Innovative Approaches for Comparative Genomics Study of Rattus Rattus and Rattus Norvegicus via Computational Biology: Statistical & Structural Functions

Agrawal SP*

Department of Zoology, D.A-V College Kanpur, India

*Corresponding author: Shashi Prabha Agrawal, Assistant Professor from Department of Zoology, D.A-V College Kanpur, India, Email: shashiprabhabioinfo@gmail.com

Abstract

Comparative genomic in large scale is integrated approach that compares species from their genetic structures to identify similarities and differences their genomes and their functional structures of the genes via this approach we identified the potential functionally of the genomes of the different species that's scientific approach to identified the differential structures prediction of two different species on behalf of computation biology. Basic objective of this study is to identify the genomic diversity between the two species (*Rattus rattus* and *Rattus norvegicus*) on behalf of statistical & functional analysis via computational biology. We also briefly used the currently available computational tools in comparative genomics such as algorithms for genome-scale sequence alignment, gene identification, and non-homology-based function prediction. Recently, comparative genomic sequence analysis has been used to help in the functional assignment of genes in a non-similarity-based manner. These comparative genomic approaches depend on the basic factors of the genes that are functionally related genes and closely associated across genomes in some form.

Keywords: *Rattus Rattus; Rattus Norvegicus*; Genomics; Computational biology

Introduction

Last few decades our generation running in an advance biological research in each and every field related to life sciences [1]. Comparative genomics is a large-scale, holistic approach that compares two or more genomes to discover the similarities and differences between the genomes and to study the biology of the individual genomes. Comparative studies can be performed at different levels of the genomes to obtain multiple perspectives about the organisms. Analysis of the global structure of genomes, such as nucleotide composition, syntenic relationships and gene ordering offer insight into the similarities and differences between genomes. Such comparisons provide information on the organization and evolution of the genomes, and highlight the unique features of individual genomes [2].

Description about *Rattus Rattus*

A native of the Indian sub-continent, R. rattus has now spread throughout the world. It is widespread in forest and woodlands as well as being able to live in and around buildings. It will feed on and damage almost any edible thing. The black rat is most frequently identified with catastrophic declines of birds on islands. It can also be very damaging to crops and plants, particularly seeds. It is very agile and often frequents tree tops searching for food and nesting there in cavities, or less frequently in bunches of leaves and twigs. A slender rat with large hairless ears, R. rattus may be grey-brown on the back with either a similarly-coloured or creamish-white belly, or it may be black all over. The uniformly-coloured tail is always longer than the head and body length combined. Its body weight is usually between 120 and 160 g but it

Innovative Approaches for Comparative Genomics Study of *Rattus Rattus* and *Rattus Norvegicus* via Computational Biology: Statistical & Structural Functions

Int J Zoo Animal Biol

can exceed 200 g. R. rattus is able to utilise most terrestrial habitat types on continents and islands. They generally avoid swimming, although they are successful swimmers and can swim distances of 300-750 m to colonise islands R. rattus can live on the ground, aboveground, and belowground. It does not require forest or substantial vertical structure, and can obtain high abundances in savannahs and on atolls with low scrub vegetation there are two forms of R. rattus that differ in chromosome number. The more widespread oceanic form has 38 chromosomes whereas the Asian form has 42 chromosomes R. rattus is an omnivorous generalist, yet can be a very selective feeder. They eat both plant and animal matter all year round. A Japanese study showed that R. rattus is primarily herbivorous, but can change its food habits when it is thirsty, or when food is in short supply In a recent review found that plant material was the dominant food in 17 out of 20 diet studies. R. rattus is a placental mammal with dependent young. Litter size averages 3-6.5 individuals, yet in laboratory trials the litter size ranges from three to 10, and averages 5-8, with frequency of litter's dependent on season and food supply. The interval between litters may be as little as 27 days. The primary economic impact of R. rattus relates to agricultural and horticultural damage. It is capable of destroying up to 30% of crops annually. Of the 60 or more species in the genus Rattus, R. rattus is likely to be the most damaging to agricultural crops globally. In addition to consuming and spoiling foods, it can cause fires and electrical interruptions by gnawing wiring in buildings [3].

