Eelgrass Zostera capensis populations in KwaZulu-Natal, South Africa, harbour distinct genomic signals despite limited geographical distance

S. Engela Smit[1], Romina Henriques[2], Megan Jackson[1], Ricky Taylor[3], Leon Vivier[4], Sophie von der Heyden[1,5,*]

[1] Department of Botany and Zoology, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa

[2] Marine Genomics Group, Department of Biochemistry, Genetics and Microbiology, University of Pretoria. Private Bag X20, Hatfield 0028, South Africa

[3] Independent ecologist, 25 Station Road, Mtunzini, 3867, South Africa

[4] Department of Zoology, University of Zululand, KwaDlangezwa, 3886, South Africa

[5] School of Climate Studies, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa

*Corresponding author: svdh@sun.ac.za

Abstract

Seagrasses are threatened by anthropogenic stressors and climate change, with numerous population declines reported. In South Africa, *Zostera capensis* is restricted to estuarine environments and has a disjunct distribution and declining status. With the majority of the distribution of *Z. capensis* on the west and south-west coasts of South Africa, the isolated, eastern-most populations in KwaZulu-Natal (KZN) are of particular interest. Following the extirpation of *Z. capensis* at Durban and St Lucia, only five populations remain, of which three (in the Amatikulu, Mlalazi, Mhlatuze) are situated <50km apart. Previous molecular analyses have showed strong population structure between *Z. capensis* populations, but this has not included these geographically close populations. In this study, using Single Nucleotide Polymorphisms (SNPs) isolated from 31 individuals collected from the three northern KZN estuaries we provide evidence for distinct population clusters, with unique evolutionary signals. *Zostera capensis* in the Mlalazi estuary have low levels of genomic diversity, likely as a result of a small, dynamic population unable to persist prolonged freshwater exposure. Our results suggest that conservation efforts need to consider unique population signals even in geographically close populations, in particular within the context of restoration, where genomic compatibility may determine persistence of restored populations.

Keywords: seagrass, restoration, evolutionary dynamics, persistence, conservation

Introduction

Seagrass populations are in decline globally (Dunic et al. 2021). In South Africa, populations of seagrasses have been extirpated at St Lucia and Durban, with decreases in abundance reported across numerous other locations (Adams 2016). This has been largely driven by changes in salinity levels, such as in St Lucia, but other anthropogenic impacts have been reported as additional factors, including eutrophication, direct habitat disturbance such as through bait digging, and pesticide pollution (Adams 2016; van Wyk et al. 2022). Declines in seagrass meadows negatively impact seagrass-associated biodiversity (Pillay et al. 2010) and have also been linked to loss of evolutionary potential (Phair et al. 2020), which in turn may decrease the resistance of Z. capensis meadows to ongoing and future environmental and climate changes. Estimating metrics such as genomic diversity and the levels of population divergence is crucial for the conservation and management of natural populations (von der Heyden 2009; Phair et al. 2021; Nielsen et al. 2023), especially for coastal species like seagrasses which are increasingly under threat. In addition, successful restoration of seagrass populations may depend on choosing appropriate donor and recipient sites, not only for environmental and ecological similarity, but also for genomic compatability (Sinclair et al. 2013; Jahnke et al. 2015; Pazzaglia et al. 2021).

The Cape eelgrass *Zostera capensis* is the most abundant seagrass species in South Africa, with a distribution ranging from the cool-temperate west coast, to warm-temperate and sub-tropical estuaries on the east coast (Adams 2016; Phair et al. 2019). Unlike other seagrasses, *Z. capensis* is only found in low flow estuarine environments, across ~62 estuaries in the region, with highly fragmented populations. This likely limits exchange of plant propagules across sites, potentially leading to increased population structuring and elevated levels of inbreeding. For example, using frequency differences of putative outlier loci (that may reflect local adaption and selection) and neutral Single Nucleotide Polymorphisms (SNPs), Phair et al. (2019) provided evidence for two distinct genomic lineages of *Z. capensis* across southern and east Africa*,* likely shaped by historical, environmental and habitat suitability. Phair et al. (2019) also showed generally low genomic heterozygosity and that populations in the eastern lineage were more distantly related than those in the western lineage, which may reflect refugial dynamics and population stability (Phair et al. 2019). Although this provided insights across the entire range of *Z. capensis*, small spatial scale dynamics remain unknown. For instance, in KZN, *Z. capensis* is mostly absent due to flow dynamics (Adams 2016), with

