**UNIVERSITY OF SAARLAND**

**Structural Bioinformatics Assignment 4**

Prof. Dr. Olga Kalinina

**THEORETICAL TASK’S ANSWERS**

1. A mutation is an alteration in the genome of a cell in a live species or a virus that is more or less irreversible and can be passed on to offspring cells or viruses. Mistakes in DNA or viral replication, mitosis, or meiosis, or other types of DNA damage, lead to inaccurate results during other forms of repair. Mutations can also occur as a consequence of transposons, which are inserting or deleting portions of DNA ("Genetic Mutation | Learn Science at Scitable", 2022). A modification in the DNA of a multicellular organism's body cell (somatic mutation) can be passed down to descendant cells by DNA replication, resulting in a region or patch of cells with defective behaviors, such as cancer. Germinal abnormalities in egg or sperm cells might result in an individual progeny with the mutation in each of their cells, which can cause major problems, like in the case of a human genetic disorder like cystic fibrosis (Genetics & Health, 2022). Types of mutations can be classified in four fundamental groups: Single-nucleotide variants (SNV) which includes subgroups; coding, non-coding region mutations, synonymous (silent), non-synonymous (missense) mutations and protein-truncating variants (nonsense) mutations (Genetics & Health, 2022). Short insertions and deletions, genome duplications, structural variants and large-scale chromosomal rearrangements can be identified in major types of mutations.
2. In order to explain this concept, focusing on bacterial resistance might offer more understandable look. Bacteria proliferate quickly and can expand in vast numbers. One cell of bacteria splits into two when it replicates. The bacterium must generate two identical copies of its genomic DNA (Woodford & Ellington, 2007).  There is indeed a potential that abnormalities, or mutations, will arise each time the bacterium goes through this procedure. These mutations occur randomly and can appear anywhere in the DNA sequence. Bacteria can develop antibiotic resistance through one of these random mutations. Some spontaneous mutations (or genes obtained from several other bacteria via horizontal gene transfer) may render the bacteria to become resistant to antibiotic ("Mutations and selection – Antibiotic resistance – ReAct", 2022). Antibiotic resistance is frequently thought to be mediated by mechanisms of horizontal gene transmission among bacterial strains or species. Only the resistant bacteria will be capable of spreading if we treat the bacterial population with that specific antibiotic; the antibiotic chooses for them (Woodford & Ellington, 2007). These bacteria can now multiply, resulting in a population of primarily resistant microorganisms.
3. It is the fact that, for two protein structure to being matched or similar to each other, they should have these two properties: Correspondence (it must be one-to-one correspondence) and alignment (there is a rigid body transform T such that the RMSD between elements in A and T(B) is smaller than a certain threshold ε (web.stanford.uni). These properties and structural similarity. In actuality, a perfect match among two proteins is unlikely; in many circumstances of interest, two proteins may be just locally similar, have different sizes, or differ physically in other ways despite requiring serious structural similarities in other aspects. In these instances, a perfect match between the two proteins is clearly unrealistic. Moreover, for measuring this issue is conducted by some scoring functions. First and major methodology for that is RMSD values. The average distance between the atoms (typically the backbone atoms) of superimposed proteins is measured by the root-mean-square deviation of atomic locations (Carugo, 2003). It's worth noting that RMSD calculations can be used on molecules that aren't proteins, such as tiny chemical compounds. The RMSD of the C atomic coordinates after optimal rigid body superposition is commonly used to quantify the similarity in three-dimensional structure in the study of globular protein conformations (Carugo, 2003). Even it has drawbacks like not considering aligning length, when it used with combination of alignment methods, crucial results can be generated. Another methodology is GDT-TS scoring. The Global Distance Test - Total Score is a method for determining how close a predicted protein structure is to a reference structure, which is usually an empirical model (foldit.fandom). The sequences of the two structures don't have to be identical. In addition, MaxSub is a brand-new, designed method that relies on and extends parts of the CASP3 evaluation methods. The goal of MaxSub is to find the biggest subset of C(alpha) atoms in a model that superimpose 'good' over the experimental structure, and to generate a single normalized score that measures the model's quality. This measurement also used when comparing two structures (Siew, Elofsson, Rychlewski & Fischer, 2000). Lastly, The TM-score tool compares two models based on their equivalence of residues (i.e., based on the residue index in the PDB file). When comparing two proteins with distinct sequences, it is usually not used. The TM-align tool is a structural alignment program for analyzing proteins with differing sequences (Depner, OxfordGroup).

REFERENCES

Carugo, O. (2003). How root-mean-square distance (r.m.s.d.) values depend on the resolution of protein structures that are compared. Journal Of Applied Crystallography, 36(1), 125-128. doi: 10.1107/s0021889802020502

Depner, L. TM-score | Oxford Protein Informatics Group. Retrieved 4 January 2022, from https://www.blopig.com/blog/2017/01/tm-score/

GDT. (2022). Retrieved 4 January 2022, from https://foldit.fandom.com/wiki/GDT

Genetics, H., & Health, V. (2022). What kinds of gene variants are possible?: MedlinePlus Genetics. Retrieved 4 January 2022, from https://medlineplus.gov/genetics/understanding/mutationsanddisorders/possiblemutations/

Genetic Mutation | Learn Science at Scitable. (2022). Retrieved 4 January 2022, from https://www.nature.com/scitable/topicpage/genetic-mutation-441/

Huang, Y., Mao, B., Aramini, J., & Montelione, G. (2022). Assessment of template-based protein structure predictions in CASP10. Retrieved 4 January 2022, from

Mutations and selection – Antibiotic resistance – ReAct. (2022). Retrieved 4 January 2022, from https://www.reactgroup.org/toolbox/understand/antibiotic-resistance/mutation-and-selection/

(2022). Retrieved 4 January 2022, from https://web.stanford.edu/class/cs273/scribing/scribe8.pdf

Woodford, N., & Ellington, M. (2007). The emergence of antibiotic resistance by mutation. Clinical Microbiology And Infection, 13(1), 5-18. doi: 10.1111/j.1469-0691.2006.01492.x