Mar Biol (2010) 157:81–89
DOI 10.1007/s00227-009-1297-8

ORIGINAL PAPER

New evidence for habitat-specific selection in Wadden Sea *Zostera marina* populations revealed by genome scanning using SNP and microsatellite markers

Katharina Oetjen · Steven Ferber · Ilka Dankert · Thorsten B. H. Reusch

Received: 20 May 2009 / Accepted: 1 September 2009 / Published online: 22 September 2009

Abstract Eelgrass *Zostera marina* is an ecosystem-engineering species of outstanding importance for coastal soft sediment habitats that lives in widely diverging habitats. Our first goal was to detect divergent selection and habitat adaptation at the molecular genetic level; hence, we compared three pairs of permanently submerged versus intertidal populations using genome scans, a genetic marker-based approach. Three different statistical approaches for outlier identification revealed divergent selection at 6 loci among 46 markers (6 SNPs, 29 EST microsatellites and 11 anonymous microsatellites). These outlier loci were repeatedly detected in parallel habitat comparisons, suggesting the influence of habitat-specific selection. A second goal was to test the consistency of the general genome scan approach by doubling the number of gene-linked microsatellites and adding single nucleotide polymorphism (SNP) loci, a novel marker type for seagrasses, compared to a previous study. Reassuringly, results with respect to selection were consistent among most marker loci. Functionally interesting marker loci were linked to genes involved in osmoregulation and water balance, suggesting different osmotic stress, and reproductive processes (seed maturation), pointing to different life history strategies. The identified outlier loci are valuable candidates for further investigation into the genetic basis of natural selection.

Communicated by S. Uthicke.

Electronic supplementary material The online version of this article (doi:10.1007/s00227-009-1297-8) contains supplementary material, which is available to authorized users.

K. Oetjen · I. Dankert
Institute for Evolution and Biodiversity,
Plant Evolutionary Ecology, University of Münster,
Hüfferstrasse 1, 48149 Münster, Germany

S. Ferber
Department of Marine Benthic Ecology and Evolution,
Centre for Ecological and Evolutionary Studies (CEES),
University of Groningen, PO Box 14,
9750 AA Haren, The Netherlands

K. Oetjen · T. B. H. Reusch (✉)
Evolutionary Ecology of Marine Fishes,
Leibniz-Institut für Meereswissenschaften (IFM-GEOMAR),
Düsternbrooker Weg 20, 24105 Kiel, Germany
e-mail: treusch@ifm-geomar.de

Introduction

Seagrasses are a polyphyletic group of angiosperms that are ecosystem-engineering species by providing three-dimensional structure on soft sediment (Duarte 2002; Orth et al. 2006). Ecosystem services of seagrass beds include the provision of nursery habitat and food source for numerous fish and invertebrate species, the stabilization of sediments and carbon and nutrient storage (Hemminga and Duarte 2000). Because these ecosystems are distributed worldwide and range among the most valuable ecosystems (Costanza et al. 1997), ongoing population declines during the last decades are of serious concern (Waycott et al. 2009). The responsible anthropogenic changes in coastal areas include changes in current regime (leading to increased suspension and turbidity and decrease in water transparency), eutrophication and mechanical disturbance (Hemminga and Duarte 2000).

Zostera marina L. (eelgrass) is the dominant seagrass on the northern hemisphere and is widely distributed across European shallow waters. In the Wadden Sea area, eelgrass is subtidally almost extinct after the wasting disease in the 1930s (den Hartog 1970). Recovery never occurred, possibly

due to ongoing pollution of coastal waters and changes in water movement and climate (van Katwijk and Hermus 2000). However, some fragments in permanently submerged creeks remained, situated on the low lying islands of the northern Wadden Sea (the 'Halligen', Reusch 2002). In contrast, there are still substantial populations in intertidal areas (Reise et al. 2005). Due to its ecological importance, Wadden Sea populations of eelgrass have been the focus of recent genetic research that provided valuable insights on genetic diversity, colonization history and genetic exchange in populations (Reusch 2002; Olsen et al. 2004). The increasing genetic information available for *Z. marina*, including an EST (Expressed Sequence Tag) database available at http://drzompo.uni-muenster.de/ (Reusch et al. 2008; Wissler et al. 2009) now allows investigating the genetic basis of physiological adaptation to extreme natural environments such as tidal flats (Reise 1985).

Wadden Sea eelgrass populations offer an outstanding opportunity to gain insights into habitat-dependent selection using two habitat-specific phenotypes that are distinguished by their life cycle. The subtidal form lives permanently submerged in creeks and reproduces asexually by forming perennial clones via rhizome expansion, in addition to sexual reproduction via seeds. An intertidal form growing on the extended tidal flats of the Wadden Sea area is subject to much higher yearly and daily fluctuation in temperature and salinity, because only a thin film of water remains when plants fall dry (Massa et al. 2009). Moreover, leaves are exposed to increased levels of UV light during low tide. Reproduction only takes place sexually via flowers and seeds, while asexual reproduction is rare (Reusch 2002).

