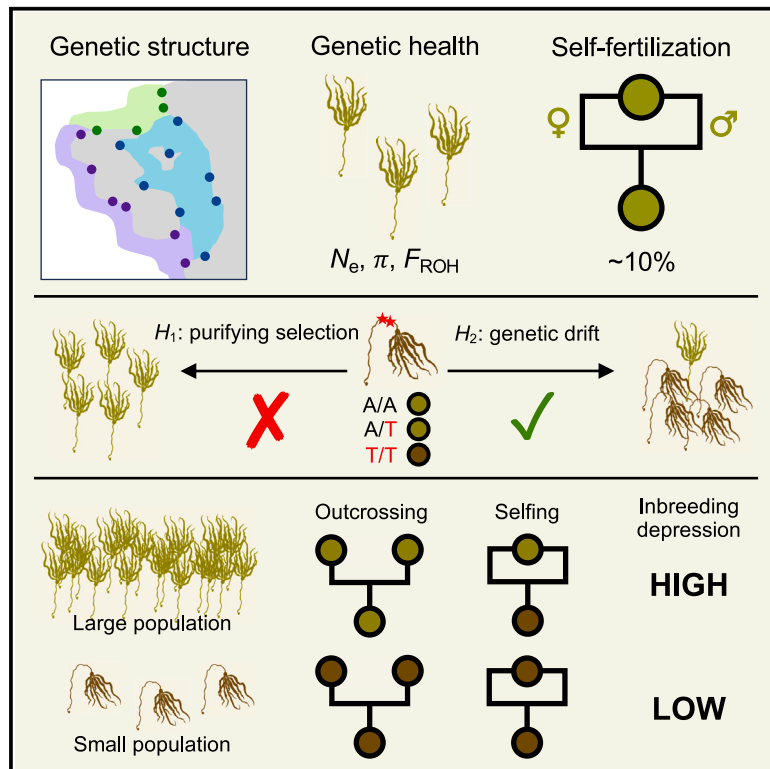# Current Biology

Population genomics reveals strong impacts of genetic drift without purging and guides conservation of bull and giant kelp

Graphical abstract

Authors

Jordan B. Bemmels, Samuel Starko,
Brooke L. Weigel, ..., Jonathan E. Page,
Christopher J. Neufeld,
Gregory L. Owens

Correspondence

jbemmels@umich.edu (J.B.B.),
grego@uvic.ca (G.L.O.)

In brief

Bemmels et al. sequence >600 whole genomes of two canopy-forming kelp species undergoing decline. Small populations experience multiple genetic health risks but no purging of harmful alleles. Instead, genetic drift is likely reducing inbreeding depression. Evolutionary dynamics and genetic structure suggest conservation strategies for these taxa.

Highlights

- Small populations of bull and giant kelp face multiple genetic health risks

- Genetic drift, but not purging, strongly impacts evolution in small populations

- Strong genetic drift has reduced inbreeding depression within small populations

- Genomics insights inform conservation and restoration of declining kelp species

CellPress

Current Biology

 CellPress

Report

Population genomics reveals strong impacts of genetic drift without purging and guides conservation of bull and giant kelp

Jordan B. Bemmels,[1,2,10,*] Samuel Starko,[1,2,3] Brooke L. Weigel,[4,5] Kaede Hirabayashi,[1,6] Alex Pinch,[1] Cassandra Elphinstone,[7] Megan N. Dethier,[4] Loren H. Rieseberg,[2,7] Jonathan E. Page,[7] Christopher J. Neufeld,[2,8] and Gregory L. Owens[1,2,9,*]

[1]University of Victoria, Department of Biology, Finnerty Road, Victoria, BC V8P 5C2, Canada
[2]The Kelp Rescue Initiative, Bamfield Marine Sciences Centre, Pachena Road, Bamfield, BC V0R 1B0, Canada
[3]University of Western Australia, School of Biological Sciences, Stirling Highway, Crawley, WA 6009, Australia
[4]University of Washington, Friday Harbor Laboratories, University Road, Friday Harbor, WA 98250, USA
[5]Western Washington University, College of the Environment, High Street, Bellingham, WA 98225, USA
[6]University of British Columbia, Michael Smith Laboratories, East Mall, Vancouver, BC V6T 1Z4, Canada
[7]University of British Columbia, Department of Botany, University Boulevard, Vancouver, BC V6T 1Z4, Canada
[8]University of British Columbia Okanagan, Department of Biology, University Way, Kelowna, BC V1V 1V7, Canada
[9]Senior author
[10]Lead contact
*Correspondence: jbemmels@umich.edu (J.B.B.), grego@uvic.ca (G.L.O.)
https://doi.org/10.1016/j.cub.2024.12.025

SUMMARY

Kelp forests are declining in many parts of the northeast Pacific.[1–4] In small populations, genetic drift can reduce adaptive variation and increase fixation of recessive deleterious alleles,[5–7] but natural selection may purge harmful variants.[8–10] To understand evolutionary dynamics and inform restoration strategies, we investigated genetic structure and the outcomes of genetic drift and purging by sequencing the genomes of 429 bull kelp (*Nereocystis luetkeana*) and 211 giant kelp (*Macrocystis* sp.) from the coastlines of British Columbia and Washington. We identified 6 to 7 geographically and genetically distinct clusters in each species. Low effective population size was associated with low genetic diversity and high inbreeding coefficients (including increased selfing rates), with extreme variation in these genetic health indices among bull kelp populations but more moderate variation in giant kelp. We found no evidence that natural selection is purging putative recessive deleterious alleles in either species. Instead, genetic drift has fixed many such alleles in small populations of bull kelp, leading us to predict (1) reduced within-population inbreeding depression in small populations, which may be associated with an observed shift toward increased selfing rate, and (2) hybrid vigor in crosses between small populations. Our genomic findings imply several strategies for optimal sourcing and crossing of populations for restoration and aquaculture, but these require experimental validation. Overall, our work reveals strong genetic structure and suggests that conservation strategies should consider the multiple health risks faced by small populations whose evolutionary dynamics are dominated by genetic drift.

RESULTS AND DISCUSSION

Bull kelp and giant kelp are the principal canopy-forming species in kelp forests of the northeast Pacific, supporting highly productive and biodiverse ecosystems[11] that generate billions of dollars annually in ecosystem services.[12] Despite their broad geographic distributions from Alaska to California (with giant kelp additionally found in northern Mexico and the southern hemisphere[11]), both species have experienced strong local and regional declines[1–4] due to factors such as marine heatwaves and urchin overgrazing.[2,3] These declines have spurred a profusion of interest in kelp forest restoration.[13,14] However, restoration and conservation strategies for bull and giant kelp are being hampered in part by lack of information about genetic structure (cf. Assis et al.[15]

and Gierke et al.[16]), patterns of local adaptation, and the genetic risks faced by small populations subject to the balance between genetic drift and natural selection. For example, such information could inform decisions about optimal regional geographic sourcing, local population selection (e.g., Wood et al.[17]), and genetic culturing methods of material used in restoration.

To address these knowledge gaps, we sequenced the genomes of 429 bull kelp and 211 giant kelp (Table S1) from British Columbia (BC), Canada, and Washington (WA), USA (hereafter, "BCWA"; Figure S1), to a mean depth of 21.0× (range: 11.5–43.0×). We identified 3,274,934 autosomal single-nucleotide polymorphisms (SNPs) with a minimum minor allele frequency of 0.01 in bull kelp and 2,341,413 such SNPs in giant kelp, or 4,327,335 SNPs when including published data[18–20] from an
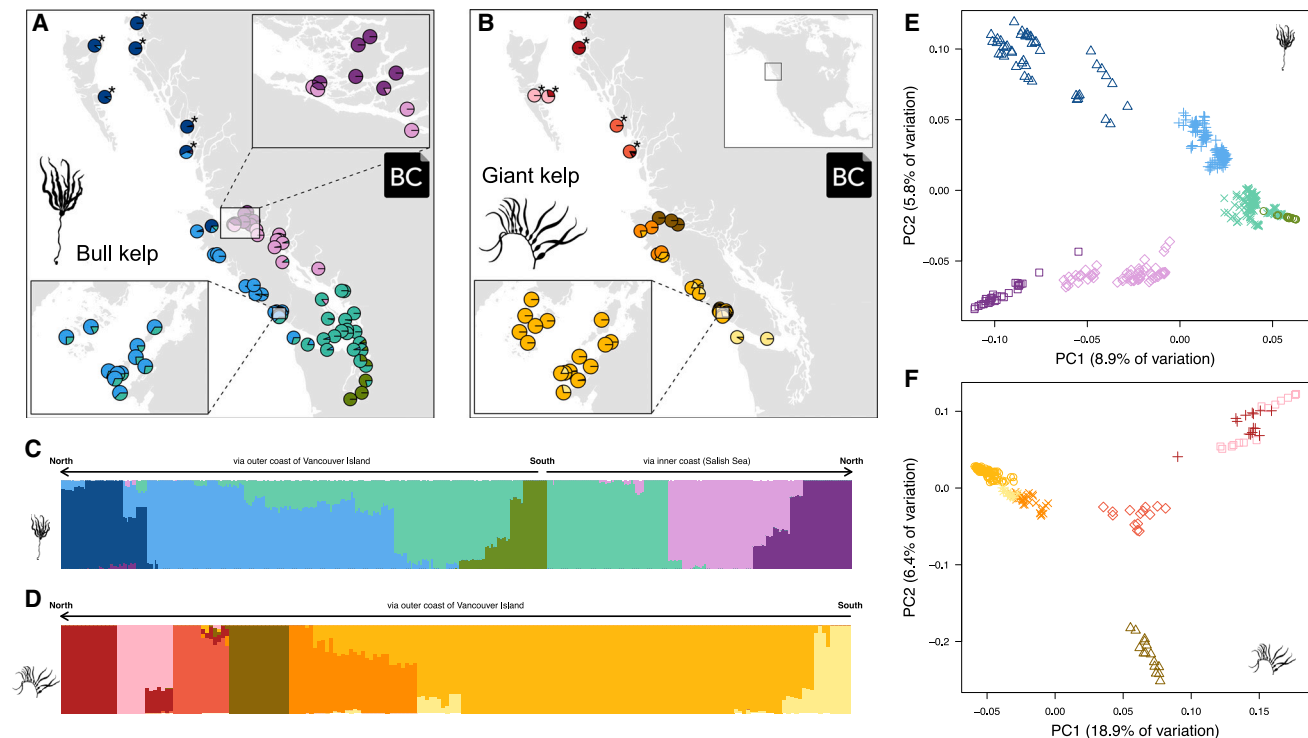
Figure 1. Genetic structure

Genetic structure of bull kelp (A, C, and E) and giant kelp (B, D, and F) in British Columbia and Washington.

(A and B) Pie charts depict the proportion of ancestry in each population belonging to different genetic clusters, with clusters represented as different colors. An asterisk indicates locations that are approximate, as coordinates were rounded to the nearest 0.5°.

(C and D) The proportion of ancestry derived from each genetic cluster in each individual, with individuals represented as vertical lines.

(E and F) Principal component analysis showing the clustering of individuals along the first two PC axes. Each point represents an individual and is colored according to the clusters from (A)–(D) and with a unique symbol for each cluster. Some samples and derived data are subject to a Biocultural (BC) Notice (see resource availability section and Table S1). See also Figures S1–S4 and Tables S1 and S2.

additional 70 individuals from California, Chile, and Australia. In light of ongoing taxonomic debate in *Macrocystis*,[21–23] we first investigated whether giant kelp from other regions were closely genetically related to BCWA. Southern and Northern Hemisphere giant kelp formed distinct genetic clusters[24] (Figures S2A and S2C) and were distinguished along the first principal component (PC) axis of genetic variation (Figure S2E). The hemispheres were highly genetically differentiated ($F_{ST} = 0.71$) and moderately genetically diverged ($d_{XY} = 0.0077$). Within North America, California formed a distinct genetic cluster (Figure S2D), was distinguished along the second PC axis (Figure S2F), and was moderately genetically differentiated and diverged from BCWA ($F_{ST} = 0.32$; $d_{XY} = 0.0039$; Figure S2). Due to the moderate to strong genetic differences between giant kelp from BCWA and other regions, we opted to focus the remainder of our analyses on BCWA only.

Within BCWA, both species exhibit strong genetic structure (Table S2). Six genetic clusters were identified in bull kelp (Figures 1A and 1C) and seven clusters in giant kelp (Figures 1B and 1D) (see also Figure S3). These clusters occupy distinct geographic regions and are largely non-overlapping along the first two PC axes of genetic variation (Figures 1E and 1F). A strong isolation-by-distance pattern of increasing genetic distance (d_{XY}) with geographic distance (Figure S4) and the presence of

populations admixed between clusters (Figures 1A–1D) suggest that adjacent clusters are connected by gene flow. Given the lack of formal kelp management zones in BCWA, guidelines are needed to inform movement of genetic material for restoration and aquaculture. We suggest that genetic clusters could be used to help define management units (MUs)[25,26] or—in combination with environmental data[27]—seed-transfer zones[28] that would delineate regions within which transfer of genetic material would pose minimal risk to the genetic integrity of local populations. Moreover, genetic clusters can be used to guide biobanking efforts and prioritize conservation investments, given that small populations in some clusters may be at risk for extirpation.

