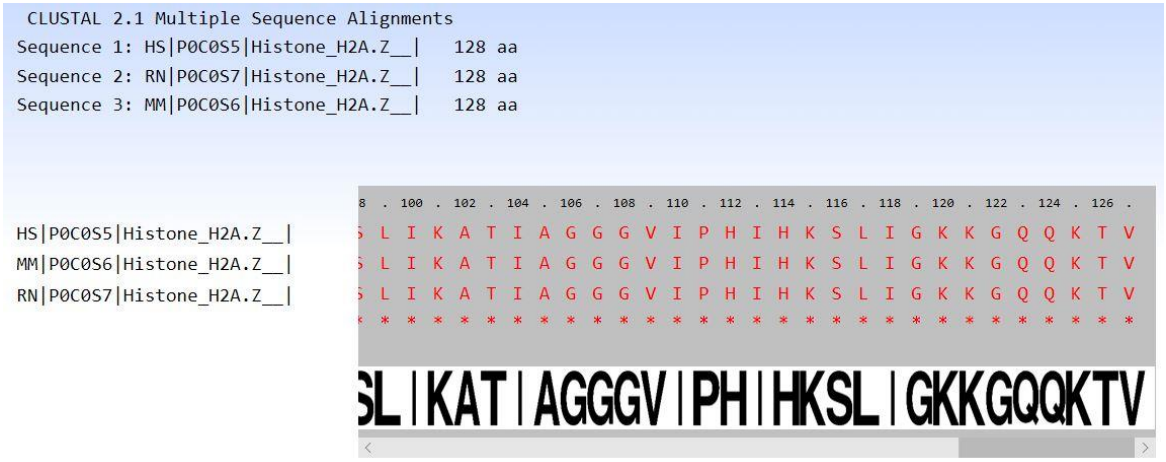
Multiple Sequence Alignment to study histone proteins within or across different species:

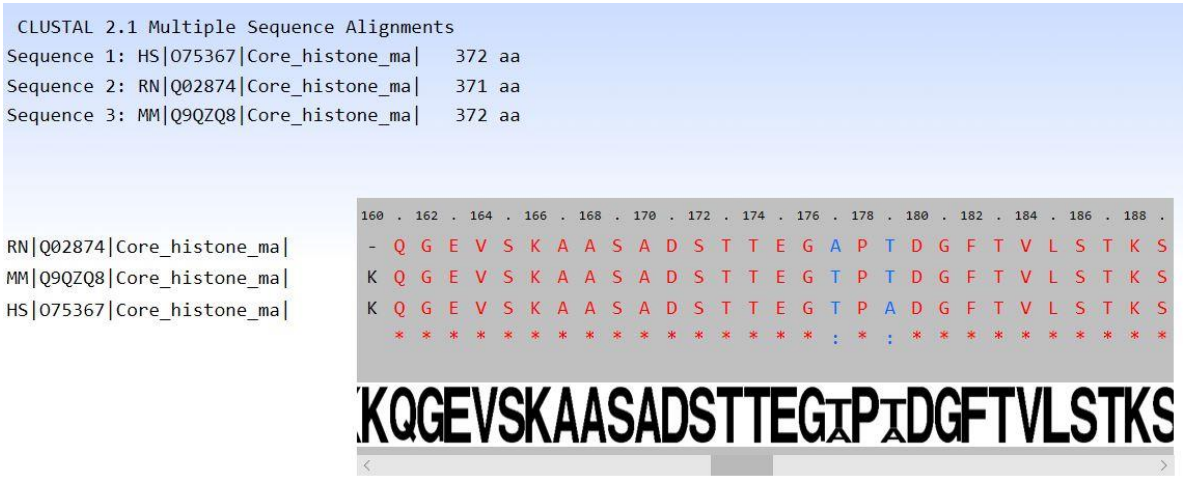
Multiple Sequence Alignment helps in aligning different proteins based on sequence similarities. Histone isoforms within species and across species are quite similar. They differ by few amino acids (1 to 3) within species. Therefore, MSA will provide information about the conservation of protein within and across species. Also, based on the algorithm, one can predict the phylogenetic distance between the species using a given histone protein sequence.

