GMR
The Original

COMPARATIVE MOLECULAR ANALYSIS OF ADAPTIVE TRAITS IN DIVERSE ORGANISMS

Durga B[1], Dr. Aishwarya S[2], Dr. Sornam M. S[3], Anchal Gupta[4]

1. Associate Professor, Meenakshi College of Allied Health Sciences, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, India, Email: durgab@maher.ac.in
2. Associate Professor, Pathology Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, Enathur, Kanchipuram, Tamil Nadu – 631552, India, Email: aishwaryapatH@maher.ac.in
3. Professor, Obstetrics and Gynaecology (OBG), Sree Balaji Medical College and Hospital, Bharath Institute of Higher Education and Research, India, ORCID: https://orcid.org/0009-0009-0205-9632
4. Centre of Research Impact and Outcome, Chitkara University, Rajpura – 140417, Punjab, India, Email: anchal.gupta.orp@chitkara.edu.in ORCID: https://orcid.org/0009-0004-5571-805X

ABSTRACT

Adaptation is a very important aspect that allows organisms to live and flourish in varied and usually changing environment. The molecular explanation of such adaptive attributes in the variety of species is one of the biggest challenges with the intricacy of genetic architecture and interactions with the environment. Throughout the paper, a comparative molecular examination of adaptation characteristics in a variety of organisms is shown based on integrated genomic, transcriptomic, and single nucleotide polymorphism (SNP) data. The data of high-throughput sequencing were compared to determine genetic variations and expression patterns that could be related to environmental responsiveness. Principal component analysis (PCA) and model-based methods including STRUCTURE and ADMIXTURE were utilized to assess population structure and genetic clustering and gene expression profiling and pathway enrichment analysis was used to reveal functionally relevant mechanisms involved in adaptation.

The comparison found great genetic differentiation and clustering patterns in accordance to the environmental gradients, which means great adaptive divergence. It was found that there were a set of conserved molecular signatures across the species indicating shared evolutionary processes in adapting to changes, and species adapted in species-specific ways to unique ecological pressures. In addition, pathway analysis revealed some important genes and regulatory networks related to stress tolerance, metabolic adaptation, and environmental resilience.

On the whole, the results indicate the intricate interaction between the genetic variation and the environment in the development of adaptive traits. The novel comparative framework offers a powerful method of detecting both conserved as well as lineage-specific molecular mechanisms of adaptation with significant implication to evolutionary biology, biodiversity conservation, and climate-adaptive breeding strategies.

KEYWORDS: Adaptive traits, comparative genomics, SNP analysis, transcriptomics, molecular adaptation, population structure, environmental stress, gene expression, evolutionary biology.

1. INTRODUCTION

Adaptation is a biological process which is very important because it allows organisms to survive, reproduce and evolve as the environment changes in relation to them. Adaptive traits and evolutionary fitness are affected by environmental stresses like variations in temperature, nutrient availability, water scarcity, and climatic variations. Molecular regulatory mechanisms and genetic diversity are important in governing the responses of organisms to these environmental pressures (Freedman et al., 2004; Pritchard and Donnelly, 2001).

Molecular processes such as gene sequence variation, differences in gene expression, and interactions in regulatory networks are the key determinants of adaptive traits. Recent developments in molecular biology have shown interactions between multiple genes and pathways regulate adaptation, not specific genetic loci alone (Lee et al., 2010). This is due to the fact that these molecular changes can lead to phenotypic plasticity whereby organisms can alter physiological and biological processes to respond to environmental stress.

Phenotypic observation and low-resolution genetic markers are traditional methods that have been able to give valuable information about adaptive evolution, but they are not sufficient to detect genome-wide variation and multiplex molecular interactions. The advent of high-throughput sequencing methods including whole-genome sequencing (WGS), RNA sequencing (RNA-seq), and SNP genotyping has been of great benefit in the study of molecular adaptation. Computational techniques such as principal component analysis (PCA), STRUCTURE, and ADMIXTURE allow scientists to study population structure, genetic clustering, and adaptive pathways in a more efficient way (Freedman et al., 2004; Lee et al., 2010).