Having delineated broad-scale genetic relationships, we next turned our attention to inferring the capacity for individual populations to persist and adapt to change.[29] Some kelp populations in BCWA have remained stable in recent decades while others have experienced strong declines.[2,4] Small populations are often subject to multiple stressors that raise the risk of extirpation.[30] We assessed three genetic health indicators in each population: (1) effective population size (N_e), with low N_e associated with higher extirpation risk due to demographic stochasticity, inbreeding, and drift[31]; (2) genetic diversity, which is required for populations to adapt to future challenges[32]; and (3) mean inbreeding coefficient, with higher inbreeding coefficients
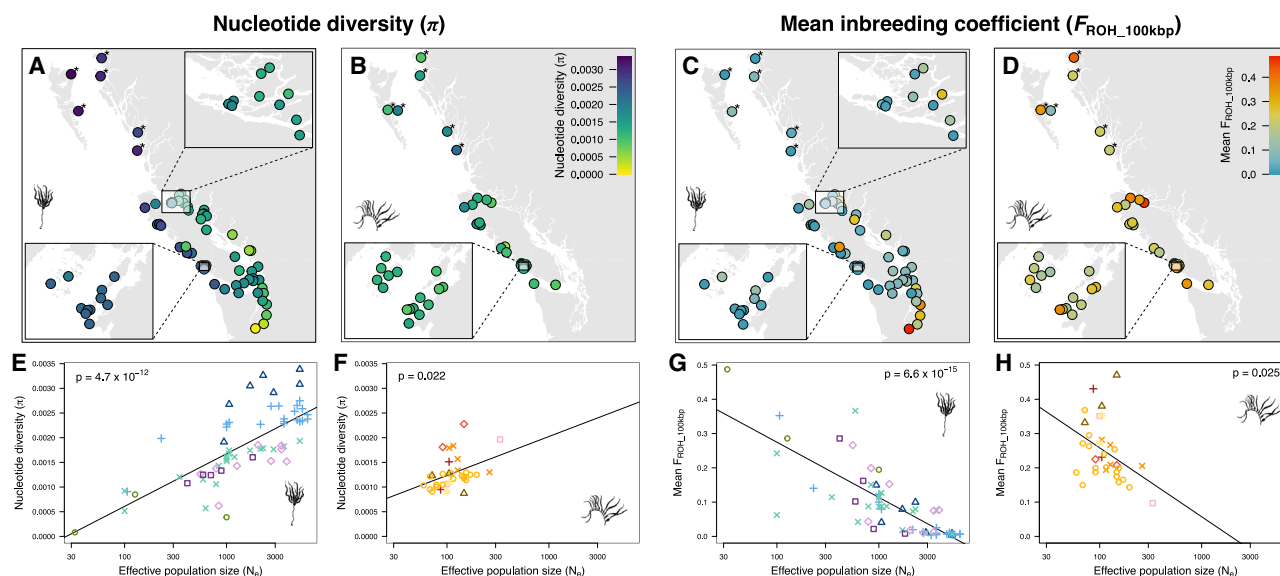
Current Biology
Report

CellPress

Figure 2. Genetic health indicators

Genetic health indicators for bull kelp (A, C, E, and G) and giant kelp (B, D, F, and H).

(A and B) Geographic distribution of nucleotide diversity (π). An asterisk indicates locations that are approximate, as coordinates were rounded to the nearest 0.5°.

(C and D) Geographic distribution of mean inbreeding coefficient ($F_{ROH\_100\ kbp}$).

(E and F) Relationship between π and effective population size (N_e).

(G and H) Relationship between $F_{ROH\_100\ kbp}$ and N_e. In (E)–(H), each point represents a population. Symbols and colors correspond to genetic clusters from Figure 1. See also Figures S5 and S6.

implying higher risks of inbreeding depression,[33] defined here as a fitness reduction of inbred relative to outbred individuals from the same population.[30] In bull kelp, N_e varied by more than two orders of magnitude (N_e = 33–6,236; Figures 2E and 2G), while nucleotide diversity (π) varied more than 40-fold between the highest-diversity populations (northern BC and northwest Vancouver Island) and the lowest-diversity population in southernmost Puget Sound (Squaxin Island, WA) (Figure 2A). A similar geographic pattern was observed in inbreeding coefficients ($F_{ROH\_100\ kbp}$; Figure 2C). In both species, N_e was positively correlated with π and negatively correlated with $F_{ROH\_100\ kbp}$ (Figures 2E–2H). At least some of the variation in inbreeding coefficients is driven by high among-population variation in the rate of selfing (Figure S5). Selfing results when separate male and female haploid spores from a single diploid adult settle on the sea floor in close proximity ($\leq$1 mm apart), a process facilitated by the limited dispersal distance of most spores.[34–36] In both species, the BCWA-wide selfing rate was 10%, in agreement with the ~10% rate predicted for giant kelp from dispersal-based models.[35] N_e was negatively correlated with selfing rate in bull kelp but not in giant kelp (Figure S6).

In contrast to the extreme among-population variation observed in bull kelp, genetic health indicators were more uniform in giant kelp. Nucleotide diversity and N_e were comparatively low to moderate (Figures 2B and 2F), and $F_{ROH\_100\ kbp}$ was comparatively moderate to high in all populations (Figure 2D). Nonetheless, the highest-diversity populations tended to be from northern BC and northwest Vancouver Island (Figure 2B), as was the case for bull kelp. Differences in dispersal biology between the two species may partly explain why giant kelp exhibited comparatively low N_e

and π and high $F_{ROH\_100\ kbp}$. Bull kelp releases spores from two heights in the water column,[37] with spores released from sori on blades near the surface dispersing orders of magnitude farther than those released from abscised sori that have sunk to the sea floor.[38] In contrast, giant kelp sporophylls are located near the base of the kelp close to the sea floor,[39] likely limiting dispersal distances and resulting in reduced local genetic connectivity. The generally lower N_e of giant kelp may partly explain why this species exhibits more genetic clusters (Figure 1) and higher F_{ST} between clusters (Table S2) over similar geographic distances and suggests that giant kelp may require conservation genetic planning at a more local scale than bull kelp (i.e., shorter translocation distances for restoration and aquaculture) if managers wish to maintain existing genetic structure.

The frequent association of low effective population size with low genetic diversity and high inbreeding coefficients in both species suggests that some populations face multiple genetic health risks that may lead to fitness declines and loss of evolutionary potential.[30,40] These populations tend to occur in Puget Sound, the Strait of Georgia, and the inner Broughton Archipelago, as well as inner reaches of smaller water bodies away from the open coast (e.g., inner Quatsino Sound and Clayoquot Sound) (Figure 2). However, high inbreeding coefficients do not necessarily imply strongly reduced fitness, as natural selection may reduce inbreeding depression by purging recessive deleterious alleles.[8–10] Specifically, in small populations where individuals are on average more closely related than in large populations[41] (and where selfing is more common in bull kelp; Figure S6), alleles are more likely to be identical by descent, resulting in increased homozygosity of rare recessive deleterious
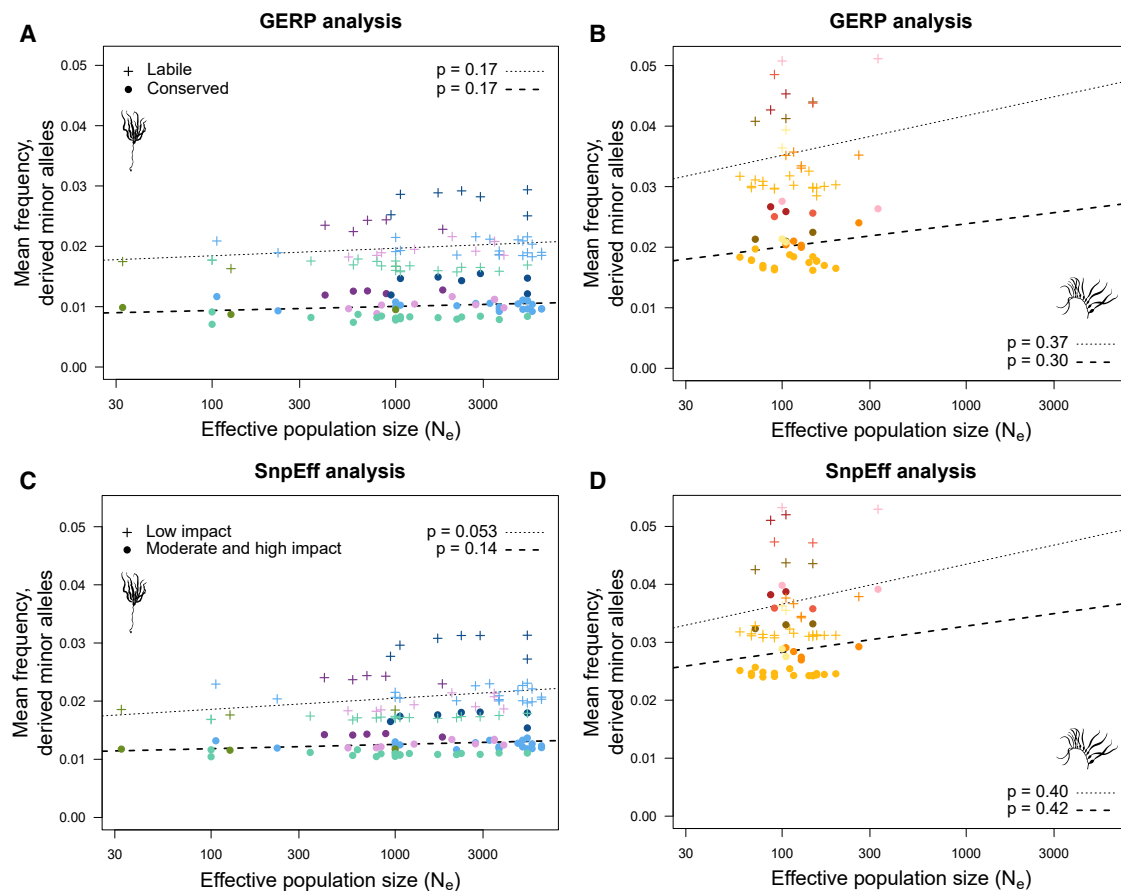
Figure 3. Lack of evidence for purging

Lack of evidence for purging by natural selection in GERP (A and B) and *SnpEff* (C and D) analyses.

(A and B) Mean frequency of derived minor alleles (DMAs) does not vary with effective population size (N_e) for either evolutionary labile (control) or conserved (putatively deleterious) sites in (A) bull kelp or (B) giant kelp.

(C and D) Mean DMA frequency does not vary with N_e for sites predicted to have either a low (control) or moderate to high (presumed deleterious) impact on proteins in (C) bull kelp or (D) giant kelp. Each point represents a population. Symbols and colors correspond to genetic clusters from Figure 1. See also Figure S7 and Table S3.

alleles. This transformation of recessive deleterious alleles from a heterozygous state ("masked load") into a homozygous state ("realized load")[42] exposes recessive deleterious alleles to natural selection. Over time, the loss of these alleles by purging can reduce the negative fitness consequences faced by inbred individuals.[8] However, inbreeding depression is unlikely to be entirely eliminated because natural selection is less effective in small populations,[43] making it difficult to purge alleles that are only mildly deleterious.[8,44,45] Given our empirical observations of high inbreeding coefficients in some populations and the contrasting theoretical outcomes of inbreeding, we next considered the likely effects of inbreeding in small kelp populations.

Southern Californian populations of giant kelp experience substantial inbreeding depression,[46–48] but to our knowledge, inbreeding depression and purging have not been evaluated in bull kelp or in giant kelp from BCWA, where giant kelp forests are smaller[49] and where N_e may sometimes be much lower (Figure 2) than in California (N_e = 50–2,500).[50] We looked for genomic signatures of purging by estimating genetic load in two ways (Figure S7). Firstly, we used Genomic Evolutionary Rate Profiling

(GERP)[51] to calculate GERP scores (S) and identify derived minor alleles (DMAs) at sites that are either evolutionarily labile or conserved, with DMAs at conserved sites considered more likely to be deleterious. Values of S indicate how many fewer substitutions were observed across a phylogeny of brown algae (Table S3) than expected based on the neutral substitution rate, and we classified sites with at least moderate evidence of a reduced substitution rate (S > 0.5) as more evolutionarily conserved and all other sites (S ≤ 0.5) as more evolutionarily labile. The frequency of DMAs was overall lower at conserved sites than labile sites (Figures 3A and 3B), which suggests greater sequence constraint at conserved sites and supports our classifications. Secondly, we annotated protein-coding genes using *SnpEff*[52] to identify DMAs whose predicted impacts on proteins are either low, moderate, or high, and we considered moderate- and high-impact sites to be more likely to be deleterious. We assumed that most putatively deleterious DMAs are likely to be at least partly recessive.[53,54]

We observed no evidence of purging in either species. We predicted that smaller populations would show a reduction in DMA
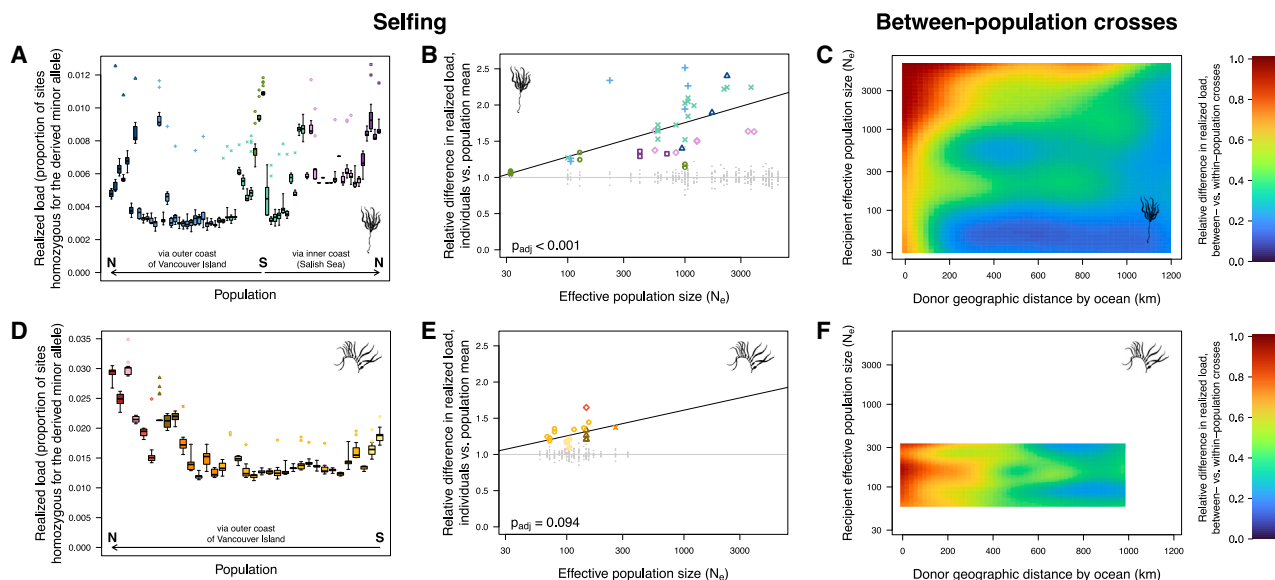
Figure 4. Predicted genetic load under different cross types, *SnpEff* analysis

Predicted changes in realized genetic load under different cross types for bull kelp (A–C) and giant kelp (D–F) suggest a penalty for selfing (A, B, D, and E) and reduced load in between-population crosses (C and F). Realized genetic load is measured as the proportion of sites homozygous for the derived allele at moderate- and high-impact sites in the *SnpEff* analysis.

