ECRONICON OPEN ACCESS

EC MICROBIOLOGY
Research Article

Peptide Properties of *Saccharomyces arboricola* H-6 Suggest Randomness in Chromosomal Organization

Daeun Roh[1,2], Shah Yunn Naing[1,2] and Maurice HT Ling[1,2,3]*

[1]*School of Life Sciences, Management Development Institute of Singapore, Singapore*

[2]*School of Applied Sciences, Northumbria University, United Kingdom*

[3]*HOHY PTE LTD, Singapore*

***Corresponding Author:** Maurice HT Ling, School of Life Sciences, Management Development Institute of Singapore and School of Applied Sciences, Northumbria University, United Kingdom and HOHY PTE LTD, Singapore.

Received: February 02, 2023; **Published:** February 28, 2023

Abstract

Eukaryotic genomes are organized into multiple chromosomes and studies have suggested that chromosomal organization is subjected to evolutionary pressure as demonstrated by non-randomness of properties across chromosomes. However, a recent study provided evidence to suggest that chromosomal organization may be more random in unicellular than multicellular eukaryotes. Here, we examine the distribution of five peptide features (length, aromaticity, instability, hydropathy, and isoelectric point) across the 16 nuclear chromosomes of *Saccharomyces arboricola* H-6, a recently identified and sequenced unicellular eukaryote. Our results show that only hydropathy is not random across the chromosomes (F = 1.914, p-value = 0.018); thereby, supporting the hypothesis that chromosomal organization may be random in unicellular eukaryotes.

Keywords: *Peptide Properties; Saccharomyces arboricola H-6; Chromosomal Organization*

Introduction

Saccharomyces arboricolus H-6 was first isolated from the bark of *Quercus fabri* in western China and phenotypically described by Wang and Bai [1] in 2008 as a novel species of yeast, which is supported by Gayevskiy and Goddard [2] using phylogenomics approach. Synonymous strain names of *S. arboricolus* H-6 are AS 2.3317 and CBS 10644 [1]. Genetic identification was carried out by Naumov's team [3,4]. However, *S. arboricolus* was renamed as *S. arboricola* in subsequent publications by Muir., *et al*. [5] and Naumov., *et al*. [6]. Both Muir., *et al*. [5] and Naumov., *et al*. [6] denote *S. arboricolus* H-6 as *S. arboricola* CBS 10644 in their publications, which corresponds to the synonymous strain names of *S. arboricolus* H-6 [1]. Hence, it can be assumed that *S. arboricola* is *S. arboricolus* if their strain names matched. *S. arboricola* H-6 has been studied for its potential in low temperature fermentation [7,8] and its genomic sequence published in 2013 [9]. Since 2016 [2], *S. arboricola* appears to be used consistently in the literature.

A volume of studies has suggested that chromosomal organization is subjected to evolutionary pressure [10-13], which can be shown as departure from random distributions of properties [10,14-16]. However, a recent study by Lim., *et al*. [17] shows that several peptide properties are randomly distributed within the chromosomes *Plasmodium falciparum* (human malaria parasite) but all peptide properties

are not randomly distributed within the chromosomes *Sarcophilus harrisii* (Tasmanian devil). This suggests that chromosomal organization may be more random in unicellular than multicellular eukaryotes.

S. arboricola H-6 is a unicellular eukaryote with 16 nuclear chromosomes [9]. In this study, we examine the distribution of five peptide features (length, aromaticity, instability, hydropathy, and isoelectric point) across the 16 nuclear chromosomes of *S. arboricola* H-6. Our results show that only hydropathy is not random across the chromosomes (F = 1.914, p-value = 0.018); thereby, supporting the findings of Lim., *et al*. [17].

