CRISPRome: Exploring Adaptive Immunity in Extremophiles

G Vartika, Sukanya Bose, Mohammad Yasir, Zahir Ali, Subha Srinivasan

Institute of Bioinformatics and Applied Biotechnology (IBAB), Bengaluru - 560100

King Abdullah University of Science and Technology, Saudi Arabia

King Abdulaziz University, Saudi Arabia

IBAB
Institute of Bioinformatics and Applied Biotechnology

ABSTRACT

CRISPR-cas system has revolutionized biological and medicinal application by simplifying gene editing technologies. It is a prevalent defense mechanism in microbes against viruses. This study explores the diversity of CRISPR-Cas systems and phages in hot spring samples from Saudi Arabia. We used metagenomic data to identify novel CRISPR sequences, cas genes and phages. Additionally, we also analyzed variations in CRISPR arrays and cas operons along with characterization of microbial communities within the samples.

BACKGROUND

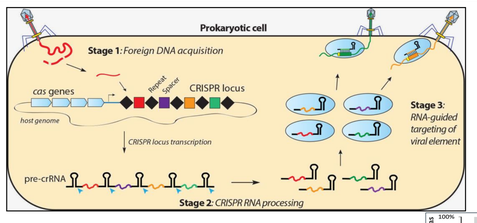


Figure 1. Mechanism of Adaptive immunity in Prokaryotic cell, Doudna Lab.

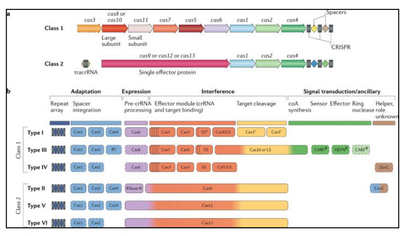


Figure 2. Modular organization of the various types of CRISPR-Cas systems, Makarova et al., 2020 .

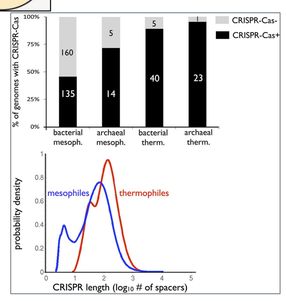
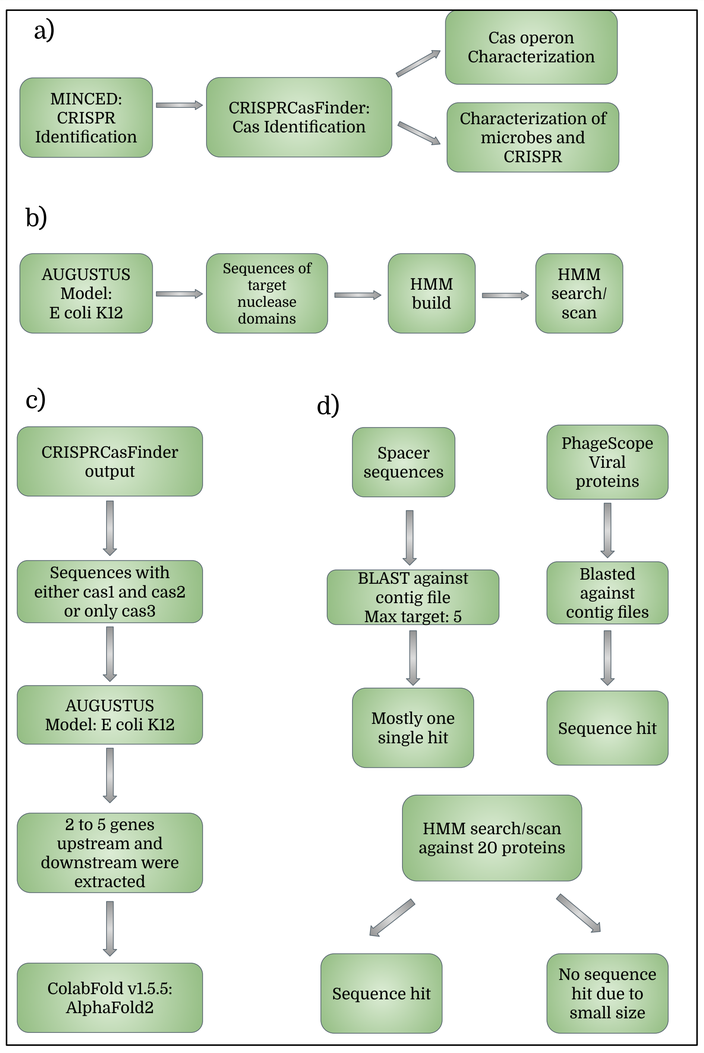


Figure 3. CRISPR abundance and CRISPR length distribution in various conditions, Ariel et al., 2012 .

