BAYESIAN ANALYSIS OF SPECIES RELATIONSHIP IN GENUS SARGASSUM: OMAN SEA AND PERSIAN GULF

Zahra Noormohammadi
Department of Biology, School of Basic Sciences, Science and Research Branch, Islamic Azad University, Tehran, Iran
marjannm@yahoo.com

ABSTRACT: The genus Sargassum C. Agardeh contains about 400 species throughout the world, but controversy exists about the number of Sargassum species in Persian Gulf. The present study considers molecular analysis (RAPD and ISSR markers) of genetic diversity and species relationship among populations of 6 Sargassum species occurring in Persian Gulf and Arabian costs in vicinity. Both RAPD and ISSR primers produced many polymorphic bands which were shown to be discriminating at species and population/locality level. Sargassum species showed high level of genetic diversity with species specific alleles which may use in characterization of species and their location (Oman Sea and Persian Gulf). Bayesian analysis by STRUCTURE program showed the presence of distinct species groups, while reticulation and coalescence analyses showed limited gene exchange among species and that gene duplication and loss are additional mechanism for bringing about genetic diversity among these species.

Key words: Coalescence, genetic admixture, ISSR, RAPD

INTRODUCTION

Sargassum C. Agardeh (1820) with about 400 species is the most species rich genus within the Sargassaceae [1]. Controversy exists about the number of Sargassum species in Persian Gulf, for example, Borgesen [2] reported the occurrence of 26 species of brown algae while, Sohrabipour and Rabii [3, 4] identified 6 Sargassum species in this region. Similarly, Gharanjik [5] reported S. assimile Harvey, S. swatzii Tuner, S. glaucescens J. Agrdh, S. cristaeolium (C.Ag.) J. Ag., S. licifolium (Tuner) C.Agrdh, S. virgatum C.Ag. and S. tenerrimum J. Agrdh in Sistan & Baloochestan seashore (Oman sea) located in southeast of Iran and latter on, Abdel-Kareem [6] reported 16 Sargassum species in Persian Gulf and Arabic (Saudi Arabia) coasts, four of which are common with those reported by Sohrabipour and Rabii [4]. Recently Noormohammad i et al. [7] studied morphological variation in 3 Sargassum species occurring in southeast of Iran (Oman Sea) and discriminated these species and populations by using quantitative and qualitative morphological characters. The same authors reported intra and inter-population genetic diversity in 3 Sargassum species of Oman Sea using RAPD and ISSR markers [8]. RAPD and ISSR markers have been extensively used in discrimination of Sargassum species and also studying population genetic structure and phylogeography of these taxa [9-14]. The aim of this work is to study Sargassum species relationship occurring in Chabahar, Tang and Quatr regions of Oman Sea seashore as well as in Bandare Lengeh of Persian Gulf. The combination of genetic diversity analysis, Bayesian clustering by STRUCTURE analysis, reticulation and coalescence analysis were used to illustrate population genetic diversity and species relationship in Sargassum species studied.

MATERIALS AND METHODS

In total 6 sargassum species were studied, 3 species namely, S. tenerrimum J. Agardh, S. glaucescens J. Agardh and S. ilicifolium C. Agardh were collected from coastal sites of Chabahar, Tang and Guatr locations (Oman Sea seashore) and 3 species namely, S. vulgare C. Agardh, S. oligocystum and S. boveanum J. Agardh were collected from Bandar Lengeh in Persian Gulf seashore (Table 1). At each site, 5 to 10 non-reproductive individuals were selected with a minimum distance of 2-3 meters at low tide or snorkeling. The specimens collected were washed thoroughly with distilled water and placed in plastic bags with silica gel beads and transported to the laboratory. Three to 5 leaf-like blades of each sample were pooled and used for DNA extraction. The total genomic DNA was extracted following the CTAB method using DNeasy Plant mini kit (Qiagen GmbH., Hilden, Germany) according to manufacturer’s instructions. The quality of DNA was examined by running on a 0.8% agarose gel.
Table 1 Sample details of *Sargassum* populations detected in the study. SST: sea surface temperature