only five populations remaining in the Amatikulu, Mlalazi, Mhlatuze, Richards Bay and Kosi estuaries, since the loss of the Durban and St Lucia populations. The closest populations of *Z. capensis* to the ones in KZN are considerable distances away, being found in Maputo Bay Mozambique and the Mnyameni Estuary in the Eastern Cape of South Africa. As such, the populations in KZN are likely geographically isolated and require additional insights to ensure that the potentially unique evolutionary dynamics of *Z. capensis* in the region are maintained.

This study used SNPS isolated through reduced representation sequencing, we examined population-level dynamics across three geographically close populations of *Z. capensis* in KZN (Amatikulu, Mlalazi, Mhlatuze; Figure 1) that are ~50km apart, representing an ideal study system to better understand genomic patterns of *Z. capensis* at small spatial scales. Notably, population sizes and stability of *Z. capensis* differ between these estuaries with more stable and larger populations in Mhlatuze and Amatikulu, and a smaller, transient, population in Mlalazi. Further, *Z. capensis* plugs were introduced to Mhlathuze in the early 1990s from St Lucia (R. Taylor, pers. comm), some of which may represent the extant population of *Z. capensis* in Mhlatuze. We hypothesise that despite their geographic proximity, locations will show some level of genomic differentiation due to a lack of contemporary gene flow with unique SNPs found in each population and that genomic diversity is reflected in population size, with Mlalazi having the lowest genomic diversity of the populations.

Materials and Methods

Sample collection

Samples were collected from the Amatikulu (n = 20), Mlalazi (n = 4) and Mhlatuze (n = 25) estuaries (Table 1), as well as from the Kosi Estuary (n = 15) (Figure 1), with plants collected between 20 - 2500 m apart to minimize collection of clones (Phair et al. 2019). Each sample was cleaned with water to remove epiphytes and placed with silica beads in labelled envelopes. Upon molecular examination of SNP data (unpubl. data), it was found that plants collected from Kosi Estuary may not be *Z. capensis* and were not included in further analyses.