Description about *Rattus Norvegicus*

R. norvegicus is globally widespread and costs primary industry hundreds of millions of dollars per year. It has caused or contributed to the extinction or range reduction of native mammals, birds, reptiles and invertebrates through predation and competition. It restricts the regeneration of many plant species by eating seeds and seedlings, eats food crops and spoils human food stores by urinating and defecating in them. Additional economic damage is caused by chewing through power cables and spreading diseases. R. norvegicus has brown fur on the back with pale grey fur on its belly. The adults normally weigh 150 - 300g, and may reach up to 500g, and are up to 390mm long. They have relatively small ears - which usually do not cover the eyes when pulled forward. The tail is shorter than the head-body length - the opposite is true for the black rat R. rattus. Females have 12 nipples. R. norvegicus probably arrived in Europe in the middle ages and certainly by the early 1700s, and from there hitch-hiked around the world on boats, leading to accidental introductions to the Americas, Australia and Africa, as well as to island groups. It is currently found in 36% of the world's island groups. Once an island group is colonised R. norvegicus is able to colonise throughout the archipelago rapidly through further hitch-hiking and natural dispersal,

including swimming. It is capable of swimming 2.5km. R. norvegicus is the primary rat used in laboratory research, but the inbred albino laboratory form is unlikely to pose any significant invasive risk. R. norvegicus is also a common pet, and this form may pose a slightly greater invasive risk, although it is unlikely to be a pet outside its current introduced range. R. norvegicus can be widespread, utilising most habitat types. In its introduced range, on continents (Europe, Americas) and large islands (UK) it is more closely associated with humans (both cities and rural environments, although also free-living in damp environments such as riverbanks) and it is dominant over other introduced rats. This situation is reversed on oceanic islands (such as New Zealand) where it is free-living in forest and wetland habitats but is dominated by black rats. R. orvegicus rarely climbs trees. At high density (such as rural UK farms) the home range of R. norvegicus is less than a hectare but in medium densities (such as forested New Zealand islands) this can increases up to 5 ha or even greater than 10 ha when invading islands. R. norvegicus is considered to be territorial throughout most of the year, but will spread when food is scarce, and migrations have been observed. From the distribution and recorded reinvasions of R. norvegicus it appears that it can cross up to 1km of water comfortably, and up to 2km of open water more rarely when conditions are suitable R. norvegicus is omnivorous and opportunistic and will eat raw or cooked meat and vegetable matter, grains and other seeds and berries as well as roots and a wide variety of vertebrate and invertebrate species, including from the inter-tidal zone. Adults require about 10% of their body weight per day in dry grain, and when on a dry diet they need to drink about 25ml of water per day. R. norvegicus in captivity has been observed to withdraw food to the nest, and sometimes store it there. R. norvegicus on Breaksea Island, New Zealand, has been reported to eat invertebrates (beetles, spiders, wetas and flies), fish, shellfish, vegetation, and birds. A Japanese study showed that R. norvegicus is essentially omnivorous, eating plant matter and animal matter (such as insects) in equal volumes. R. norvegicus has also been known to attack and kill young rabbits [4].

Materials and Methodology

The genome sequences of the species were retrieved from National Centre for Biological Information (NCBI) (www.ncbi.nlm.nihgov). Statistical analysis was done via Oligonucleotide Properties Calculator then function study done via mVISTA and kinefold and other online available tools for computational study [5].

Results and Discussions

Comparative genomics study done on behalf of two modes: Statistical & Functional. Statistical comparative genomics

Agrawal SP. Innovative Approaches for Comparative Genomics Study of *Rattus Rattus* and *Rattus Norvegicus* via Computational Biology: Statistical & Structural Functions. Int J Zoo Animal Biol 2020, 3(5): 000245.

