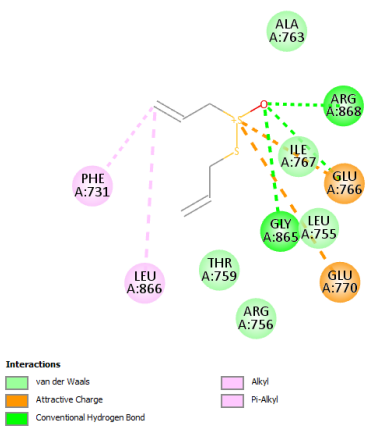
7. GitHub account link - <https://github.com/RasithaArafaRafiudeen>

Molecular Docking (DAY 8 and 9)

8. Protein Name: **Crystal structure of the Kinase domain of Human HER2 (erbB2)**
9. Protein ID – **3PP0**
10. PDB – DOI: <https://doi.org/10.2210/pdb3PP0/pdb>

Ligand Name	Ligand ID	Energy value	Dock Image - 2D
<i>Aspirin</i>	2244	-6.7	  <p>Interactions</p> <ul style="list-style-type: none"> van der Waals Conventional Hydrogen Bond Pi-Sigma Pi-Alkyl
<i>Curcumin</i>	638024	-9.6	  <p>Interactions</p> <ul style="list-style-type: none"> van der Waals Conventional Hydrogen Bond Carbon Hydrogen Bond Pi-Sigma Pi-Pi T-shaped Alkyl Pi-Alkyl
<i>Piperine</i>	969516	-10	  <p>Interactions</p> <ul style="list-style-type: none"> van der Waals Pi-Sigma Alkyl Pi-Alkyl
<i>Lapatinib</i>	969516	-8.2	<p>No ligand found</p> 

Ascorbic acid	54670067	-5.8	 
Caffeine	2519	-6.1	<p>No ligand found</p> 
Pancratistatin	441597	-7.5	 
Allicin	65036	-4.1	 

Ajoene	5386591	-4.1	 
---------------	----------------	-------------	--

Cancer therapy (DAY 10)

Cancer type	Hallmarks	Drug	Mechanism of drug
Mesothelioma	Cancer affecting the lining of the lungs, abdomen, or heart	Pemetrexed and Cisplatin	Pemetrexed inhibits multiple enzymes involved in DNA and RNA synthesis, disrupting cancer cell growth. Cisplatin cross-links DNA and induces apoptosis
Cholangiocarcinoma (Bile Duct Cancer)	Uncontrolled cell growth in the bile ducts	Gemcitabine and Cisplatin	Gemcitabine inhibits DNA synthesis, preventing cell division, while cisplatin cross-links DNA, leading to DNA damage and cell death
Breast Cancer	Uncontrolled growth of breast cells	Tamoxifen	It is a selective estrogen receptor modulator (SERM) that blocks the estrogen receptors in breast cells which inhibits the growth-promoting effects of estrogen and helps prevent estrogen-driven breast cancer growth

Colorectal Cancer	Uncontrolled growth of cells in the colon or rectum	Oxaliplatin, Fluorouracil (5-FU), and Leucovorin	Oxaliplatin is a platinum-based chemotherapy drug that induces DNA damage, preventing cell division. Fluorouracil (5-FU) is an antimetabolite that interferes with DNA and RNA synthesis, while Leucovorin enhances the efficacy of 5-FU
GIST (Gastrointestinal Stromal Tumors)	Abnormal growth of specialized cells in the gastrointestinal tract	Imatinib	It is a tyrosine kinase inhibitor that targets mutated forms of the KIT protein, which is often present in GIST. By inhibiting KIT, imatinib prevents cell growth and promotes apoptosis
Merkel Cell Carcinoma	Aggressive skin cancer originating from Merkel cells	Avelumab.	It is a monoclonal antibody that targets and inhibits the programmed cell death-ligand 1 (PD-L1) on cancer cells. By blocking PD-L1, avelumab enhances the immune system's ability to recognize and eliminate cancer cells
Prostate Cancer	Cancer originating from the prostate gland	Enzalutamide	Enzalutamide is an androgen receptor antagonist. It blocks the binding of androgens (such as testosterone) to the androgen receptor, inhibiting the signalling pathways that promote prostate cancer cell growth