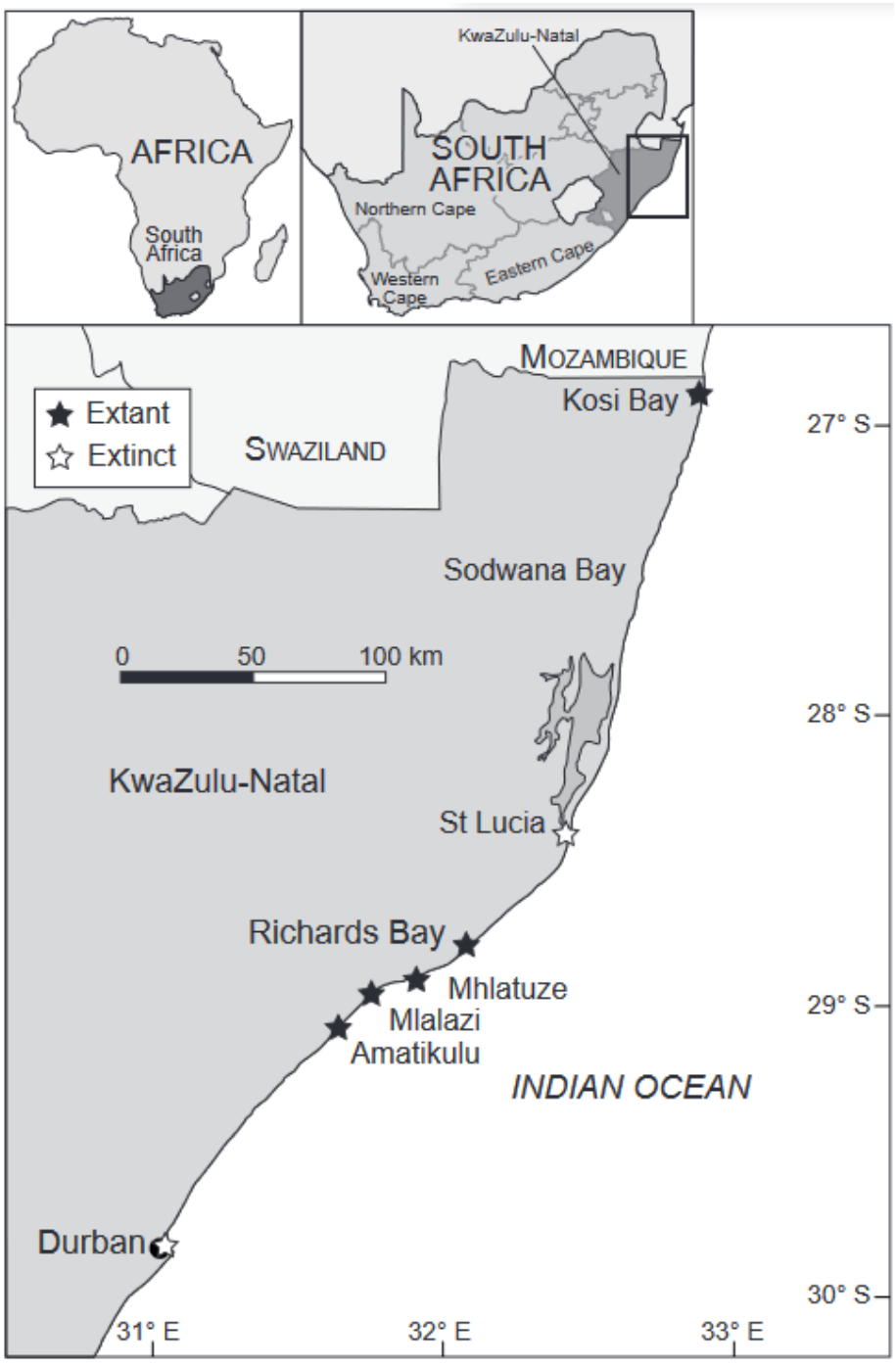

Figure 1. Map showing sampling locations of eelgrass Zostera capensis (a) within South Africa, and (b) in KwaZulu-Natal Province. Symbols refer to extant and extinct population.

Table 1: Sampling locations with sample numbers (*n*) and summary statistics: number of SNPs (Nb of SNPs), number of private SNPs (Private SNPs), expected heterozygosity (H_e), observed heterozygosity (H_o), nucleotide diversity (π), and inbreeding coefficient (F_{IS}) for each site.

Sampling site	*n*	Nb of SNPs	Private SNPs	H_e	H_o	π	F_{IS}
Mhlatuze (HLA)	12	1675	105	0.32(±0.004)	0.68 (±0.004)	0.34 (±0.004)	-0.030
Mlalazi (MLA)	4	452	2	0.12 (±0.005)	0.88 (±0.005)	0.16 (+0.007)	-0.818
Amatikulu (AMT)	15	1689	213	0.31 (±0.004)	0.69 (±0.004)	0.32 (±0.004)	0.042*

*Significantly different from zero

DNA extraction, library preparation and sequencing

All samples were individually treated with liquid nitrogen, and a mortar and pestle were used to disrupt tissues; DNA was extracted using the E.Z.N.A. Plant DNA DS Kit (Omega Bio-tek), following modifications to the protocol outlined in Jackson (2022). The quality and integrity of DNA extractions were checked using agarose gels and Qubit fluorometry. DNA extracts were sent to LGC Biosearch Technologies (Berlin, Germany) for library preparation and Illumina sequencing, where a genotyping by sequencing (GBS) approach was used to generate 150 bp paired-end reads using the PstI and MseI enzymes, with sequencing carried out in one lane of an Illumina NextSeq 500/550.