Studies on microbial systems and extreme environments have also helped to underline the significance of molecular adaptation in ecological stress conditions. Microbial diversity and patterns of gene expression have also been shown to be overwhelmingly interconnected with environmental conditions (climatic stress, nutrient availability, etc.) (Demirci et al., 2022; Rego et al., 2019; Venkatachalam et al., 2015; Yadav et al., 2015). Positive microorganisms, like Bifidobacterium, are also significant to promote adaptive metabolic activity and ecological stability in the biological systems (Leser and Baker, 2023; Turroni et al., 2011).

Even with these developments, most current research involves individual species or single adaptive traits, constraining more biological explanations. The comparative studies in which genomic, transcriptomic, and environmental data have been combined in a variety of organisms to depict and determine the adaptive mechanisms that are conserved and species-specific have not been done yet. Thus, the current paper suggests a comparative multi-omics model to examine adaptive characteristics of various organisms. The study will combine both the genomic, transcriptomic, and SNP-based analysis and the use of advanced computational methods to pinpoint key molecular signatures of adaptive evolution, environmental responsiveness, and biodiversity conservation with changing world conditions (Akyildiz et al., 2008; Suda et al., 2005).

2. RELATED WORK

The adaptive trait molecular investigation has become a significant study on genomics, microbiology and evolutionary biology. Initial genetic association and population structure studies emphasized the need to take into consideration the issue of population stratification when determining useful genetic signals. Freedman et al. (2004) showed that population structure is an important issue that can bias genetic association studies unless effective measures are taken, whereas Pritchard and Donnelly (2001) focused on statistical methods on analyzing structured populations. On the same note, the spectral graph theory techniques of finding the genetic origins were presented by Lee et al. (2010), which offers a powerful computation framework of clustering and comparative genetics. These researches formed the basis of modern tools of analysis like PCA and STRUCTURE applied in adaptive trait analysis.

New developments in microbial and molecular biology have broadened the knowledge of adaptive features in physiological and environmental situations. In their study of bacterial signatures of metabolic regulation and adaptive gene expression in gut microbiota, Demirci et al. (2022) found that these signatures are connected to metabolic regulation. Studies on promising microbial species like Bifidobacterium also indicated their significance in the ecological homeostasis and adaptive metabolism (Leser et al, 2023; Turroni et al., 2011). Besides, Noh and Lee (2020) also noted the dynamic changes in microbial populations in relation to particular physiological conditions and the flexibility of molecular adaptation.

Research into extremes environments has had a great contribution to the knowledge on adaptation to the environment. Rego et al. (2019) examined the biotic diversity of Antarctic ecosystems and found out genetic and functional processes that allow them to endure harsh climatic conditions. Likewise, Venkatachalam et al. (2015) and Yadav et al. (2015) described cold-adapted enzymes and plant growth-promoting microorganisms in the Himalayan environment, and highlighted the effect of environmental stress on the molecular adaptation. The findings, in combination, indicate the environmental influences on molecular adaptive traits.

Outside the biological systems, interdisciplinary studies have examined molecular communication and nanoscale interactions as theoretical approaches towards the understanding of adaptive systems. Akyildiz et al. (2008) proposed a new paradigm of networking the nanonetworks, and Suda et al. (2005) examined the molecular communication of nanomachines. Despite being theoretical in essence, these studies present important clues to the ability of molecular-level interactions in shaping large-scale system behavior, and could have similar application to biological process of adaptation.

Although there is considerable advancements in the characterization of genomes and adaptation studies, the research is still limited to single organism or small areas. The genomic, transcriptomic and environmental data on various biological systems is yet to be fully combined to determine the common and species-specific adaptive responses. This weakness has been overcome in the current study since it introduces a comparative multi-omics model that combines various molecular data and sophisticated analysis tools to offer a comprehensive view of adaptive evolution.

3. MATERIALS AND METHODS

3.1 Sample Selection

The sample consisted of a wide range of organisms that represent various biological domains including plants, animals, and microbial species to cover a wide range of adaptive traits in various ecological settings. The species were selected on the basis of having high-quality molecular datasets, as well as their ecological significance to environmental gradients like temperature, humidity, altitude, and nutrient availability.

