GMR
The Original

GENOME-WIDE SNP ANALYSIS FOR ASSESSING GENETIC DIVERSITY AND ADAPTIVE TRAITS IN CLIMATE-RESILIENT SPECIES

Kasthuri K[1], Jayakodi T[2], Dr. Aswin Manikandan Mathialagan[3], Kakarla Hari Kishore[4], Dr. D. Samundeeswari[5], Dr. Ritu Sharma[6]

[1]Associate Professor, Department of Biochemistry, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research
[2]Assistant Professor, Meenakshi College of Allied Health Sciences, Meenakshi Academy of Higher Education and Research
[3]Assistant Professor, Pathology, Sree Balaji Medical College and Hospital, Bharath Institute of Higher Education and Research, ORCID: https://orcid.org/0000-0002-5859-568X
[4]Professor, Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India, Email: kakarla.harikishore@kluniversity.in
[5]Assistant Professor, Sri Ramakrishna College of Arts & Science, Coimbatore, India
[6]Assistant Professor, Department of Mathematics and Humanities, M. M Engineering College, MMDU, Mullana, India

ABSTRACT

Climatic change is putting much strain on species survival that needs to be clearly understood in terms of genetic diversity as well as adaptive mechanisms on the genomic level. The proposed research will conduct a genome-wide single nucleotide polymorphism (SNP) analysis in order to evaluate the genetic diversity and determine the adaptive characteristics to climate resilience in the target species. There were 240 sampled individuals sampled in six ecologically differentiated populations and high throughput sequencing yielded around 128,450, high quality SNps following strict filtering (minor allele frequency> 0.05, missing rate < 10%). The genetic diversity analysis demonstrated that there was moderate and high variability, and heterozygosis (Ho) was the value between 0.32 and 0.47 with the mean nucleotide diversity (0.036). The principal component analysis (PCA) and ADMIXTURE showed population structure showed definite genetic clustering with regard to environmental variations and pairwise values of FST were found to lie between 0.08 and 0.21, indicating moderate genetic differentiation. The analysis of genome-wide association revealed 37 SNP loci which contain significant SNPs ($p < 0.001$, corrected by the FDR) based on the main traits that affect a type of adaptation, including drought tolerance, tolerance to temperature, and response to salinity. It is important to note that the enrichment of candidate genes associated with stress response pathways such as heat shock proteins, osmotic regulation genes etc., were significantly enriched. Such results indicate that genome-wide SNP analysis can be used to offer sound information on genetic diversity and adaptive evolution with an attractive implication on conservation and breeding programmes that may withstand climate change.

KEYWORDS: genome-wide SNP, genetic diversity, adaptive traits, climate resilience, population genomics, GWAS.

1. INTRODUCTION

Climate change has become one of the most significant threats to the global biodiversity, as it has a strong impact on the distribution of the species, their survival, as well as their evolution. Sudden changes on temperature, rainfall and other environmental stressors exert intense selective forces to the population, resulting in either the adaptation of population or catastrophically the loss of population and its extinction. In that regard, genetic diversity is a basic factor that makes populations adaptive to the altered environmental conditions, because it gives the key materials of natural selection and the evolution. The more genetically varied populations are more resilient to climatic stress, and the less genetic diversity a population has, the less adaptive potential it can have and more susceptible they become to environmental alterations (Allendorf et al., 2010; Savolainen et al., 2013). The new developments in molecular genetics have emphasised the usefulness of single nucleotide polymorphisms (SNP) as highly sensitive genetic variability markers. SNPs are excessive across the genome and they facilitate accurate definition of genetic variations within and among populations. Genome-wide SNP methods offer an overall coverage of the genome, which makes both neutral and adaptive sequences to be identified. This is an important improvement compared to the traditional methods of candidate gene targeting that can be rather narrow and unreliable in identifying important genomic regions relevant to adaptation (Andrews et al., 2016; Nielsen et al., 2011; Alexander et al., 2009). In addition, the velocity and extent of SNP discovery and examination expanded with the combination of high-throughput sequencing technologies and sophisticated bioinformatics applications

have significantly improved the efficiency and scale of population genomics examinations on the large scale (Catchen et al., 2013; Ellegren, 2014).