Bioinformatic analyses

Data was analysed in STACKS (Catchen et al. 2011, 2013) following the same approach and parameter optimization protocols as in Jackson (2022). This was done in two ways: 1) mapping reads against the genome of *Zostera marina* and 2) using a *de novo* assembly in order to better understand the impact of mapping to a cross-species reference genome. As some individuals lost up to 90% of their sequencing data when mapped to *Z. marina* (Supplementary Materials Table 1), all analyses were carried out using the *de novo* assembly. SNP calling was done through the *populations* programme in STACKS with the following filtering parameters: minimum minor allele count = 3, minimum 50% of individuals across a population had to possess a locus for it to be processed and only one random SNP per locus was selected. The

populations programme was used to calculate expected heterozygosity (H_E), observed heterozygosity (H_O), nucleotide diversity (π), the inbreeding coefficient (F_{IS}), population differentiation (F_{ST}) and to generate SNP data in a Variant Call Format (VCF) which was used in downstream analyses. As there were uneven sample sizes (with only four individuals obtained from Mlalazi) some of the analyses, such as F_{ST} and Principal Components Analysis (PCA), were also performed using a subset of the data with four individuals from each population to account for differences in sample size.

Population differentiation and outlier loci detection

A Principal Component Analysis (PCA) was carried out with the *PCAdapt* package (Luu et al. 2017) in R v4.1.1 (R Core Team, 2021) to determine divergence between populations based on the estimated genetic distances. In addition, pairwise F_{ST} values were estimated in *populations* and levels of clonality determined in the R package *poppr* v2.9.3 (Kamvar et al. 2015). To estimate admixture coefficients, a sparse Non-Negative Matrix Factorization (sNMF) algorithm, using the R package *LEA* (Frichot and Francois 2015), was employed. This function estimates the ancestry proportions of each individual, which relates to a specific number of ancestral gene pools (Boehm et al. 2015). Lastly, outlier loci were identified using the package *PCAdapt* (Luu *et al.* 2017). Putative outlier loci were removed from the dataset and a PCA was carried out to determine if the same patterns found in the presence of outlier loci were observed in the neutral dataset.

Results

A total of 158 440 758 raw reads were retrieved, with 109 708 057 reads retained post-filtering. A total of 6 430 194 reads aligned to the *Z. marina* genome, with 581 SNPs retained, and a mean depth coverage of 1.3X. In contrast, a total of 10 992 252 reads mapped to the *de novo* assembly, with 7 758 SNPs retained, and a mean depth coverage of 18.4X (Supplementary Table S1. As such, the *de novo* assembly was used for all downstream analyses.

No clones were detected. The Mlalazi population showed the lowest diversity metrics (H_e, nucleotide diversity) compared to seagrasses in the other two locations. Positive but low F_{IS} values were observed for Mhlatuze and Amatikulu, in contrast to plants from Mlalazi with a negative F_{IS} close to zero (Table 1). The F_{ST} analysis revealed significant differentiation

between all populations and ranged from 0.07 to 0.152, with the PCA confirming that individuals from each population cluster together and separately per location (Figure 2).). The sNMF analysis provided almost equal support for two or three clusters (Supplementary Figure S1), but with some admixture between populations detected (Figure 3). In all analyses, the Mlalazi population was the most distinct, but all populations had private SNPs (2 to 213; Table 1). Putative outlier analyses identified 33 loci as outliers, hereby denominated the "adaptive" dataset. However, comparison of the full data set with the "adaptive" dataset revealed no difference in the patterns of population structure (Supplementary Materials Figure 2), nor did they match the previously identified private SNPs (Phair et al. 2019). Results from the reduced dataset with only four individuals per population did not differ to those from the full data set with 31 individuals (Supplementary Materials Figure 2).