(A and D) Realized load is higher in selfed individuals (point symbols, one point per individual) than in non-selfed individuals (distribution in boxplots, with boxplot whiskers extending to the most extreme values).

(B and E) Relative difference in realized load of selfed individuals (colored symbols, one point per individual) compared with the mean realized load of non-selfed individuals, as a function of effective population size (N_e). Non-selfed individuals are shown in small gray points for comparison. Adjusted p values are calculated from 1,000 randomizations of N_e. Symbols and colors correspond to genetic clusters from Figure 1.

(C and F) Realized genetic load is predicted to be equal or lower (≤ 1.0) in between-population crosses (recipient × donor population) relative to within-population crosses (recipient × recipient). The predicted relative load (color scale) is plotted as a function of recipient N_e and geographic distance by ocean to the donor population. The colored two-dimensional surface was interpolated by kriging raw data points (Figure S13). See also Figures S8–S13 and Table S4.

frequency at evolutionarily conserved sites (GERP analysis) due to increased homozygosity and exposure to selection, yet DMA frequency was uncorrelated with population size (Figures 3A and 3B). Similarly, we predicted but failed to find evidence for a positive correlation between population size and frequency of moderate- to high-impact DMAs (*SnpEff* analysis; Figures 3C and 3D). Although purging has been empirically demonstrated in small populations of many species,[10,55,56] in many other cases it is not detected,[57] such as in recently bottlenecked populations where there has been insufficient time for purging to remove deleterious alleles[54] or in extremely small populations where drift completely overwhelms natural selection.[58] Alternatively, if an ancestral population experienced a prolonged ancient bottleneck, most deleterious variants could have already been purged.[59] Such a bottleneck could conceivably have occurred during the Last Glacial Maximum, especially if recolonization of BCWA occurred from one or more small ice-free refugia, e.g., off the coast of Haida Gwaii as previously inferred for bull kelp.[16] Likewise, purging of alleles expressed during the haploid gametophyte life stage[60] in all populations could reduce the signal of additional purging in diploid sporophytes that is brought about by small population size.

Although we did not find evidence that natural selection is reducing inbreeding depression through purging, genetic drift may also reduce inbreeding depression within small populations.[5,7,41,61] Though rarely empirically tested (reviewed in Willi

et al.[7]; see also Spigler et al.[5] and Pekkala et al.[62]), this counter-intuitive prediction's theoretical foundations date back to Kimura et al.,[43] who noted that the more pronounced effects of genetic drift[63] and reduced efficacy of natural selection in small populations would cause many mildly recessive deleterious alleles to become lost or fixed. With few recessive deleterious alleles segregating within a population, the relative difference in realized load between inbred and outbred individuals is expected to be very small.[5,7,41,61] Our results strongly support the expectations of this scenario in bull kelp. Firstly, the number of genomic sites with at least one copy of a DMA present in the population was reduced in small populations (Figures S8A and S9A), consistent with the expected loss of many DMAs. Loss of DMAs could occur through either drift or purging, yet the predictions of purging were not supported (see above). Secondly, remaining DMAs were present at higher frequency (Figures S8C and S9C) and more likely to be fixed (Figures S8E and S9E) in small populations. Thirdly, although the realized load of putatively deleterious DMAs was higher in selfed than outbred individuals in all populations (Figures 4A and S10A), the relative increase in realized load in selfed individuals was strongly positively correlated with population size (Figures 4B and S10B). These trends suggest that genetic drift has greatly reduced the relative fitness penalty for inbreeding relative to outcrossing within small populations. In contrast to the strong support for these predictions in bull kelp, trends were typically not significant for giant kelp,

although slopes of relationships were often very similar to those for bull kelp, and p values often marginally non-significant (Figures 4 and S8–S10), suggesting that lack of significance could relate to the minimal variation observed in population size.

While our results suggest that the loss and (near) fixation of many recessive deleterious alleles may reduce the strength of inbreeding depression in small populations of bull kelp, the fixation of such alleles due to genetic drift is also expected to result in an overall fitness reduction known as "drift load."[5–7] A large drift load could cause all individuals in small populations to exhibit reduced performance relative to individuals from large populations.[5,7,41,61] We found strong support for increased drift load in small populations of bull kelp, as effective population size was positively correlated with the realized load of putatively deleterious DMAs in non-selfed individuals (Figures S11A and S11C). The relationship was not significant in giant kelp (Figures S11B and S11D).

Given that small populations of bull kelp likely suffer from increased drift load, we considered whether crossing populations would reduce the genetic load of offspring, which could make crossing populations a useful strategy in restoration and aquaculture contexts. In particular, heterosis (i.e., hybrid vigor, or increased performance of between-population crosses relative to crosses within parental populations) may be expected in crosses between small populations that have independently and randomly fixed a different set of recessive deleterious alleles, due to the increased heterozygosity of such alleles expected in the offspring.[41,64] Increased heterosis in small populations relative to large populations has been empirically demonstrated in numerous species,[7,41,62,65,66] though not universally.[67] Likewise, heterosis has been observed in some but not all Chilean populations of giant kelp, but an effect of population size was not tested.[68] Consistent with the expectations of heterosis, we frequently observed a lower realized genetic load in simulated between-population crosses (recipient × donor population) relative to simulated within-population crosses (recipient × recipient population) of both species (Figures 4C, 4F, S10C, S10F, S12, and S13). As predicted, the realized genetic load was more strongly reduced in crosses in which the recipient population was smaller (bull kelp only; Table S4). In addition, realized genetic load was more strongly reduced when the donor population was more geographically distant (both species; Table S4), likely because geographically distant populations are expected to share few high-frequency deleterious alleles, and so F_1 homozygosity will be minimal.

Collectively, our analyses suggest a lack of purging in small populations of either species (Figure 3). Instead, genetic drift has caused the fixation of many putatively deleterious alleles (Figures S8 and S9) in small populations of bull kelp, increasing the drift load of all individuals (Figure S11), reducing the difference in realized load between selfed and non-selfed individuals (Figures 4 and S10), and leading to the potential for heterosis in between-population crosses (Figures 4, S10, S12, and S13). Our findings result in several interesting evolutionary predictions that also have important implications for kelp conservation. These predictions are derived from genomic signatures and are based on the standard assumption that realized load is a reasonable proxy for fitness,[8] but we caution that we have not measured

the fitness of any individuals, and experimental validation will be needed. Nonetheless, we predict that:

(1) All else being equal, individuals from small populations may perform more poorly in restoration or aquaculture than individuals from large populations of bull kelp. This prediction arises from the increased drift load observed in small populations (Figure S11) due to fixation of recessive deleterious alleles (Figures S8 and S9) and presents an additional reason that large populations may be preferred as a genetic source over small populations above and beyond the higher genetic diversity of large populations (Figures 2E and 2F). However, a small population could still be a preferred genetic source if it is locally adapted to a historical environment that is a better match to the current or future environment of an outplanting site.[69,70] A meta-analysis of transplant studies has suggested that the benefits of local adaptation outweigh the negative effects of drift load in some but not all contexts.[71] In BCWA kelp, the extent of local adaptation is unknown, though Mexican giant kelp appear locally adapted[72] and South American populations have shown mixed results.[73,74]

(2) Offspring of crosses between small populations may perform better than offspring of crosses within a single small population. Any potential benefits of this predicted heterosis would need to outweigh the risks of outbreeding depression[75] (e.g., through disruption of coadapted gene complexes[76] or local adaptation[77,78]). The risks of outbreeding are poorly empirically known[75] and often delayed until the F_2 generation,[75,76] suggesting the need for long-term experimental monitoring. Nonetheless, the risk of outbreeding depression is generally minimal when populations occupy similar environments and have been recently fragmented (<500 years).[79] Divergence times and rates of gene flow between kelp populations in BCWA are unknown. However, moderate to high F_{ST} values between genetic clusters in both species (Table S2) suggest that the potential for outbreeding depression in crosses between clusters should not be dismissed. In South American giant kelp, outbreeding depression and heterosis have both been observed depending on the geographic region,[74] highlighting the context-specific nature of the tradeoff between these phenomena.

(3) Selfing may be an effective means of reproduction in small populations of bull kelp. In contrast, the fitness penalty for selfing is likely to remain high in large populations where selfing substantially increases the realized load relative to outbred individuals (Figures 4B and S10B). This situation could imply a shift from an effectively primarily outcrossing system to a mixed mating system as populations decrease in size and accumulate drift load (Figure S11), such as has been observed in leading-edge plant populations following range expansion.[65] Importantly, this shift could occur without the need to hypothesize explicit benefits to selfing, such as reproductive assurance[80] or the perpetuation of genotypes suited to local environments.[81] Due to the predicted fitness penalty

Current Biology
Report

CellPress

for selfing in large populations, actions should be taken to minimize selfing (e.g., increasing the number of parents) in restoration cultures sourced from a large population. While outcrossing is still preferred when parents are sourced from a small population, avoiding selfing may be of lower priority, especially if obtaining high-quality reproductive material from numerous parents is difficult.

In summation, we have demonstrated strong genetic and geographic clustering in bull and giant kelp from BCWA that could aid in designating MUs or seed-transfer zones. Small populations face multiple genetic health risks but show no evidence of purging. Instead, allele frequency changes in small populations appear to be dominated by genetic drift. Our genomic analyses have revealed fundamental insights into the evolutionary dynamics of small populations and imply several strategies that could be cautiously applied (pending experimental validation) to conservation and restoration of these at-risk and declining[1–4] kelp species.

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jordan Bemmels (jbemmels@umich.edu).

Materials availability
This study did not generate new, unique reagents.

Data and code availability
- Raw DNA sequences generated in this study have been deposited at NCBI (NCBI SRA: PRJNA1164249) and are publicly available as of the date of publication. Some of the raw DNA sequences are derived from samples (identified in Table S1) subject to a Biocultural Notice[82]: "The Biocultural (BC) Notice is a visible notification that there are accompanying cultural rights and responsibilities that need further attention for any future sharing and use of this material or data. The BC Notice recognizes the rights of Indigenous peoples to permission the use of information, collections, data, and digital sequence information (DSI) generated from the biodiversity or genetic resources associated with traditional lands, waters, and territories. The BC Notice may indicate that BC Labels are in development and their implementation is being negotiated. For more information about the BC Notice, see localcontexts.org/notice/bc-notice/."
- All original code has been deposited at Zenodo and is publicly available at https://doi.org/10.5281/zenodo.14454933 as of the date of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - DNA extractions and sequencing
  - Alignment to reference genome
  - SNP genotyping
  - Derived minor allele classification
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Genetic structure
  - Genetic health indicators and selfing rate
  - Tests for purging and genetic drift
  - Genetic load under different cross types

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2024.12.025.

REFERENCES

1. Berry, H.D., Mumford, T.F., Christiaen, B., Dowty, P., Calloway, M., Ferrier, L., Grossman, E.E., and VanArendonk, N.R. (2021). Long-term changes in kelp forests in an inner basin of the Salish Sea. PLoS One 16, e0229703. https://doi.org/10.1371/journal.pone.0229703.

2. Starko, S., Timmer, B., Reshitnyk, L., Csordas, M., McHenry, J., Schroeder, S., Hessing-Lewis, M., Costa, M., Zielinksi, A., Zielinksi, R., et al. (2024). Local and regional variation in kelp loss and stability across coastal British Columbia. Mar. Ecol. Prog. Ser. 733, 1–26. https://doi.org/10.3354/meps14548s.

3. Rogers-Bennett, L., and Catton, C.A. (2019). Marine heat wave and multiple stressors tip bull kelp forest to sea urchin barrens. Sci. Rep. *9*, 15050. https://doi.org/10.1038/s41598-019-51114-y.

4. Mora-Soto, A., Schroeder, S., Gendall, L., Wachmann, A., Narayan, G., Read, S., Pearsall, I., Rubidge, E., Lessard, J., Martell, K., et al. (2024). Back to the past: long-term persistence of bull kelp forests in the Strait of Georgia, Salish Sea, Canada. Front. Mar. Sci. *11*, 1446380. https://doi.org/10.3389/fmars.2024.1446380.

5. Spigler, R.B., Theodorou, K., and Chang, S.M. (2017). Inbreeding depression and drift load in small populations at demographic disequilibrium. Evolution *71*, 81–94. https://doi.org/10.1111/evo.13103.

6. Whitlock, M.C. (2000). Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. Evolution *54*, 1855–1861. https://doi.org/10.1111/j.0014-3820.2000.tb01232.x.

7. Willi, Y., Griffin, P., and Van Buskirk, J. (2013). Drift load in populations of small size and low density. Heredity *110*, 296–302. https://doi.org/10.1038/hdy.2012.86.

8. Dussex, N., Morales, H.E., Grossen, C., Dalén, L., and Van Oosterhout, C. (2023). Purging and accumulation of genetic load in conservation. Trends Ecol. Evol. *38*, 961–969. https://doi.org/10.1016/j.tree.2023.05.008.

9. Glémin, S. (2003). How are deleterious mutations purged? Drift versus nonrandom mating. Evolution *57*, 2678–2687. https://doi.org/10.1111/j.0014-3820.2003.tb01512.x.

10. van der Valk, T., de Manuel, M., Marques-Bonet, T., and Guschanski, K. (2019). Estimates of genetic load suggest frequent purging of deleterious alleles in small populations. Preprint at bioRxiv. https://doi.org/10.1101/696831.

11. Wernberg, T., Krumhansl, K., Filbee-Dexter, K., and Pedersen, M.F. (2019). Status and trends for the world's kelp forests. In World Seas: an Environmental Evaluation (Elsevier), pp. 57–78. https://doi.org/10.1016/B978-0-12-805052-1.00003-6.