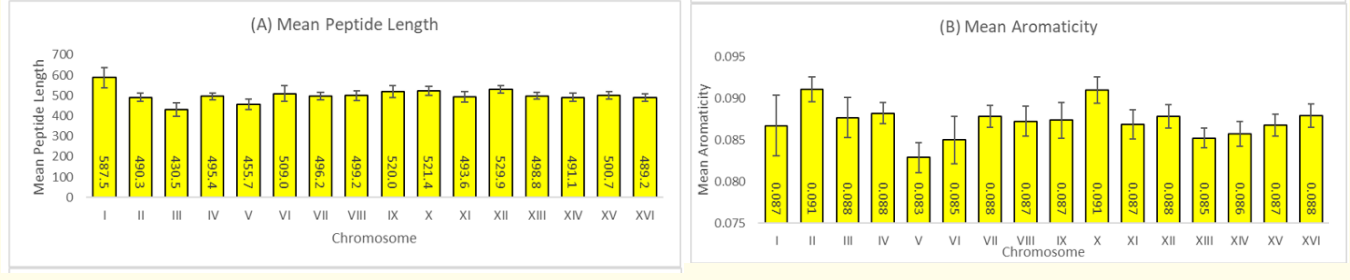
Materials and Methods

Sequences: The complete set of nuclear peptide sequences of *Saccharomyces arboricola* H-6 (Assembly SacArb1.0) [9] were downloaded from NCBI Genome by chromosomes; namely, Chromosome I (Accession CM001563.1), Chromosome II (Accession CM001564.1), Chromosome III (Accession CM001565.1), Chromosome IV (Accession CM001566.1), Chromosome V (Accession CM001567.1), Chromosome VI (Accession CM001568.1), Chromosome VII (Accession CM001569.1), Chromosome VIII (Accession CM001570.1), Chromosome IX (Accession CM001571.1), Chromosome X (Accession CM001572.1), Chromosome XI (Accession CM001573.1), Chromosome XII (Accession CM001574.1), Chromosome XIII (Accession CM001575.1), Chromosome XIV (Accession CM001576.1), Chromosome XV (Accession CM001577.1), and Chromosome XVI (Accession CM001578.1).

Analysis of peptide properties: Peptide length refers to the number of amino acids in the peptide. Aromaticity refers to the relative abundance of aromatic amino acids in a peptide [18]. Hydropathy (GRAVY) refers to the overall hydrophobic/hydrophilic properties of a peptide [19]. Isoelectric point (pI) is the pH where a peptide is electrical neutrality [20]. Instability index refers to the stability of the peptide where high instability score suggests shorter half-life [21]. All five methods are available in Biopython library [22] through SeqProperties [23]. Analysis of the five peptide sequences were carried out based on previous studies [24,25] as follows: (a) One-way ANOVA by chromosome was carried out to for each peptide sequence, (b) Pearson's correlation between properties, and (c) Significance of Pearson's correlation was carried out using ANOVA. P-values of less than 0.05 were considered significant.

Results and Discussion

Our results (Figure 1, table 1 and 2) show that the mean length (F = 0.851, p-value = 0.621), mean aromaticity (F = 1.112, p-value = 0.339), mean isoelectric point (pI; F = 0.796, p-value = 0.683), and mean instability (F = 1.176, p-value = 0.283) of peptides are not significantly different across the 16 chromosomes. Only the mean hydropathy is significant (F = 1.914, p-value = 0.018) across the 16 chromosomes. This result is consistent with the findings of Lim., *et al*. [17].