DATASET

The dataset comprises of four hot spring soil samples (AH1, AK, AW5, WM) collected from non-volcanic regions across Saudi Arabia. These samples were provided by Dr. Yasir Mohammad, King Abdulaziz University. The temperature of these hot springs varies significantly, ranging from 90°C (AH1) to 56°C (AW5 and WM), offering a unique opportunity to investigate the influence of temperature on the microbial communities and CRISPR-Cas systems within these extreme environments.

METHODS



ACKNOWLEDGEMENT

This study was supported by Dept. of IT, BT, Govt. of Karnataka and Govt. of India.

RESULTS

a) Characterizing the Microbial communities by Culturomics

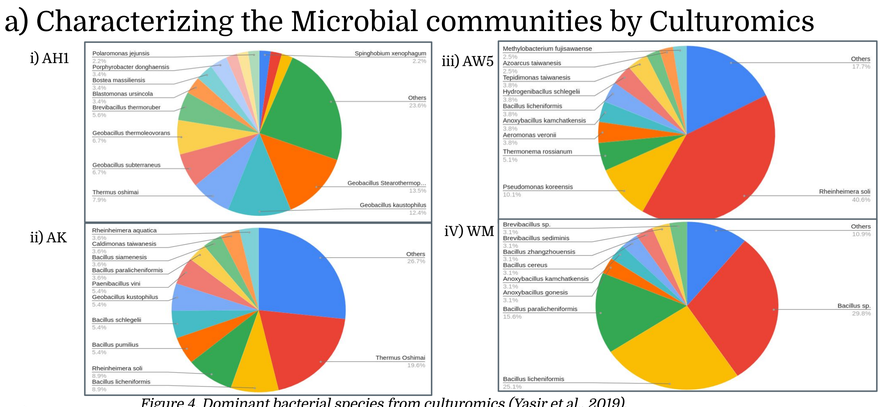


Figure 4. Dominant bacterial species from culturomics (Yasir et al., 2019).

b) 16S rRNA sequencing

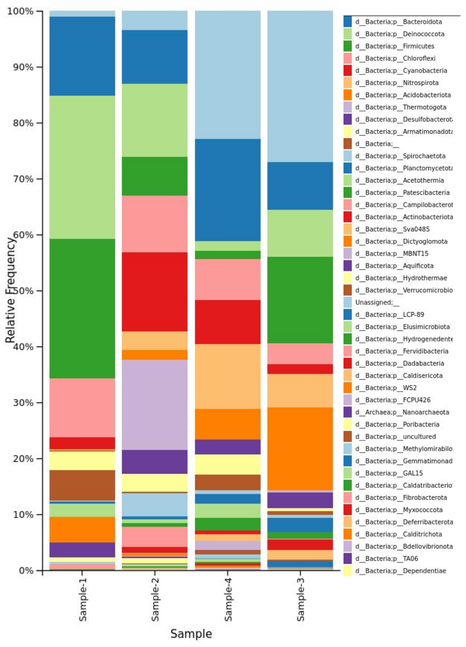


Figure 5. Relative frequency of identified taxa .

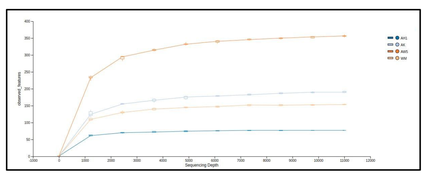


Figure 6. Alpha diversity using sequencing depth .

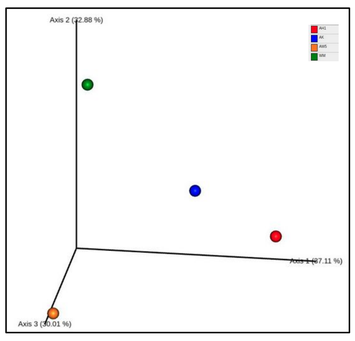
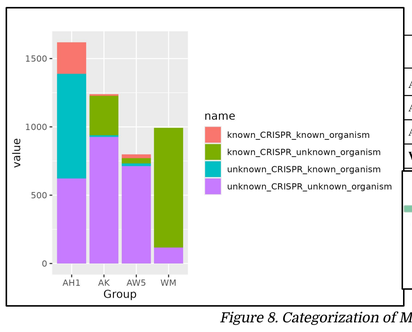