12. Eger, A.M., Marzinelli, E.M., Beas-Luna, R., Blain, C.O., Blamey, L.K., Byrnes, J.E.K., Carnell, P.E., Choi, C.G., Hessing-Lewis, M., Kim, K.Y., et al. (2023). The value of ecosystem services in global marine kelp forests. Nat. Commun. *14*, 1894. https://doi.org/10.1038/s41467-023-37385-0.

13. Eger, A.M., Marzinelli, E.M., Christie, H., Fagerli, C.W., Fujita, D., Gonzalez, A.P., Hong, S.W., Kim, J.H., Lee, L.C., McHugh, T.A., et al. (2022). Global kelp forest restoration: past lessons, present status, and future directions. Biol. Rev. Camb. Philos. Soc. *97*, 1449–1475. https://doi.org/10.1111/brv.12850.

14. Morris, R.L., Hale, R., Strain, E.M.A., Reeves, S.E., Vergés, A., Marzinelli, E.M., Layton, C., Shelamoff, V., Graham, T.D.J., Chevalier, M., et al. (2020). Key principles for managing recovery of kelp forests through restoration. BioScience *70*, 688–698. https://doi.org/10.1093/biosci/biaa058.

15. Assis, J., Alberto, F., Macaya, E.C., Castilho Coelho, N., Faugeron, S., Pearson, G.A., Ladah, L., Reed, D.C., Raimondi, P., Mansilla, A., et al. (2023). Past climate-driven range shifts structuring intraspecific biodiversity levels of the giant kelp (*Macrocystis pyrifera*) at global scales. Sci. Rep. *13*, 12046. https://doi.org/10.1038/s41598-023-38944-7.

16. Gierke, L., Coelho, N.C., Khangaonkar, T., Mumford, T., and Alberto, F. (2023). Range wide genetic differentiation in the bull kelp *Nereocystis luetkeana* with a seascape genetic focus on the Salish Sea. Front. Mar. Sci. *10*, 1275905. https://doi.org/10.3389/fmars.2023.1275905.

17. Wood, G., Marzinelli, E.M., Vergés, A., Campbell, A.H., Steinberg, P.D., and Coleman, M.A. (2020). Using genomics to design and evaluate the performance of underwater forest restoration. J. Appl. Ecol. *57*, 1988–1998. https://doi.org/10.1111/1365-2664.13707.

18. Gonzalez, S.T., Alberto, F., and Molano, G. (2023). Whole-genome sequencing distinguishes the two most common giant kelp ecomorphs. Evolution *77*, 1354–1369. https://doi.org/10.1093/evolut/qpad045.

19. Molano, G., Diesel, J., Montecinos, G.J., Alberto, F., and Nuzhdin, S.V. (2022). Sporophyte stage genes exhibit stronger selection than gametophyte stage genes in haplodiplontic giant kelp. Front. Mar. Sci. *8*, 774076. https://doi.org/10.3389/fmars.2021.774076.

20. Iha, C., Layton, C., Amancio, C.E., Flentje, W., Lenton, A., Fraser, C.I., Johnson, C., and Willis, A. (2023). Population genomic analysis reveals genetic structure and thermal-tolerant genotypes in remnant Tasmanian giant kelp populations. Preprint at bioRxiv. https://doi.org/10.1101/2023.10.08.561451.

21. Lindstrom, S.C. (2023). A reinstated species name for north-eastern Pacific *Macrocystis* (Laminariaceae, Phaeophyceae). Not. Algarum *290*, 1–2.

22. Macaya, E.C., and Zuccarello, G.C. (2010). DNA barcoding and genetic divergence in the giant kelp *Macrocystis* (Laminariales). J. Phycol. *46*, 736–742. https://doi.org/10.1111/j.1529-8817.2010.00845.x.

23. Demes, K.W., Graham, M.H., and Suskiewicz, T.S. (2009). Phenotypic plasticity reconciles incongruous molecular and morphological taxonomics: the giant kelp, *Macrocystis* (Laminariales, Phaeophyceae), is a monospecific genus. J. Phycol. *45*, 1266–1269. https://doi.org/10.1111/j.1529-8817.2009.00752.x.

24. Raj, A., Stephens, M., and Pritchard, J.K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics *197*, 573–589. https://doi.org/10.1534/genetics.114.164350.

25. Moritz, C. (1994). Defining 'Evolutionarily Significant Units' for conservation. Trends Ecol. Evol. *9*, 373–375. https://doi.org/10.1016/0169-5347(94)90057-4.

26. Palsbøll, P.J., Bérubé, M., and Allendorf, F.W. (2007). Identification of management units using population genetic data. Trends Ecol. Evol. *22*, 11–16. https://doi.org/10.1016/j.tree.2006.09.003.

27. Massatti, R., Shriver, R.K., Winkler, D.E., Richardson, B.A., and Bradford, J.B. (2020). Assessment of population genetics and climatic variability can refine climate-informed seed transfer guidelines. Restor. Ecol. *28*, 485–493. https://doi.org/10.1111/rec.13142.

28. Ying, C.C., and Yanchuk, A.D. (2006). The development of British Columbia's tree seed transfer guidelines: purpose, concept, methodology, and implementation. For. Ecol. Manag. *227*, 1–13. https://doi.org/10.1016/j.foreco.2006.02.028.

29. Hohenlohe, P.A., Funk, W.C., and Rajora, O.P. (2021). Population genomics for wildlife conservation and management. Mol. Ecol. *30*, 62–82. https://doi.org/10.1111/mec.15720.

30. Willi, Y., Kristensen, T.N., Sgrò, C.M., Weeks, A.R., Ørsted, M., and Hoffmann, A.A. (2022). Conservation genetics as a management tool: the five best-supported paradigms to assist the management of threatened species. Proc. Natl. Acad. Sci. USA *119*, e2105076119. https://doi.org/10.1073/pnas.2105076119.

31. Harmon, L.J., and Braude, S. (2010). Conservation of small populations: effective population sizes, inbreeding, and the 50/500 rule. In An Introduction to Methods and Models in Ecology, Evolution, and Conservation Biology, S. Braude, and B.S. Low, eds. (Princeton University Press), pp. 125–138. https://doi.org/10.2307/j.ctvcm4gbm.19.

32. Hoffmann, A.A., Sgrò, C.M., and Kristensen, T.N. (2017). Revisiting adaptive potential, population size, and conservation. Trends Ecol. Evol. *32*, 506–517. https://doi.org/10.1016/j.tree.2017.03.012.

33. Crnokrak, P., and Roff, D.A. (1999). Inbreeding depression in the wild. Heredity *83*, 260–270. https://doi.org/10.1038/sj.hdy.6885530.

34. Edwards, M.S. (2022). It's the little things: the role of microscopic life stages in maintaining kelp populations. Front. Mar. Sci. *9*, 871204. https://doi.org/10.3389/fmars.2022.871204.

35. Gaylord, B., Reed, D.C., Raimondi, P.T., and Washburn, L. (2006). Macroalgal spore dispersal in coastal environments: mechanistic insights revealed by theory and experiment. Ecol. Monogr. *76*, 481–502. https://doi.org/10.1890/0012-9615(2006)076[0481:MSDICE]2.0.CO;2.

Current Biology
Report

 CellPress

36. Reed, D.C. (1990). The effects of variable settlement and early competition on patterns of kelp recruitment. Ecology 71, 776–787. https://doi.org/10.2307/1940329.

37. Amsler, C.D., and Neushul, M. (1989). Diel periodicity of spore release from the kelp *Nereocystis luetkeana* (Mertens) Postels et Ruprecht. J. Exp. Mar. Biol. Ecol. 134, 117–127. https://doi.org/10.1016/0022-0981(90)90104-K.

38. Burnett, N.P., Ricart, A.M., Winquist, T., Saley, A.M., Edwards, M.S., Hughes, B., Hodin, J., Baskett, M.L., and Gaylord, B. (2024). Bimodal spore release heights in the water column enhance local retention and population connectivity of bull kelp, *Nereocystis luetkeana*. Ecol. Evol. 14, e70177. https://doi.org/10.1002/ece3.70177.

39. Neushul, M. (1963). Studies on the giant kelp, *Macrocystis*. II. Reproduction. Am. J. Bot. 50, 354–359. https://doi.org/10.1002/j.1537-2197.1963.tb07203.x.

40. Frankham, R., Bradshaw, C.J.A., and Brook, B.W. (2014). Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. Biol. Conserv. 170, 56–63. https://doi.org/10.1016/j.biocon.2013.12.036.

41. Keller, L.F., and Waller, D.M. (2002). Inbreeding effects in wild populations. Trends Ecol. Evol. 17, 230–241. https://doi.org/10.1016/S0169-5347(02)02489-8.

42. Bertorelle, G., Raffini, F., Bosse, M., Bortoluzzi, C., Iannucci, A., Trucchi, E., Morales, H.E., and Van Oosterhout, C. (2022). Genetic load: genomic estimates and applications in non-model animals. Nat. Rev. Genet. 23, 492–503. https://doi.org/10.1038/s41576-022-00448-x.

43. Kimura, M., Maruyama, T., and Crow, J.F. (1963). The mutation load in small populations. Genetics 48, 1303–1312. https://doi.org/10.1093/genetics/48.10.1303.

44. Kardos, M., Taylor, H.R., Ellegren, H., Luikart, G., and Allendorf, F.W. (2016). Genomics advances the study of inbreeding depression in the wild. Evol. Appl. 9, 1205–1218. https://doi.org/10.1111/eva.12414.

45. Ohta, T. (1973). Slightly deleterious mutation substitutions in evolution. Nature 246, 96–98. https://doi.org/10.1038/246096a0.

46. Johansson, M.L., Raimondi, P.T., Reed, D.C., Coelho, N.C., Serrão, E.A., and Alberto, F.A. (2013). Looking into the black box: simulating the role of self-fertilization and mortality in the genetic structure of *Macrocystis pyrifera*. Mol. Ecol. 22, 4842–4854. https://doi.org/10.1111/mec.12444.

47. Raimondi, P.T., Reed, D.C., Gaylord, B., and Washburn, L. (2004). Effects of self-fertilization in the giant kelp, *Macrocystis pyrifera*. Ecology 85, 3267–3276. https://doi.org/10.1890/03-0559.

48. San Miguel, R.A. (2017). Gametophyte fitness and costs of self-fertilization in the giant kelp Macrocystis pyrifera. Master's thesis (California State University Monterey Bay).

49. Graham, M.H., Vásquez, J.A., and Buschmann, A.H. Global ecology of the giant kelp *Macrocystis*: from ecotypes to ecosystems. Oceanogr. Mar. Biol. Annu. Rev. 45, 39–88.

50. Alberto, F., Raimondi, P.T., Reed, D.C., Coelho, N.C., Leblois, R., Whitmer, A., and Serrão, E.A. (2010). Habitat continuity and geographic distance predict population genetic differentiation in giant kelp. Ecology 91, 49–56. https://doi.org/10.1890/09-0050.1.

51. Cooper, G.M., Stone, E.A., Asimenos, G., NISC; Comparative; Sequencing Program, Green, E.D., Batzoglou, S., and Sidow, A. (2005). Distribution and intensity of constraint in mammalian genomic sequence. Genome Res. 15, 901–913. https://doi.org/10.1101/gr.3577405.

52. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 6, 80–92. https://doi.org/10.4161/fly.19695.

53. Bosse, M., Megens, H.J., Derks, M.F.L., De Cara, Á.M.R., and Groenen, M.A.M. (2019). Deleterious alleles in the context of domestication, inbreeding, and selection. Evol. Appl. 12, 6–17. https://doi.org/10.1111/eva.12691.

54. Taylor, R.S., Manseau, M., Keobouasone, S., Liu, P., Mastromonaco, G., Solmundson, K., Kelly, A., Larter, N.C., Gamberg, M., Schwantje, H., et al. (2024). High genetic load without purging in caribou, a diverse species at risk. Curr. Biol. 34, 1234–1246.e7. https://doi.org/10.1016/j.cub.2024.02.002.

55. Khan, A., Patel, K., Shukla, H., Viswanathan, A., Van Der Valk, T., Borthakur, U., Nigam, P., Zachariah, A., Jhala, Y.V., Kardos, M., et al. (2021). Genomic evidence for inbreeding depression and purging of deleterious genetic variation in Indian tigers. Proc. Natl. Acad. Sci. USA 118, e2023018118. https://doi.org/10.1073/pnas.2023018118.

56. Crnokrak, P., and Barrett, S.C.H. (2002). Perspective: Purging the genetic load: a review of the experimental evidence. Evolution 56, 2347–2358. https://doi.org/10.1111/j.0014-3820.2002.tb00160.x.

57. Byers, D.L., and Waller, D.M. (1999). Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. Annu. Rev. Ecol. Syst. 30, 479–513. https://doi.org/10.1146/annurev.ecolsys.30.1.479.

58. López-Cortegano, E., Moreno, E., and García-Dorado, A. (2021). Genetic purging in captive endangered ungulates with extremely low effective population sizes. Heredity 127, 433–442. https://doi.org/10.1038/s41437-021-00473-2.

59. Jamieson, I.G., Wallis, G.P., and Briskie, J.V. (2006). Inbreeding and endangered species management: is New Zealand out of step with the rest of the world? Conserv. Biol. 20, 38–47. https://doi.org/10.1111/j.1523-1739.2005.00282.x.

60. Barner, A.K., Pfister, C.A., and Wootton, J.T. (2011). The mixed mating system of the sea palm kelp *Postelsia palmaeformis*: few costs to selfing. Proc. Biol. Sci. 278, 1347–1355. https://doi.org/10.1098/rspb.2010.1928.

61. Bataillon, T., and Kirkpatrick, M. (2000). Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. Genet. Res. 75, 75–81. https://doi.org/10.1017/S0016672399004048.

62. Pekkala, N., Knott, K.E., Kotiaho, J.S., Nissinen, K., and Puurtinen, M. (2014). The effect of inbreeding rate on fitness, inbreeding depression and heterosis over a range of inbreeding coefficients. Evol. Appl. 7, 1107–1119. https://doi.org/10.1111/eva.12145.