We thus hypothesized that different habitat conditions have resulted in local adaptation visible not only at the phenotypic level, apparent as different life history strategies, but also at the molecular genetic level. To test this, genome scans recently became one of the most promising molecular genetic approaches. Genome scans extend traditional population genetics approaches to large marker numbers coupled with statistical tests in order to identify genetic loci influenced by selection (Schlötterer 2002b; reviewed by Luikart et al. 2003; Stinchcombe and Hoekstra 2008). The principle of detection is simple. If molecular markers are physically linked to functionally important and polymorphic genes, divergent selection acting on such genes can be detected via genetic hitchhiking (Lewontin and Krakauer 1973; Smith and Haigh 1974; Harr et al. 2003; Schlötterer 2003; Storz 2005). Whereas random or demographic effects affect all loci in the same way, effects of divergent selection lead to increased differentiation at certain target loci (Nielsen 2005), quantified as Wright's F_{ST}. A critical step is to test the statistical significance of the numerical value to be outside neutral expectations. Here, coalescent simulation–based methods have made great progress that compare population differentiation at single loci with a model distribution and allow detection in patterns not consistent with neutral evolution (Beaumont and Nichols 1996; Vitalis et al. 2001). Another approach uses empirical data to construct the null distribution and is able to detect a reduction in genetic diversity at certain loci, which is caused by a 'selective sweep' (Schlötterer 2002a).

Here, we simultaneously use three model approaches to examine the mutual consistency of the different methods. Even more importantly, we assessed three repeated habitat contrasts in our system. If one locus is detected erroneously in one habitat contrast, it should not appear as an outlier in the other two contrasts (Wilding et al. 2001; Luikart et al. 2003; Campbell and Bernatchez 2004; Storz 2005; Vasemagi et al. 2005; Bonin et al. 2006).

For our genome scan, we employed three classes of genetic markers: anonymous microsatellites, gene-linked microsatellites derived from an EST (expressed sequence tag) library, and SNPs (single nucleotide polymorphisms). Microsatellites are among the most widely used molecular markers for population genetic assays and combine high information content per marker locus with relatively easy and cost-effective development and screening. Anonymous markers are derived from standard enrichment libraries for microsatellite motifs and do not provide any information with respect to the vicinity of the marker. In contrast, gene-linked microsatellites are located close to a gene (often in the untranslated regions) and are typically derived from an EST (expressed sequence tag) library (Bouck and Vision 2007). In genome scans, gene-linked markers represent a very useful means to assign functional genes possibly responsible for the effect seen at an outlier locus (Vasemagi et al. 2005; Namroud et al. 2008).

SNPs represent the simplest genetic polymorphism possible, the substitution of one nucleotide by another. Such sites occur in very high numbers in genomes of higher organisms. Recently, they have been regarded as the new genetic marker of choice, due to high frequencies, simple mutation mechanism and low error rates (Ryynanen et al. 2007). Information concerning the utility of SNPs for population genetic and evolutionary studies in nonmodel organisms is scarce, in particular in genome scans (but see Rengmark et al. 2006).

An important additional aim was to test the consistency of genome scan approaches via re-examining a given set of contrasting populations with an extended number of genetic markers. On one hand, neutral loci may be classified as being under selection due to incorrect models of demographic history (e.g. island vs. stepping stone models, see Akey et al. 2004) or ascertainment bias (Thornton and Jensen 2007). On the other hand, some loci influenced by

selection might be missed by genome scanning (false negatives, Teshima et al. 2006). We will thus for the first time retest an earlier genome scan in order to evaluate the usefulness of genome scans for detection of selection.

Materials and methods

Collection of samples and DNA extraction

Three *Z. marina* population pairs from the North Frisian Wadden Sea (Northern Germany) were sampled. We extended the analysis of samples described in Oetjen and Reusch (2007) by reanalysing six populations sampled in 2005. Samples from three subtidal populations and three intertidal populations were reused, and no further sampling was done for this study. Sampling sites were located on three islands: Hallig Hooge, Hallig Langeness and Sylt island (for coordinates refer to Oetjen and Reusch 2007). While intertidal populations were located at the shores of the islands in the tidal flat area (Langeness: ca 1 km from shoreline, Hooge: ca 100 m, Sylt: ca 100 m), subtidal populations grew in tidal creeks inland the islands. Each population pair comprised one subtidal location and one intertidal location, thus spanning a habitat contrast while being maximally 4 km distant from each other. Maximal distances between sampling sites was 53 km. We collected samples at $\geq$1 m random points by areal sampling on 40 m $\times$ 20 m on tidal flats, and by sampling transects along the creeks which was dictated by their elongated shape. By this sampling strategy, we avoided sampling clones more than once. As earlier research had shown, on tidal flats, plants grow mainly individually and not in clone patches and clones in creeks normally do not exceed 1 m in diameter (T. B. H. Reusch, personal observation). Sample sizes ranged between 39 (Sylt subtidal) and 50 individuals (Langeness intertidal). In total, 284 individuals were collected at the 6 locations, none of which were clone members. Seagrass samples were taken and total genomic DNA was extracted as described by Reusch (2002).

Genetic markers

In total, 37 single nucleotide polymorphisms and 40 microsatellite markers were initially tested in a genomic scan approach. SNP markers were developed on a transoceanwide panel of genotypes including Pacific and Atlantic populations by S. Ferber. PCR conditions and genotyping procedure are described in Ferber et al. (2008). Initial testing revealed that of the 37 SNP markers, only 6 were polymorphic at the spatial scale of our study (Table 1).

The microsatellite loci included 29 gene-linked and 11 anonymous markers (EMBL nucleotide database, accession no. AM408830-AM408843, AJ249303-AJ249307, AJ009898, AJ009900, AJ009901 and AJ009904) (Oetjen and Reusch 2007). Of the gene-linked microsatellites (Table 1, accession numbers FN4353336-FN435350), 15 were newly developed from an EST database (http://drzompo.uni-muenster.de/). The database was screened for microsatellite motifs using the PERL script MISA (http://pgrc.ipk-gatersleben.de/misa/) under the prerequisite that dinucleotide motifs had to include at least seven repeats, tri- and tetranucleotide motifs at least five repeats.