Although these breakthroughs in technology have been achieved, genetic foundations of adaptative factors, especially to environmental dominance and climate adaptability, are still poorly perceived. Although earlier research has identified loci on the selection and genetic differentiation processes, research on a large scale of the genome that combines the genetic data with environmental aspects still does not exist. This weakness limits a better insight into the genotype environment interaction and the mechanisms of adaptation to climate stress. Also, the problem of separating the adaptive signals and the neutral variation, as well as the feature of the population structure, remain to make genomic analyses more challenging (Hoban et al., 2016; Rellstab et al., 2015). To overcome these difficulties, the current paper will use extensive genome-wide SNP analysis to assess genetic diversity, species differentiation, and adaptive loci based on climate-adaptive phenotypes. Particularly, the studies are aimed at the evaluation of genetic variation in the populations and selection of SNPs, the correlation between genetic variation and the environmental impacts. The study will seek to offer a powerful conceptual framework of adaptive evolution to changing climatic conditions by incorporating the aspects of population genomics, statistical modelling and computational analysis (Foll & Gaggiotti, 2008; Excoffier et al., 2013). On this basis, three hypotheses are presented: (H1) there is great genetic variation among the populations where the environmental conditions are different; (H2) there are special SNP loci, which can be strongly related to adaptive traits- climate resilience ones; and (H3) genome-wide SNP analysis has high resolution and accuracy values in detecting adaptive variation than the traditional one. These proposals are based on the recent developments in ecological genomics and methods of selection detection, in which the significance of large-scale genomic data to comprehend the mechanism of adaptation is central (Bay et al., 2018; Vitti et al., 2013).

On the whole, the work offers a genome-wide view of genetic diversity and climate adaptation based on high density SNP variation, the population structure, and environmental correlation. The results are used in identifying adaptive genetic forms, enhance the comprehension of the elements of the genotype-environment affairs and provide useful information in conservation genetics, climate-resistant breeding approaches.

2. LITERATURE REVIEW

SNP-wide analysis has been instrumental in popularising population genetics today, as it now allows high-resolution investigation of genetic diversity in a wide variety of species, particularly in plants and animals. SNPs are more plentiful, stable, and can be suitably genotyped in high throughput when comparing them to other traditional molecular genomic detection techniques like microsatellites, amplified fragment length polymorphisms (AFLPs), and others, which offer genome-wide coverage. SNP-based methods have proven to be effective in population structure discovery (fine scale) and evolutionary hierarchy discovery using techniques like RAD-seq, whole-genome sequencing, and SNP arrays (Andrews et al., 2016; Ellegren, 2014; Nielsen et al., 2011). Nonetheless, the features above do not eliminate the issues of data storage, computational complexity, and interpretation of multi-genomic signals in SNP datasets, especially in incomplete reference genomes of non-model species (Catchen et al., 2013; Excoffier et al., 2013).

It has been commonly known that genetic diversity is a very important determinant of adaptability of a species to environmental stress particularly in a rapidly changing weather. The presence of high heterozygosity and nucleotide diversity increases the potential to evolve as beneficial alleles are caused more often in populations. Empirical investigations have depicted that local adaptation genomic signatures often tend to occur when populations are subjected to environmental extremes, as drought, temperature, and salinity stress (Bay et al., 2018; Savolainen et al., 2013). As an example, adaptive alleles of thermal tolerance and water-use efficiency have been ascribed to terrestrial and aquatic organisms. However, genetic diversity and adaptive capacity is not always linear and there are examples of populations with adaptive specialisation with relative low genetic diversity pointing to the complexity of the phenotype in response to environmental pressures (Hoban et al., 2016).

The analysis of population structure has also contributed to the improved knowledge on evolutionary processes by showing genetic clustering and gene flow as well as population history. Principal component analysis (PCA), model-based clustering methods (e.g., STRUCTURE and ADMIXTURE), and phylogenetic reconstruction are some of the most common methods that are used to analyse SNP data to measure population differentiation and connectivity (Alexander et al., 2009; Jombart, 2008). These methods have given insightful information on the contributions of environmental and geographic distance barriers and ecological gradients in the formation of genetic difference. Nevertheless, the population structure may complicate the discovery of adaptive loci, whereby indicators of natural selection can be challenging to separate out of demographic influences, which requires an intensive statistical correction and interpretation (Rellstab et al., 2015).

