



Morphometric methods for the analysis and classification of gastropods: a comparison using *Littorina littorea*

Darragh Doyle, Martin P. Gammell and Róisín Nash

Marine and Freshwater Research Centre, Department of Natural Sciences, Galway-Mayo Institute of Technology, Dublin Road, Galway, Ireland

Correspondence: D. Doyle; e-mail: darragh.doyle@research.gmit.ie

(Received 21 August 2017; editorial decision 12 February 2018)

ABSTRACT

The study of morphology is a common means of biological grouping and classification. In recent years, morphometric studies have been dominated by quantitative geometric-morphometric methods of data extraction such as outline or landmark-based analysis. These methods are often used in conjunction with various classification methods such as linear discriminant analysis (LDA) and random forests (RF) in order to achieve inter- and intraspecific grouping based on environmental factors. Despite numerous studies incorporating these data-extraction and classification methods, comparisons of the effectiveness of these methods are largely lacking, especially for species which display low morphological variation. The aim of this study was to compare the effectiveness of two data-extraction methods, elliptic Fourier analysis (EFA) and generalized Procrustes analysis, and two classification methods, LDA and RF, using *Littorina littorea* as the study organism. The results show that the principal component scores derived from EFA, provided the optimal data input for classification while the greatest percentage of successfully classified individuals was achieved using LDA. However, based on this study RF is the recommended classification method as it is resistant to overfitting, makes no assumptions about the data, is well suited to morphometric data and produces similar rates of classification to LDA. The results are discussed in a biological context for *L. littorea*, based on the environmental factors of zonation and shore exposure.

INTRODUCTION

The study of form and morphology has always been vital to taxonomic classification. Even with recent advances in genetics, morphological assessment is still the dominant means of species grouping and classification. As such, methodological comparisons for the extraction and analysis of morphological data are essential. In recent decades, the study of morphology has been advanced by the ‘geometric-morphometric revolution’, in reference to a family of methods that use landmark- or outline-based methods to capture morphology as a set of Cartesian coordinates or outline contours. This morphological information can then be compared between populations or species using multivariate statistics (Rohlf & Marcus, 1993; Adams, Rohlf & Slice, 2004). These methods have been adopted for morphological analysis of numerous marine invertebrate groups, such as gastropods (Primost, Bigatti & Márquez, 2016), bivalves (Sherratt *et al.*, 2016) and polychaetes (Glasby & Glasby, 2006).

Since the advent of this revolution, the literature has been dominated by landmark-based Procrustes methods (Rohlf & Slice, 1990) such as generalized Procrustes analysis (GPA), which use biologically homologous points to describe shape differences between specimens. Even in the present day, far less attention is paid to outline-based methods such as Fourier analysis and its successor elliptic Fourier analysis (EFA). The fact that EFA has not attracted the same widespread use can be attributed to a number

of reasons. First, the mathematical foundation underpinning EFA is believed to be quite complex in contrast to that of Procrustes-based methods (Schmittbuhl *et al.*, 2003) and this perceived complexity may have deterred early practitioners (Caple, Byrd & Stephan, 2017). Second, early critics of the method pointed to the fact that outline methods disregard biologically homologous points and instead give equal weight to the entire structure. However, both of these points have been tackled in recent years and with the advent of user-friendly proprietary software (Iwata & Ukai, 2002), EFA has come into more widespread use, though still not to the same degree as Procrustes-based landmark methods. The early claim that EFA is an extremely powerful tool for morphological studies (Rohlf & Archie, 1984) has never been disproven. A brief summary of both GPA and EFA is given below.

GPA uses points, or ‘landmarks’, which are biologically homologous between specimens, in order to capture shape (Rohlf & Marcus, 1993). These points are then compared with their counterparts on each specimen in order to determine how shapes vary. As such, the derived shape depends entirely on the chosen landmark positions (Webster & Sheets, 2010). Hence, there is a need to choose landmarks that are not only biologically meaningful, but which can be placed with precision from specimen to specimen. For some structures, this is extremely difficult. For example, most gastropod species tend to have few discernible landmarks on their shell surface. A solution to this has been to use ‘semilandmarks’,

which are usually placed on a curve and optimally slid to achieve minimum bending energy. In this way, they capture a point or points which “acquire geometric homology or correspondence” from specimen to specimen (Mitteroecker & Gunz, 2009: 242). Another method is to superimpose a grid over each specimen and anchor the extremities of the grid to biologically homologous points, in order to provide consistent landmark positions. This grid should be rescaled from specimen to specimen to ensure that the same relative positions on each shell are being recorded (Maddux & Franciscus, 2009; Vaux *et al.*, 2017).

In contrast to GPA, EFA entails the decomposition of an object outline into a sum of harmonically related ellipses (or harmonics; Tracey, Lyle & Duhamel, 2006). This provides a set of four coefficients for each harmonic: the trigonometric (sine and cosine) amplitudes of the X and Y increments (Haines & Crampton, 2000; Van Bocxlaer & Schultheiß, 2010). The more harmonics that are used, the better the constructed outline adheres to that of the original object (Kuhl & Giardina, 1982). As such, complex shapes require more harmonics to be reconstructed than do simple objects. Crampton (1995) recommended using the first eight harmonics to capture the shape of a bivalve shell. However, the author also cautioned that the number of harmonics to be used should be carefully considered and that the use of unnecessary or statistically insignificant harmonics may add ‘noise’ to the outline.