Genotyping

Taking SNP and microsatellite markers together, we assessed 46 genetic markers in total. Genotyping was carried out on a 3130xl ABI automated sequencer (Applied Biosystems). No clone was sampled twice, so we proceeded analysing all samples. Of the 284 individuals genotyped, only 13 samples lacked information at one locus. For the microsatellites, we used standard fluorescent PCR with 6-Fam and Hex-labelled forward PCR primers. For the SNP scoring, we used the SNAPshot kit (Applied Biosystems) following the protocols in Ferber et al. (2008). The size calling and allele binning were performed with the software GeneMapper (Applied Biosystems) against ROX 350 (microsatellites) and Liz 120 (SNP) size standards. The calculation of allele frequency, expected and observed heterozygosities and population differentiation (F_{ST}) was performed with the software Microsatellite Analyzer (MSA; Dieringer and Schlötterer 2003). Tests for null alleles were carried out using the Microchecker 2.2.3 software (Van Oosterhout et al. 2004) with a 99% significance threshold and revealed no homozygote excess at any of the loci except for the anonymous microsatellite GA35 that was not detected as an outlier in any analysis.

Population structure

As baseline data prior to genome scan analyses, we carried out several nested analyses of molecular variance (AMOVA) to determine the component of genetic variation (estimated as F_{ST}) due to habitat contrasts. Genetic variation was partitioned among sampling sites and among population pairs nested within sites. Consequently, we grouped populations according to sampling sites, each group containing one subtidal and one intertidal population. In doing so, we summarized the effect of differentiation across the three population pairs sharing the same habitat contrast. Thus, nested AMOVA is capable of revealing a potential signal of divergent selection that might not be traceable else wise. Comparison of AMOVA results regarding the group of loci identified as outliers before and the remaining (neutral) loci gives insights into the effect of habitat differences

Table 1 Eelgrass *Zostera marina*: 15 newly developed microsatellites and six SNPs loci first used in a genome scan

Locus name	Acc. No.	Type	Repeat motif	Primer sequences	Allel size (range)	No. of alleles	He	PCR multiplex	Primer conc. (μM)	No. cycles
CL766Contig1	FN435336	Msat	(GA)9	F: *FAM*-GAACGTTTCCCGGTCATTT R:GGAATCGGTCAAGCAAAAAC	122–128	4	0.066	Pool F	0.02	29
CL11Contig1	FN435337	Msat	(AGC)5	F:*FAM*-GTGGAGGAAAGTGTGGGTGT R:CTTGCATCCACCTTCATTTG	294–300	3	0.022	Pool G	0.5	30
CL559Contig1	FN435338	Msat	(AG)10	F:*FAM*-CCACTTCCGTAGTTGCTGTT R:CGATGAGGACGATGAGGAAT	171–177	4	0.153		0.14	
ZME02125	FN435339	Msat	(CCT)5 (CTT)5	F:*FAM*-CGTTCAACTCAACACGCATT R:GGTGACGAAAAGAAGCGAAG	103–112	4	0.043		0.05	
ZMF02381	FN435340	Msat	(ATC)8	F:*HEX*-GTGCAGGCGATCGAGTTATC R:AAATTCGAGCTCTCAACTTCAA	121–154	11	0.671		0.25	
CL202Contig1	FN435341	Msat	(TA)7	F:*FAM*-TTGAAAAGATTAATTATTGGTGGTG R:TCAAGTCCGGATAAATTCGAT	200–206	4	0.121	Pool H	0.45	29
CL380Contig1	FN435342	Msat	(CATC)5	F:*FAM*-CCGCCTTCTTCTTCGTTAGA R:TGTTGTTCTTGGAAAAGAATCAGT	121–133	5	0.069		0.1	
CL805Contig1	FN435343	Msat	(AG)8	F:*HEX*-GGGGAGGTTTCCGAATACTTT R:TGGAAGATGTTGGACATGGA	184–208	4	0.034		0.35	
CL172Contig1	FN435344	Msat	(TGGC)8	F:*FAM*-CTCCTGGACGCAGAAATATG R:GACAAACGATTAATTCAGAAACAAAA	184–208	7	0.297	Pool I	0.15	27
CL53Contig1	FN435345	Msat	(GAT)6	F:*FAM*-AACTCCTGGCGCAACTACTG R:CTTCGTTTGCGGTTGCTT	90–93	2	0.011		0.15	
ZME06302	FN435346	Msat	(GCA)5	F:*FAM*-TCTAGCTTGTCGATGGCTGA R:CCGTCAAATGTTTCCAAGGT	291–297	2	0.022	Pool J	0.2	28
ZMC05062	FN435347	Msat	(CT)8	F:*HEX*-GAAGCCAACTTAATTCAACATCG R:TTAATATAAATCCGAGACACAGACTC	96–100	3	0.007	Pool K	0.225	29
ZMC19062	FN435348	Msat	(GAC)5	F:*FAM*-CACTCTCCTCTTTCCGTTCG R:CAGGGGCCTTCCTCTTACTC	293–296	2	0.003		0.125	
ZME05315	FN435349	Msat	(AC)8	F:*FAM*-AAACGAGATGGTGGTTCCAT R:TGCGAGCAGCTAACTAAGTCC	174–185	7	0.057		0.125	
ZME02369	FN435350	Msat	(TTC)5	F:*FAM*-AAGTCGAAATGGGGATACCA R:TCGTCGGAAGAAAAAGAAGC	99–102	2	0.007	Pool L	0.15	30
287CT623*		SNP					0.089			
71CA329*		SNP					0.403			
20CT465*		SNP					0.403			
71TC178*		SNP					0.398			
95GT380*		SNP					0.492			
98AT518*		SNP					0.330			