<table>
<thead>
<tr>
<th>Population</th>
<th>Species</th>
<th>Sample site</th>
<th>SST/Salinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chabahar-Oman Sea</td>
<td><em>S. glaucescens</em> (G1-Ch) <em>S. ilicifolium</em> (11-Ch)</td>
<td>Chabahar, Sistan &amp; Balochestan province; N 61° 39´, E 25° 17´</td>
<td>22.3°-31°/35ppt-39ppt</td>
</tr>
<tr>
<td>Quatr- Oman Sea</td>
<td><em>S. tenerrimum</em> (T3-Qu) <em>S. glaucescens</em> (G3-Qu) <em>S. ilicifolium</em> (I3-Qu)</td>
<td>Guatr, Hormozgan province, N 61° 30´, E 25° 10´</td>
<td>20°-30.5°/ 35ppt-40ppt</td>
</tr>
<tr>
<td>Bandar Lengeh- Persian Gulf</td>
<td>Sargassum vulgare (Vul)</td>
<td>Bandar Lengeh, Hormozgan province, E 54° 53´, N 26° 33´</td>
<td>30-32°/36ppt-39ppt</td>
</tr>
<tr>
<td>Bandar Lengeh- Persian Gulf</td>
<td>Sargassum oligocystum (Olig)</td>
<td>Bandar Lengeh, Hormozgan province, E 54° 53´, N 26° 33´</td>
<td>30-32°/36ppt-39ppt</td>
</tr>
<tr>
<td>Bandar Lengeh- Persian Gulf</td>
<td>Sargassum boveanum (Bov)</td>
<td>Bandar Lengeh, Hormozgan province, E 54° 53´, N 26° 33´</td>
<td>30-32°/36ppt-39ppt</td>
</tr>
</tbody>
</table>

**Molecular markers amplification**

Thirty homo-RAPD primers of different Operon kits (A, C, I, M, R) from Operon Technologies, Calif., USA as well as 6 hetero-RAPD primers were used. The Inter Simple Sequence Repeats (ISSR) primers used in the present study were selected in a set of four homo-primers; UBC807, UBC810, UBC811, UBC823 UBC834 and UBC849, commercialized by UBC (the University of British Columbia) used by Zhao et al. (2008) as well as six hetero- ISSR primers. All PCR reactions as well as thermal programs for both RAPD and ISSR markers were set up based on Noormohammadi et al. [8] protocols. Amplification products were visualized by running on 2% agarose gel in 0.5 X TBE buffer system, followed by ethidium bromide (0.5 µg mL⁻¹) staining. Fragment size was estimated by using a 100 base pairs (bp) molecular size ladder (Fermentas, Germany).

