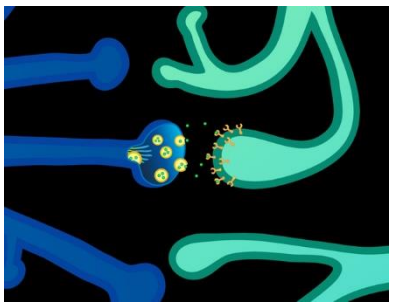
Mohammad Hadi Azarabad – Jan 2024 – Bioinformatic Project

Human Genetic Disorder[i]

A genetic disorder is a disease caused in whole or in part by a change in the DNA sequence away from the normal sequence. Genetic disorders can be caused by a mutation in one gene (monogenic disorder), by mutations in multiple genes (multifactorial inheritance disorder), by a combination of gene mutations and environmental factors, or by damage to chromosomes (changes in the number or structure of entire chromosomes, the structures that carry genes).

Parkinson's Disease[ii]

[iii]Parkinson's disease (PD) is a neurological condition characterized by tremor and stiffness in movement, predominantly affecting individuals over the age of 60. While most cases are sporadic, some families exhibit a genetic predisposition, especially when three or more family members are affected, particularly at an early age.



Our genetic material, organized into genes, dictates various aspects of our traits and susceptibility to conditions. Genes come in pairs, inherited from each parent, and variations in genes can influence our likelihood of developing certain conditions. Genetic factors play a role in Parkinson's disease, with identified genes such as **SNCA**, UCHL1, **LRRK2**, PARK2, PARK7, and PINK1[iv] associated with either dominant or recessive inheritance patterns.

Role Of SNCA[v]

The primary pathological features involve neuronal loss and the accumulation of Lewy bodies (LB) in surviving neurons, with alpha-synuclein (α-syn) being a key component of LB.

Mutations in the alpha-synuclein gene (SNCA) were the first identified genetic cause of PD. Patients with SNCA mutations often exhibit early-onset parkinsonism with severe non-motor symptoms, including cognitive decline. Additionally, various SNCA polymorphisms have been associated with non-motor manifestations, although the functional role of these polymorphisms is not fully understood.

Genomes for SNCA

I selected 4 genomes to create a phylogenetic tree for SNCA of Humans.

1. NC_000004.12 of humans which is on Chr. 4 which SNCA is located on
2. NC_051339.1 of rodents because ortholog gene of SNCA in rodents is on this Chr.
3. NC_052535.1 of birds because ortholog gene of SNCA in birds is on this Chr.
4. NC_041758.1 of primates because ortholog gene of SNCA in primates is on this Chr.

SNCA orthologs[vi]

Role Of LRRK2[vii]

Dysfunction of LRRK2 can influence the accumulation of α-synuclein, leading to alterations in cellular functions and signaling pathways. The accumulation of α-synuclein stimulates microglial activation, contributing to neuroinflammation and neuronal death in PD.

Genomes for LRRK2

I selected 3 genomes to create a phylogenetic tree for LRRK2 of Humans.

1. NC_000012.12 of humans which is on Chr. 12 which LRRK2 is located on
2. NC_007136.7 of fishes because ortholog gene of LRRK2 in fishes is on this Chr.
3. NW_010224851.1 of birds because ortholog gene of LRRK2 in birds is on this Chr.

LRRK2 orthologs[viii]

Why SNCA and LRRK2

The Leucine-rich repeat kinase 2 (LRRK2) gene and α-synuclein gene (SNCA) play crucial roles in Parkinson's disease (PD). Understanding the complex relationship among LRRK2, α-synuclein, and microglia could pave the way for targeted clinical therapies for PD. Regarding the relation between SNCA and LRRK2, I want to analyze and trace them across different species with the phylogenetic tree separately.

Procedure

1. Ortholog Class

   This class is responsible to retrieve and store data from ortholog csv file and also fetch sequence of genes from NCBI or load from local if there is any.

2. Loading Orthologs for each gene (SNCA and LRRK2)
3. Creating multiple alignment of proteins for phylogenetic tree
   a. Locating the best possible location on genome for aligning the gene

      I used local align with match score 1 and gap score -10 to demonstrate gaps and also gene window 100 (because genes are about 150 AA and window of 100 would be a good proportion for alignment) and also genome window 1000.

   b. Translating the DNA to Protein
4. Performing the multi sequence alignment

   I performed the multi sequence alignment using ClustalW.

5. Creating Phylogenetic Tree and visualizing

   In order to create phylogenetic tree, I used PhyML V3.1 and I round all branch lengths to 2 decimals and random color for clades.