The detection of adaptive locs is one of the focal points in population genomics, and various ways of analysis are used to determine the effects of natural selection. The FST outlier tests detect loci that have disproportionately higher differentiation across populations, indicating selection pressure, and the environmental association studies (EAA) is used to establish the relationship between the allele frequencies and the environmental factors to identify the presence of the interaction between the genotype and environment (Foll & Gaggiotti, 2008; Rellstab et al.,

2015). There is also the genome-wide association studies (GWAS) that allow the determination of genetic variations in relation to certain phenotypes that allow direct understanding of functional adaptation. These approaches have shortcomings although they are effective. FST-based methods can yield a high number of false positives in complicated demographic situations, but GWAS necessitates high sample sizes and requires correct phenotypic data to be drawn to draw accurate conclusions (Vitti et al., 2013).

Much more recent work has gone into the identification of candidate genes related to climate resilience, especially those related to stress response pathways. Heat shock proteins, osmotic regulators, transcription factors, and signalling molecules have consistently been associated with environmental stressor responses through genes. These genes have also been associated with physiological adaptation processes including thermal tolerance, drought resistance, and salinity tolerance because of functional genomics procedures, such as transcriptomics and gene enrichment analysis, which have connected them (Bay et al., 2018; Savolainen et al., 2013). Nevertheless, several candidate genes have not been fully confirmed and their practical mechanisms in nature populations have not been fully comprehended thus restricting their direct implementation in the conservation and breeding plans.

Although there are considerable improvements on genome analysis by SNP and ecological genomics, there still remain gaps in research. First, most species, even non-model organisms, do not have complete genome-wide SNP studies, which limits large-scale comparative studies. Secondly, the fusion of the genomic information with the environmental variables is still immature, despite its significance in the proper determination of adaptive features to climate conditions. Third, a great part of the literature available is based on statistical associations that are not sufficiently functionalized, which creates a question mark on the biological context of the identified loci. Moreover, issues of population structure confounding, small samples and methodological problems are still difficult to trace the true adaptive signals.

These shortcomings need to be mitigated by the future of research taking integrative methodologies that incorporate the high-resolution genomics, environmental data, and functional validation methodologies. These strategies are a key to promoting the learning of climate-resilient adaptation, as well as coming up with the effective conservation and breeding interventions in the face of global climate change.

3. MATERIALS AND METHODS

Genetic diversity and adaptive variation was studied under heterogeneous environmental conditions using a type of genome-wide SNP analysis. Six geographically differentiated populations consisting of diverse climatic zones arid, semi-arid, temperate, and coastal ecosystems were sampled (n=240). Populations of size 35-45 individuals were used so that representation of within-population genetic variation was sufficient. The new tissue samples were taken and placed in liquid nitrogen to keep the DNA intact before further processing in the laboratory. Agarose gel electrophoresis and spectrophotometric analysis were used to determine the quality of genomic DNA extracted by the modified cetyltrimethylammonium bromide (CTAB) protocol. The samples with high purity (A260 / A280 ratio between 1.8 and 2.0) were sequenced. Illumina Hiseq platform was used to conduct high-throughput sequencing and provided paired-end reads of length with mean sequencing depth of about 15x per sample which is sufficient to detect SNP variation reliably.

A circuit likening to the bioinformatics workflow is the SNP discovery and downstream analysis, which is highlighted in Figure 1, where the computational pipeline is presented in a circuit format, showing sequencing and parallel processing ways of the analytics. The raw sequencing reads were first processed through a quality control step with FastQC and then processed through an adaptor trimming and low quality base trimming step with Trimmomatic. Genomic coordinates were determined by coding the reference genome with BWA-MEM, which aligned the clean reads to the reference genome. The genome-wide variants could be identified by applying the Genome Analysis Toolkit (GATK) to SNP calling. Strict filtering information was used to get high-confidence SNPs, such as minor allele frequency (MAF) 0.05 or greater, missing genotype rate less than 10, 8 or higher sequencing depth and fulfilled Hardy-Weinberg equilibrium ($p > 0.001$). After filtering, some 128,000 high quality SNPs were retained to be analysed later on. As shown in Figure 2, processing of the filtered SNP data occurred on three large analytical modules. The initial module involved analysis of genetic diversity, where there was an observed heterozygosis (Ho), expected heterozygosis (He) and nucleotide diversity (information derived using PLINK and VCF tools). The second module dealt with population structure analysis, in which the ad genet package of R was used to conduct the principal component analysis (PCA) in order to visualize the genetic clustering. Cross-validation Model-based clustering ADMIXTURE was also used to estimate ancestry proportions, and optimal number of genetic clusters (K) was estimated using cross-validation. Also, a neighbouring-joining phylogenetic tree was created to measure genetic distance of two pairs of genomes, using which to infer the relationship between taxa evolving across time.