63. Wright, S. (1931). Evolution in Mendelian populations. Genetics 16, 97–159. https://doi.org/10.1093/genetics/16.2.97.

64. Whitlock, M.C., Ingvarsson, P.K., and Hatfield, T. (2000). Local drift load and the heterosis of interconnected populations. Heredity 84, 452–457. https://doi.org/10.1046/j.1365-2540.2000.00693.x.

65. Koski, M.H., Layman, N.C., Prior, C.J., Busch, J.W., and Galloway, L.F. (2019). Selfing ability and drift load evolve with range expansion. Evol. Lett. 3, 500–512. https://doi.org/10.1002/evl3.136.

66. Oakley, C.G., and Winn, A.A. (2012). Effects of population size and isolation on heterosis, mean fitness, and inbreeding depression in a perennial plant. New Phytol. 196, 261–270. https://doi.org/10.1111/j.1469-8137.2012.04240.x.

67. Paland, S., and Schmid, B. (2003). Population size and the nature of genetic load in Gentianella germanica. Evolution 57, 2242–2251. https://doi.org/10.1111/j.0014-3820.2003.tb00236.x.

68. Westermeier, R., Patiño, D.J., Müller, H., and Müller, D.G. (2010). Towards domestication of giant kelp (*Macrocystis pyrifera*) in Chile: selection of haploid parent genotypes, outbreeding, and heterosis. J. Appl. Phycol. 22, 357–361. https://doi.org/10.1007/s10811-009-9466-1.

69. Havens, K., Vitt, P., Still, S., Kramer, A.T., Fant, J.B., and Schatz, K. (2015). Seed sourcing for restoration in an era of climate change. Nat. Areas J. 35, 122–133. https://doi.org/10.3375/043.035.0116.

70. O'Neill, G.A., and Gómez-Pineda, E. (2021). Local was best: sourcing tree seed for future climates. Can. J. For. Res. 51, 1432–1439. https://doi.org/10.1139/cjfr-2020-0408.

71. Bontrager, M., Usui, T., Lee-Yaw, J.A., Anstett, D.N., Branch, H.A., Hargreaves, A.L., Muir, C.D., and Angert, A.L. (2021). Adaptation across

geographic ranges is consistent with strong selection in marginal climates and legacies of range expansion. Evolution *75*, 1316–1333. https://doi.org/10.1111/evo.14231.

72. Ladah, L.B., and Zertuche-González, J.A. (2022). Local adaptation of juvenile giant kelp, Macrocystis pyrifera, from their southern limit in the northern hemisphere explored using reciprocal transplantation. Eur. J. Phycol. *57*, 357–366. https://doi.org/10.1080/09670262.2021.2007543.

73. Becheler, R., Haverbeck, D., Clerc, C., Montecinos, G., Valero, M., Mansilla, A., and Faugeron, S. (2022). Variation in thermal tolerance of the giant kelp's gametophytes: suitability of habitat, population quality or local adaptation? Front. Mar. Sci. *9*, 802535. https://doi.org/10.3389/fmars.2022.802535.

74. Solas, M., Correa, R.A., Barría, F., Garcés, C., Camus, C., and Faugeron, S. (2024). Assessment of local adaptation and outbreeding risks in contrasting thermal environments of the giant kelp, *Macrocystis pyrifera*. J. Appl. Phycol. *36*, 471–483. https://doi.org/10.1007/s10811-023-03119-4.

75. Edmands, S. (2007). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. Mol. Ecol. *16*, 463–475. https://doi.org/10.1111/j.1365-294X.2006.03148.x.

76. Lynch, M. (1991). The genetic interpretation of inbreeding and outbreeding depression. Evolution *45*, 622–629. https://doi.org/10.1111/j.1558-5646.1991.tb04333.x.

77. Aitken, S.N., and Whitlock, M.C. (2013). Assisted gene flow to facilitate local adaptation to climate change. Annu. Rev. Ecol. Evol. Syst. *44*, 367–388. https://doi.org/10.1146/annurev-ecolsys-110512-135747.

78. Schiltroth, B. (2016). Effects of climate change on two species of foundational brown algae, *Nereocystis luetkeana* and *Fucus gardneri*, within the Salish Sea. Master's thesis (Simon Fraser University).

79. Frankham, R., Ballou, J.D., Eldridge, M.D.B., Lacy, R.C., Ralls, K., Dudash, M.R., and Fenster, C.B. (2011). Predicting the probability of outbreeding depression. Conserv. Biol. *25*, 465–475. https://doi.org/10.1111/j.1523-1739.2011.01662.x.

80. Kalisz, S., Vogler, D.W., and Hanley, K.M. (2004). Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. Nature *430*, 884–887. https://doi.org/10.1038/nature02776.

81. Levin, D.A. (2010). Environment-enhanced self-fertilization: implications for niche shifts in adjacent populations. J. Ecol. *98*, 1276–1283. https://doi.org/10.1111/j.1365-2745.2010.01715.x.

82. Liggins, L., Hudson, M., and Anderson, J. (2021). Creating space for Indigenous perspectives on access and benefit-sharing: encouraging researcher use of the Local Contexts Notices. Mol. Ecol. *30*, 2477–2482. https://doi.org/10.1111/mec.15918.

83. Alberto, F., Toft, J., Allen, B., Marimuthu, M.P.A., Nguyen, O., Beraut, E., Escalona, M., Miller, B., Raimondi, P., and Shaffer, H.B. (2023). Genome assembly of the bull kelp (Nereocystis luetkeana), uoNerLuet1.0. GenBank Assembly GCA_031213475.1.

84. Diesel, J., Molano, G., Montecinos, G.J., DeWeese, K., Calhoun, S., Kuo, A., Lipzen, A., Salamov, A., Grigoriev, I.V., Reed, D.C., et al. (2023). A scaffolded and annotated reference genome of giant kelp (*Macrocystis pyrifera*). BMC Genomics *24*, 543. https://doi.org/10.1186/s12864-023-09658-x.

85. Druehl, L.D., Collins, J.D., Lane, C.E., and Saunders, G.W. (2005). An evaluation of methods used to assess intergeneric hybridization in kelp using Pacific Laminariales (Phaeophyceae). J. Phycol. *41*, 250–262. https://doi.org/10.1111/j.1529-8817.2005.04143.x.

86. Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics *34*, i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

87. Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at arXiv. https://doi.org/10.48550/arXiv.1303.3997.

88. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., et al. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008. https://doi.org/10.1093/gigascience/giab008.

89. Broad Institute (2024). Picard toolkit (Broad Institute). https://broadinstitute.github.io/picard/.

90. Van der Auwera, G.A., and O'Connor, B.D. (2020). Genomics in the Cloud: Using Docker, GATK, and WDL in Terra, First Edition (O'Reilly Media).

91. Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

92. Hanghøj, K., Moltke, I., Andersen, P.A., Manica, A., and Korneliussen, T.S. (2019). Fast and accurate relatedness estimation from high-throughput sequencing data in the presence of inbreeding. GigaScience *8*, giz034. https://doi.org/10.1093/gigascience/giz034.

93. Fox, E.A., Wright, A.E., Fumagalli, M., and Vieira, F.G. (2019). ngsLD: evaluating linkage disequilibrium using genotype likelihoods. Bioinformatics *35*, 3855–3856. https://doi.org/10.1093/bioinformatics/btz200.

94. Bushnell, B. (2014). BBMap: a fast, accurate, splice-aware aligner. https://sourceforge.net/projects/bbmap/.

95. Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.-M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics *26*, 2867–2873. https://doi.org/10.1093/bioinformatics/btq559.

96. Shumate, A., and Salzberg, S.L. (2021). Liftoff: accurate mapping of gene annotations. Bioinformatics *37*, 1639–1643. https://doi.org/10.1093/bioinformatics/btaa1016.

97. Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., and Weir, B.S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics *28*, 3326–3328. https://doi.org/10.1093/bioinformatics/bts606.

98. Zheng, X., Gogarten, S.M., Lawrence, M., Stilp, A., Conomos, M.P., Weir, B.S., Laurie, C., and Levine, D. (2017). SeqArray—a storage-efficient high-performance data format for WGS variant calls. Bioinformatics *33*, 2251–2257. https://doi.org/10.1093/bioinformatics/btx145.

99. Kawai, H., Hanyuda, T., Draisma, S.G.A., Wilce, R.T., and Andersen, R.A. (2015). Molecular phylogeny of two unusual brown algae, Phaeostrophion irregulare and Platysiphon glacialis, proposal of the Stschapoviales ord. nov. and Platysiphonaceae fam. nov., and a re-examination of divergence times for brown algal orders. J. Phycol. *51*, 918–928. https://doi.org/10.1111/jpy.12332.

100. Goudet, J. (2005). hierfstat, a package for R to compute and test hierarchical F-statistics. Mol. Ecol. Notes *5*, 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x.

101. Van Etten, J. (2017). R package gdistance: distances and routes on geographical grids. J. Stat. Softw. *76*, 1–21. https://doi.org/10.18637/jss.v076.i13.

102. R Core Team (2023). In R: a Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/.

103. Zeitler, L., and Gilbert, K.J. (2024). Using runs of homozygosity and machine learning to disentangle sources of inbreeding and infer self-fertilization rates. Genome Biol. Evol. *16*, evae139. https://doi.org/10.1093/gbe/evae139.

104. Contreras-Moreira, B., Filippi, C.V., Naamati, G., García Girón, C., Allen, J.E., and Flicek, P. (2021). K-mer counting and curated libraries drive efficient annotation of repeats in plant genomes. Plant Genome *14*, e20143. https://doi.org/10.1002/tpg2.20143.

105. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

Current Biology
Report

CellPress

106. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575. https://doi.org/10.1086/519795.

107. Renaud, G., Hanghøj, K., Korneliussen, T.S., Willerslev, E., and Orlando, L. (2019). Joint estimates of heterozygosity and runs of homozygosity for modern and ancient samples. Genetics *212*, 587–614. https://doi.org/10.1534/genetics.119.302057.

108. Koch, D. (2022). snapKrig. R package. https://github.com/deankoch/snapKrig.

109. Bringloe, T.T., Zaparenkov, D., Starko, S., Grant, W.S., Vieira, C., Kawai, H., Hanyuda, T., Filbee-Dexter, K., Klimova, A., Klochkova, T.A., et al. (2021). Whole-genome sequencing reveals forgotten lineages and recurrent hybridizations within the kelp genus *Alaria* (Phaeophyceae). J. Phycol. *57*, 1721–1738. https://doi.org/10.1111/jpy.13212.

110. Lopez, B.R., Hernandez, J.-P., Bashan, Y., and de-Bashan, L.E. (2017). Immobilization of microalgae cells in alginate facilitates isolation of DNA and RNA. J. Microbiol. Methods *135*, 96–104. https://doi.org/10.1016/j.mimet.2017.02.005.

111. Fox, C., and Swanson, A. (2007). Nested PCR detection of microscopic life-stages of laminarian macroalgae and comparison with adult forms along intertidal height gradients. Mar. Ecol. Prog. Ser. *332*, 1–10. https://doi.org/10.3354/meps332001.

112. Waples, R.K., Albrechtsen, A., and Moltke, I. (2019). Allele frequency-free inference of close familial relationships from genotypes or low-depth sequencing data. Mol. Ecol. *28*, 35–48. https://doi.org/10.1111/mec.14954.

113. Westermeier, R., Murúa, P., Patiño, D.J., Muñoz, L., Ruiz, A., Atero, C., and Müller, D.G. (2013). Utilization of holdfast fragments for vegetative propagation of *Macrocystis integrifolia* in Atacama, Northern Chile. J. Appl. Phycol. *25*, 639–642. https://doi.org/10.1007/s10811-012-9898-x.

114. Choi, S.-W., Graf, L., Choi, J.W., Jo, J., Boo, G.H., Kawai, H., Choi, C.G., Xiao, S., Knoll, A.H., Andersen, R.A., et al. (2024). Ordovician origin and subsequent diversification of the brown algae. Curr. Biol. *34*, 740–754.e4. https://doi.org/10.1016/j.cub.2023.12.069.

115. Bringloe, T.T., Starko, S., Wade, R.M., Vieira, C., Kawai, H., De Clerck, O., Cock, J.M., Coelho, S.M., Destombe, C., Valero, M., et al. (2020). Phylogeny and evolution of the brown algae. Crit. Rev. Plant Sci. *39*, 281–321. https://doi.org/10.1080/07352689.2020.1787679.

116. Cánovas, F.G., Mota, C.F., Serrão, E.A., and Pearson, G.A. (2011). Driving south: a multi-gene phylogeny of the brown algal family Fucaceae reveals relationships and recent drivers of a marine radiation. BMC Evol. Biol. *11*, 371. https://doi.org/10.1186/1471-2148-11-371.

117. Silberfeld, T., Leigh, J.W., Verbruggen, H., Cruaud, C., de Reviers, B., and Rousseau, F. (2010). A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): investigating the evolutionary nature of the "brown algal crown radiation.". Mol. Phylogenet. Evol. *56*, 659–674. https://doi.org/10.1016/j.ympev.2010.04.020.

118. Starko, S., Soto Gomez, M., Darby, H., Demes, K.W., Kawai, H., Yotsukura, N., Lindstrom, S.C., Keeling, P.J., Graham, S.W., and Martone, P.T. (2019). A comprehensive kelp phylogeny sheds light on the evolution of an ecosystem. Mol. Phylogenet. Evol. *136*, 138–150. https://doi.org/10.1016/j.ympev.2019.04.012.

119. Li, H. (2012). HTSbox. https://github.com/lh3/htsbox.

120. Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput. Biol. *6*, e1001025. https://doi.org/10.1371/journal.pcbi.1001025.

121. Krasovec, M., Hoshino, M., Zheng, M., Lipinska, A.P., and Coelho, S.M. (2023). Low spontaneous mutation rate in complex multicellular eukaryotes with a haploid–diploid life cycle. Mol. Biol. Evol. *40*, msad105. https://doi.org/10.1093/molbev/msad105.

122. Weir, B.S., and Cockerham, C.C. (1984). Estimating F-statistics for the analysis of population structure. Evolution *38*, 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x.