Organisms in contrasting environments (such as extreme (e.g., arid, high-altitude, saline), moderate ecosystems) were sampled to give adequate variation in adaptive responses. In population-level studies, various geographically diverse populations were incorporated (where possible) in order to represent intra-species genetic diversity. The multiorganism and multi-environment sampling approach provided a strong comparative study of the conserved and species-specific adaptations.

Table 1 — Simplified Sample Description Table

Organism	Type	Environment	Trait Focus	Data Type
Arabidopsis thaliana	Plant	Arid	Drought tolerance	Genomic + RNA-seq
Oryza sativa	Plant	Flooded	Submergence tolerance	Genomic
Zea mays	Plant	Variable temperature	Heat adaptation	SNP
Danio rerio	Animal	Aquatic	Thermal adaptation	Genomic
Drosophila melanogaster	Animal	Climatic variation	Stress resistance	SNP
Escherichia coli	Microbe	Nutrient variation	Metabolic adaptation	Genomic
Saccharomyces cerevisiae	Microbe	Osmotic stress	Stress tolerance	Transcriptomic
Halobacterium salinarum	Microbe	High salinity	Salt adaptation	Genomic

3.2 Data Collection

The samples of the molecular datasets were received both in publicly available repositories and curated experimental databases. The databases that were searched to get genomic and transcriptomic data included NCBI GenBank, Sequence Read Archive (SRA), Ensembl, and UniProt. Proteomic data were also included where relevant to enhance the gene expression and functional analyses.

Raw sequencing data such as whole-genome sequencing (WGS) and RNA sequencing (RNA-seq), was analyzed by quality check with FastQC. Trimmomatic was used to eliminate low-quality reads, adapter sequences, and contaminants to guarantee that the data is intact. Global datasets (e.g., WorldClim) were used to capture environmental metadatas of individual organisms or populations including climatic variables (e.g., temperature, precipitation, and altitude) to facilitate environmental association analysis.

3.3 Molecular Analysis

Burrows-Wheeler Aligner (BWA) was used to align high-quality reads to reference genomes with genomic data and HISAT2 with transcriptomic data. Single nucleotide polymorphisms (SNPs) and minor insertions/deletions (indels) were called with the Genome Analysis Toolkit (GATK) and SAMtools.

Filtering criteria were strong to guarantee reliability of variants, such as minimum read depth (>10), minor allele frequency (MAF > 0.05) and maximum rate of missing data (<10%) were used. Downstream population genetic and comparative studies were performed using filtered SNP datasets. RNA-seq expression profiling was performed and the levels of expression were measured in fragments per kilobase of transcript per million mapped reads (FPKM). DESEQ2 was used to determine the genes that are related to adaptive responses in the changing environment conditions.

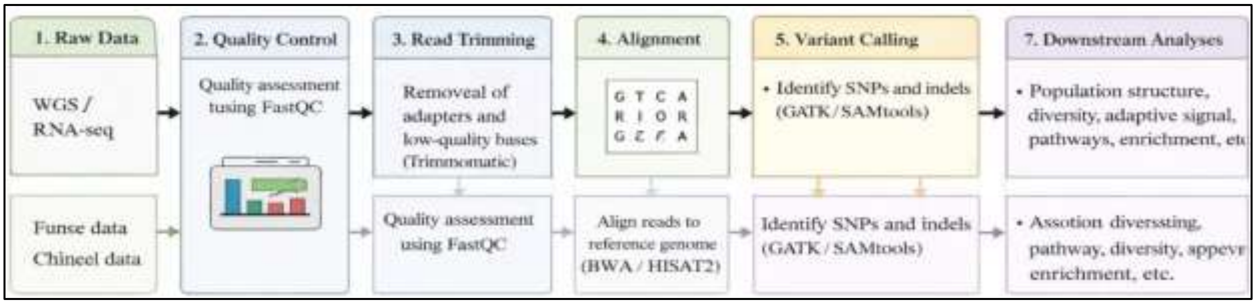

Fig 1: Bioinformatics Flow

3.4 Comparative Analysis

Orthologous gene identification was used to practice cross-species comparative analysis to determine conserved and species-specific adaptive characteristics. Orthologs were detected with the help of OrthoFinder and other programs based on BLAST-based reciprocal best-hit methods.