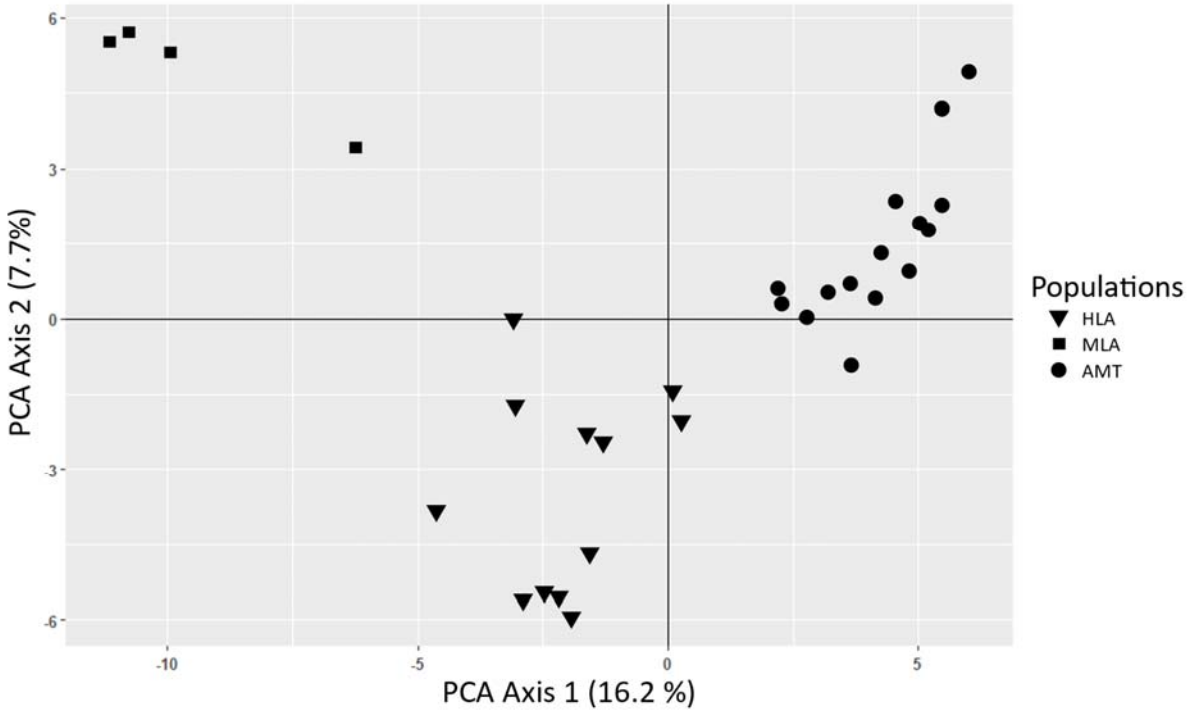

Figure 2: Clustering Principal Component Analysis plots showing distinct population structure of *Z. capensis* in KwaZulu-Natal. Refer to Table 1 for sample site abbreviations.

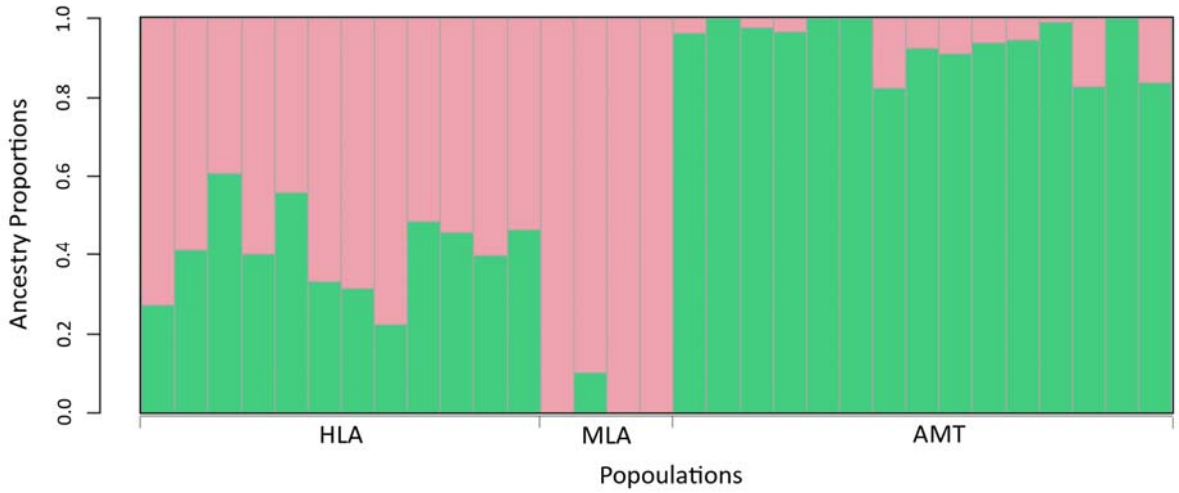

Figure 3: Admixture plot, where each line represents one individual, with individuals grouped by population. Refer to Table 1 for sample site abbreviations.