123. Nei, M., and Li, W.H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA *76*, 5269–5273. https://doi.org/10.1073/pnas.76.10.5269.

124. Korunes, K.L., and Samuk, K. (2021). pixy: unbiased estimation of nucleotide diversity and divergence in the presence of missing data. Mol. Ecol. Resour. *21*, 1359–1368. https://doi.org/10.1111/1755-0998.13326.

125. GeoBC Branch. (2002). NTS BC coastline polygons 1:250,000 – digital baseline mapping (NTS). B.C.'s Map Hub, ID 5ebbc1fd1d7e456684b8-d1a3f10186e4. governmentofbc.maps.arcgis.com.

126. Hijmans, R.J. (2023). raster: geographic data analysis and modeling. R package version 3.6-26. https://rspatial.org/raster.

127. Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. Cancer Res. *27*, 209–220.

128. Girgis, H.Z. (2015). Red: an intelligent, rapid, accurate tool for detecting repeats de-novo on the genomic scale. BMC Bioinformatics *16*, 227. https://doi.org/10.1186/s12859-015-0654-5.

129. Contreras-Moreira, B., Naamati, G., Rosello, M., Allen, J.E., Hunt, S.E., Muffato, M., Gall, A., and Flicek, P. (2022). Scripting analyses of genomes in Ensembl Plants. In Plant Bioinformatics: Methods and Protocols Methods in Molecular Biology, D. Edwards, ed. (Springer), pp. 27–55. https://doi.org/10.1007/978-1-0716-2067-0.

130. Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics *123*, 585–595. https://doi.org/10.1093/genetics/123.3.585.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
449 bull kelp and 364 giant kelp tissue samples	This paper	Table S1
Critical commercial assays		
Magnetic beads	Sergi Lab Supplies	Cat# 1040
2X KAPA Hifi HotStart ReadyMix	Roche	Cat# KK2602
xGen DNA Library Prep EZ Kit	IDT	Cat# 10009821
xGen Deceleration Module	IDT	Cat# 10009823
Deposited data		
Whole-genome sequencing data for 449 bull kelp and 239 giant kelp from BCWA	This paper	NCBI SRA: PRJNA1164249; Table S1
Whole-genome sequencing data for 75 giant kelp from California and Chile	Gonzalez et al.[18]	NCBI SRA: PRJNA938791; Table S1
Whole-genome sequencing data for 49 giant kelp from California	Molano et al.[19]	NCBI SRA: PRJNA661280; Table S1
Whole-genome sequencing data for 1 giant kelp from Australia	Iha et al.[20]	NCBI SRA: PRJEB55054; Table S1
Bull kelp reference genome	Alberto et al.[83]	NCBI Datasets: GCA_031213475.1; Table S3
Giant kelp reference genome	Diesel et al.[84]	JGI PhycoCosm: *Macrocystis pyrifera* CI_03; Table S3
Reference genomes for 27 additional brown algae	Table S3	Table S3
Oligonucleotides		
Bull kelp ITS primers	Fox and Swanson	F and R2
Giant kelp ITS primers	Druehl et al.[85]	KG4 and "*Macrocystis intergrifolia*"
xGen 10-nucleotide UDI Primer Plates 1-4	IDT	Cat# 10008052
Software and algorithms		
fastp v.0.23.2	Chen et al.[86]	https://github.com/OpenGene/fastp
bwa-mem v.0.7.17-r1188	Li[87]	https://github.com/lh3/bwa
SAMtools v.1.17	Danecek et al.[88]	https://github.com/samtools/samtools
Picard v.2.26.3	Broad Institute[89]	https://github.com/broadinstitute/picard
GATK v.3.8	Van der Auwera and O'Connor[90]	https://gatk.broadinstitute.org/
BCFtools v.1.11-1.19	Danecek et al.[91]	https://github.com/samtools/bcftools
ngsRelate v.2	Hanghøj et al.[92]	https://github.com/ANGSD/NgsRelate
VCFtools v.0.1.1678	Danecek et al.[91]	https://vcftools.github.io/
ngsLD v.1.2.1	Fox et al.[93]	https://github.com/fgvieira/ngsLD
Scripts for running GERP analyses	This paper	Zenodo: https://doi.org/10.5281/zenodo.14454933
BBmap v.39.06	Bushnell[94]	https://sourceforge.net/projects/bbmap/
HTSBox v.r345	Li[95]	https://github.com/lh3/htsbox
Modified gerpcol script v.2023/11/20	Taylor et al.[54]	https://github.com/BeckySTaylor/Phylogenomic_Analyses/
SnpEff v.5.2a	Cingolani et al.[52]	https://pcingola.github.io/SnpEff/
Liftoff v.1.6.3	Shumate and Salzberg[96]	https://github.com/agshumate/Liftoff
Scripts for processing SnpEff variant annotations	This paper	Zenodo: https://doi.org/10.5281/zenodo.14454933

(Continued on next page)

Current Biology
Report

![CellPress logo]

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
SNPRelate v.1.38.0	Zheng et al.[97,98]	https://github.com/zhengxwen/SNPRelate
fastSTRUCTURE v.1.0	Raj et al.[24]	https://github.com/rajanil/fastStructure
hierfstat v.0.5-11	Goudet[99]	https://github.com/jgx65/hierfstat
pixy v.1.2.7.beta199	Korunes and Samuk[94]	https://pixy.readthedocs.io/
raster v.3.6-26	Hijmans[100]	https://github.com/rspatial/raster
gdistance v.1.6.4	Van Etten[101]	https://github.com/AgrDataSci/gdistance
R v.4.2.3	R Core Team[102]	https://www.r-project.org/
roh-selfing v.2024/02/23	Zeitler and Gilbert[103]	https://github.com/LZeitler/roh-selfing
Ensembl Plants v.1.2	Contreras-Moreira et al.[104]	https://github.com/Ensembl/plant-scripts
BEDTools v.2.30.0	Quinlan and Hall[105]	https://github.com/arq5x/bedtools2
Scripts for identifying short runs of heterozygosity	This paper	Zenodo: https://doi.org/10.5281/zenodo.14454933
PLINK v.1.90b6.21	Purcell et al.[106]	https://www.cog-genomics.org/plink/
ROHan v.1.0.1	Renaud et al.[107]	https://github.com/grenaud/ROHan
Scripts to test for purging and genetic drift	This paper	Zenodo: https://doi.org/10.5281/zenodo.14454933
Scripts to calculate realized genetic load of observed individuals	This paper	Zenodo: https://doi.org/10.5281/zenodo.14454933
Scripts to simulate realized genetic load in offspring of crosses	This paper	Zenodo: https://doi.org/10.5281/zenodo.14454933
snapKrig v.0.0.2113	Koch[108]	https://github.com/deankoch/snapKrig
Other		
Composite phylogenies of brown algae	This paper	Zenodo: https://doi.org/10.5281/zenodo.14454933
Processed rasters of BCWA coastline	This paper	Zenodo: https://doi.org/10.5281/zenodo.14454933

EXPERIMENTAL MODEL AND SUBJECT DETAILS

This study generated sequencing data for 449 bull kelp and 239 giant kelp. Sample sites and sample sizes are provided in Table S1. Each sample site corresponds to a single kelp forest (hereafter, "population"), sampled between 2021 and 2023. For up to 15 individuals from each population, a cutting of blade tissue was collected and dried using silica gel. Effort was made to collect from blades a minimum of 3 m apart, as assessed from the water surface, but the spacing of holdfasts along the ocean floor could not typically be determined for either species and 3 m spacing could not be ensured for some of the physically smallest bull kelp populations.

Giant kelp is the common name applied to all kelp of the genus *Macrocystis*, which has been widely recognized in recent decades as a single species *M. pyrifera*[22,23] but was recently split into a minimum of four species[21]: *M. pyrifera* in southern Chile, *M. integrifolia* in northern Chile, *M. tenuifolia* in North America north of Point Conception, California, and a yet-unnamed species in North America south of Point Conception. We refer to all *Macrocystis* as "giant kelp" and refer to this taxon as a "species" in the present study but acknowledge that the taxonomic situation is in flux. Importantly, regardless whether giant kelp from our main study region of BCWA are recognized as *M. pyrifera* sensu lato[22,23] or *M. tenuifolia*,[21] there are no species boundaries currently proposed within BCWA.

Related to but separate from the issue of species delineation, there are several described *Macrocystis* ecomorphs that differ in holdfast morphology.[18,22] We could not assess ecomorph identity at the time of sampling because samples were collected from the surface without access to the holdfast, but assume that all of our samples (from BCWA) are of the 'integrifolia' ecomorph based on previous description of the geographic distribution of each ecomorph.[22] The remaining global *Macrocystis* samples we analyzed from previously published data represent a combination of 'integrifolia' and 'pyrifera' ecomorphs (Figure S2).

METHOD DETAILS

DNA extractions and sequencing

We extracted DNA from dried blade tissue using a cetyltrimethylammonium bromide (CTAB) protocol, substantially modified from one developed by.[109] We disrupted approximately 0.5 cm^2 of tissue using a *Qiagen* TissueLyser, then added 1 mL of citrate wash buffer (0.055 M sodium citrate dihydrate, 0.030 M EDTA, 0.150 M NaCl, adjusted to pH 8.0)[110] and incubated with agitation for 15 min at room temperature. We then centrifuged at 10,000 G for 5 min (this and all subsequent centrifugation steps were performed

at room temperature), discarded the supernatant, and repeated the wash step a second time but without the 15-min incubation. To the resulting precipitate, we added 700 uL of CTAB isolation buffer (1% CTAB, 5 M NaCl, 0.5 M ethylenediaminetetraacetic acid [EDTA], 1% polyvinylpyrrolidone [PVP] mol wt 10,000, 10 mM Tris-HCl), 14 uL proteinase-K (20 mg/mL) and 10 uL of RNAse A (100 mg/mL), and incubated at 55°C, inverting the sample tube every 5-10 min for the first ~40 min before leaving it to incubate overnight. The next day, we centrifuged for 10 min at 13,300 G, resulting in the formation of an aqueous clear upper layer and a green lower layer. We transferred the upper layer (~650 uL) to a new tube and added 0.2-0.3 volumes of 100% ethanol very slowly with constant stirring to precipitate residual polysaccharides without precipitating DNA. To this solution, we then added 650 uL of 24:1 chloroform:isoamyl alcohol, mixed briefly by inversion, and centrifuged at 12,000 G for 5 min. We then transferred 600 uL of the resulting upper layer to a new tube (leaving ~50 uL behind to prevent contamination of the transferred portion with the bottom layer), added 600 uL of isopropanol to the transferred upper layer, vortexed the solution, and centrifuged at 13,300 G for 30 min at room temperature. After removing the supernatant, a small white DNA pellet remained, which we washed twice with 250 uL of 70% ethanol. We air-dried the pellet and dissolved it in 50 uL of nuclease-free water. We then cleaned the DNA using magnetic beads (*Sergi Lab Supplies*) at 0.8X following the manufacturer's protocol.

Initial testing indicated that the functionality of cleaned extractions in downstream applications could not be reliably predicted from standard quality checks (i.e., using a *Thermo Scientific* NanoDrop spectrophotometer and a *Thermo Scientific* Qubit fluorometer), suggesting undetected contaminants in some samples. We therefore ran a test polymerase chain reaction (PCR) on each sample to rule out the presence of PCR inhibitors. We used primers for the internal transcribed spacer (ITS) locus for each species (bull kelp: *N. luetkeana* F and *N. luetkeana* R2 from[111]; giant kelp: KG4 reverse and *Macrocystis integrifolia* forward primers from[85]). PCRs were performed in a 5-uL reaction volume using 2.5 uL of *Roche* 2X KAPA Hifi HotStart ReadyMix, 0.15 uL of 10 uM forward primer, 0.15 uL of 10 uM reverse primer, and 1.7 uL of nuclease-free water. Reactions were denatured at 95°C for 3 min; followed by 35 cycles of 98°C for 20s, 62°C for 15s and 72°C for 30s; and a final extension of 72°C for 2 min. We visualized samples on a 2% agarose gel to check for the presence of a 575-bp band in bull kelp[111] and a 912-bp band in giant kelp.[85] Samples without a strong band were re-extracted or a different individual from the same population was selected instead.

We prepared libraries for whole-genome sequencing (WGS) using *Integrated DNA Technologies (IDT)* xGen DNA Library Prep EZ Kits, following the manufacturer's protocol with the following options and minor modifications: 165 ng of input DNA topped up to 16 uL in nuclease-free water; use of the *IDT* xGen Deceleration Module to slow enzymatic fragmentation, with reaction times reduced by 1 min relative to the timing indicated for each specific lot of each *IDT* kit (typically 12 min reduced from an indicated 13 min); the use of *IDT* xGen 10-nucleotide primers; no use of the optional *IDT* xGen Normalase Module; and six PCR cycles to account for the additional cycling requirements of the Deceleration Module. We performed bead-cleaning steps using *Sergi Lab Supplies* magnetic beads at the same ratios indicated in the protocol for AMPure XP beads (*Beckman Coulter*). We combined between 62 and 66 barcoded samples per library pool (except for initial test libraries of fewer samples) across a total of 12 library pools. We sent samples to Génome Québec (Montreal) or Canada's Michael Smith Genome Sciences Centre (Vancouver) for 150-bp paired-end sequencing on an *Illumina* NovaSeq 6000, targeting 2.5 billion paired-end reads per library pool.

Surprisingly, many of our library pools failed quality control at the sequencing core due to the presence of a large, unexpected secondary peak of short fragments (ca. 310–330 bp; target peak size typically >700 bp) that was not detectable in gel electrophoresis on a manual agarose gel but appeared on the results of automated electrophoresis machines at the sequencing core. This secondary peak was repeatedly detected across multiple independent sample submissions, was present in multiple library pools prepared months apart, could not be removed by further bead clean-up at stringent bead:sample ratios, and could not be visualized in house on any manual agarose gel under any visualization settings, suggesting that it may have represented an unknown contaminant rather than short DNA fragments. Test sequencing of affected libraries on an *Illumina* MiSeq revealed no apparent negative impacts on the sequencing reaction and a size distribution of sequenced fragments corresponding to the primary (target) peak only. We therefore proceeded with sequencing all of our library pools and could detect no impacts on the quality or quantity of the final data we obtained.