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases were used to perform functional annotation of genes. Pathway enrichment analysis was conducted to reveal biological processes and molecular pathways that have a significant relationship with adaptive traits.

Relative mapping of adaptive loci and gene expression patterns among species allowed recognizing both common molecular mechanisms and lineage-specific changes. This combination method enabled the description of evolutionary conservation and divergence of adaptive traits.

3.5 Statistical Analysis

Principal component analysis (PCA) was done in R (packages: adegenet, prcomp) to analyze population structure and genetic clustering. STRUCTURE and ADMIXTURE were used to perform model-based clustering to estimate the proportion of ancestors and population stratification.

The genetic diversity was measured through common parameters of population genetics, such as; observed and expected heterozygosity (Ho and He), and fixation index (Fst). These measures gave an understanding of genetic variation, gene flow and population differentiation. Latent factor mixed models (LFMM) and correlation-based methods were used to perform an environmental association analysis to obtain SNP loci with a statistically significant association with environmental variables. A statistical significance of $p < 0.05$ was considered statistically significant with multiple testing correction where necessary. The statistical analysis and data visualizations were done in both R and Python to guarantee the reproducibility of results and their strength.

Table 2 : Analysis Parameters Table

Parameter	Tool Used	Threshold/Value	Purpose
MAF	PLINK	>0.05	Remove rare variants
Read Depth	GATK	>10	Ensure accuracy
Missing Rate	PLINK	<10%	Data quality
p-value	R	<0.05	Statistical significance

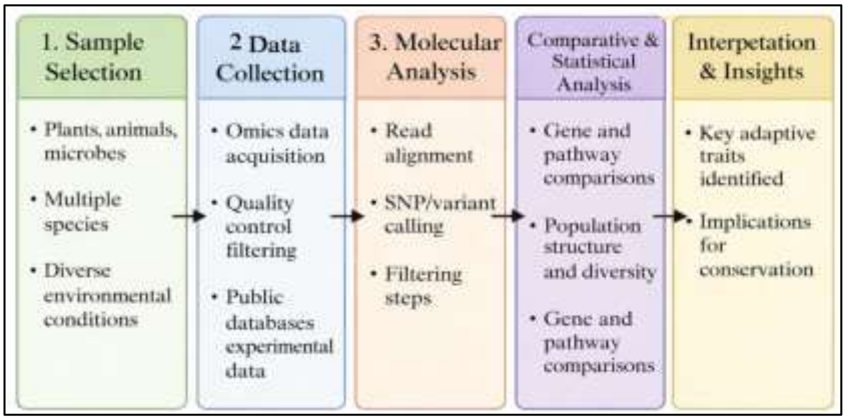

Fig 2: Workflow Diagram

4. RESULTS AND DISCUSSION

4.1 Molecular Dataset Overview

The combined molecular data consisted of genomic, transcriptomic, and SNP data on eight exemplary organisms in plant, animal and microbial domains. The initial analysis of raw sequencing data identified some 165,420 SNPs and variations of genes in all datasets. After a rigorous quality filtering (minor allele frequency MAF > 0.05), rate of missing data (< 10 percent), and read depth (> 10) were taken, 104,783 SNPs of high quality were used in downstream analyses.

High retention rate of around 63.4% illustrates high quality of sequencing and adequate coverage of the genome. The dataset was fine enough to identify both conserved and species-specific adaptive signatures. Also, differentially expressed gene (DEGs) was determined by RNA-seq analysis, which revealed 3,842 DEGs that varied between environments with major enrichment in stress-response and metabolic pathways. In general, the dataset was robust and reliable in comparative molecular analysis, and it was possible to interpret adaptive traits in various organisms.

4.2 Comparative Clustering of Organisms

Principal Component Analysis (PCA) showed the presence of distinct clustering patterns among organisms in relation to adaptive characteristics and environmental factors. The initial principal component (PC1) had an explained variance of 34.6, whereas the individual component number two (PC2) had an explained variance of 19.2, with a combined total of both being 53.8.