Discussion

In South Africa, population genetic structuring has been reported for many coastal species (Dalongeville et al. 2022), which often aligns with regional biogeographic patterns (Teske et al. 2011). This is not surprising, given that strong environmental gradients influence the coastal environment, resulting in a dynamic interplay of biotic and abiotic factors that shape distributions of diverse local communities (Griffiths et al. 2010). These factors include contemporary and historical aspects of habitat availability (Toms et al. 2014; Phair et al. 2019), as well as water temperature and salinity (Teske et al. 2019; Nielsen et al. 2020). However, known evolutionary dynamics stem primarily from large-scale studies that consider individuals sampled hundreds of kilometres apart, with finer-scale levels of population genomic differentiation poorly understood. Using the endangered seagrass Z. capensis as a model, we show population differentiation across only ~50 km of distribution, with two distinct population clusters identified in the Amatikulu and the Mlalazi estuaries, with admixture in the Mhlatuze Estuary. This pattern is probably not only driven by unique genomic signals, as there were few private SNPs at each population, but is likely largely due to frequency differences at both neutral and putative outlier loci between populations. This most likely reflects historical divergence that is maintained by a lack of ongoing contemporary gene flow, given the uncertainty of whether the thin blades of seagrass can survive transport in

inshore marine conditions. Other studies globally have also shown small-scale differences in genomic structure among seagrasses (e.g. Halophila ovalis: Liu and Hsu 2021; Z. marina: Kamel et al. 2012; Z. muelleri: Sherman et al. 2016), suggesting that small spatial genomic structure is prevalent in this group, despite the potential for high dispersal, either through vegetative matter or even seeds. This finding may also be linked to the smaller sample sizes from Mlalazi Estuary, both for missing data and the number of individuals included in the study, although the patterns of distinct population structure remained when only four individuals per population were analysed.

A second important finding of our work is the different levels of genomic diversity across estuaries, with Mlalazi showing significantly lower levels of diversity than those elsewhere. At the time of collection, only four plants were evident in the estuary (R. Taylor, pers. comm), with only a few small beds of *Z. capensis* in general (Adams 2016). The population in the Mlalazi is particularly dynamic and may not be detected for a few years at a time when conditions are dominated by freshwater. As such, any population increase may only come from a few propagules, with the low genomic diversity a signal of a population bottleneck comprising of closely related individuals. Genomic diversity has long been considered important for the persistence of species and their ability to adapt to changing environmental conditions (DeWoody et al. 2021; Nielsen et al. 2023) and in seagrasses, it has been shown to increase both resilience and resistance to change (Phair et al. 2020). In South Africa, genomic erosion has been documented for some seagrass populations, thereby threatening the evolutionary dynamics and persistence of these keystone species in the region (Phair et al. 2020). Further, *Z. capensis* exposed to thermal stress show a variety of responses in gene expression (Ndhlovu & von der Heyden 2022), further highlighting the need for maintaining genomic diversity to help natural populations adapt to changing conditions. However, disentangling population declines from natural dynamics, where seagrass cover fluctuates in response to, for example, salinity and turbidity in response to rainfall and freshwater flow regimes, is difficult, particularly for estuarine systems with temporary mouth closures (Adams 2016; Mokumo et al. 2023).