**Data analysis**

RAPD and ISSR bands obtained were treated as binary characters and coded accordingly (presence = 1, absence = 0). Bayesian clustering method was performed on all 360 plants to elucidate the populations genetic structure by using STRUCTURE v. 2.3 [15]. The program STRUCTURE implements a model-based clustering method for inferring population structure using genotype data consisting of unlinked markers. For this reason we first performed linkage disequilibrium test for SSR loci as implemented in POPGENE ver. 1.32 (2000). The model applied in STRUCTURE analysis assumes the existence of K clusters [16]. The Markov chain Monte Carlo simulation was run 20 times for each value of K (6) for 10⁶ iterations after a burn-in period of 10⁵. All other parameters were set at their default values. Data were entered as suggested by Falush et al. [17] and data sample provided in STRUCTURE home page. STRUCTURE Harvester web site [18] was used to visualize the STRUCTURE results and also to perform Evanno method [19], to identify proper number of K [19]. The choice of the most likely number of clusters (K) was carried out comparing log probabilities of data [Pr(X|K)] for each value of K [16], as well as by calculating an ad hoc statistic ΔK based on the rate of change in the log probability of data between successive K values, as described by Evanno et al. [19]. Genetic differentiation of population subgroups was determined by Fst determined by STRUCTURE. For other analyses, to reduce data and get a better picture, data of plant replications in each genotype were pooled and used for clustering and other ordination methods implemented. Dice as well as Nei’s genetic distance [20-22] determined among species and populations were used for the grouping of the genotypes by unweighted paired group method with arithmetic average (UPGMA) and Neighbor Joining (NJ) clustering methods after 100 times bootstrapping [22]. Similarly ordination plot based on principal co-ordinate analysis (PCoA), (after 999 times permutation), Discriminate analysis (DA), and Multidimensional scaling (MDS) [23], by using PAST ver. 2.17 (2012) [24] and DARwin ver. 5 (2012). Dentreated correspondence analysis (DCA) and Principal components analysis (PCA) with biplot were used to check distribution of SSR loci and their role in discrimination of species also by the same software. Genetic diversity parameters including percentage of allelic polymorphism, allele diversity [21], Nei’s gene diversity (H), Shannon information index (I) [21, 22], number of effective alleles and percentage of polymorphism were determined. AMOVA (Analysis of molecular variance) test (with 1000 permutations) as performed in GenAlex 6.4 [25], was used to show molecular difference among the populations.
Reticulation was performed by T-REX (Tree and Reticulogram Reconstruction) ver. 3 (2000), and DARwin ver. 5 (2012) which infer reticulogram from distance matrix. For reticulation, we first built a supporting phylogenetic tree using Neighbor Joining (NJ), Followed by a reticulation branch that minimizes the least-squares at each step of the algorithm [26]. Furthermore, coalescence analysis was performed for molecular data after 1000 times reiteration [27], as suggested for SSR polymorphic data by Wilson & Balding [28] and performed in Mesquite [29]. Gene tree heterogeneity and discordance with the population tree was checked by parameters provided in Mesquite [30], including deep coalescence, gene duplication and extinction. Canonical correspondence analysis (CCA) was performed to show *Sargassum* species divergence in response to longitude and altitude.

**RESULTS**

**Genetic diversity analysis**

Eight of 30 homo and hetero-RAPD primers produced reproducible bands. In total 130 polymorphic bands obtained with mean value of genetic polymorphism 100% among 6 *Sargassum* species.

Out of 130 RAPD polymorphic bands obtained, 11 bands were species specific. The highest specific bands occurred in *S. ilicifolium* (5 bands), while no specific band was observed in *S. boveanum* samples (Table 2). Some bands were only observed in *Sargassum* species collected from Persian Gulf but absent in the species collected from Oman Sea. OPA4-1250 bands was only specific band for Oman Sea *Sargassum* species (Table 2).

Six homo-ISSR and 2 hetero-ISSR loci produced 68 polymorphic bands, out of which, 16 species specific bands were observed. *S. glaucescens* showed the highest number of specific bands (6), while, *S. tenerimum* had no bands (Table 2). UBC849 – 370 was unique ISSR band observed in *Sargassum* species collected from Persian Gulf, while UBC811-833 band was unique ISSR band in Oman Sea species (Table 2).