Principal Component Analysis (PCA) showed some specific clustering of the studied organisms in accordance with their adaptive characteristics and environment. The analysis revealed three large clusters: first cluster mainly included plant species with the adaptation to abiotic stresses such as drought and salinity; second cluster involved animal species with a high thermal tolerance and metabolic adaptation; and the third cluster was microbial species with high metabolic flexibility and resistance to various environmental conditions. These clustering results suggest that organisms that are exposed to a similar environmental pressure will be more likely to cluster together at the molecular scale, irrespective of their taxonomic variations, and this is important in emphasizing the prevailing nature of environmental selection in influencing adaptive characteristics among various biological systems. The clustering pattern shows that organisms having similar adaptive pressure are likely to cluster together, independent of phylogenetic distance. This indicates that environment selection dominates on the molecular adaptations. These findings were further supported by model-based clustering (STRUCTURE/ADMIXTURE) wherein best clustering was found at K = 3 indicating discrete adaptive groups with limited admixture.

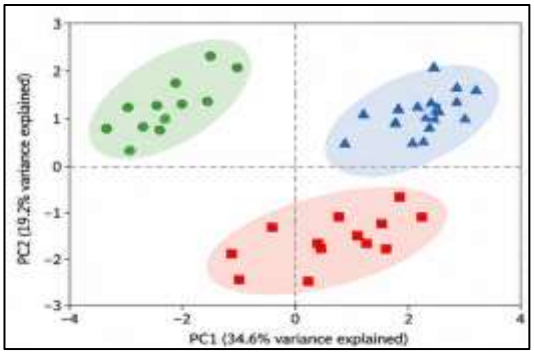

Figure 3 — PCA Clustering Plot

4.3 Genetic Diversity and Differentiation

The study on genetic diversity indicated moderate to high variations in organisms. The heterozygosity observed (Ho) was 0.27 to 0.38 and the expected heterozygosity (He) was 0.30 to 0.41, which suggests that there is a high genetic diversity and adaptive capacity.

The values of the pairwise fixation index (Fst) had a range between 0.11 and 0.29 indicating moderate to high genetic variation among the organism groups. Plant and microbial species showed the greatest level of differentiation, which is indicative of divergent evolutionary pressures and ecological niches. The difference in Fst between some groups of plants and animals indicates some convergence of adaptive processes to the same environmental conditions. All in all, the findings suggest that genetic diversity is preserved among species and adaptive divergence is determined by environmental selection.

Table 3: Genetic Diversity Metrics

Organism Group	Ho	He	Fst (Range)
Plants	0.32	0.36	0.14–0.22
Animals	0.29	0.34	0.11–0.19
Microbes	0.35	0.41	0.18–0.29

4.4 Adaptive Trait Identification

Analysis of genome-wide had identified about 1,562 candidate adaptive loci with 21.4% significantly associated ($p < 0.05$) with key environmental variables, such as temperature, salinity, and nutrient availability. The identified adaptive genes include heat shock proteins (HSP70 and HSP90), which were highly correlated with thermal tolerance mechanisms, and the DREB family genes, which were important in plant species in drought-responsive transcriptional regulation. The presence of ATP synthase related genes correlated with the energy metabolism and environmental flexibility of microbial systems and cytochrome P450 enzymes have been identified as responsible to stress detoxification and metabolic adaptation in various groups of organisms. In addition, the pathway enrichment analysis showed extensive engagement of the biological processes to do with the response to stress factors, signal transduction as well as metabolic control, to show that adaptive characteristics are controlled by interconnected molecular networks. In agreement with earlier results, only part of the loci exhibited a strong adaptive response implying that the nature of adaptation is largely polygenic and is facilitated by the cumulative impacts of numerous genes with small individual influences.