For the microsatellites (*Msat*), the primer sequences and the repeat motifs are given. Two to five primer pairs were pooled. PCR were performed as follows: 0.2 μl DNA template in a 10 μl reaction; 1 μl 5× reaction buffer (Promega); 0.2 mM of each dNTP; 1.5 mM $MgCl_2$ and 0.25 units of Taq polymerase (Promega). Thermocycling programme: initial denaturation 94°C for 3 min, followed by 27–30 cycles of 1-min denaturation 94°C, 1-min annealing at 56°C and 1-min extension at 72°C, followed by a final extension of 10 min at 72°C. * Indicates SNP which were assessed according to Ferber et al. (2008)

on genetic variation in the two loci groups. All AMOVAs were carried out with the software Arlequin 3.11 (Excoffier et al. 2005).

Testing for divergent selection

We applied three different tests for diverging selection to the three population pairs on tidal flat or in subtidal creeks. Due to the spatial proximity of the population pairs, we concentrated on the detection of diversifying (divergent) selection which is defined here as selection that acts in contrasting directions in two populations. The first test was implemented in the software Fdist2 (http://www.rubic.reading.ac.uk/~mab/software/fdist2.zip) by Beaumont and Nichols (1996) compares F_{ST} estimates of single loci in population pairs with a joint distribution based on weighted mean F_{ST} of all loci. The distribution is constructed using coalescent simulations. From this distribution, individual V values for every single locus are calculated. We regarded loci that lie outside the 0.95% quantile of the distribution as outliers, i.e. corresponding to a P-value <0.05 (one-sided test). We applied an infinite allele model that corresponds better to the allelic states and distribution of our data than the alternative, a stepwise mutation model. Settings of the

model also comprised assumption of 100 demes (the maximum number), being the most conservative version after testing, number of populations was set to 2, and median sample size corresponded to the respective population pair (ranging from 98 to 85, as the model uses haploid data).

The second model approach is based on different assumptions concerning population history, albeit using population differentiation parameters to assess locus behaviour also. It was developed by Vitalis et al. (2001) and implemented in the software DetSel 1.0 (Vitalis et al. 2003). The coalescent simulations are based on the assumption that the two populations of a pair originate from one population that has split at a time point in the past. We applied three mutation rates ($\mu = 0.001$, 0.005 and 0.01) that were realistic as former investigations of pedigree data had shown (TBH Reusch, unpublished data). We used the maximum population size before split ($N_0 = 500$) and simulated ancestral effective population sizes (N_e) of 500, 1,000 and 10,000 individuals. Two other parameters were modulated until the achieved allele distribution resembled our original data, which is recommended by the authors of the model. T_0 (time since an assumed bottleneck) was set to 50, 100 and 200 generations, and t (=time since the population split) to 50 generations. As allele numbers for all SNPs were restricted to two, this may distort the fit of the model data because microsatellites typically have >5 alleles. Hence, we repeated all analyses with this model excluding the 6 SNP loci from the data. 100,000 coalescent simulations created the distribution of the population differentiation parameters *F1* and *F2*. Identification of outlier loci was done by comparing the distribution conditioned on the number of alleles ($k = 2$, 3, 4, 5 and ≥ 6) with the *F1*/*F2* values of single loci.

As a third approach we used lnRH test that compares the variability of single loci within a larger sample of marker loci genome (Kauer et al. 2003). This is done by calculating the ratio of gene diversity (H, heterozygosity) in two populations of all loci. As such, the lnRH test should be most influenced by the specific group, and sample size of markers chosen. It has been shown that lnRH is approximately normally distributed under neutrality (Kauer et al. 2003). Constructing a normal distribution from the data obtained of our 47 loci, we tested whether the lnRH values of the single loci were located outside the 95% confidence interval.

As this study is an explorative one, we did not apply a strict Bonferroni correction to the significance levels of outlier tests, in line with other such studies in the field (Vasemagi et al. 2005; Bonin et al. 2006).

Functional assignment

The function of genes (according to BLASTX hits) linked to loci detected in outlier tests was categorized using gene ontology (GO) functional groups. The gene ontology system has been developed as a standardized vocabulary to characterize genetic functions and to allow comparison of genes across species, organisms, tissues or cells (Ashburner et al. 2000; Harris et al. 2006). The system comprises three main categories, 'molecular function', 'biological process' and 'cellular component', for all of which entries for a certain gene can exist. In each of the categories, subcategories with partly hierarchical order comprise bigger or smaller groups of genes under broad- to fine-scale terms.