Figure 7. Beta diversity in principle component plot.

c) CRISPR Spacer Characterization and Diversity Analysis from WGS



	Type I	Type II	Type III	Type IV	No of CRISPR	No of sequences
AH1	231	NA	766	622	1659	1619
AK	12	289	12	926	1213	1194
AW5	28	39	19	713	817	799
WM	NA	876	NA	117	1021	993

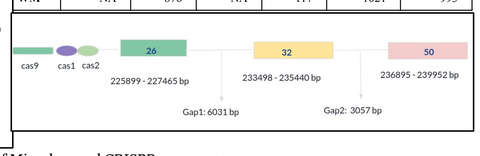


Figure 8. Categorization of Microbes and CRISPR sequences.

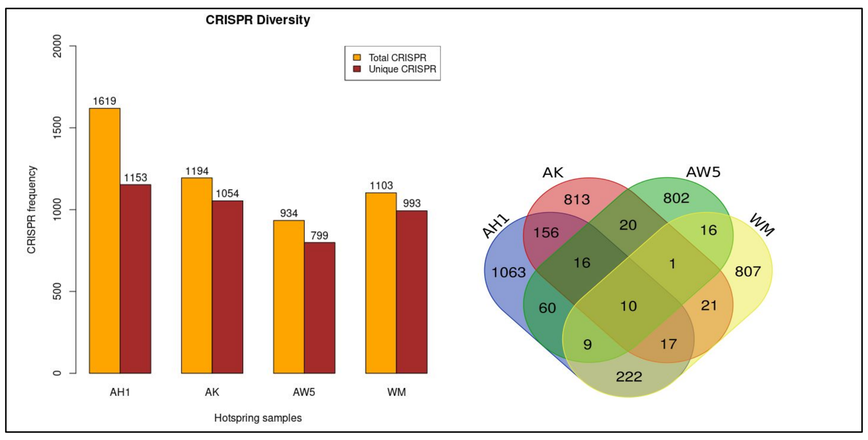


Figure 9. CRISPR diversity among samples .

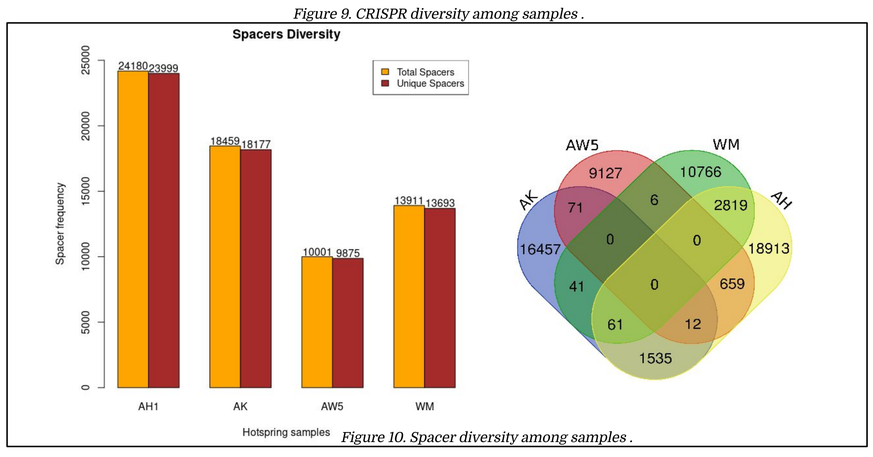
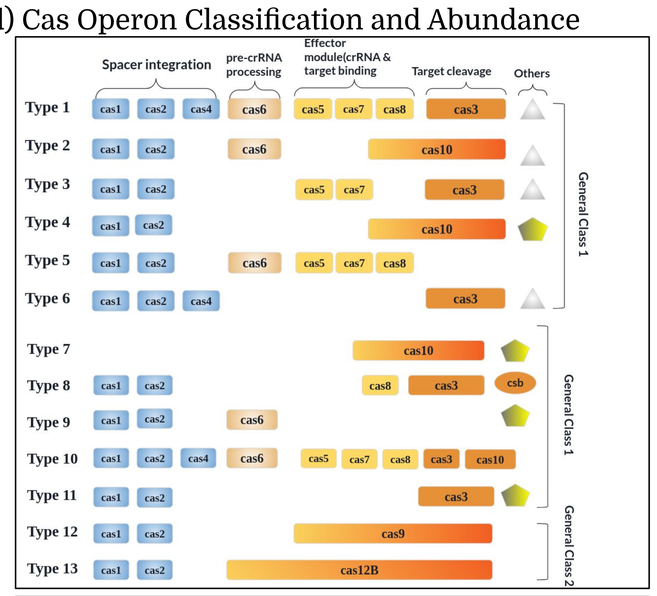


Figure 10. Spacer diversity among samples .