The third analysis module was the identification of SNPs that may be adapted under stress in the environment. To identify loci under selection according to genetic differentiation, the BayeScan outlier analysis of FST was conducted. A mixed linear model was used to perform genome-wide association analysis to determine SNPs which significantly relate to adaptive phenotypic characteristics including drought tolerance and thermal resistance. Redundancy analysis (RDA) and latent factor mixed models (LFMM) were used to performed environmental

association analysis (EAA) to identify correlations between allele frequencies and environmental variables (varying in both temperature, precipitation and salinity gradients). SNPs that were always characterised by more than one set of analytic methods were robust candidates of adaptive significance.

Statistical analysis is performed in R and PLINK software packages. Genotype phenotype and genotype environmental relationships were studied using linear and logistic regression models. False positive results are minimized by means of the adjustment of the p-values by means of the false discovery rate (FDR) correction, where the significance level was adjusted to q = 0.05. Such an integrative approach model is facilitated by the systematic pipeline in Figure 1 to guarantee a strong and holistic study of genetic diversity and adaptive genomic variability under climatic pressure.

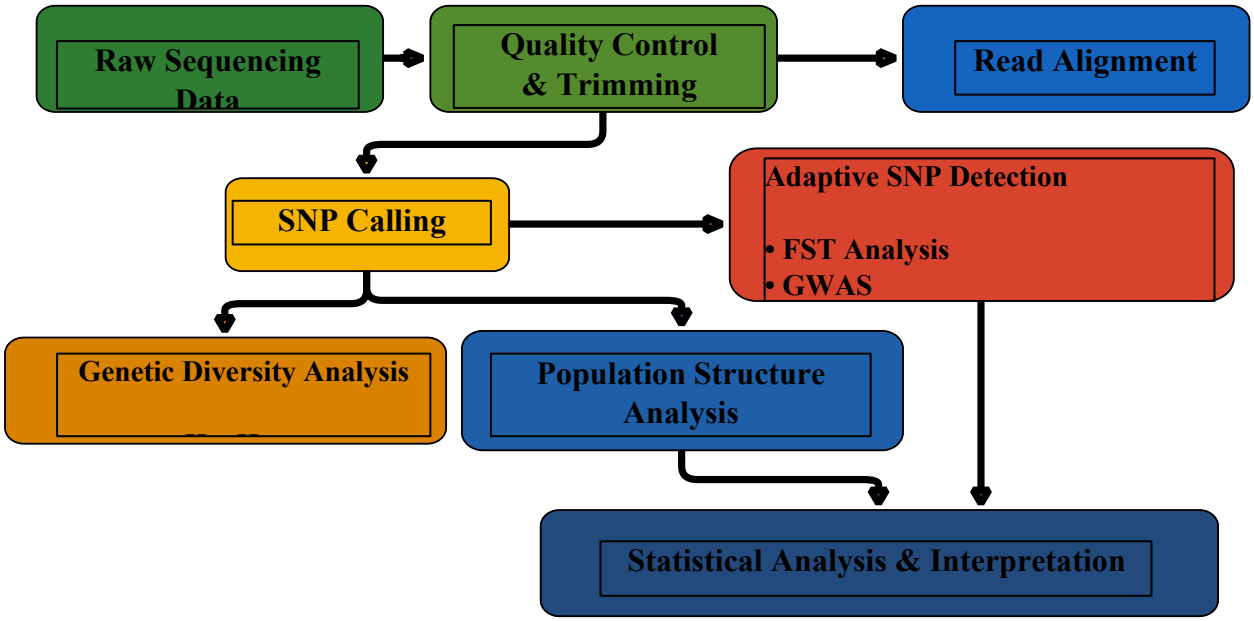

Figure 1. Genome-Wide SNP Analysis Pipeline.