Results

Population differentiation and genetic diversity

Among the six Wadden Sea *Z. marina* populations, the level of genetic polymorphism varied considerably among the 46 loci. The expected heterozygosity of microsatellites ranged between 0.003 (locus ZMC19062) and 0.897 (locus GA35). As expected, heterozygosity values for the biallelic SNP loci were lower and ranged between 0.08 (287CT623) and 0.49 (95GT380). The number of alleles of microsatellites varied between 2 (CL53) and 29 (GA35). The global F_{ST} estimate according to Weir and Cockerham (1984) across all populations indicated low but significant genetic differentiation ($F_{ST} = 0.017$, $P < 0.0001$). Pairwise genetic differentiation between the three population pairs was also low but nevertheless significant. F_{ST} values for the contrasting population pairs on Hallig Hooge were 0.008, and 0.004 and 0.014 for Sylt island and Hallig Langeness (all markers). Excluding the six polymorphic SNP loci changed F_{ST} values very little.

A comparison of habitat types revealed slightly lower average allelic richness in subtidal (4.03) than in intertidal habitats (4.42), an effect was almost significant (two-tailed t-test, $P = 0.08$). Expected heterozygosity was similar between subtidal (0.36) and intertidal (0.37) habitats (ns, $P = 0.37$). In an AMOVA, most of the genetic variation was explained by the within and among individual component (98.16%), while only 0.86% of the genetic variance was explained by habitat differences.

Tests for divergent selection

The three model approaches we used for identification of loci influenced by selection revealed partly consistent results (see Table 2 for an overview on candidate loci). The approach by Beaumont and Nichols detected only few loci which is consistent with earlier results (Vitalis et al. 2001; Vasemagi et al. 2005; Oetjen and Reusch 2007). The DetSel model developed by Vitalis et al. (2001) detected several loci as being subject to selection. Reassuringly, all three loci that had been detected in an earlier study with far

Table 2 Microsatellite loci in eelgrass *Zostera marina* detected as statistical outliers in comparisons among tidal flat and subtidal populations either by more than one test procedure (FDist, DetSel, lnRH), or in more than one habitat contrast

Outlier locus	Habitat comparison/population pair									Homology/best BLASTX Hit
	Hooge			Langeness			Sylt			
	Fdist	DetSel	lnRH	Fdist	DetSel	lnRH	Fdist	DetSel	lnRH	
ZMC19089				**0.014**	**<0.01**					*Arabidopsis thaliana* hypothetical protein CAB81404
CL734	**0.019**	<0.02								*Oryza sativa* unknown protein NP_913907
GA20		**<0.01**			**<0.01**			**<0.01**		Anonymous microsatellite
ZMC01058		<0.02			<0.03			**<0.01**		*Gossypium hirsutum* vacuolar H^+-ATPase AAA82977
ZMC19066		<0.03			**<0.01**	**0.018**				*Arabidopsis thaliana* acid phosphatase NP_194655
CL412		**<0.01**						**<0.01**	0.0280	*Oryza sativa* putative nodulin 3 XP_465955
ZME05315		**<0.01**	**0.0188**					<0.03		*Arabidopsis thaliana* putative ubiquitin-conjugating enzyme 16 Q9fwt2
CL380		**<0.01**	**0.0004**							*Arabidopsis thaliana* nucleotide sugar epimerase-like protein Q9m0b6
CL559			**0.013**			**0.0125**				*Oryza sativa* putative 24-kda seed maturation protein Q8s2k0
CL766					**<0.01**	**0.0002**				*Oryza sativa* putative ddt domain-containing protein Q6zi78
ZME2125								**<0.01**	**0.0002**	*Oryza sativa* putative myb-related protein Q9aut3

Significant (bold figures) or marginally significant *P* values are reported, according to criteria given in "Material and methods" section (Fdist: P values <0.025, DetSel: P values <0.01, lnRH: P values <0.025). Those loci detected in at least two of three independent habitat comparisons are indicated in boldface

fewer loci (Oetjen and Reusch 2007; GA20, ZMC19066, ZMC01058) were also detected with the increased number of markers. The locus GA20, an anonymous microsatellite, was found to be an outlier in all three population pairs. Locus ZMC19066 (linked to acid phosphatase gene) was an outlier in the population pair on Hallig Langeness ($P < 0.01$) and on Hallig Hooge ($P < 0.03$) in our study. Locus ZMC01058 (linked to vacuolar H^+-ATPase) proved to be an outlier in the population pair on Sylt island ($P < 0.01$), on Hallig Hooge ($P < 0.02$) and on Hallig Langeness ($P < 0.03$) also. One locus, that had been detected as an outlier in two population pairs (Hallig Hooge and Sylt island, both $P < 0.01$) earlier, CL412 (linked to putative nodulin 3), showed the same pattern in our recent study. A newly developed microsatellite (ZME05315, linked to putative ubiquitin-conjugating enzyme) was a significant outlier ($P < 0.01$) in population pair Hallig Hooge and on Sylt island ($P < 0.03$). Weaker evidence for selection was found for two other loci that were detected in only one of the three population pairs (newly developed microsatellite CL559, linked to putative seed maturation protein, and SNP 71TC178, both in Langeness population pair with $P < 0.01$). Excluding all SNPs from the data set yielded qualitatively similar results.

The lnRH test detected eight outliers with *P* values significant or marginally significant at the 5% level. Most of them were detected in only one population pair (CL380/$P = 0.0004$; ZME05315/$P = 0.0188$ in pop pair Hallig Langeness, CL412/$P = 0.0283$ and ZME02125/$P < 0.0001$ in pop pair Hallig Hooge, CL766/$P = 0.0002$ and ZMC19066/$P = 0.018$ in pop pair Sylt island), except locus CL559, which occurred in two comparisons (Hallig Hooge, $P = 0.013$, Sylt island, $P = 0.0125$). All of these loci were detected by DetSel in the respective population pair as well.