Alignment to reference genome

Prior to aligning raw reads to reference genomes, we trimmed adapter sequences, merged overlapping paired-end reads, and performed basic quality and read-length filtering using *fastp* v.0.23.2[86] with default parameters. We added a unique read-group to each sample and aligned filtered, adapter-trimmed reads to reference genomes for either bull kelp (467.6 Mb; NCBI Datasets: GCA_031213475.1)[83] or giant kelp (537.5 Mb; JGI PhycoCosm: *Macrocystis pyrifera* CI_03 v1.0)[84] using *bwa-mem* v.0.7.17-r1188[87] with default parameters. We then sorted and merged aligned reads (paired, unpaired, and merged) into a single file for each sample using SAMtools v.1.17,[88] and removed read duplicates using *Picard* v.2.26.3.[89] We removed one giant kelp individual that had an extremely low read alignment rate (1.4%) and for which the tissue sample was noted to have been visibly degraded and contaminated with symbionts.

To correct potential misalignments around indels that could lead to the identification of false variants, we next realigned the sorted, merged, and duplicate-removed reads using *GATK* v.3.8.[90] We first identified target intervals representing putative indels using the *RealignerTargetCreator* tool from *GATK*, identifying targets from all individuals in a single command for each species. We then used the *IndelRealigner* tool from *GATK* to realign reads around indels.

Current Biology
Report

CellPress

SNP genotyping

We performed SNP genotyping steps separately for datasets of bull kelp from BCWA only, giant kelp from BCWA only, and giant kelp from all global samples. We identified Single Nucleotide Polymorphisms (SNPs) and called genotypes using *BCFtools* v.1.11-1.19.[88] We created binary Variant Call Format (VCF) files of genotype likelihoods using *bcftools mpileup*, requiring minimum base and mapping qualities of 30 (-Q 30 -q 30), and called genotypes with *bcftools call* using the multiallelic caller (-m) and outputting variant sites only (-v). We filtered raw genotype calls with *bcftools view* to include only SNPs (-v snps) that were biallelic (-m 2 -M 2) and with both alleles observed in our samples and not only in the reference genome (-q 0.000001:minor). We retained only nuclear SNPs by removing the (-t ^) mitochondrial (mtDNA) and chloroplast (cpDNA) assemblies, as identified in the reference genomes.

We next filtered raw SNP calls to generate a set of high-quality sites for downstream analyses. We required a minimum depth of 10 to retain a SNP call by using the +*setGT* plugin of *BCFTools* to set genotypes to missing at sites where depth was less than 10. We then used *bcftools view* to exclude sites with abnormally high (>75%) heterozygosity across all samples, which suggests multiple regions aligning to the same location in the reference genome and may represent gene-duplication events or errors in genome assembly or read alignment. We also removed sites that failed any of five additional quality tests flagged by *BCFTools* in the INFO column of the VCF files. Specifically, we retained only sites where the following was true: log-likelihood ratio of Segregation Based Metric (SGB) $\geq$ 2, p-value of Mapping Quality Bias (MQB) $\geq$ 0.05, p-value of Mapping Quality vs Strand Bias (MQSB) $\geq$ 0.05, p-value of Read Position Bias (RPB) $\geq$ 0.05, and p-value of Variant Distance Bias (VDB) $\geq$ 0.05. These quality metrics and their cutoff values were selected based on inspecting heterozygous SNP calls in selected regions identified as putative Runs of Homozyosity (ROHs) from visual inspection of sliding windows of heterozygosity in unfiltered SNP datasets (i.e., from inspecting presumed erroneous SNPs). We additionally excluded sites above the 98th percentile of depth of coverage across all samples combined.

After applying SNP filters, we removed seven bull kelp and nine giant kelp individuals from downstream analyses due to high missing data ($\geq$ 50% missing). The mean depth of coverage of all retained samples was $\geq$ 9.5X. We then recalculated allele frequencies (-t AN,AC) using the +*fill-tags* plugin of *BCFTools*. For giant kelp we recalculated allele frequencies and created datasets for separate subsets of individuals (rangewide, North American, and BCWA samples, respectively).

We then used *BCFTools* to perform additional site filtering, retaining sites with a minimum minor allele frequency of 0.01 (-q 0.01:minor) and maximum 20% missing data. We also retained only putative autosomal SNPs by removing sites on putative sex chromosomes (JARUPZ010000001.1 in *Nereocystis* and scaffold_2 in *Macrocystis*; sex chromosomes were identified as scaffolds were large regions contained approximately half the depth of coverage of the remainder of the genome across all individuals, suggestive of a haploid region). To retain only a high-quality nearly scaffold-level assembly, we removed scaffolds and contigs smaller than 1.5 Mbp (cutoff selected by visual inspection of scaffold lengths), resulting in retention of 94% and 87% of the putatively autosomal *Macrocystis* and *Nereocystis* genomes, respectively.

Because some downstream applications required the removal of closely-related individuals, we inferred relatedness between all pairs of individuals using *ngsRelate* v.2.[92] We subsetted our autosomal dataset to a minimum minor allele frequency (MAF) of 0.05 (*bcftools view -q 0.05:minor*) and thinned to a minimum distance of 10 kbp between SNPs using *VCFtools* v.0.1.16[95] and used the resulting VCF file to run *ngsRelate* with default parameters, using genotype likelihoods rather than called genotypes. We used the KING-robust kinship estimator (calculated from the 2D site-frequency spectrum[112]) to identify individuals that were genetically identical or first-degree relatives (i.e., parent-offspring and full-sib pairs), with thresholds for delimiting relatedness categories following.[95] We then selected the minimum set of individuals that needed to be excluded to create two sets of samples with no genetically identical individuals and additionally no first-degree relatives, respectively, retaining the individual with the higher depth of coverage from pairs where the choice of individual to exclude was arbitrary. From the un-thinned files with minimum MAF of 0.01, we created multiple datasets excluding 13 genetically identical and three additional first-degree relatives for bull kelp respectively, and 23 genetically identical and 16 additional first-degree relatives for giant kelp, respectively. As expected, all pairs of close relatives were from the same sampling site. Genetically identical individuals may reflect asexual reproduction, which has been reported in giant kelp,[49,113] or errors in sampling in which the same individual was sampled more than once (despite a minimum distance implemented between individuals) because it was sometimes difficult to tell from the surface of the water which blades belonged to which individual.

Finally, after removing close relatives, we thinned SNPs to a minimum distance of 10 kbp using *VCFtools*. We confirmed that 10 kbp was an appropriate thinning distance based on visual inspection of linkage-disequilibrium (LD) decay curves showing that LD had substantially declined from its maximum and 10 kbp was near the inflection point in the curve. We calculated LD using a 2% random sample of pairwise comparisons (--rnd_sample 0.02) using *ngsLD* v.1.2.1,[93] and plotted LD decay curves (with LD measured as r^2) using the *fit_LDdecay.R* script from *ngsLD*.

The above methods describe the biallelic SNP dataset(s) used in the majority of analyses, yet some analyses required information about all sites in the genome. We therefore also called genotypes at invariant and triallelic (or quadrallelic) sites using the same procedure as for our biallelic SNPs with minor modifications. Specifically, we repeated our initial *bcftools call* command with the variant flag (-v) removed to call genotypes at all sites in the genome. We identified invariant sites as those with only one allele observed in our samples (*bcftools view -Q 0.000001:nonmajor*) and excluded indels (-V indels). Subsequent filtering steps were the same as for biallelic SNPs except that the heterozygosity filter and variant quality filters (SGB, MQB, MQSB, RPB, and VDB) were not applied, nor was filtering for a minimum MAF. We identified triallelic (or quadrallelic) sites as SNP variants with a minimum of 3 alleles (*bcftools view -m 3 -v snps*), and implemented all filters except for a minimum MAF. Finally, we also reprocessed our biallelic SNPs in the same way as previously described, except without any minimum MAF, to generate a set of all biallelic SNPs in the genome regardless

of frequency. We then used *bcftools concat* to combine the separately-filtered invariant, biallelic, and triallelic (or quadrallelic) sites into a single dataset representing all high-quality filtered sites in the genome.

Derived minor allele classification

We identified and classified derived minor alleles (DMAs) into different categories of deleteriousness using (1) Genomic Evolutionary Rate Profiling (GERP)[51] and (2) the genetic variant annotator *SnpEff*[52] (Figure S7).

(1) Our GERP analysis closely followed.[54] GERP consists of inferring the rates of evolution at individual sites in the genome across a phylogeny of closely related species, to identify sites that are relatively conserved across taxa (where new mutations are more likely to be deleterious) and sites that evolve relatively rapidly (where new mutations are less likely to be deleterious).[51] We compiled a time-calibrated composite phylogeny of brown algae (Phaeophyceae) from existing sources, using[114] as a backbone, primarily for deep divergences within the class (e.g., between orders). We added additional branches and divergence times to the backbone using phylogenies that focused on subsets of brown algae.[99,115–118] For kelp specifically, we used the divergence times and branching structure from Starko et al.[118] All other branch times added to the composite brown algae phylogeny were also as given in original publications, except that divergences between *Choristocarpus* and *Discospor-angium* and between *Halopteris* and *Sphacelaria* from Kawai et al.[99] were proportionally rescaled relative to the earliest split within Phaeophyceae (Discosporangeales vs. others) according to,[114] to account for the substantially older early divergence times in Choi et al.[114] relative to Kawai et al.[99]

We downloaded reference genomes for species in the brown algae phylogeny (accession numbers in Table S3) and split the genomes into shorter fragments of 500 bp using the *reformat.sh* script of *BBmap* v.39.06.[94] We then aligned fragments to the focal species' (bull kelp or giant kelp) genome following[54] using *bwa-mem* v.0.7.17-r1188[87] with modified mismatch penalty (*-B 3*) and gap opening penalty (*-O 4,4*); used the *view* command in SAMtools v.1.17[88] to filter by mapping quality (*-q 2*) and remove reads aligning to multiple locations (*-F 2048*); and converted alignments to fasta format using the *pileup* command in *HTSBox* v.r345,[119] requiring minimum mapping and base qualities (-q 30 -Q 30), length (-l 35) and depth (-s 1), and printing a random allele (-R). If multiple alignments were available that mapped to the same branch on the phylogeny (e.g., multiple species in the same genus), the alignment with the highest number of primary mapped reads was retained and other alignments discarded. We then used a custom script heavily modified from[54] to combine all species into a single alignment for each scaffold or contig, with the focal species' genome excluded to reduce bias.

For each site in the genome, GERP involves calculating N, the neutral rate or expected number of substitutions across the phylogeny of outgroup species (based on the total branch length of taxa with aligned data at that site), and then S, the GERP score. Positive S indicates fewer empirical substitutions have occurred across the phylogeny than expected (i.e., higher evolutionary constraint), whereas negative S indicates more substitutions have occurred than expected (i.e., lower evolutionary constraint). To calculate N and S for each site, we used the modified *gerpcol* script (v.2023/11/20) from[54] to run *GERP++*.[120] We used our previously-calculated (see above) species-specific transition:transversion ratios (*-r*) and a brown algae substitution rate of 8.135×10^{-4} mutations per million years (i.e., per unit of branch length; *-s 0.0008135*). This substitution rate was derived by averaging substitution rates for *Ectocarpus* and *Scytosiphon* (4.07×10^{-10} and 1.22×10^{-9} substitutions per generation, respectively[121]) and assuming a generation time of one year. We also printed the alleles for each site (*-e*) at the three closest outgroups (either bull or giant kelp, *Saccharina japonica* and *Laminaria digitata*) for each species. Finally, we classified sites as being more evolutionarily conserved ($S > 0.5$) or more evolutionarily labile ($S \leq 0.5$), retaining only sites with $N > 0.5$. We excluded sites where $N \leq 0.5$ (following[51]) because it is not possible to detect evolutionary constraint when the expected number of substitutions (N) is close to zero, which occurs at sites with extensive missing data in the outgroup alignments. The choice of a threshold for binning sites according to S will depend on the available dataset and study goals. We chose a threshold of $S = 0.5$ to classify our sites (instead of $S = 0$, for example) because when $N > 0.5$, then all sites where $S \leq 0.5$ (including sites where $0 < S \leq 0.5$) will have at least one substitution in the outgroup phylogeny and we considered such sites to be better described as evolutionarily labile rather than evolutionarily conserved.

To incorporate the GERP scores with our SNP datasets, we used custom R scripts to determine the ancestral and derived alleles at each SNP site, and calculate the derived allele frequency globally and in each population. Using the alleles printed for the three closest outgroups, we considered an allele to be derived if it was not present in any of the outgroups (allowing for up to two outgroups to have missing data). If neither or both alleles were present in the three outgroups, ancestral and derived alleles could not be defined. Because genetic load is only relevant to putatively deleterious mutations, we excluded sites with a global derived allele frequency greater than 0.5 as these sites are unlikely to be deleterious. We obtained a total of 18,905 SNPs (9,396 labile and 9,509 conserved) for downstream analyses in bull kelp and 9,074 SNPs (4,884 labile and 4,190 conserved) in giant kelp.

(2) For our second method of classifying sites, we used *SnpEff* v.5.2a[52] to predict the impacts of derived mutations in protein-coding genes. *SnpEff* classifies predicted protein impacts as low, moderate, or high, with an additional modifier category for non-coding variants.[52] Gene annotations were available in the reference genome for giant kelp but not bull kelp. We transferred gene annotations from giant kelp to bull kelp using *Liftoff* v.1.6.3[96] using the options *-infer_genes -copies -a 0.95 -s 0.95 -d 5.0 -flank 0.8 -polish*, and retained only annotations flagged with a valid open reading frame. For both species, we built *SnpEff* databases from gene annotation files (*-gff3*) using the *build* command, and then added variant annotation to VCF files of SNP datasets with only identical individuals removed using the *ann* command with default parameters. We used custom

Current Biology
Report

CellPress

scripts to calculate global and population-specific ancestral and derived allele frequencies of annotated variants, using the same ancestral and derived definitions as determined above in the GERP analysis. We obtained a total of 115,549 non-modifier SNPs (51,988 with low impact and 63,561 with moderate or high impact) for downstream analysis in bull kelp and 254,297 SNPs (109,642 with low impact and 144,655 with moderate or high impact) in giant kelp.