Copyright© Agrawal SP.

study of *Rattus rattus* and *Rattus norvegicus* was done via online nucleotides properties calculator Statistical analysis on behalf of different parameters. Physical constants (Physical analysis of genome sequence), thermodynamic constants (We can assume that the concentration of DNA and the concentration of the DNAprimer complex are equal so this simplifies the equation considerably. If the two strands are in equal concentration, the effective concentration is 0.25 the total concentration of oligonucleotide, melting temperature), Melting Comparative Genomics Study of *Rattus rattus* and *Rattus norvegicus* via temperature calculations are based on the thermodynamic relationship between entropy, enthalpy, free energy and temperature [5] (Table 1).
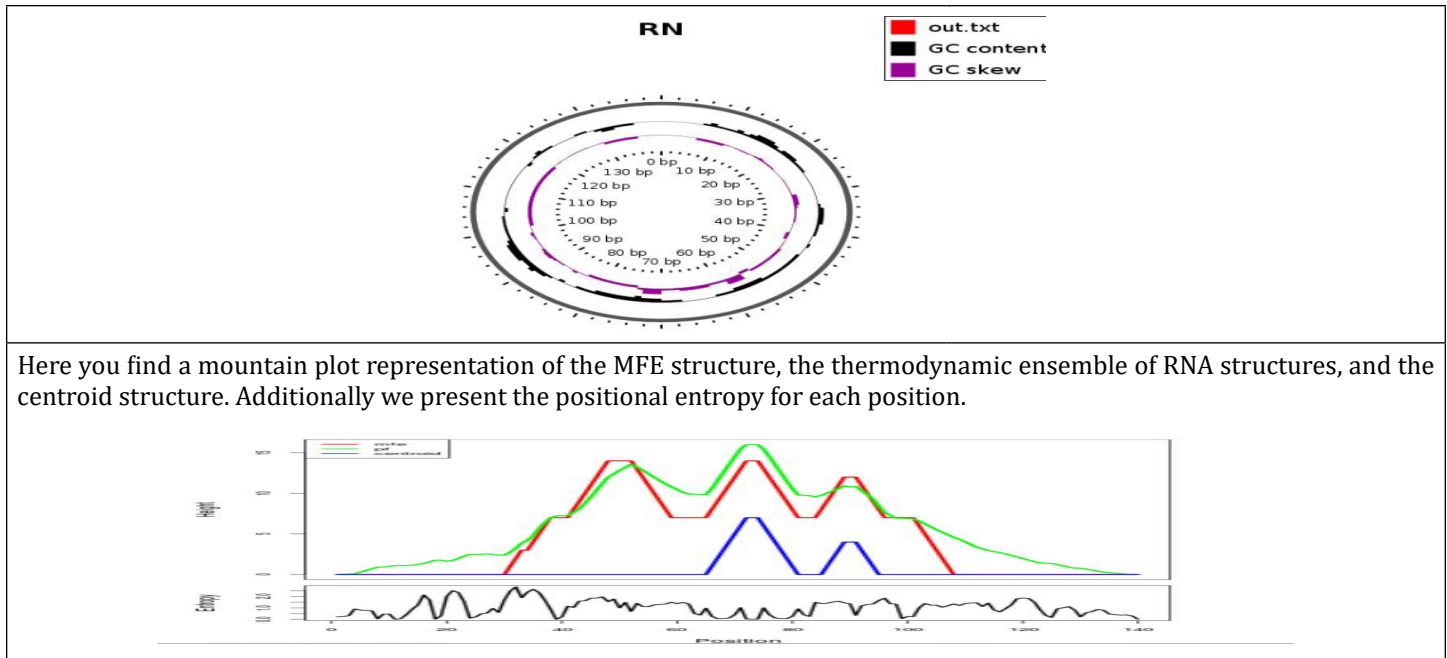
S. No	Annotated name of selected part of genome sequence	Selected part of sequence transcription (DNA to RNA)	Selected part of sequence translation (RNA to Protein)
1	>RNSeq1 (Nucleotide) ACTATCAAACCCTAT-GTCCTGATCAATTCTAGTAGTTCCCAAAATAT-GACTTATATTTTAGTACTTGTAAAAATTTTA-CAAAATCATGTTCCGTGAACCAAAACTCTAAT-CATACTCTATTACGCAATAAACATTAACAA	>RNSeq1 transcription ACUAUCAAACCCUAU-GUCCUGAUCAAUUCUAGUAGUUCCCAAAAU-AUGACUUAUAUUUUAGUACUUGUAAAAAU-UUUACAAAAUCAUGUUCCGUGAACCAAAACUCU-AAUCAUACUCUAUUACGCAAUAAACAUUAACAA	>RN Protein sequence TIKPYV-LINSSSSQNMTYIL-VLVKILQNHVPTK-TLIILYYAININ
2	>RRSeq2(Nucleotide) TTAAAAGGCTTCTC-CATTCTAGTAGACCACAAAATTTTAACT-TATCTTAGCATTGGTAAAATTTCCCGACA-CAAAATCTTTCCTTCTAACTAACCCTCTTTACTT-GCCTACCTCGAAAATTCCACATACACCAAA	>RRSeq 2 transcription UUAAAAGGCUU-CUCCAUUCUAGUAGACCACAAAAU-UUUAACUUAAAUCUUAGCAUUGGUAAAAU-UUCCCGACACAAAAUCUUUCCUUCUAACU-AAACCCUCUUUACUUGCCUACCCUCAGAAAAU-UCCACAUACACCAAA	>RRProtein Seq LKGFSILVDHKILTI-LALVKFPDTKS-FLLTKPSLLAY-PQKIPHTP