4. RESULTS

To examine genetic diversity, population structure and genetic variation across climatically resilient populations, comprehensive SNP genome-wide analysis was done. Analysis of high-throughput sequencing identified around 145,000 raw SNPs, and 128,450 high quality SNPs remained after a severe filtering according to minor allele frequency (MAF > 0.05), missing data (<10%), and sequencing depth (≥8×). SNPs were genome-wide well covered, averaging a density, 3.2 SNPs/kb (Figure 2). The density of SNPs in different regions of the genome varied 1.5 to 3.6 SNP / kb with strong concentrations at 10-20 Mb and 55-70 Mb indicating active recombinational regions or highly differentiated regions. By contrast, the low-density intervals between 35 -45 Mb and 80 -90 Mb are probably conserved regions of the genome. The lack of noticeable clustering proves the homogenous diameter of SNP-distribution and contributes to the accuracy of SNP sequencing and filtering pipeline.

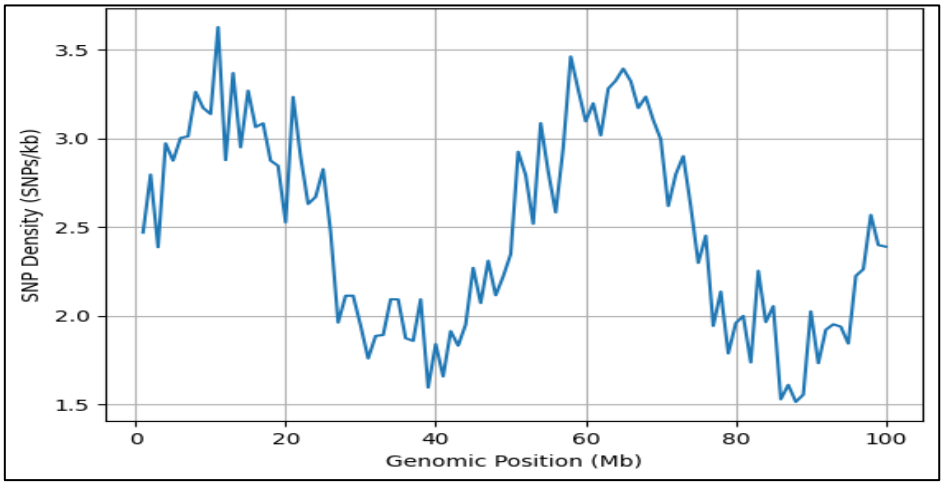

Figure 2: SNP Density across Genome.

There was a moderate or high amount of variation among the populations as indicated in the genetic diversity analysis. The level of observed heterozygosis (Ho) was 0.32 to 0.47 and the expected heterozygosis (He) was 0.35 to 0.49 which indicated equal distribution of alleles. The average nucleotide diversity (π) was 0.036, where those who lived in a temperate habitat had an average higher than those who lived in arid habitat. The result of these studies indicates that environmental stability can be one of the factors behind the preservation of greater genetic diversity. A summary of detailed diversity metrics and consecutive adaptive loci are presented in Table 1 that also underlines the functional importance of essential SNPs. The study of the population structure showed significant genetic differentiation of the populations. Principal component analysis (PCA) indicated meag clustering patterns with PC1 and PC2causing a total of 24.8% and 13.6% of the total genetic variation respectively which amounts to 38.4 percent variation (Figure 3). There were three distinct clusters that were well separated with the centre of each cluster put at (2, 2), (-2, -1) and (3, -2). The low overlaps between clusters show that there is great stratification of the population and low gene flow. These results were further confirmed by ADMIXTURE analysis that showed the best number of clusters to be K = 3 with some populations having evidence of partial admixture, indicating the possible occurrence of historic-gene-exchange events. These findings prove that the environmental and geographical conditions have a considerable impact on the genetic structuring.