<table>
<thead>
<tr>
<th>Species specific band (RAPD loci)</th>
<th>Specific band in Persian Gulf species</th>
<th>Specific band (Oman Sea species)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. tenerimum</em></td>
<td>OPA04 (1750bp)</td>
<td>OPA04 (1250 bp)</td>
</tr>
<tr>
<td><em>S. glaucescens</em></td>
<td>OPA13 (1550bp), OPA02 (750bp)</td>
<td>OPA04 (1250 bp)</td>
</tr>
<tr>
<td><em>S. ilicifolium</em></td>
<td>OPA03 (1660bp), OPA02 (450, 550,870 and 1100 bps)</td>
<td>OPA04 (1250 bp)</td>
</tr>
<tr>
<td><em>S. vulgare</em></td>
<td>OPA13 (530 bp)</td>
<td>OPA13 (890 and 1650 bps), OPA03 (4100 and 300 bps), OPA02 (1250 bp)</td>
</tr>
<tr>
<td><em>S. oligocystum</em></td>
<td>OPA03 (1660 and 1750 bps)</td>
<td>OPA13 (890 and 1650 bps), OPA03 (4100 and 300 bps), OPA02 (1250 bp)</td>
</tr>
<tr>
<td><em>S. boveanum</em></td>
<td>---</td>
<td>OPA13 (890 and 1650 bps), OPA03 (4110 and 300 bps), OPA02 (1250 bp)</td>
</tr>
<tr>
<td>Species specific band (ISSR loci)</td>
<td>Specific band in Persian Gulf species</td>
<td>Specific band (Oman Sea species)</td>
</tr>
<tr>
<td><em>S. tenerimum</em></td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><em>S. glaucescens</em></td>
<td>UBC834 (350, 1300 bps), UBC849 (330 bp), UBC807/823 (250, 650 and 1250 bps)</td>
<td>---</td>
</tr>
<tr>
<td><em>S. ilicifolium</em></td>
<td>UBC810 (390 and 410 bps), UBC811 (750 bp)</td>
<td>---</td>
</tr>
<tr>
<td><em>S. vulgare</em></td>
<td>UBC810 (370 bp), UBC807/823 (450 and 550 bps)</td>
<td>UBC849 (370bp)</td>
</tr>
<tr>
<td><em>S. oligocystum</em></td>
<td>UBC810 (100bp), UBC823/849 (450bp)</td>
<td>UBC849 (370bp)</td>
</tr>
<tr>
<td><em>S. boveanum</em></td>
<td>UBC810 (330 bp), UBC811 (330 bp)</td>
<td>UBC849 (370bp)</td>
</tr>
</tbody>
</table>
Dentrented correspondence analysis plot obtained including Convex hulls as well as 95% Ellipses methods showed scattered distribution of ISSR loci in the plot, indicating they are well scattered in the genome and possibly are not correlated to each other (Fig. 1A). These statistical methods identified more effective (discriminating) loci in genetic analysis of *Sargassum* species (Fig 1B); the loci placed closer to the center of Convex hulls and 95% Ellipses are more reliable loci to discriminate *Sargassum* species and those loci placed farther from center are less effective for this purpose. Therefore, loci number, 52, 62, 76, 47, 37, 40, 28, 26, 7, 6 and 51 are discriminating loci.

When PCA biplot was sketched (Fig. 2), it was seen that the effects of these loci differ from each other. For example, ISSR loci 47 (UBC811-750) and 37 (UBC834-1000) were acting opposing each other and separating different species (*S. ilicifolium* and *S. tenerrimum* respectively). The same was true for loci No. 26 and 28, and 47 and 62, etc. Genetic diversity analysis revealed a higher degree of genetic polymorphism and expected heterozygosity for *S. ilicifolium* (37.17% and 0.15 respectively), while the lowest values for the same parameters occurred in *S. tenerrimum* (11% and 0.048 respectively). STRUCTURE analysis (will be discussed in coming paragraphs) grouped *Sargassum* species in 4 distinct genetic groups. Therefore, genetic diversity and AMOVA analyses were performed among these 4 groups (Table 3). AMOVA test showed significant genetic difference (p<0.01) between four species groups and revealed that 61% of total genetic variation is due to among group difference, while 39% is due to within group genetic variation. This indicates the presence of great genetic difference among *Sargassum* species.

**Table 3. AMOVA analysis based on RAPD and ISSR data among 6 Sargassum species**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>Est. Var</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among Pops</td>
<td>3</td>
<td>300.939</td>
<td>100.313</td>
<td>29.745</td>
<td>61%</td>
</tr>
<tr>
<td>Within Pops</td>
<td>7</td>
<td>134.333</td>
<td>19.190</td>
<td>19.190</td>
<td>39%</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>435.273</td>
<td>48.935</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Stat Value</td>
<td></td>
<td>P(rand =&gt; data)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PhiPT</td>
<td>0.608</td>
<td>0.010</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Genetic affinity and grouping of Sargassum species

Test for linkage disequilibrium performed for RAPD and ISSR loci as implemented in POPGENE, did not show any significant association between them. Therefore, these molecular markers were considered to meet the assumptions for applying the Bayesian method implemented in the program STRUCTURE to assign individuals to species groups. Q-matrix plot of STRUCTURE analysis (Fig. 3), showed presence of 4 subgroups. Evanno method and ad hoc statistic $\Delta K$ also showed $k = 4$ as the best number of population subgroups (Fig. 4). STRUCTURE plot based on all 6 species, showed low gene exchange and admixture among species groups studied (different colors intermixed in each group). STRUCTURE plot also revealed that each species group contains particular allelic composition (dominant color in each group), which comes from genetic background of Sargassum species forming that group. For example, in S. ilicifolium with allelic composition (blue color) is different from other groups having S. glaucescens (yellow), or S. tenerrimum (pink) and Persian Gulf Sargassum (light blue) species.