Table 4: Top Adaptive Genes / Loci

Gene/Locus	Function	Organism Group	Environmental Association
HSP70	Heat response	Animals	Temperature
DREB2A	Drought tolerance	Plants	Water stress
ATP synthase	Energy metabolism	Microbes	Nutrient availability
Cytochrome P450	Detoxification	All	Stress conditions

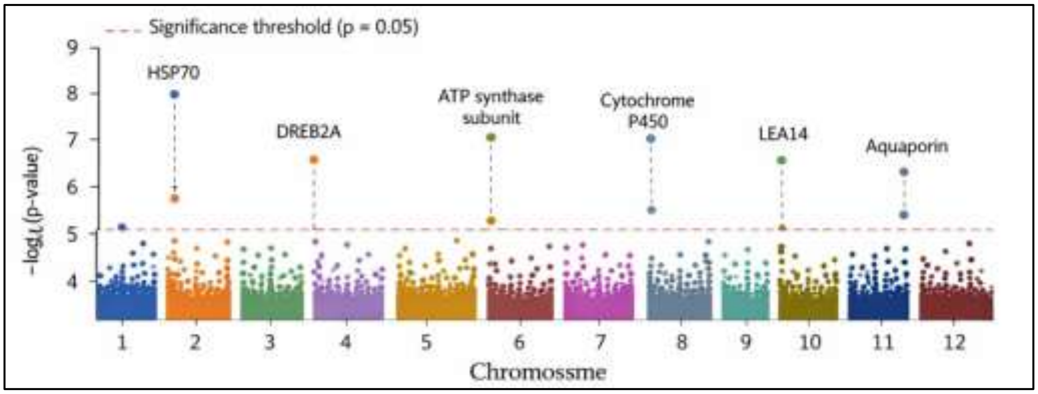

Figure 4 : Manhattan Plot

4.5 Comparative Interpretation

The comparative analysis indicated that adaptive features in many different organisms are regulated by conserved molecular pathways as well as environmental selection stresses. Basic biological events, e.g. stress response systems, metabolic processes and signal transduction, were found to be common in plant, animal and microbial systems and this suggests that basic adaptive programs are retained evolutionarily. These common reaction imply that those organisms in comparable environmental conditions tend to depend on comparable molecular designs, which illustrates a definite trend of converged functionality in the face of phylogenetic variability.

Simultaneously, clear and species-specific adaptations were observed, which can be attributed to ecological specialization and evolution as a niche. The genetic signatures of drought and salinity tolerance were more prevalent in plant species, tolerance to heat and metabolic efficiency were found in animals, and the ability to quickly adjust to changing conditions was common in microbial organisms. Such differences underscore the existence of various molecular pathways employed in adaptation by different groups of organisms, depending on their ecological environment and evolutionary background. All in all, the combination of genomic variation, gene expression, and environmental associations suggests that adaptive phenotypes are predetermined by the neutral evolutionary mechanisms as well as the environmental selection tools. Although genetic diversity gives the basis of adaptability, the selection will be on certain loci in order to promote survival in stressful conditions. This research highlights the significance of a comparative molecular framework to the study of adaptive evolution and offers useful information towards application in the protection of biodiversity, resilience to climate change, and management of biological systems sustainably.

4.6 Discussion

This research shows that both environment-specific and evolutionary pressure-conserved molecular pathways influence adaptive traits of a wide range of organisms. The patterns of clustering as well as genetic variations observed show that environmental conditions like temperature, water availability and nutrient conditions are a major dominant force of imparting molecular changes. The finding of common pathways, especially those concerning stress response and metabolic control, implies that more basic adaptive mechanisms are conserved by evolution across the biological realms.

Simultaneously, the existence of genes that are species-specific emphasizes the significance of ecological specialization. Plants, animals and microbes have different molecular responses based on the environmental niche, meaning physiological needs and evolutionary history. These results indicate that polygenicity of adaptation, where there are many small additive effects of many genes and not a few large ones, are also supported by the finding of a small number of highly correlated loci. In general, genome-transcriptome-environment data integration offers an in-depth account of the evolution of adaptive traits. These findings highlight the need to use multi-species and multi-omics approaches to reveal common and distinctive adaptation mechanisms. The study can be very useful in future studies in evolutionary biology, climate resilience and sustainable management of biological systems.