REFERENCES

- Yasir, M., Qureshi, A. K., Khan, I., Bibi, F., Rehan, M., Khan, S. B.,& Azhar, E. I. (2019). Culturomics-Based Taxonomic Diversity of Bacterial Communities in the Hot Springs of Saudi Arabia. OMICS: A Journal of Integrative Biology, 23(1), 17-27. doi: 10.1089/omi.2018.0176
- Couvin, D., Bernheim, A., Toffano-Nioche, C., Touchon, M., Michalik, J., Néron, B. Vergnaud, G. (2018). CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrated search for Cas proteins. Nucleic Acids Research, 46(W1), W546-W551. doi: 10.1093/nar/gky409

d) Cas Operon Classification and Abundance



	AH1	AK	AW5	WM	Known Types
TYPE 1	15	12	8	21	Type IA
TYPE 2	8	8	6	5	Type IIIA
TYPE 3	7	0	0	0	Type IE
TYPE 4	3	0	0	0	Novel
TYPE 5	3	2	2	1	Novel
TYPE 6	3	0	0	0	Type ID
TYPE 7	9	0	0	1	Type IIIA
TYPE 8	3	0	0	0	Novel
TYPE 9	7	0	0	0	Novel
TYPE 10	8	2	2	5	Novel
TYPE 11	0	0	0	4	Novel
TYPE 12	5	2	5	6	Type II
TYPE 13	2	0	0	7	Type V

Figure 11. Classification of cas operons beside CRISPR array.

e) Distinct configuration

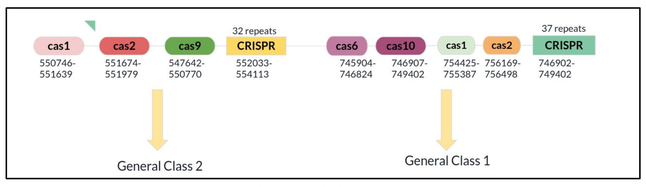


Figure 12. WM Node (1056004 bp) at 56 °C .

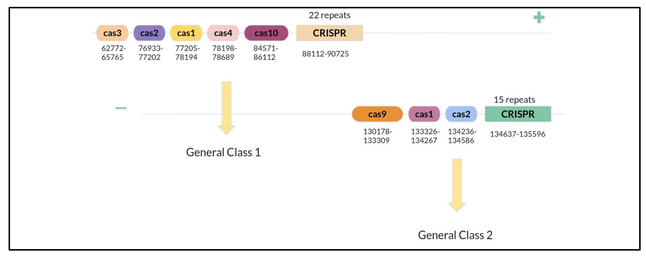


Figure 13. AW5 Node (257978 bp) at 56 °C .

f) Putative Cas-like Proteins

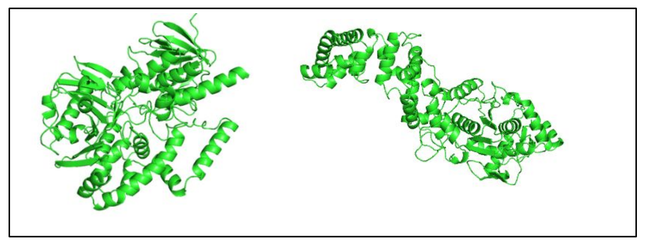


Figure 14. Folded structures from ColabFold v1.5.5: AlphaFold2 .

CONCLUSION AND FUTURE PROSPECTS

Conclusion:

- The comparative analysis of the four hotspring sample shows diversity in CRISPR arrays and cas operons architecture
- Despite the presence of Bacteroidota, cas13 is absent in the samples
- We found cas operons with more than one cas3 genes and also in some operons both cas3 and cas10 target cleavage genes are present
- We found putative cas-like proteins
- tracrRNA sequences were found residing in General class II Type II and Type V

Future prospects:

- Identification of PAM (Protospacer adjacent motif) following PhagOme analysis
- Analyzing the cas protein domains to well characterize any more cas-like proteins
- Characterization of CRISPR in different hotspring samples from variety of geographical conditions