Our findings have significant implications for both the management and potential restoration of Z. capensis in KZN. First, from a conservation perspective, the finding of distinct population clusters suggests that each Z. capensis population should be managed separately rather than as a combined management unit, and that each estuary, with its distinct threats

and pressures, should be evaluated individually to ensure the persistence of this seagrass. Second, maintaining and allowing for natural processes to potentially increase genomic diversity of seagrasses (such as through somatic mutations, as documented in Z. marina [Yu et al. 2020]), by preventing loss from anthropogenic disturbances, is critical. Given that Z. capensis can survive challenging environmental conditions and exhibits dynamic responses to environmental changes, even if in only small populations, conservation agencies should wait several years before declaring Z. capensis extinct in any particular estuary—and only then consider the introduction of plants from another estuary. For example, St Lucia has been without any detected Z. capensis for several years and may be a candidate for restoration once conditions are suitable. However, this would require careful consideration of donor sites (potentially the Mhlatuze Estuary, which may still harbour genomic signals from St Lucia transplants in the 1990s, although there are no records of where seagrasses were transplanted to and whether they survived, to the best of our knowledge). Overall, the monitoring of seagrasses in South Africa needs to be strengthened considerably, which will allow us to better understand individual population dynamics and make stronger evaluations of their threat status.

Third, restoration, particularly the movement of donor plant material to other sites, must consider not only ecological and environmental but also genomic aspects, given the importance of genomic similarity in transplant success (e.g. Sinclair et al. 2013; Jahnke et al. 2015). Restoration trials are not yet common for Z. capensis, and given the limited success of transplant experiments within (Mokumo et al. 2023) and between (Adams 2016) populations in South Africa, as well as the unique evolutionary dynamics of Z. capensis populations as characterised to date (Phair et al. 2019; Jackson 2022), genomic profiling of potential donor and recipient sites will be crucial to maximise transplant success and thus safeguard the long-term persistence of this valuable ecosystem engineer.

Acknowledgements

We thank Dr A. Ndhlovu for supporting the bioinformatic analyses. We also thank Dr Nasreen Peer, Jamila Janna, Courtney Fagg and Jacques Joubert for their assistance with sample collection in KZN. This project was funded by the NRF South Africa through a Marine and Coastal Research grant, UID 116048, awarded to SvdH. Thanks to the Department of Fisheries,

Forestry, and the Environment for providing the collection permit (RES 2021/68). Thanks to TransNet Port Authority for granting access to the Mhlathuze Estuary.

Data availability statement

Raw reads and associated code are available via Data Dryad at https://datadryad.org/stash/share/nTww7slTPVNzFSxUPosVvsDz343fggcZ3T2ObBSjw.

References

Adams JB. 2016. Distribution and status of *Zostera capensis* in South African estuaries — A review. *South African Journal of Botany* 107: 63-73.

Boehm JT, Waldman J, Robinson J D, Hickerson MJ. 2015. Population genomics reveals seahorses (Hippocampus erectus) of the western mid-Atlantic coast to be residents rather than vagrants. *PloS One*, 10(1), e0116219.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping Loci de novo from short-read sequences. G3 (Bethesda, Md.), 1(3): 171–182.

Dalongville A, Nielsen ES, Teske PR, von der Heyden S. 2022. Comparative phylogeography in a marine biodiversity hotspot provides novel insights into evolutionary processes across the Atlantic-Indian Ocean transition. *Diversity and Distributions* 28**:** 2622-2636

DeWoody JA, Harder AM, Mathur S, Willoughby JR. 2021. The long-standing significance of genetic diversity in conservation. *Molecular Ecology* 30: 4147-4154

Dunic JC, Brown CJ, Connolly RM, Turschwell MP, Côté IM. 2021. Long-term declines and recovery of meadow area across the world's seagrass regions. *Global Change Biology* 27: 4096-4109

Griffiths CL, Robinson TB, Lange L, Mead A. 2010. Marine biodiversity in South Africa: an evaluation of current states of knowledge. *PLoS One* 5(8), e12008.

Frichot E, Francois O. 2015. LEA: An R Package for Landscap and Ecological Association Studies. *Universit´e Grenoble-Alpes.*

Jackson MJ. 2022. Investigating fine-scale population dynamics in the threatened Cape Dwarf eelgrass, *Zostera capensis*, using RADseq. *MSc thesis, University of Stellenbosch, South Africa.*

Jahnke M, Olsen JL, Procaccini G. 2015. A meta-analysis reveals a positive correlation between genetic diversity metrics and environmental status in the long-lived seagrass *Posidonia oceanica*. Mol Ecol, 24: 2336-2348. https://doi.org/10.1111/mec.13174

Kamel SJ, Hughes AR, Grosberg RK, Stachowicz JJ. 2012. Fine-scale genetic structure and relatedness in the eelgrass *Zostera marina*. *Marine Ecology Progress Series* 447: 127-137

Kamvar ZN, Brooks JC, Grünwald NJ. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics* 6: 208.