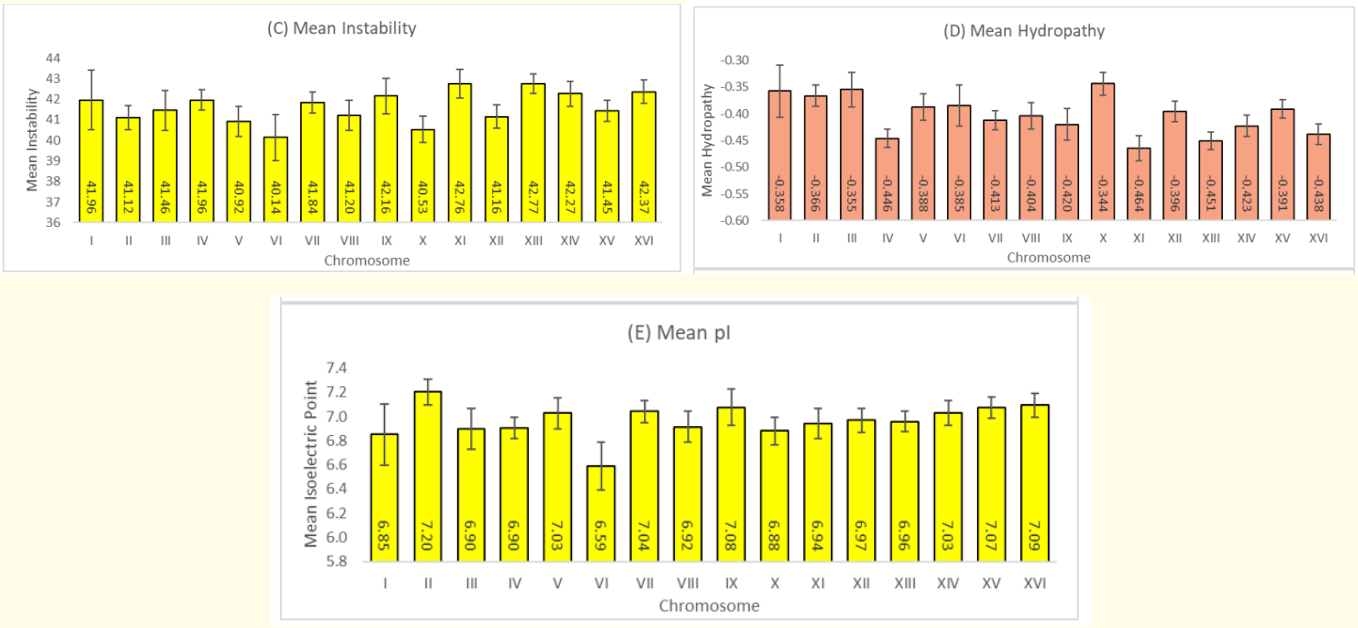


Figure 1: *Mean peptide properties by chromosome. Error bar denotes standard error. Panels A, B, C, D, and E denote mean peptide length, mean aromaticity, mean instability, mean hydropathy, and mean isoelectric point (pI) across 16 chromosomes respectively. One-way ANOVA analysis results is given in table 2. Of the 5 features, only mean hydropathy (Panel D) is significant between chromosomes (F = 1.914, p-value = 0.018).*

Chromosome and Peptide Count	Means and Standard Deviations of Peptide Properties				
	Peptide Length	Aromaticity	Instability	Hydropathy	Isoelectric Point
I (44)	587.5 (332.65)	0.0867 (0.02403)	41.96 (9.632)	-0.358 (0.3254)	6.85 (1.691)
II (262)	490.3 (370.56)	0.0910 (0.02820)	41.12 (10.340)	-0.366 (0.3731)	7.20 (1.879)
III (100)	430.5 (307.31)	0.0877 (0.03119)	41.46 (9.876)	-0.355 (0.3967)	6.90 (1.865)
IV (364)	495.4 (339.75)	0.0882 (0.02801)	41.96 (10.421)	-0.446 (0.4203)	6.90 (1.848)
V (173)	455.7 (310.78)	0.0829 (0.02824)	40.92 (9.722)	-0.388 (0.4158)	7.03 (1.911)
VI (72)	509.0 (348.92)	0.0850 (0.02515)	40.14 (9.808)	-0.385 (0.3925)	6.59 (1.755)
VII (332)	496.2 (344.48)	0.0878 (0.02999)	41.84 (9.933)	-0.413 (0.4005)	7.04 (1.890)

VIII (172)	499.2 (381.34)	0.0872 (0.02630)	41.20 (9.834)	-0.404 (0.3291)	6.92 (1.802)
IX (123)	520.0 (380.43)	0.0874 (0.03010)	42.16 (9.708)	-0.420 (0.4232)	7.08 (1.838)
X (225)	521.4 (379.15)	0.0910 (0.02962)	40.53 (9.304)	-0.344 (0.3823)	6.88 (1.793)
XI (190)	493.6 (431.40)	0.0868 (0.03039)	42.76 (10.463)	-0.464 (0.3806)	6.94 (1.858)
XII (294)	529.9 (463.02)	0.0878 (0.03028)	41.16 (9.500)	-0.396 (0.3516)	6.97 (1.863)
XIII (405)	498.8 (347.24)	0.0852 (0.02769)	42.77 (10.216)	-0.451 (0.3945)	6.96 (1.847)
XIV (258)	491.1 (329.92)	0.0857 (0.02639)	42.27 (10.335)	-0.423 (0.3624)	7.03 (1.813)
XV (349)	500.7 (343.20)	0.0868 (0.03021)	41.45 (10.382)	-0.391 (0.3712)	7.07 (1.787)
XVI (288)	489.2 (315.65)	0.0879 (0.02626)	42.37 (10.097)	-0.438 (0.3620)	7.09 (1.796)
Genome (3651)	498.2 (361.52)	0.0873 (0.02857)	41.75 (10.061)	-0.411 (0.3837)	7.00 (1.837)

Table 1: *Summary of peptide properties for each chromosome. The number of peptides in each chromosome is given in brackets. The standard deviations of peptide properties are given in brackets.*