Compared to an earlier study, the results of the present study were mostly consistent. The locus ZMC19089 was newly identified as an outlier in the population pair (subtidal vs. intertidal) on Hallig Langeness ($P = 0.014$). Another locus with unknown function (CL734) was became significantly different from neutral expectation ($P = 0.019$) in the population pair on Hallog Hooge in this study. A microsatellite locus (ZMA04093) linked to a purple acid phosphatase (BLAST hit: BAC55157 *Nicotiana tabacum*) is marginally significant in the recent analysis ($P = 0.0126$) and is also closer to significance than in the result of the previous scan ($P = 0.0129$).

An AMOVA revealed that when only the six outlier loci were regarded, the component of variance between

contrasting habitat pairs approximately doubled from 0.86 to 1.64%. For the remaining 40 loci, variation between habitat contrasts decreased to 0.78, and the effect of variation between sampling sites increased to 1.09%.

Functional role of candidate loci

The strongest marker locus candidates were linked to a broad spectrum of genes. While GA20 was an anonymous microsatellite, ZMC01058 is linked to a vacuolar H^+-ATPase, hence a gene that is involved in osmotic regulation. Even more interesting, ZMC19066 is located next to an acid phosphatase gene which is associated with hyperosmotic stress according to gene ontology annotation. CL412 next to a putative nodulin gene which forms water channels in cell membranes and is hence also involved in osmoregulation (Luu and Maurel 2005). ZME05315 is linked to a putative ubiquitin-conjugating enzyme and CL559 to a putative seed maturation protein. The latter gene may be responsible for the much higher seed output of intertidal populations compared to subtidal ones.

Discussion

The present analysis extends previous genome scans in *Z. marina* by doubling the number of markers employed and adding a third outlier test. For the first time in a genome scan in a marine plant, we included SNPs as a novel marker type. As a goal of this study, we assessed how 2 years of additional marker development would affect the reliability of results collected with relatively few markers. Several genome scan studies have reanalysed existing data sets with different outlier tests (Vitalis et al. 2001; Beaumont and Balding 2004) or assessed the validity of outlier tests by using more than one test statistic (Vasemagi et al. 2005; Kane and Rieseberg 2007; Tsumura et al. 2007; Egan et al. 2008; Mäkinen et al. 2008). In contrast, extending a genome scan by doubling the number of markers has been done here for the first time. Reassuringly, we found that earlier results of the genome scan were largely confirmed. Compared to previous analyses, the increased number of markers in our study was able to detect selection more reliably and to confirm most of the outliers detected before. All three loci that seemed to be influenced by habitat-dependent selection based on a smaller amount of data (ZMC19066, ZMC01058, GA20) were confirmed by the extended approach. Additionally, two new candidate loci could be detected. ZME05315, linked to a putative ubiquitin-conjugating enzyme, and CL559, linked to a putative seed maturation protein, revealed strong evidence for influence of selection.

It is well known that even under relatively high gene flow, selection gradients, if strong enough, will counteract gene flow (Silander 1979). Although the three population pairs are farther away from each other than the two habitat types, we cannot exclude that gene flow across similar habitats by tidal currents transports alleles under selection within the three populations within a habitat type. Nevertheless, if true, homogenizing gene flow across the pairs at any one location is probably higher and would erase the signal of selection, with or without parallel gene flow.

Among the six SNPs loci tested in the genome scan, there were no further candidates. Only at one SNP (71TC178, detected in only one population pair), we found weak evidence for selection. This is most likely due to the low number of polymorphic SNPs included in the study. Although we originally targeted for approximately the same number of SNPs than microsatellites, most SNPs developed on a global panel of *Z. marina* genotypes (Ferber et al. 2008) were not polymorphic at the spatial scale chosen in this study. The reason for this is most likely the close proximity and genetic relationship of the populations that do not comprise a deeper phylogeographic signal that would be needed to detect polymorphism for this marker type. Note that the SNPs used here were detected based on a genotype panel that comprised widely distant populations from Pacific and Atlantic populations. Despite being valuable tools for population genetic analyses at larger geographical scales (Ferber et al. 2008), development of SNPs is costly where no in silico screening can be done and screening effort is higher as the information content per SNP is low. For example, about five times more SNPs than microsatellites are needed to determine relationships between populations (Glaubitz et al. 2003). Nevertheless, it should be possible to use them for genome scanning in *Z. marina* also in the future, especially as a significant number of SNPs are ready to use with good amplification and screening properties. In silico search for novel SNPs is also easily feasible with a sound database, namely ESTs (http://drzompo.uni-muenster.de/; Wissler et al. 2009), already available and a *Zostera* genome project to be finished in the near future.

Despite the overall consistency of the two genome scans, we found some signals of selection to become weaker which sends a cautionary message to all studies using genome scans with relatively few marker loci, as results definitely have to be regarded as preliminary in such cases. This is even more important as the outlier test that detects the highest number of outlier loci (the DetSel test) has been criticized before, especially when the outlier findings are compared to the results by the Fdist test (see Vitalis et al. 2001; Vasemagi et al. 2005). Congruency of these two test statistics seems to be poor. Our study revealed close congruency of the DetSel test and the lnRH method, as all nine loci detected by lnRH were also detected by DetSel in the respective population pair. This is particularly encouraging

because both approaches work differently, one by using an empirical null distribution (lnRH), the other a demographic model (DetSel).