Fig. 3. Q-matrix of STRUCTURE analysis. 1- S. ilicifolium (Chabahar), 2- S. ilicifolium (Tang), 3- S. ilicifolium (Quatr) 4- S. glaucescens (Chabahar), 5- S. glaucescens (Tang), 6- S. glaucescens (Quart), 7- S. tenerrimum (Chabahar) 8- S. tenerrimum (Quatr), 9- S. vulgare, 10- S. oligocystum, 11- S. Boveanum

Fig. 2. PCA biplot discrimination of two Sargassum species by two characters: UBC811-750 and UBC834-1000. Species names are according to Table 1.
Nei’s genetic distance determined among population groups varied from 0.280 between *S. tenerrimum* and Persian Gulf species to 0.510 between *S. tenerrimum* and *S. glaucescens*. NJ and UPGMA trees after bootstrapping produced similar results. Therefore NJ tree is discussed here (Fig. 5). In general 4 major clusters are formed. First cluster comprised of *S. glaucescens* populations while second cluster grouped *S. ilicifolium* populations. *S. tenerrimum* populations formed third cluster which are placed close to forth cluster including 3 species of Persian Gulf (*S. vulgare*, *S. oligocystum* and *S. Boveanum*). Nei’s genetic distance also proved low distance between these two groups. PCoA ordination plot of GENEALEX (Fig. 6) based on 999 reiteration, also separated Sargassum species. This plot also produced 4 distinct groups similar to NJ tree. Three species located in Oman Sea seashore were placed with some distance from each other as well as 3 Persian Gulf species due to their genetic differences. While, *S. Vulgare, S. Oligocystum* and *S. boveanum* located in Persian Gulf were placed close to each other due to their genetic similarity.

We considered K=6 in STURCTURE analysis; therefore Fst analysis was performed on 6 species groups. The Fst values obtained ranged from 0.06 in *S. glaucescens* to 0.73 in *S. oligocystum*. Species groups; *S. ilicifolium*, *S. glaucescens* and *S. tenerrimum* had the lowest Fst values (0.06, 0.08 and 0.34 respectively) indicating low distance and high gene exchange among *Sargassum* species occurring in Oman Sea. The other Species groups (Persian Gulf species) had high Fst values (>0.60), indicating a good genetic distinctness among these species. Reticulogram (Fig. 7), showed genes exchange/shared common genetic loci among all *Sargassum* species groups, such as between *S. glaucescens* and 3species of Persian Gulf, between *S. ilicifolium* and *S. glaucescens*, as well as between *S. tenerrimum* and *S. oligocystum*. These results support STRUCTURE analysis results.

Coalescence analysis produced gene trees which grouped *Sargassum* species almost similar to NJ tree and when gene trees were contained in population tree, the best result obtained showed deep coalescence cost of 7 with 3 gene duplication and 13 gene extinction (Fig. 8). Based on gene tree, Gene duplications and deletions mostly happened in divergence of *S. ilicifolium* and *S. glaucescens* in comparison to *Sargassum* species in Persian Gulf seashore (Fig. 8).

The Mantel test performed between genetic distance and geographical distance of 6 *Sargassum* species showed no significant correlation between the two distance matrix data (R² = 0.002, P>0.05, Fig. 9). Therefore, no isolation by distance exists among *Sargassum* species.

CCA plot (Fig. 10) showed that both longitude and altitude are affecting *Sargassum* species divergence. Longitude factor was more effective than altitude on *S. ilicifolium* and *S. glaucescens* divergence while 3 Sargassum species of *S. vulgare, S. oligocystum* and *S. boveanum* located in Persian Gulf were mostly affected by altitude factor.
Fig 5. NJ tree based on molecular markers. Sample numbers are according Figure 3.