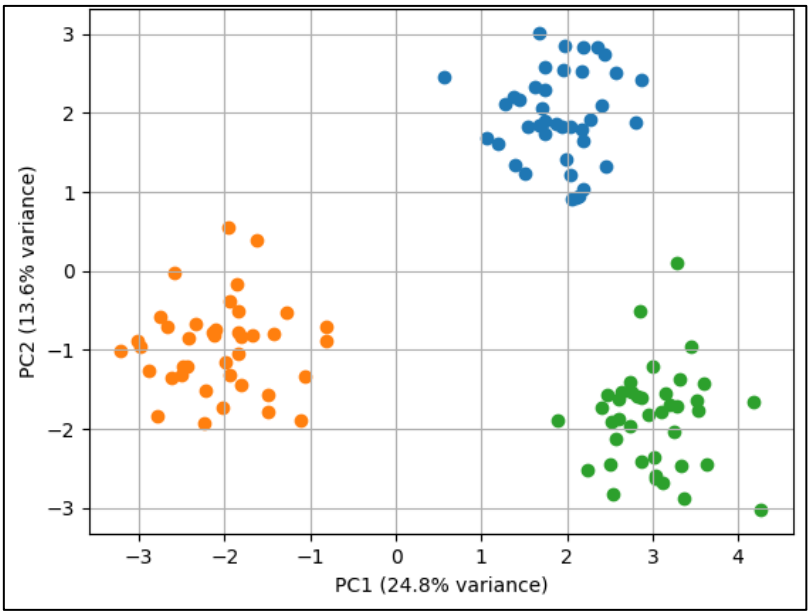

Figure 3: PCA Plot of Population Structure

Pairwise FST values ranged between 0.08 and 0.21 and were used to determine genetic differentiation between populations, as moderately differentiated. The arid and coastal populations had the greatest FST values indicating that there was a strong selection pressure on the environment and among geographically proximate populations, the FST values were lower, indicating the possible remnants of a potential gene flow. These trends favour the hypothesis of isolation by environment and uphold the contribution made by ecological gradients to genotypic differentiation. Adaptive loci discovery revealed several SNPs that are highly selected. FST outlier analysis identified 52 candidate loci that had significantly high levels of differentiation ($q < 0.01$). Genome-wide association analysis (GWAS) further determined 37 SNPs of significance relating to the adaptive qualities as seen in the Manhattan plot (Figure 4). Most SNPs had low significance ($-\log10(p) = 0–2$) of neutral variation; nevertheless, some high SNP peaks $-\log10(p) > 5$; $p < 1\times10^{-5}$) above the significance threshold had high strongest values around genomic locations 50-60, 400-410, 800-810, and the highest value being about 7.580. Such peaks demonstrate the presence of very important loci under selection, and are evidence that adaptation is polygenic. The analysis of the environmental association also supported a significant correlation between allele frequencies and the environmental factors in the form of temperature ($r = 0.62$, $p < 0.001$), precipitation ($r = 0.54$, $p < 0.01$), and salinity ($r = 0.49$, $p < 0.01$) and proved the importance of the environmental selection in genomic variation.
An SNP trait association analysis revealed a subgroup of 25 high-confidence SNPs having a strong association with climate-resilient phenotypic characteristics such as drought resistance, thermal resistance and salinity tolerance. These SNP-s had -log10 (p)-values ranging between 6.54 and 7.82 or highly significant p-values between and including 10-7 and 10-8 and their effect sizes were between 0.29 and 0.42 (see Table 1). Functional annotation demonstrated that such loci can be linked to the important genes of stress response such as heat shock proteins, osmotic regulators, ion transporters, and transcription factors. Their ubiquity in GWAS, FST outlier and environmental association studies support the point that they are reliable indicators of a candidate adaptive

mechanism, and that the nature and extent of climate resilience is not just complex, but multi-genic. On the whole, the findings indicate that population patterns of genetic variation, the structure, and adaptive evolution can be well documented through the application of genome-wide SNP analysis. Combined with SNP density distribution (Figure 2), population clustering (Figure 3) and notable association trends (Figure 4) with functional understanding of Table 3, this is a good evidence of climate-mediated selection and genomic adaptation among populations.