Peptide Properties	$MS_{Chromosome}$	MS_{Error}	F-statistic	p-value
Peptide Length	111236	130775	0.851	0.621
Aromaticity	0.000907	0.000816	1.112	0.339
Hydropathy	0.2808	0.1467	1.914	0.018
Isoelectric Point	2.689	3.377	0.796	0.683
Instability	118.9	101.1	1.176	0.283

Table 2: *One-way ANOVA analysis of peptide properties by chromosome.*

Pairwise correlations of peptide properties at genome level (Figure 2) show that seven correlations are significantly different from zero (p-value < 3.96E-2). The three non-significant correlations are (a) aromaticity against peptide length (Figure 2A, p-value = 0.251), (b) hydropathy against peptide length (Figure 2C, p-value = 0.658), and (c) pI against hydropathy (Figure 2J, p-value = 0.217). Of the ten pairwise correlations, eight were consistent with the findings of Maitra and Ling [24]. The two correlations showing inconsistent results are (a) aromaticity against peptide length (Figure 2A), and (b) pI against hydropathy; both showing non-significant correlation in this study but significant correlation in Maitra and Ling [24].

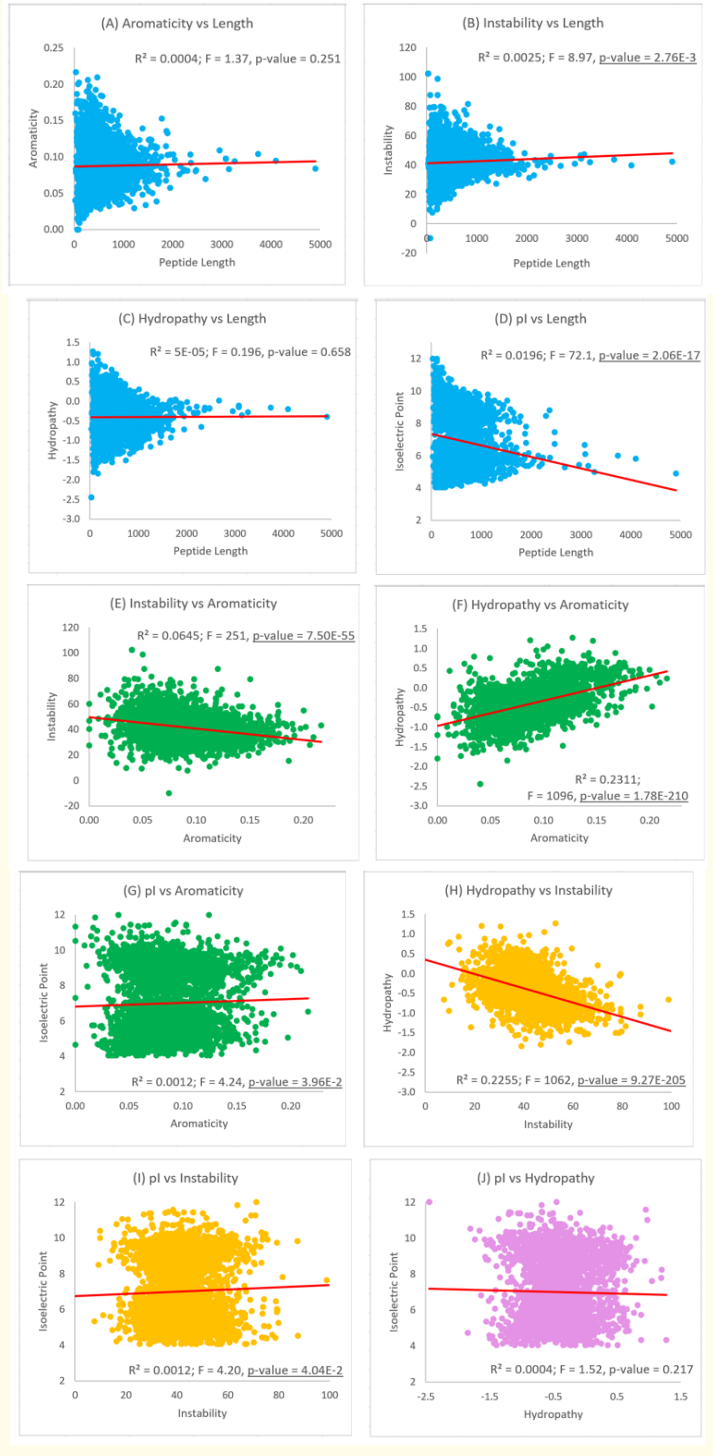


Figure 2: *Correlations between peptide properties. Underlined p-values denote significant correlations. Panels A, B, C, and D denote aromaticity, instability, hydropathy, and pI against peptide length. Panels E, F, and G denote instability, hydropathy, and pI against aromaticity. Panels H, and I denote hydropathy, and pI against instability. Panel J denotes pI against hydropathy.*