Example of MSA Analysis:

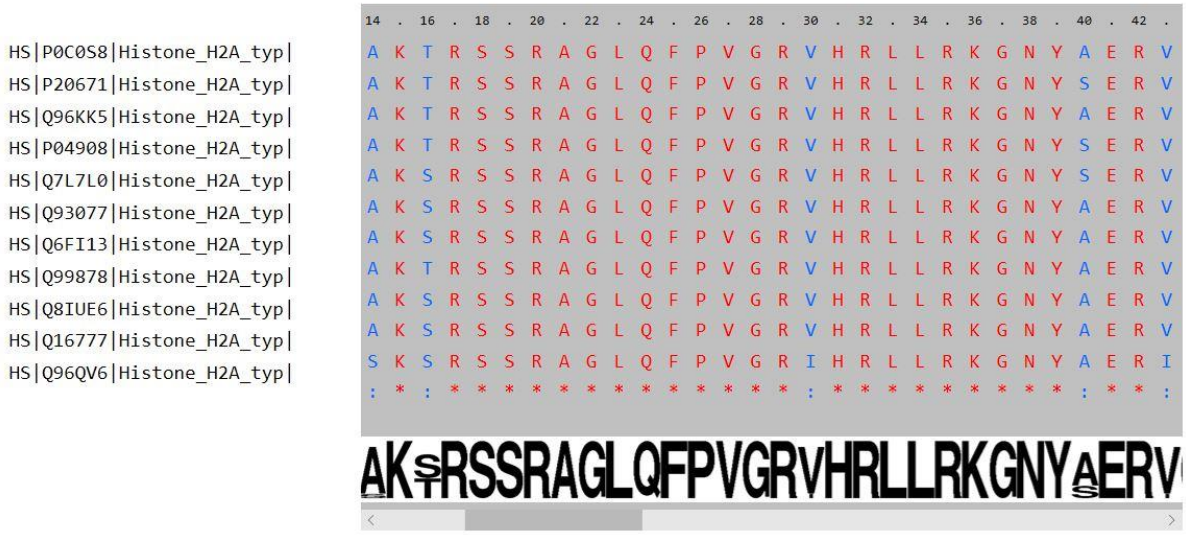
1. Histone variant H2A.Z: H2A.Z has a cell cycle independent expression and play a role in transcription initiation. Many reports have shown that H2A.Z is involved in epithelial to mesenchymal transition process. Considering its importance, one can be interested in looking at the conservation across species. MSA cofirms that H2A.Z is 100% identical in Human, Rat, and Mouse and suggest that the protein may have a similar role in all species.



2. Histone Variant, macroH2A.1: Involved in stable X chromosome inactivation in female. Also, important for transcription repression and shown to be upregulated in multiple cancers. However, the sequence similarity analysis of protein in different species have shown a difference at 8 positions. This suggests that the protein may be providing unique stability to nucleosome in different species.



3. Histone H2A isoforms within the same organism:

Histone isoform genes are present in clusters on the genome. Isoform proteins are incorporated in the nucleosome during DNA replication. MSA analysis of different human H2A isoforms shows that these proteins mostly differ at amino acid position numbers 16, 40, 51 and 99. Based on variation in protein sequences, they may be interacting with H2B protein with different affinities and thereby altering the nucleosome stability. Recent studies have also shown non-redundant cellular functions of highly conserved histone isoforms (Singh et al. 2015; Bhattacharya et al. 2017).

```
 CLUSTAL 2.1 Multiple Sequence Alignments
Sequence 1: HS|P0C0S8|Histone_H2A_typ|   130 aa
Sequence 2: HS|Q96QV6|Histone_H2A_typ|   131 aa
Sequence 3: HS|P04908|Histone_H2A_typ|   130 aa
Sequence 4: HS|Q93077|Histone_H2A_typ|   130 aa
Sequence 5: HS|P20671|Histone_H2A_typ|   130 aa
Sequence 6: HS|Q96KK5|Histone_H2A_typ|   128 aa
Sequence 7: HS|Q99878|Histone_H2A_typ|   128 aa
Sequence 8: HS|Q6FI13|Histone_H2A_typ|   130 aa
Sequence 9: HS|Q8IUE6|Histone_H2A_typ|   130 aa
Sequence 10: HS|Q16777|Histone_H2A_typ|   129 aa
Sequence 11: HS|Q7L7L0|Histone_H2A_typ|   130 aa
```