Table 1: Genome Sequence.

This table gives detailed information about the genome sequence that are compared of Rattus rattus and Rattus norvegicus in which first coloum gives information about the nucleotide sequence taken from databse, second coloum contains the transcript sequence produce to obtain protein sequence in third coloum.

Results for minimum free energy prediction	
The optimal secondary structure in dot-bracket notation with a minimum free energy	-8.40 kcal/mol
Results for thermodynamic ensemble prediction	
The free energy of the thermodynamic ensemble is	-11.56 kcal/mol
The frequency of the MFE structure in the ensemble is	0.59 %
The ensemble diversity is	29.67
he centroid secondary structure in dot-bracket notation with a minimum free energy	-3.90 kcal/mol
A note on base-pairing probabilities: The structure below is colored by base- pairing probabilities. For unpaired regions the color denotes the probability of being unpaired	
MFE secondary structure	Centroid secondary structure

Agrawal SP. Innovative Approaches for Comparative Genomics Study of *Rattus Rattus* and *Rattus Norvegicus* via Computational Biology: Statistical & Structural Functions. Int J Zoo Animal Biol 2020, 3(5): 000245.

Copyright© Agrawal SP.

Here you find a mountain plot representation of the MFE structure, the thermodynamic ensemble of RNA structures, and the centroid structure. Additionally we present the positional entropy for each position.