QUANTIFICATION AND STATISTICAL ANALYSIS

Genetic structure

We performed principal component analysis using *SNPRelate* v.1.38.0[97,98] with default parameters, using the SNP datasets with up to first-degree relatives removed. We also performed genetic clustering analyses with *fastSTRUCTURE* v.1.0,[24] using the SNP datasets with up to first-degree relatives removed. We used simple priors and varied the number of clusters (K) from two to 10. For each K-value, we ran the program 100 times and selected the run with the highest likelihood. We used these 10 highest-likelihood runs to select the optimal range of K-values using the *chooseK.py* script distributed with *fastSTRUCTURE*. For bull kelp, the optimal value was K=8 by both reported criteria, but two of the clusters represented only one or two low-diversity populations in a small, isolated geographic area. As these two clusters may have represented differentiation due to recent bottlenecks rather than long-term regional genetic structure, we opted to use K=6 as the best model to represent regional genetic structure across BCWA. For giant kelp from BCWA only, the optimal value was K=7; from North America only, optimal K=7; from all global samples, optimal K was between 5 and 6 but we selected K=5 because K=6 contained one virtually unused cluster for which no individual had >1% ancestry.

We used the above-identified clusters to group populations for calculating pairwise genetic differentiation and divergence within BCWA for each species, and between the Southern and Northern Hemispheres and between BCWA and California in giant kelp. We calculated pairwise genetic differentiation (Weir and Cockerham's[122] F_{ST}) using *hierfstat* v.0.5-11,[100] using the SNP datasets with up to first-degree relatives removed. We calculated pairwise genetic divergence (d_{XY}, the average number of nucleotide differences between two random individuals from different populations[123]) using *pixy* v.1.2.7.beta1,[124] using the SNP datasets with only identical individuals removed and containing both variant and invariant sites.

To test for a pattern of isolation by distance, we plotted d_{XY} against the geographic distance between populations by the shortest ocean route (km). The geographic distance was calculated by converting a polygon of the BCWA coastline[125] into a raster at 1-millidegree resolution using the *rasterize()* function in the R package *raster* v.3.6-26.[126] This high resolution was required to accurately represent connectivity along BC's complex coastline, but calculations between very distant populations became computationally intractable. We therefore split the coastline into two regions: (1) the Salish Sea and Vancouver Island and (2) northern BC. In northern BC we coarsened the resolution to 2-millidegrees to aid computation. We manually inspected both rasters to ensure that population sampling locations were accurately rasterized as ocean and not land and that narrow passages between islands through which kelp might disperse were fully passable, manually converting pixels from land to ocean if needed. We used the R package *gdistance* v.1.6.4[101] to calculate an 8-directional transition matrix for each raster using the *transition()* function; to correct the transition matrix, to account for the fact that degrees of latitude and longitude are not equal in distance, using *geoCorrection*() with type "c" correction; and to calculate the least-cost path between populations using *costDistance()*. For population pairs where one population was located in each of the two rasters, we calculated the distance between each population and an intermediate coastal point shared between the two rasters (Cape Caution, BC; 51.165°N, 127.797°W) and then summed the two distances. This method provided computational tractability and also forced populations to disperse along the central coast of BC rather than across the open waters of Queen Charlotte Sound, which is likely a biologically realistic representation given the assumption of stepping-stone dispersal between populations. After obtaining geographic distances, we performed linear regressions of the relationship between d_{XY} and geographic distance using the *lm()* function in *R* v.4.2.3.[102] Unless stated otherwise, all simple linear regressions in this study were also performed with *lm()*. To obtain adjusted p-values for the relationship between d_{XY} and geographic distance, we performed Mantel tests[127] by permuting the geographic distance matrix 1,000 times.

Genetic health indicators and selfing rate

We calculated three genetic health indicators for each population:

(1) We estimated effective population size (N_e) using *roh-selfing* v.2024/02/23.[103] Because *roh-selfing* requires information on runs of homozygosity (ROHs) as input, we masked repetitive regions of the genome to ensure that potential misidentified SNPs in these regions would not prevent accurate inference of ROHs. We identified repetitive regions in the reference genomes with *Red*[104,128] using the *Red2Ensembl.py* script distributed with *Ensembl Plants* v.1.2,[129] and then masked them in the SNP datasets with only identical individuals removed using the *intersect* command in *BEDTools* v.2.30.0.[105] We then further masked potentially problematic regions of the genome by excluding short heterozygous regions surrounded by ROHs. To do so, we estimated an initial round of ROHs for each individual in each population from the repeat-masked SNP datasets using the *roh* command of *BCFTools* v.1.19,[88] with allele frequencies calculated automatically for populations of ≥ 4 individuals by *BCFTools*, or else using global allele frequencies across all individuals (provided with the –AF-file flag) for populations with <4 individuals. We then used a custom *R* script to identify short runs of heterozygosity (ROHets; the inverse of ROHs) ≤10 kbp in length and surrounded by ROHs in each individual. We also used a custom *R* script to identify 10-kbp

windows of the genome that contained a short ROHet in greater than n individuals, where n was equal to the 99.99th percentile of a Poisson distribution with parameter λ equal to the mean number of individuals containing a short ROHet across all 10-kbp windows of the genome. Finally, we masked these windows (in addition to the mask for repetitive regions) in our SNP datasets by re-running *Red* and *BEDTools* as described above. Using these masked SNP datasets, we then estimated a final round of ROHs for each individual in each population using the *roh* command of *BCFTools* as described above, retaining results only for populations of ≥ 4 individuals.

In addition to ROHs, *roh-selfing* requires estimates of inbreeding coefficient (F) and Tajima's[130] D as input. For each population, we calculated F using *PLINK* v.1.90b6.21[106] and Tajima's D using *VCFtools* v.0.1.16.[91] We then ran the *RF-sequential* model (model ID 202310021917048AtJy) of *roh-selfing*[103] on each population. We generated a single estimate of N_e for each population by taking the mean of $\log_{10}(N_e)$ for each autosomal scaffold or contig.

(2) We calculated nucleotide diversity (π)[123] for each population using *pixy* v.1.2.7.beta1,[124] using the SNP datasets with only identical individuals removed and containing both variant and invariant sites.

(3) We calculated mean inbreeding coefficients (F_{ROH}) for each population, representing the proportion of the genome that is in ROHs. The estimation of ROHs using *BCFTools* was described above. We then calculated F_{ROH_100kbp} for each individual as the summed length of ROHs ≥ 100 kbp divided by the total length of all scaffolds and contigs for which ROHs were inferred, and calculated the mean across individuals within each population.

In addition to these three genetic health indices, we also calculated the observed selfing rate in each population from long ROHs (≥ 500 kbp), with selfed individuals expected to have approximately 50% of their autosomal genomes in long ROHs. We used *BCFTools* above to calculate ROHs for *roh-selfing* as the program is trained on *BCFTools* output, but were concerned that any false SNP regions not removed by our SNP filtering and masking procedures might break up long ROHs, making it difficult to infer accurate selfing rates. We therefore identified long ROHs using *ROHan* v.1.0.1,[107] which classifies ROH status in large windows, does not rely on called genotypes, and can accommodate a background heterozygosity rate in putative ROHs. For each individual, we ran *ROHan* using indel-realigned BAM files (described in the alignment to reference genome section), restricting the analysis to large autosomal scaffolds (*–auto*), using 500-kbp windows (*–size 500000*), with an expected background heterozygosity rate in non-ROH regions (*–rohmu*) of 5.0×10^{-4} for bull kelp and 6.0×10^{-4} for giant kelp, and supplying a species-specific transition:transversion (TSTV) ratio (*–tstv*). The TSTV ratio was 1.45 for bull kelp and 1.75 for giant kelp, calculated using *bcftools stats* from the SNP datasets with up to first-degree relatives removed. The background heterozygosity rate was determined heuristically for each species by running *ROHan* on several individuals strongly suspected of being selfed (based on visual inspection of Manhattan plots of observed heterozygosity, described below) and plotting histograms of the heterozygosity inferred by *ROHan* across all 500-kbp windows (with at least 80% of the sites having data). For selfed individuals, the distribution of heterozygosity is expected to be bimodal, with two large peaks corresponding to windows in ROHs and not in ROHs, respectively. The background heterozygosity rate was selected to approximately correspond to the maximum value of the first peak of this distribution. After running *ROHan* with optimized parameters for each individual, the mean F_{ROH_500kbp} was calculated for each individual as the proportion of 500-kbp windows inferred to be in ROHs using *ROHan*'s mid-value estimates of heterozygosity.

Selfing rate was then calculated as the proportion of individuals in each population (using non-identical individuals only) with $F_{ROH\_500kbp} > 0.3536$. The expected F_{ROH} of selfed individuals is 0.5, while parent-offspring pairs and full siblings are expected to have $F_{ROH} = 0.25$. We used 0.3536 as the threshold for binary classification following a proposed inference criterion for distinguishing 0.5 and 0.25 kinship coefficients.[95] In two bull kelp populations (NL-PS-06 and NL-PS-07), selfing rate could not be reliably determined from F_{ROH_500kbp} because *ROHan* was unable to classify most segments of the genome as either ROH or non-ROH given that the entire genome had extremely low genetic diversity. For these populations, we additionally used the criterion $F_{ROH\_100kbp} > 0.3536$ from *BCFTools* to identify individuals that could potentially be selfed. We then confirmed the inference of selfing by visual inspection of Manhattan plots of observed heterozygosity (H_o) in 100-kbp windows across the genome, calculated using *pixy* as described above for calculating π. We visually confirmed the expected presence of ROHs spanning entire chromosomes or the majority of chromosomes in selfed individuals. In addition, in one giant kelp population (MP-NC-02), all individuals had high F_{ROH_500kbp}, suggesting that non-selfed individuals could potentially exceed the 0.3536 threshold. We reclassified four individuals from this population as non-selfed after visual inspection of Manhattan plots suggested that the genomes had numerous smaller ROHs, but few ROHs approaching chromosome length.

Tests for purging and genetic drift

For both GERP and *SnpEff* analyses, we used custom scripts to test for purging and examine the effects of genetic drift in small populations. We expected that purging would remove putatively deleterious alleles from small populations but have no effect on frequencies of alleles in less deleterious categories. We considered derived minor alleles (DMAs) at evolutionarily conserved sites to be putatively deleterious in GERP analyses and DMAs at moderate- and high-impact sites to be putatively deleterious in *SnpEff* analyses. We calculated the mean frequency of DMAs in each allele category from sites genotyped at a minimum of three individuals in each population, and expected a positive relationship between DMA frequency and N_e.

After determining that there was no evidence of purging, we tested for the expected putative signatures of genetic drift on DMA frequency. To facilitate comparisons among populations that contained different numbers of individuals, for each population we

Current Biology
Report

CellPress

sampled three non-missing genotypes per site (n = 100 sampling replicates). We calculated the mean (across sampling replicates) number of sites with at least one derived allele present, the mean frequency of derived alleles that were present, and the mean fixation rate of derived alleles that were present. We used simple linear regressions to test the expectations that genetic drift would reduce the number of sites with a derived allele present in small populations relative to large populations, and that remaining derived alleles in small populations would have higher mean frequency and be more likely to be fixed. We also estimated the realized load from the resampled datasets as the proportion of genotypes homozygous for the DMA (at evolutionarily conserved sites and moderate- to high-impact sites only) and tested for an expected negative correlation with N_e.

Genetic load under different cross types

We examined the effects of selfing on realized genetic load using empirical genetic load estimates from selfed and non-selfed individuals. We first used custom scripts to estimate the realized genetic load for all individuals at putatively deleterious DMAs for both the GERP and *SnpEff* analyses. Using the definitions of which individuals were selfed or non-selfed (described above), we then calculated the relative difference in realized genetic load between each selfed individual and the mean realized genetic load of all non-selfed individuals in the corresponding population. We used linear regressions to test for a relationship between this relative difference in realized load and N_e. Because data points were not independent (i.e., in some cases multiple selfed individuals were compared to the same population mean), we calculated adjusted p-values for the relationship by permuting N_e 1,000 times and taking the proportion of permuted *t*-statistics that were greater than the empirical *t*-statistic.

We also predicted the effects of different cross types on realized genetic load by comparing simulated crosses within and between populations. We used custom scripts to randomly sample one individual per population for each pairwise combination of populations. For within-population crosses we ensured that first-degree relatives were not sampled and the same individual was not sampled twice, so that we would not simulate any highly inbred or selfed individuals that could confound comparisons. For each sampled pair of individuals, we randomly selected one allele at each site and calculated the realized genetic load of DMAs in the simulated offspring. Sampling of individuals was repeated 100 times for each pair of populations, and we calculated the mean realized load across all replicates of each pair.

To test whether outcrossing between populations reduced realized genetic load relative to crossing within populations, we calculated the relative difference in the mean realized load (across the 100 sampling replicates) of between-population crosses relative to that of within-population crosses for each recipient population. The recipient population was defined as the population used as the comparison in the within-population cross and the donor population as the other population. For example, considering recipient population A and donor population B, the relative difference in realized genetic load was calculated for cross A x B relative to cross A x A. Switching the definition of the recipient and donor populations results in a second comparison of crosses B x A and B x B for the same population pair.

To test our prediction that the reduction in mean realized load upon between-population outcrossing would be greater when the recipient population was small and the donor population was far away, we performed multiple linear regressions using the *lm()* function in *R* v.4.2.3[102] with the relative realized genetic load as the response variable and the N_e of the recipient population and the geographic distance between populations as predictor variables. Because the data points were not statistically independent, we calculated adjusted p-values by permuting either N_e or the matrix of geographic distances while holding all other values constant and taking the proportion of permuted *t*-statistics more extreme than the empirical *t*-statistic for the variable of interest. To visualize the three-dimensional relationship between N_e, geographic distance, and relative realized load, we treated the predictor variables as a two-dimensional landscape and performed smoothing and interpolation of the response variable across this landscape using *snapKrig* v.0.0.2,[108] with a grid of 51 x 51 cells for kriging and default parameters to select the maximum likelihood model.