The number of markers detected as outliers in this study was slightly higher (6 (13%) out of 46) compared to other recent studies (1.4–9.5%; Campbell and Bernatchez 2004; Vasemagi et al. 2005; Bonin et al. 2006; Kane and Rieseberg 2007). It is to be expected that for gene-linked (EST) markers, more signals of diverging selection can be detected than for anonymous markers (e.g. 21% in Scotti-Saintagne et al. 2004; 12% in Vasemagi et al. 2005). This is because for the latter marker type, chances are high that markers to reside in intronic and intergenic regions physically distant from a functionally important gene, thus missing the 'hitchhiking' signal.

Because the number of marker loci tested is still rather small, we cannot test whether or not the 6 outliers are over – or underrepresented in terms of their function (defined by GO—gene ontology categories) compared to the overall spectrum of genes. Notwithstanding, some loci merit attention, such as CL412, which is linked to a putative nodulin class 3. Nodulins belong to the protein class of aquaporins that form trans-membrane water channels in cellular membranes (Luu and Maurel 2005). This locus, as well as ZMC01058 (linked to a vacuolar H^+-ATPase), is likely to be involved in osmoregulation and water balance. It is conceivable that changes in either structure or expression of such genes are needed given the drastically different environmental conditions on tidal flats compared to a permanently submerged habitat (Reise 1985). Another interesting locus is CL559, linked to a putative seed maturation protein. As populations inhabiting both habitat types differ in their reproductive strategy—predominantly annual versus perennial, this gene points to adaptive divergence at the molecular level. The ripening process of seeds is an important factor in the life history of plants influencing perennial persistence or annual growth from new seeds. Above genes may serve as valuable starting points for further investigation into the genetic basis of natural selection and habitat adaptation in *Z. marina*.

For *Z. marina* as an ecological-engineering species with outstanding ecological importance, the prospects for studying the genetic basis to ecological adaptation are promising, given that a *Zostera* genome project has been initiated by the Joint Genome Institute (JGI). The limitations of current genome scan studies brought along by moderate marker numbers, and restriction to EST sequences will be obsolete. In particular for SNP markers that allow a much higher throughput than microsatellites, tailored polymorphic sites in genes selected by their functionally relevance (e.g. osmoregulation) will revolutionize the possibilities to detect natural selection in natural settings. Because *Z. marina* grows under very different environmental conditions locally and spans a wide biogeographic range from the White Sea to Portugal in Europe alone, it will become a model for how genetic and ecological information can be merged in a common framework of ecological genetics and genomics (Ouborg and Vriezen 2007).

Acknowledgments We thank B. Hasert for technical assistance. Many thanks to Prof. Karsten Reise from AWI Sylt for help and support, as well as to Andreas Zipperle for help collecting samples. This work was funded by Max-Planck society and the DFG (grant no. Re 1108/7-1).

References

Akey JM, Eberle MA, Rieder MJ, Carlson CS, Shriver MD, Nickerson DA et al (2004) Population history and natural selection shape patterns of genetic variation in 132 genes. PLOS Biol 2:1591–1599

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM et al (2000) Gene ontology: tool for the unification of biology. Nat Genet 25:25–29

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. Mol Ecol 13:969–980

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc Biol Sci 263:1619–1626

Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). Mol Biol Evol 23:773–783

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. Mol Ecol 16:907–924

Campbell D, Bernatchez L (2004) Genomic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. Mol Biol Evol 21:945–956

Costanza R, d'Arge R, de Groot R, Farber S, Grasso M, Hannon B et al (1997) The value of the world's ecosystem services and natural capital. Nature 387:253–260

den Hartog C (1970) The seagrasses of the world. Verh K Ned Akad We. Afd Natuurkd II 59:1–275

Dieringer D, Schlötterer C (2003) MICROSATELLITE ANALYSER (MSA): a platform independent analysis tool for large microsatellite data sets. Mol Ecol Notes 3:167–169

Duarte CM (2002) The future of seagrass meadows. Environ Cons 29:192–206

Egan SP, Nosil Patrik, Funk DanielJ (2008) Selection and genomic differentiation during ecological speciation: isolating the contributions of host association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. Evolution 62:1162–1181

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

Ferber S, Reusch TBH, Stam WT, Olsen JL (2008) Characterization of single nucleotide polymorphism markers for eelgrass (*Zostera marina*). Mol Ecol Resour 8:1429–1435

Glaubitz JC, Rhodes OE, Dewoody JA (2003) Prospects for inferring pairwise relationships with single nucleotide polymorphisms. Mol Ecol 12:1039–1047

Harr B, Kauer M, Schlötterer C (2003) Hitchhiking mapping: a population-based fine-mapping strategy for adaptive mutations in *Drosophila melanogaster* (vol 99, pg 12949, 2002). Proc Natl Acad Sci USA 100:3004

Harris MA, Clark JI, Ireland A, Lomax J, Ashburner M, Collins R et al (2006) The gene ontology (GO) project in 2006. Nucleic Acids Res 34:D322–D326

Hemminga M, Duarte C (2000) Seagrass ecology. Cambridge University Press, Cambridge

Kane NC, Rieseberg LH (2007) Selective sweeps reveal candidate genes for adaptation to drought and salt tolerance in common sunflower, *Helianthus annuus*. Genetics 175:1823–1834

Kauer MO, Dieringer D, Schlötterer C (2003) A microsatellite variability screen for positive selection associated with the "Out of Africa" habitat expansion of *Drosophila melanogaster*. Genetics 165:1137–1148