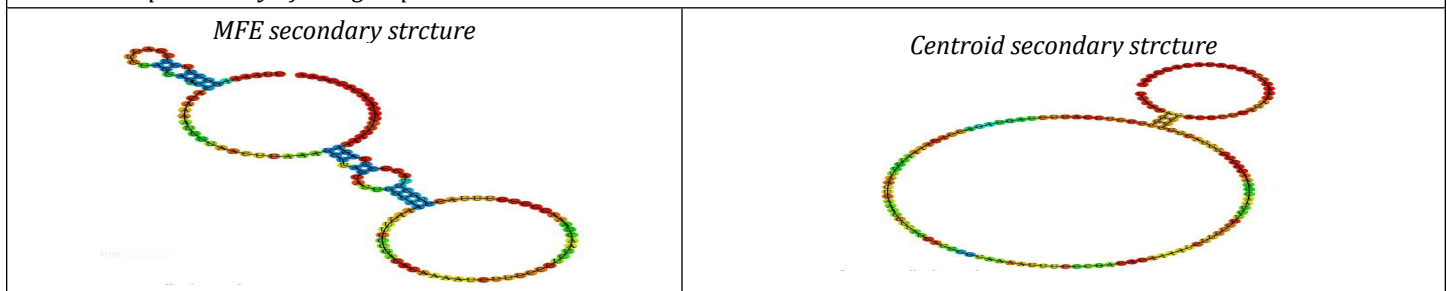**Analysis results for : >RNSeq1 transcription.**

In above analysis of >RNSeq1 transcription hairpin structures of predicted transcription using RNAfold software. a: MFE structure drawing encoding base-pair probabilities. b: a mo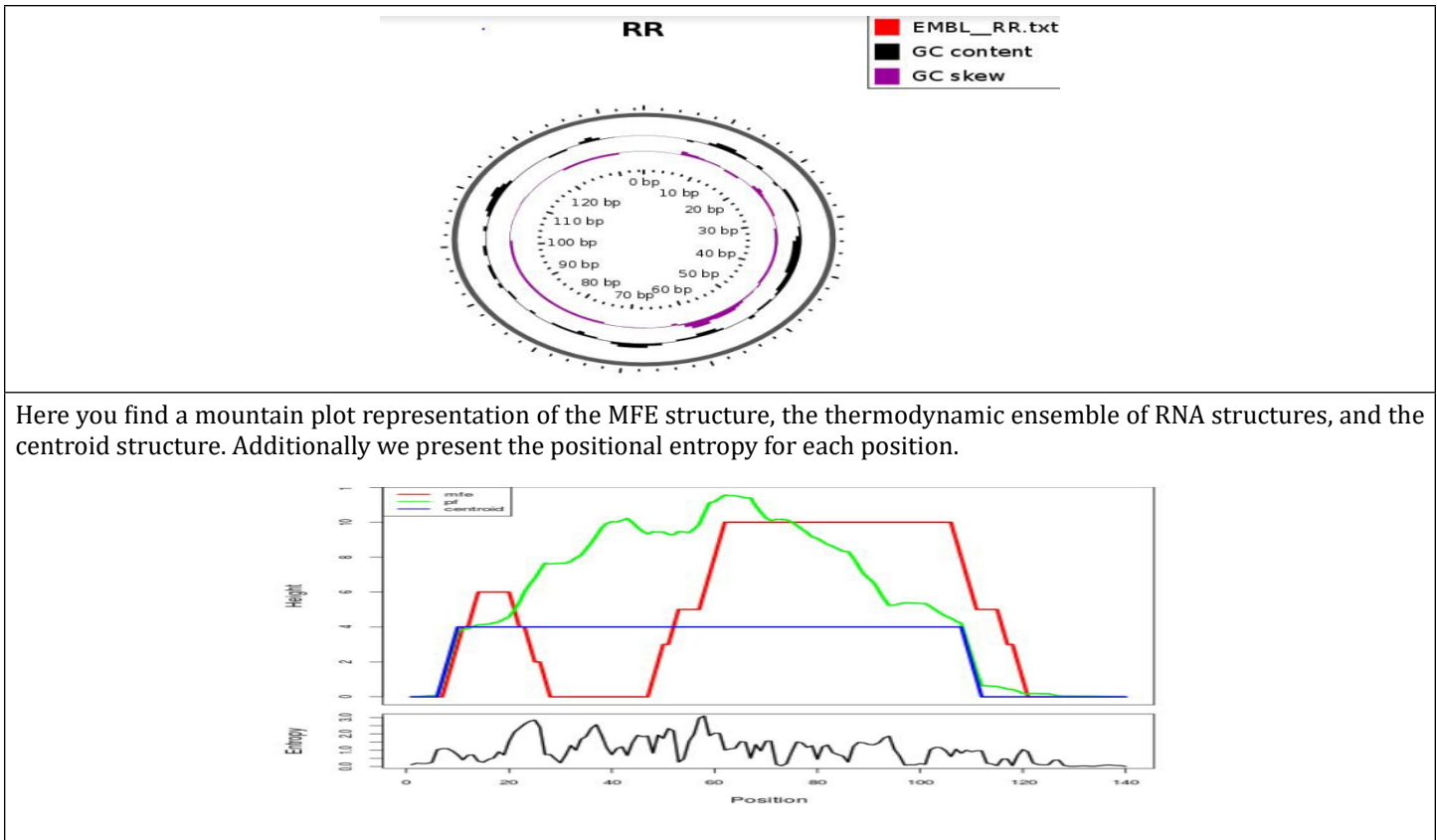untain plot representation of the MFE structure, the thermodynamic ensemble of RNA structures, and the centroid structure, the positional entropy for each position was also presented