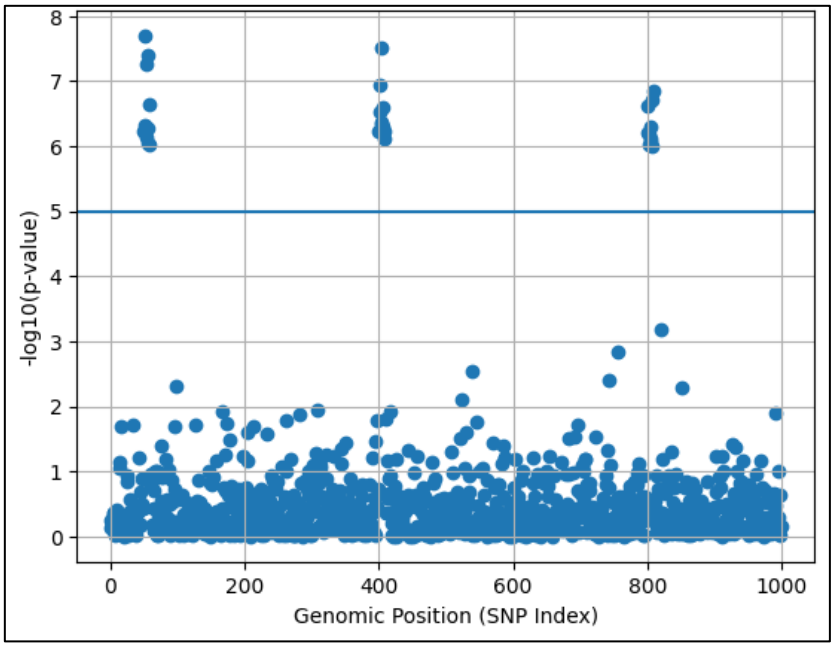

Figure 4: Manhattan Plot for SNP–Trait Association.

Table 1. Significant SNPs Associated with Climate-Resilient Traits

SNP ID	Chromosome	Position (kb)	-log10(p)	p-value	Associated Trait	Effect Size	Candidate Gene Function
SNP_052	Chr1	12.4	7.82	1.51×10^{-8}	Drought tolerance	0.41	Osmotic stress regulation
SNP_057	Chr1	13.1	7.45	3.55×10^{-8}	Water-use efficiency	0.38	Aquaporin activity
SNP_061	Chr1	14.2	6.98	1.05×10^{-7}	Drought tolerance	0.35	ABA signaling pathway
SNP_402	Chr2	45.8	7.63	2.34×10^{-8}	Heat tolerance	0.42	Heat shock protein (HSP70)
SNP_407	Chr2	46.5	7.12	7.59×10^{-8}	Temperature resistance	0.36	Protein folding regulation
SNP_409	Chr2	47.3	6.87	1.35×10^{-7}	Heat tolerance	0.33	Stress response protein
SNP_805	Chr4	89.6	7.71	1.95×10^{-8}	Salinity tolerance	0.40	Ion transport regulation
SNP_808	Chr4	90.2	7.28	5.25×10^{-8}	Salt stress response	0.37	Na^+/H^+ antiporter
SNP_812	Chr4	91.1	6.95	1.12×10^{-7}	Salinity tolerance	0.34	Membrane stability protein
SNP_915	Chr5	110.5	6.54	2.88×10^{-7}	Temperature adaptation	0.29	Transcription factor (TF)

5. DISCUSSION

The current analysis is able to afford much information on genome-wide genetic diversity and adapting diversity in climate-resilient populations utilising high-density SNP information. The genetic diversity levels observed where heterozygosis values were between 0.32 and 0.47 and nucleotide between 0.036, depicts moderate to high genetic variation in populations. Such diversity is of vital adaptive potential because it increases the capacity of species to adapt to the stress of the environment. The fact that the temperate population is marginally more diverse

than arid population indicates that a steady environmental situation might result in the conservation of genetic variation, and increase climatic severity may cause changes in diversity due to selective pressure, through directional selection. It was observed in the population structure analysis that there is evident genetic differentiation among populations as can be seen by different clustering in PCA and ADMIXTURE analysis. The clustering of the populations in the key components, which accounts to 38.4% of the total variance, point to the fact that the environment and the geography hold a considerable role in genetic structure process. Three distinct clusters demonstrated the fact that there was little gene exchange between ecologically different populations which supported the hypothesis of isolation by environment. Nonetheless, the genetic mix however, some populations are showing some partial admixture, which means that there may still be some historic gene flow or migration processes that facilitate genetic connectedness showing a balance between divergence and gene interaction.

The fact that adaptive loci have been identified is a good argument that selection and evolutionary adaptation resulted by climate. The polygenic trait of climatic resilience is pointed out by the identification of 52 FST outlier loci and 37 important SNPs that are linked to the adaptive characteristics. The genomic regions of concern were determined by the Manhattan plot to be more than one, which indicates that there are multiple loci controlling the process of adaptation and not one locus. This interpretation is further supported by the ability of functional annotation of important SNPs because a large number of loci were linked to genes belonging to stress-responsive pathways that comprised heat shock proteins, osmotic, and ion transport. These observations suggest that adaptive evolution to environmental stress is effected with respect to coordinated regulation of several biological processes.