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms. Genetics 74:175–195

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4:981–994

Luu D-T, Maurel C (2005) Aquaporins in a challenging environment: molecular gears for adjusting plant water status. Plant Cell Environ 28:85–96

Mäkinen HS, Cano JM, Merilä J (2008) Identifying footprints of directional and balancing selection in marine and freshwater threespined stickleback *Gasterosteus aculeatus* populations. Mol Ecol 17:3565–3582

Massa SI, Arnaud-Haond S, Pearson GA, Serrao E (2009) Temperature tolerance and survival of intertidal populations of the seagrass *Zostera noltii* (Hornemann) in southern Europe (Ria Formosa, Portugal). Hydrobiologia. doi:10.1007/s10750-008-9609-4

Namroud M-C, Beaulieu Jean, Juge Nicolas, Laroche Jerome, Bousquet Jean (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. Mol Ecol 17:3599–3613

Nielsen R (2005) Molecular signatures of natural selection. Annu Rev Genet 39:197–218

Oetjen K, Reusch TBH (2007) Genome scans detect consistent divergent selection among subtidal vs. intertidal populations of the marine angiosperm *Zostera marina*. Mol Ecol 16:5156–5157

Olsen JL, Stam WT, Coyer JA, Reusch TBH, Billingham M, Bostrom C et al (2004) North Atlantic phylogeography and large-scale population differentiation of the seagrass *Zostera marina* L. Mol Ecol 13:1923–1941

Orth RJ, Carruthers TJB, Dennison WC, Duarte CM, Fourqurean JW, Heck KL et al (2006) A global crisis for seagrass ecosystems. Bioscience 56:987–996

Ouborg NJ, Vriezen WH (2007) An ecologist's guide to ecogenomics. J Ecol 95:8–16

Reise K (1985) Tidal flat ecology: an experimental approach to species interactions. Springer, Berlin

Reise K, Jager Z, De Jong D, Van Katwijk M, Schanz A (2005) Seagrass—Wadden Sea ecosystem no. 19. Common Wadden Sea Secretariat, Wilhelmshaven

Rengmark AH, Slettan A, Skaala O, Lie O, Lingaas F (2006) Genetic variability in wild and farmed Atlantic salmon (*Salmo salar*) strains estimated by SNP and microsatellites. Aquaculture 253:229–237

Reusch TBH (2002) Microsatellites reveal high population connectivity in eelgrass (*Zostera marina*) in two contrasting coastal areas. Limnol Oceanogr 47:78–85

Reusch TBH, Veron AS, Preuss C, Weiner J, Wissler L, Beck A et al (2008) Comparative analysis of expressed sequence tag (EST) libraries in the seagrass *Zostera marina* subjected to temperature stress. Mar Biotechnol 10:297–309

Ryynanen HJ, Tonteri A, Vasemagi A, Primmer CR (2007) A comparison of Biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic Salmon (*Salmo salar*). J Hered 98:692–704

Schlötterer C (2002a) A microsatellite-based multilocus screen for the identification of local selective sweeps. Genetics 160:753–763

Schlötterer C (2002b) Towards a molecular characterization of adaptation in local populations. Curr Opin Genet Dev 12:683–687

Schlötterer C (2003) Hitchhiking mapping—functional genomics from the population genetics perspective. Trends Genet 19:32–38

Scotti-Saintagne C, Mariette S, Porth I, Goicoechea PG, Barreneche T, Bodenes K et al (2004) Genome scanning for interspecific differentiation between two closely related oak species *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. Genetics 168:1615–1626

Silander JA (1979) Microevolution and clone structure in *Spartina patens*. Science 203:658–660

Smith JM, Haigh J (1974) Hitch-hiking effect of a favorable gene. Genet Res 23:23–35

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity 100:158–170

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. Mol Ecol 14:671–688

Teshima KM, Coop G, Przeworski M (2006) How reliable are empirical genomic scans for selective sweeps? Genome Res 16:702–712

Thornton KR, Jensen JD (2007) Controlling the false-positive rate in multilocus genome scans for selection. Genetics 175:737–750

Tsumura Y, Kado T, Takahashi T, Tani N, Ujino-Ihara T, Iwata H (2007) Genome scan to detect genetic structure and adaptive genes of natural populations of *Cryptomeria japonica*. Genetics 176:2393–2403

van Katwijk MM, Hermus DCR (2000) Effects of water dynamics on *Zostera marina*: transplantation experiments in the intertidal Dutch Wadden Sea. Mar Ecol Prog Ser 208:107–118

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Vasemagi A, Nilsson J, Primmer CR (2005) Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). Mol Biol Evol 22:1067–1076

Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. Genetics 158:1811–1823

Vitalis R, Dawson K, Boursot P, Belkhir K (2003) DetSel 1.0: a computer program to detect markers responding to selection. J Hered 94:429–431

Waycott M, Duarte CM, Carruthers TJB, Orth RJ, Dennison WC, Olyarnik S et al (2009) Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proc Natl Acad Sci USA 106:12377–12381

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. Evolution 38:1358–1370

Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. J Evol Biol 14:611–619

Wissler L, Dattolo E, Moore AD, Reusch TBH, Olsen JL, Migliaccio M, Bornberg-Bauer E, Procaccini G (2009) Dr. Zompo: an online data repository for *Zostera marina* and *Posidonia oceanica* ESTs. Database 2009: bap009