The consensus view is that multi-chromosomal organization is likely to originate from single-chromosome [26,27]. Furthermore, there is evidence of an ancient whole genome duplication in the history of *Saccharomyces* genus [28-30]. This is supported by recent studies to combine the chromosomes of *Saccharomyces cerevisiae* into a single mega-chromosome [31,32], suggesting tolerance from multi-chromosomality back to single-chromosomality. This is supported by similar assemblies with *Schizosaccharomyces pombe* [33]. While reduced growth and tolerance [34] in single chromosome strains may suggest the benefits of multi-chromosomal organization [35], it appears that these benefits may be rudimentary. This may suggest multi-chromosomal unicellular eukaryotes and multi-chromosomal prokaryotes may be at the preliminary transition from single-chromosomality to multi-chromosomality. As such, chromosomal organization may be random in unicellular organisms compared to multicellular organisms. Our results showing no significant differences in four of the five peptide properties and the findings of Lim., *et al*. [17] support this hypothesis.

Conclusion

One of the five peptide features is not random (F = 1.914, p-value = 0.018) across the 16 nuclear chromosomes of *Saccharomyces arboricola* H-6, a recently identified and sequenced unicellular eukaryote; thus, providing supportive evidence that chromosomal organization may be random in unicellular eukaryotes.

Supplementary Materials

Data files for this study can be downloaded at https://bit.ly/H6_peptides.

Conflict of Interest

The authors declare no conflict of interest.

Bibliography

1. Wang S-A and Bai F-Y. "*Saccharomyces arboricolus* sp. nov., A Yeast Species from Tree Bark". *International Journal of Systematic and Evolutionary Microbiology* 58.2 (2008): 510-514.

2. Gayevskiy V and Goddard MR. "*Saccharomyces eubayanus* and *Saccharomyces arboricola* Reside in North Island Native New Zealand Forests". *Environmental Microbiology* 18.4 (2016): 1137-1147.

3. Naumov GI. "Hybridization Analysis of the New Biological Species *Saccharomyces arboricolus* Wang et Bai". *Doklady Biological Sciences: Proceedings of the Academy of Sciences of the USSR, Biological Sciences Sections* 426 (2009): 247-249.

4. Naumov GI., *et al*. "Genetic Identification of New Biological Species *Saccharomyces arboricolus* Wang et Bai". *Antonie Van Leeuwenhoek* 98.1 (2010): 1-7.

5. Muir A., *et al*. "A Multiplex Set of Species-Specific Primers for Rapid Identification of Members of the Genus *Saccharomyces*". *FEMS Yeast Research* 11.7 (2011): 552-563.

6. Naumov GI., *et al*. "Molecular Genetic Diversity of the *Saccharomyces* Yeasts in Taiwan: *Saccharomyces arboricola, Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii*". *Antonie Van Leeuwenhoek* 103.1 (2013): 217-228.

7. Magalhães F., *et al*. "Frozen-Dough Baking Potential of Psychrotolerant *Saccharomyces* Species and Derived Hybrids". *Food Microbiology* 94 (2021): 103640.

8. Nikulin J., *et al*. "Alternative *Saccharomyces* Interspecies Hybrid Combinations and Their Potential for Low-Temperature Wort Fermentation". *Yeast* 35.1 (2018): 113-127.

9. Liti G., *et al*. "High Quality *De Novo* Sequencing and Assembly of the *Saccharomyces arboricolus* Genome". *BMC Genomics* 14 (2013): 69.

10. Danchin A., *et al*. "Mapping the Bacterial Cell Architecture into the Chromosome". *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 355.1394 (2000): 179-190.

11. Stauffer D and Cebrat S. "Haplotype Complementarity Under Mutational Pressure". *Frontiers in Bioscience* 3.2 (2011): 408-415.

12. Balzano E and Giunta S. "Centromeres under Pressure: Evolutionary Innovation in Conflict with Conserved Function". *Genes* 11.8 (2020): 912.