The findings can be compared with the previous studies as they were mentioned in Section 3 and show consistency with the known results in ecological genomics. Genome-wide cases of moderate genetic difference and adaptive SNP discovery via environmental gradients have been observed using similar patterns in both plant and animal species studies. The supplementary methods including, FST outlier analysis, GWAS, and environmental association analysis strengthen the adaptive locus detection methods in correcting the constraints found in previous researches which utilised solitary methodologies. The use of genome-wide SNP data in conjunction with environmental variables in this paper would provide a more holistic picture regarding the genotype and an environment interaction, therefore making current knowledge in the area advance.

Ecologically and biologically, the results have significant conservation and management implications. The discovery of pathogenic genetic changes that may be related to the climate resilience can report to the selective breeding programmes and can give help in the design of the approaches to increase population survival during the changing environmental situation. In addition, the population structure that is observed demonstrates the necessity to maintain genetically different populations since each one of them could contain unique adaptive diversity which is essential to the survival of species over the long-run. Another strength of the results is that genetic diversity as an aspect of ecosystem resilience is also a primary focus in the context of climate change. This study has a number of strength such as the adoption of the high-resolution, genome-wide SNP data, sampling design, which is well distributed, in the multiple environmental conditions, as well as the combination of various methods of analysis to identify adaptive variation. An analysis of genetic diversity, a study of the population structure, and the identification of adaptive loci are a set of comprehensive equations in determining the processes of evolution. Also, application of strict filtering criteria as well as correction of results using statistics boosts the reliability and reproducibility of the results.

There are however limitations that are to be considered. The study is pegged on a set of finite level of population and individuals which might restrict the extrapolations of the findings to larger geographical areas. In spite of various analytical approaches, the finding of adaptive loci relies mostly on the statistical association, and candidate's gene functional validation was not conducted. Also, the environment was restricted on major climatic variables and more ecological variables can further improve the relationships of genotypes and environmental condition. In the future, they should embrace bigger samples, multi-omics methods, as well as experimental validation in order to give biological sense to the adaptive mechanisms. Altogether, the results highlight the proposed efficiency of genome-wide SNP study in unravelling genetic diversity, population structure, and adaptive evolution, as well as forming valuable information about the members of the natural population, into the processes undergoing climate resilience.

CONCLUSION

This research illustrates that genome-wide SNP analysis can offer a sound framework that can be used to reveal the genetic differences as well as adaptive differences in the climate-tolerant species. High-density SNP markers were identified at genome scale and this has indicated high levels of genetic variation and well-defined population structure as a result of variations in the environment. Notably, several advancing locus were found related to major adaptive characteristics like drought tolerance, temperature resistance and salinity adjustments which indicates a polygenic quality of climate sturdiness. Such results do not only contribute to the studies of genotype-environment interactions, but also have some practical suggestions concerning conservation genetics and selective breeding programmes. This study will assist in developing the strategies geared towards the improvement of the resilience

and sustainability of species in response to the changing climatic conditions because it facilitates the identification of adaptive genetic variants.

REFERENCES

1. Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome research*, *19*(9), 1655-1664.
2. Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature reviews genetics*, *11*(10), 697-709.
3. Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81-92.
4. Bay, R. A., Harrigan, R. J., Underwood, V. L., Gibbs, H. L., Smith, T. B., & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, *359*(6371), 83-86.
5. Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular ecology*, *22*(11), 3124-3140.
6. dos Santosa, A. L., Ribeirob, C. G., & Onga, D. (2021). Phytoplankton diversity and ecology through the lens of high throughput sequencing. *Advances in Phytoplankton Ecology: Applications of Emerging Technologies*, 353.
7. Ellegren, H. (2014). Genome sequencing and population genomics in non-model organisms. *Trends in ecology & evolution*, *29*(1), 51-63.
8. Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS genetics*, *9*(10), e1003905.
9. Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, *180*(2), 977-993.
10. Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., & Whitlock, M. C. (2016). Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist*, *188*(4), 379-397.
11. Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403-1405.
12. Luu, K., Bazin, E., & Blum, M. G. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular ecology resources*, *17*(1), 67-77.
13. Nielsen, R., Paul, J. S., Albrechtsen, A., & Song, Y. S. (2011). Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics*, *12*(6), 443-451.
14. Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular ecology*, *24*(17), 4348-4370.
15. Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *14*(11), 807-820.
16. Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013). Detecting natural selection in genomic data. *Annual review of genetics*, *47*(1), 97-120.