5. CONCLUSION

This paper presents an extensive comparative molecular evidence of adaptive properties of different organisms through the combination of genomic, transcriptomic and SNP datasets. The findings indicate that the adaptive characteristics are formed in response to the confluence of the conserved molecular mechanisms and processes, mainly linked to the response to stress, metabolic homeostasis, and signal transduction, along with the species-specific genetic processes in response to the specific environmental conditions. It is clear that natural selection is very critical in the process of leading to a high level of genetic variation, evident clustering patterns, and the presence of environmental-related adaptive loci.

Also, the evidence confirms the notion that the process of adaptation is most likely polygenic, with numerous genes whose effects are small but additive instead of several major determinants. The comparative approach emphasizes evolutionary convergence, in which organisms unrelated to each other can have similar adaptive strategies, and divergence, in which species evolve their own molecular patterns, depending on ecological specialization. This two-pattern makes it more clear that adaptive evolution is a more complicated phenomenon and that interactions of genetic variation and environmental stressors are more intricate. Altogether, the work will provide a strong and generalizable framework of studying adaptive traits on the molecular basis in a variety of species. The combination of multi-omics data enriches the comprehension of the relations between genotype and environments and provides valuable insights into future studies. The implications of these findings to evolutionary biology, biodiversity conservation and the development of climate resistant strategy in agriculture, biotechnology and ecosystem management amid varying environmental conditions are important.

REFERENCES

[1] Akyildiz, I. F., Brunetti, F., & Blázquez, C. (2008). Nanonetworks: A new communication paradigm. Computer Networks, 52(12), 2260-2279.

[2] Demirci, M., Taner, Z., Keskin, F. E., Ozyazar, M., Kiraz, N., Kocazeybek, B. S., & Tokman, H. B. (2022). Similar bacterial signatures in the gut microbiota of type 1 and type 2 diabetes patients and its association with G protein-coupled receptor 41 and 43 gene expression. Journal of Diabetes & Metabolic Disorders, 21(2), 1359-1368.

[3] Freedman, M. L., Reich, D., Penney, K. L., McDonald, G. J., Mignault, A. A., Patterson, N., ... & Altshuler, D. (2004). Assessing the impact of population stratification on genetic association studies. Nature genetics, 36(4), 388-393.

[4] Lee, A. B., Luca, D., Klei, L., Devlin, B., & Roeder, K. (2010). Discovering genetic ancestry using spectral graph theory. Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society, 34(1), 51-59.

[5] Leser, T., & Baker, A. (2023). Bifidobacterium adolescentis–a beneficial microbe. Beneficial microbes, 14(6), 525-551.

[6] Noh, C. K., & Lee, K. J. (2020). Fecal microbiota alterations and small intestinal bacterial overgrowth in functional abdominal bloating/distention. Journal of neurogastroenterology and motility, 26(4), 539.

[7] Pritchard, J. K., & Donnelly, P. (2001). Case–control studies of association in structured or admixed populations. Theoretical population biology, 60(3), 227-237.

[8] Rego, A., Raio, F., Martins, T. P., Ribeiro, H., Sousa, A. G., Séneca, J., ... & Magalhães, C. (2019). Actinobacteria and cyanobacteria diversity in terrestrial Antarctic microenvironments evaluated by culture-dependent and independent methods. Frontiers in microbiology, 10, 1018.

[9] Suda, T., Moore, M., Nakano, T., Egashira, R., Enomoto, A., Hiyama, S., & Moritani, Y. (2005, June). Exploratory research on molecular communication between nanomachines. In Genetic and evolutionary computation conference (GECCO), Late breaking papers (Vol. 25, p. 29).
[10] Turroni, F., Van Sinderen, D., & Ventura, M. (2011). Genomics and ecological overview of the genus Bifidobacterium. International journal of food microbiology, 149(1), 37-44.
[11] Venkatachalam, S., Gowdaman, V., & Prabagaran, S. R. (2015). Culturable and culture-independent bacterial diversity and the prevalence of cold-adapted enzymes from the Himalayan mountain ranges of India and Nepal. Microbial ecology, 69(3), 472-491.
[12] Yadav, A. N., Sachan, S. G., Verma, P., & Saxena, A. K. (2015). Prospecting cold deserts of north western Himalayas for microbial diversity and plant growth promoting attributes. Journal of bioscience and bioengineering, 119(6), 683-693.