Liu SYV, Hsu C-H. 2023. Genetic analyses reveal fine-scale genetic structure between lagoon and open water spoon seagrass (*Halophila ovalis*) populations around Dongsha Island. *Aquatic Botany* 174: 103421

Luu K, Bazin E, Blum MG. 2017. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular ecology resources*:*17*(1), 67–77. https://doi.org/10.1111/1755-0998.12592

Mokumo MF, Adams JB, von der Heyden S. 2023. Investigating transplantation as a mechanism for seagrass restoration in South Africa. *Restoration Ecology*, in press

Ndhlovu A & von der Heyden S. 2022. *De novo* assembly and annotation of the transcriptome of the endangered seagrass *Zostera capensis*: insights from differential gene expression under thermal stress. *Marine Genomics* 66: 100984

Nielsen ES, Henriques R, Beger M, Toonen RJ, von der Heyden S. 2020. Multi-model seascape genomics identifies distinct drivers of selection among sympatric marine species. *BMC Evolutionary Biology* 20: 121

Nielsen ES, Hanson JO, Carvalho SB, Beger M, Henriques R, Kershaw F, von der Heyden S. 2023. Molecular ecology meets systematic conservation planning. *Trends in Ecology and Evolution*, 38**:** 143-155

Pazzaglia J, Nguyen HM, Santillán-Sarmiento A, Ruocco M, Dattolo E, Marín-Guirao L, Procaccini G. 2021. The Genetic Component of Seagrass Restoration: What We Know and the Way Forwards. *Water: 13*, 829.

Phair NL, Toonen RJ, Knapp I, von der Heyden S. 2019. Shared genomic outliers across two divergent population clusters of a highly threatened seagrass. *PeerJ*, **7:**e6806

Phair NL, Toonen RJ, Knapp IS, von der Heyden S. 2020. Anthropogenic pressures negatively impact genomic diversity of the vulnerable seagrass Zostera capensis. *Journal of Environmental Management* 255: 109831

Phair NL, Nielsen ES, von der Heyden S. 2021. Applying genomic data to seagrass conservation. *Biodiversity and Conservation* 30: 2079-2096

Pillay D, Branch GM, Griffiths CL, Williams C, Prinsloo A. 2010. Ecosystem change in a South African marine reserve (1960–2009): role of seagrass loss and anthropogenic disturbance. *Marine Ecology Progress Series* 415:35-48.

Scherman CDH, York PH, Smith TM, Macreadie PI. 2016. Fine-scale patterns of genetic variation in widespread clonal seagrass species. *Marine Biology* 163: 82

Sinclair EA, Verduin J, Krauss SL, HardingeJ, Anthony J, Kendrick GA. 2013. A genetic assessment of a successful seagrass meadow (*Posidonia australis*) restoration trial. *Ecological Management & Restoration*, 14: 68-71

Teske PR, von der Heyden S, McQuaid CD, Barker NP. 2011. A review of marine phylogeography in southern Africa. *South African Journal of Science 107, Art #514*

Teske PR, Sandoval-Castillo J, Rao Golla T, Emami-Khoyi A, Tine M, von der Heyden S, Beheregaray LB. 2019. Thermal selection as a driver of marine ecological speciation. *Proceedings of the Royal Society B*, 286: 20182023

Toms JA, Compton JC, Smale M, von der Heyden S. 2014. Variation in paleo-shorelines explains contemporary population genetic patterns of rocky shore species. *Biology Letters,* 10: 20140330

van Wyk JW, Adams AB, von der Heyden S. 2022. Conservation implications of herbicides on seagrasses: sublethal glyphosate exposure decreases fitness in the endangered *Zostera capensis*. *PeerJ*, 10:e14295

von der Heyden S. 2009. Why do we need to integrate population genetics into South African Marine Protected Area planning? *African Journal of Marine Science* 31: 263-269