13. Tomasch J., *et al*. "Connection Between Chromosomal Location and Function of CtrA Phosphorelay Genes in Alphaproteobacteria". *Frontiers in Microbiology* 12 (2021): 662907.

14. Ling MHT., *et al*. "Conserved Expression of Natural Antisense Transcripts in Mammals". *BMC Genomics* 14 (2013): 243.

15. Keng BM., *et al*. "Codon Usage Bias is Evolutionarily Conserved". *Asia Pacific Journal of Life Sciences* 7.3 (2014): 233-242.

16. Wang M., *et al*. "The Non-Random Patterns of Genetic Variation Induced by Asymmetric Somatic Hybridization in Wheat". *BMC Plant Biology* 18.1 (2018): 244.

17. Lim GZ., *et al*. "Significant Differences in Nucleotide and Peptide Features Between Chromosomes Suggesting Sequence Non-randomness Across Chromosomes". *Acta Scientific Microbiology* 4.4 (2021): 23-28.

18. Lobry JR and Gautier C. "Hydrophobicity, Expressivity and Aromaticity are the Major Trends of Amino-Acid Usage in 999 *Escherichia coli* Chromosome-Encoded Genes". *Nucleic Acids Research* 22.15 (1994): 3174-3180.

19. Kyte J and Doolittle RF. "A Simple Method for Displaying the Hydropathic Character of a Protein". *Journal of Molecular Biology* 157.1 (1982): 105-132.

20. Bjellqvist B., *et al*. "The Focusing Positions of Polypeptides in Immobilized pH Gradients can be Predicted from Their Amino Acid Sequences". *Electrophoresis* 14.10 (1993): 1023-1031.

21. Guruprasad K., *et al*. "Correlation Between Stability of a Protein and its Dipeptide Composition: A Novel Approach for Predicting In Vivo Stability of a Protein from its Primary Sequence". *Protein Engineering* 4.2 (1990): 155-161.

22. Cock PJA., *et al*. "Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics". *Bioinformatics* 25.11 (2009): 1422-1423.

23. Ling MHT. "SeqProperties: A Python Command-Line Tool for Basic Sequence Analysis". *Acta Scientific Microbiology* 3.6 (2020): 103-106.

24. Maitra A and Ling MH. "Codon Usage Bias and Peptide Properties of Pseudomonas balearica DSM 6083T". *MOJ Proteomics and Bioinformatics* 8.2 (2019): 27-39.

25. Kim JH and Ling MH. "Proteome Diversities Among 19 Archaebacterial Species". *Acta Scientific Microbiology* 2.5 (2019): 20-27.

26. Das B and Chattoraj DK. "Commentary: Functionality of Two Origins of Replication in *Vibrio cholerae* Strains with a Single Chromosome". *Frontiers in Microbiology* 10 (2019): 1314.

27. Jha JK., *et al*. "Chromosome Dynamics in Multichromosome Bacteria". *Biochimica Et Biophysica Acta* 1819.7 (2012): 826-829.

28. Kellis M., *et al*. "Proof and Evolutionary Analysis of Ancient Genome Duplication in the Yeast *Saccharomyces cerevisiae*". *Nature* 428.6983 (2004): 617-624.

29. Wolfe KH and Shields DC. "Molecular Evidence for an Ancient Duplication of the Entire Yeast Genome". *Nature* 387.6643 (1997): 708-713.

30. Wolfe KH. "Origin of the Yeast Whole-Genome Duplication". *PLoS Biology* 13.8 (2015): e1002221.

31. Shao Y., *et al*. "Creating a Functional Single-Chromosome Yeast". *Nature* 560.7718 (2018): 331-335.

32. Luo J., *et al*. "Karyotype Engineering by Chromosome Fusion Leads to Reproductive Isolation in Yeast". *Nature* 560.7718 (2018): 392-396.

33. Gu X., *et al*. "Single-Chromosome Fission Yeast Models Reveal the Configuration Robustness of a Functional Genome". *Cell Reports* 40.8 (2022): 111237.

34. Doughty TW., *et al*. "A Single Chromosome strain of *S. cerevisiae* Exhibits Diminished Ethanol Metabolism and Tolerance". *BMC Genomics* 22.1 (2021): 688.

35. Bochkareva OO., *et al*. "Genome Rearrangements and Selection in Multi-Chromosome Bacteria *Burkholderia* spp". *BMC Genomics* 19.1 (2018): 965.

Volume 19 Issue 3 March 2023