Results for minimum free energy prediction	
The optimal secondary structure in dot-bracket notation with a minimum free energy	-4.77 kcal/mo
Results for thermodynamic ensemble prediction	
The free energy of the thermodynamic ensemble is	-7.32 kcal/mol
The frequency of the MFE structure in the ensemble is	1.61 %
The ensemble diversity is	25.29
The centroid secondary structure in dot-bracket notation with a minimum free energy	-2.42 kcal/mol

Graphical output

A note on base-pairing probabilities: The structure below is colored by base-pairing probabilities. For unpaired regions the color denotes the probability of being unpaired.

MFE secondary strcture	Centroid secondary strcture

Agrawal SP. Innovative Approaches for Comparative Genomics Study of *Rattus Rattus* and *Rattus Norvegicus* via Computational Biology: Statistical & Structural Functions. Int J Zoo Animal Biol 2020, 3(5): 000245.

Copyright© Agrawal SP.

Here you find a mountain plot representation of the MFE structure, the thermodynamic ensemble of RNA structures, and the centroid structure. Additionally we present the positional entropy for each position.

Analysis results for : >RRSeq 2 transcription

Conclusion

In above analysis of >RNSeq2 transcription hairpin structures of predicted transcription using RNA fold software. a: MFE structure drawing encoding base-pair probabilities. b: a mountain plot representation of the MFE structure, the thermodynamic ensemble of RNA structures, and the centroid structure, the positional entropy for each position was also presented. In above tables there are mfe secondary structure and centroid secondary structure present of rna transcript of rat so In comparison with the minimum free energy (MFE) structure, the centroid of the ensemble makes 30.0% fewer prediction errors as measured by the positive predictive value with marginally improved sensitivity. The Boltzmann ensemble can be separated into a small number of clusters. The "best cluster centroid" as determined by comparison to the known structure simultaneously improves PPV by 46.5% and sensitivity by 21.7%. These results tells that the energy well containing the MFE structure under the current incomplete energy model is often different from the one for the unavailable complete model that presumably contains the unique native structure [6].

References

1. Alka D (2014) Morphological Studies of Five Known Nematode Species Via Taxa-Informatics Approaches. Science Innovation 2(1): 7-10.

2. Wei L, Liu Y, Dubchak I, Shona J, Park J (2002) Comparative genomics approaches to study organism similarities and differences 35(2): 142-150.

3. (2019) Rattus rattus (black rat). Invasive Species Compendium.

4. (2019) Rattus norvegicus(brown rat). Invasive Species Compendium.

5. Dubey A, Bawari M (2016) Comparative Genomics Study of Rhinoceros unicornis and Diceros bicornis via Bioinformatics Approaches: Statistical & Functional. Advances in Bio and Medico Sciences pp: 76-81.

6. Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL (2008) The Vienna RNA websuite. Nucleic Acids Res 36: 70-74.