Fig 7. Reticulogram of Sargassum species. Sample numbers are according Figure 3.
Fig. 8. Coalescence tree of *Sargassum* species contained within species tree. Sample names are according to Table 1.

Fig. 9. Mantel test between geographical distance (GGD) and genetic distance (GD).

Fig. 10. Canonical correspondence analysis (CCA) plot based on geographical altitude and longitude of 6 *Sargassum* species. Species names are according to table 1.
DISCUSSION

This study represents the first attempt to study Bayesian analysis of species relationship in genus Sargassum situated in different geographical locations in Oman Sea and Persian Gulf seashores. Results of present study indicate the presence of high genetic diversity among 6 Sargassum species as well as between Oman Sea Sargassum species and species located in Persian Gulf. Both RAPD and ISSR data showed high level of allelic polymorphism indicating usefulness of these molecular markers for discriminating Sargassum species. Similar to our results, Zhao and co-workers [12, 13] could discriminate S. thunbergii and S. muticum populations by using single ISSR loci. Wang et al [31] suggested that different reproductive mode and different growth environment might have critical roles in the genetic differentiation of S. thunbergii populations.

Noormohammadi et al. [8] reported high genetic diversity between Sargassum species populations while low genetic variation were obtained within each population (Chabhar, Tang and Quatr) by using RAPD and ISSR markers. The occurrence of molecular mutations/recombinations in these species populations may be the reason for such high genetic difference. Since, even single base change at the primer annealing site is manifested as appearance or disappearance RAPD and ISSR bands, these bands may indicate the occurrence of genetic changes in the genome of species either through the loss or rearrangement of some of their nucleotides. Chromosomal crossing over during meiosis may result in loss of primer attachment pair sites in the offspring leading to novel molecular marker pattern in species [32].

The interesting results was obtaining species specific bands (Table 2). According to these results S. ilicifolium and S. glaucescens had the highest number of RAPD and ISSR species specific bands and are discriminated from the other species studied. This is also supported by NJ tree (Fig. 5). On the other hand, some specific bands are only present in special geographical locations. For example, 6 bands occurred only in Sargassum species located in Bandar-Lengeh (Persian Gulf seashore), while they were absent in species of Oman Sea seashore (Table 2). In general, these bands are informative for identification and differentiation of species as it has been supported by DCA analysis (Fig. 1). DCA analysis also identified more effective (discriminating) loci in genetic analysis of Sargassum species. For example, ISSR loci 47 (UBC811-750) which introduced by DCA analysis as discriminative allele, is species specific band for S. ilicifolium (Fig. 2, Table 2). STRUCTURE plot showed low degree of gene exchange and admixture among species groups studied and revealed that each species group contains particular allelic composition which is also supported by reticulation analysis. Coalescence analysis showed close similarity between gene tree and species trees. Genetic rearrangement like duplication, insertion/deletion and recombination may consider as a main tools in divergence of Sargassum species studied. As gene tree depicted in Fig. 8 most of duplication and extinction alleles are happened in species located in Oman Sea seashore than Persian Gulf species. This result may be under influence of number of populations studied here.

CCA analysis suggested that longitude and altitude may affect on divergence of species. As above mentioned longitude affected discriminate most of Sargassum species in Oman Sea while altitude is effective factor in species located in Persian Gulf. Although in our previous study [8] two physicochemical factors including SST (sea surface temperature) and effects of salinity levels (Table 1) did not show significant differences between three collection sites and genetic data (data not shown), more environmental factors need to assess their effects on species diversity. However, Chang et al. [33] suggested possible influence of gradual change in sea surface temperature (SST) on some morphological characters. Up to now no reports have been published to evaluate correlation between environmental factors and genetic data.

At the end we may summarize our findings as: 1- Sargassum species show high level of genetic diversity with species specific alleles which can used in characterization of species and their location (Oman Sea and Persian Gulf), 2- Admixuture by STRUCTURE showed the presence of distinct species groups, 3- Reticulation and coalescence analyses showed gene exchange among species and gene duplication and loss as additional mechanism for bringing about genetic diversity and 4- the effectiveness of geographical altitude and longitude on divergence of species.

REFERENCES