Mohammad Hadi Azarabad – Jan 2024 – Bioinformatic Project

LRRK2 Result

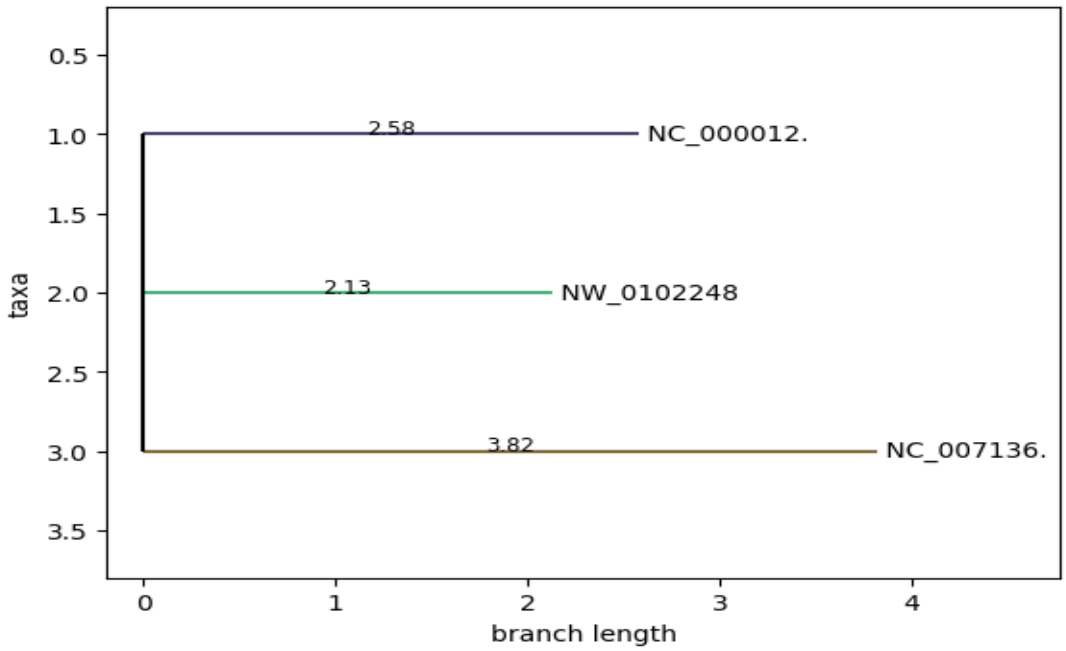


Figure 1 LRRK2 Result

- From Phylogenetic tree of LRRK2 we can see LRRK2 gene from human has more similarity to NW_0102248 which is for bird's genome not human and That might be due to window sizing or Gap score. Also, there is a lead that I can find more informative phylogenetic tree if I use ortholog of birds instead of using humans in order to find ancestors with lower distance clades and better clades.
- The Common ancestor for all of them is not known and I assume that by adding more genomes, I can find common ancestor or some ancestor for some of the genomes.
- From this phylogenetic tree, it seems Human and bird are more likely to have common ancestor.
- For researches, studying on birds seem to be a good alternative instead of humans regarding to its similarity.

SNCA Result

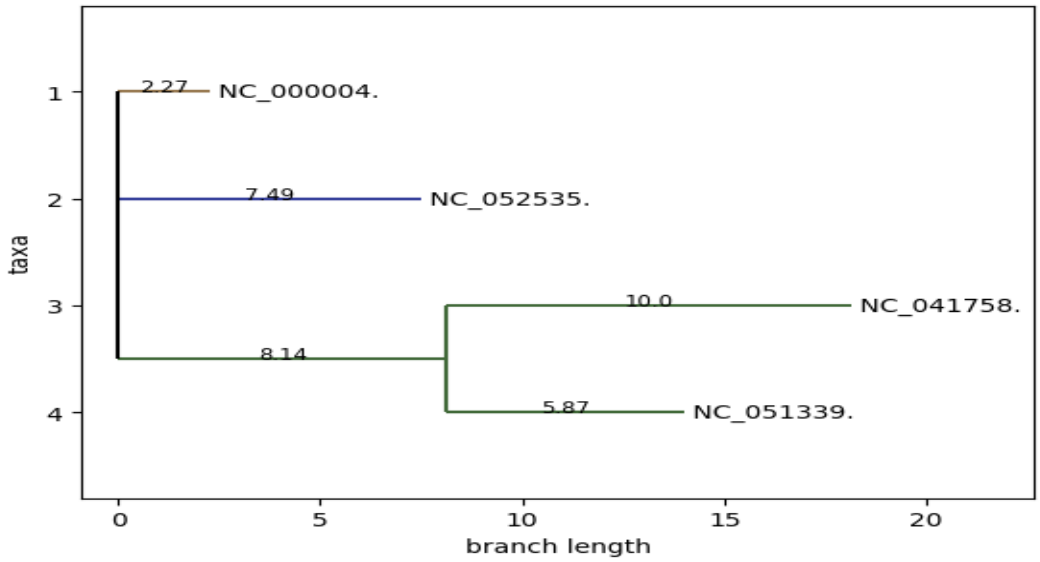


Figure 2 SNCA Result

- This tree is showing that Rodents and Primates have common ancestor for SNCA while Humans and birds have theirs.
- SNCA gene from human is more similar to human genome, and birds have the most similarity at the second place.
- From this Tree and their distances, we can assume we need more genome in order to say more confident about whole evolutionary line but it for now we can say rodents and primates had more evolutionary background.
- Also, this tree notes that, Humans are far from primates while it's expected to see more similarity between primates and humans

Conclusion of comparing both Trees

- In general, birds have the most similarity to human genomes, Therefor, they can be a good alternative for studying human related genes.
- Both trees show birds have not common evolutionary background with others.
- Humans are not sharing any ancestor with others.
- If we have a gene of human, it's not mandatory to have the closest similarity with human genome and that can be a good point for researches on human related diseases.

References

[i] https://www.genome.gov/For-Patients-and-Families/Genetic-Disorders
[ii] https://www.genome.gov/Genetic-Disorders/Parkinsons-Disease
[iii] https://medlineplus.gov/genetics/condition/parkinsons-disease/#causes
[iv] https://www.genome.gov/Genetic-Disorders/Parkinsons-Disease
[v] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8401994/
[vi] https://www.ncbi.nlm.nih.gov/gene/6622/ortholog/?scope=7776
[viivii] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6251048/
[viii] https://www.ncbi.nlm.nih.gov/gene/120892/ortholog/?scope=7742