Once morphometric data have been extracted from the specimens and compiled, a central aim of most morphometric-based studies is to implement statistical procedures that find the greatest spatial differences between groups within the data. These spatial differences are then used to split the groups according to shape and to provide input for a classification or confusion matrix (Conde-Padin, Grahame & Rolan-Alvarez, 2007; Van Bocxlaer & Schultheiß, 2010).

A method commonly employed to split groups based on shape differences is linear discriminant analysis (LDA; Fisher, 1936), producing axes that minimize the ratio of between-class and within-class variation (Swets & Weng, 1996). Despite the vast amount of morphometric-based literature implementing LDA as a tool for classification (e.g. Valenzuela *et al.*, 2004; Urra, Oliva & Sepúlveda, 2007), very little attention has been given to the assumptions of the test itself (Rexstad *et al.*, 1990). Certain assumptions, such as multivariate normality and equal covariance matrices for each class, are deemed difficult to achieve (Van Bocxlaer & Schultheiß, 2010). Moreover, the number of samples within each group must exceed the dimensionality of each data vector (number of measurements of each sample) in order for the covariance estimates to have full rank and thus be inverted. This problem is difficult to avoid in smaller morphometric datasets (for example, restricting EFA to the first 10 harmonics results in 40 variables per specimen) and methods of ordination such as principal component analysis (PCA) or factor analysis are frequently used for dimensionality reduction. Despite these issues, LDA is generally regarded as being relatively robust to violations of certain assumptions of the test, namely multivariate normality and equal population-covariance matrices (Lachenbruch & Goldstein, 1979; Li, Zhu & Ogihara, 2006).

A solution to the problem of nonparametric classification is to use a machine-learning tool. Machine learning refers to computer-based methods that automate analytical modelling. One of the most popular ensemble learning tools, random forests (RF) (Liaw & Wiener, 2002), utilizes an ensemble of classification or regression trees to predict the dependent variable as a result of majority vote or average assignment across trees (Breiman, 2001; Strobl, Malley & Gerhard, 2009). Once appropriately tuned, RF allows correlated predictor variables to obtain unbiased predictions and estimates of variable importance, and to achieve group classification (Dub *et al.*, 2013).

Numerous methods of data extraction and classification have been used to study a variety of marine molluscs (e.g. Monnet *et al.*,

2009; Sherratt, Serb & Adams, 2017). However, the extraction of morphological data is a more straightforward task for some taxa as opposed to others. For example, smooth-shelled caenogastropods have few identifiable and homologous points that can be compared across individuals, with the exception of the protoconch-teleconch boundary and gerontic features, which leave a record on the shell (Johnston, Tabachnick & Bookstein, 1991). This is a problem when attempting to compare morphological features across species or ecotypes. Even more difficult is the morphological comparison of species that display low levels of interspecific variation.

In this study, *Littorina littorea* (Linnaeus, 1758) is used as the subject, for two reasons. First, the species shows a low level of genetic and ecophenotypic variation (Fevolden & Garner, 1987; Reid, 1996) and so the sensitivity of the method used to detect morphological differentiation can be determined by its ability to discriminate this species into groups reliably, based on shore exposure and vertical zonation. Second, despite the vast body of literature concerning *L. littorea* morphology (Kemp & Bertness, 1984; Cummins *et al.*, 2002; Cotton, Rundle & Smith, 2004), a thorough morphometric study exploring the effects of shore exposure and vertical zonation is lacking for the species.

Littorina littorea is a dioecious, intertidal caenogastropod. The species has an almost ubiquitous presence on North Atlantic rocky shores, where it is often the dominant macroalgal grazer (Lubchenco, 1983). Shell polymorphism between populations of *L. littorea* has been studied extensively in the past (review by Reid, 1996) and even over distances of hundreds of kilometres the species has been found to exhibit only very slight morphological variation (Johannesson, 1992). This lack of clear differentiation between allopatric populations is attributed to the planktotrophic mode of reproduction and prolonged larval dispersal phase (Johannesson, 1988). This results in high gene flow between populations (Yamada, 1987). However, morphological differences are still achievable through the implementation of a plastic phenotype (Hollander *et al.*, 2006). These morphological differences are known to depend on a variety of environmental factors such as shore exposure and zonation, both of which will be explored here.

The aims of this study are (1) to compare the traditional classification method of LDA with the ensemble-learning method of RF to determine which method has greater discrimination success, (2) to provide a comparison of morphometric data extraction for the analysis of a gastropod species with low levels of sympatric and allopatric morphological variation and (3) to provide the first thorough geometric morphometric study of *L. littorea*, taking into account the factors of shore exposure, zonation and sex.

MATERIAL AND METHODS

Sample site and collection

Littorina littorea specimens were collected from Blackhead (53°9′5.0004″N, 9°16′6.9996″W) and Flaggy Shore (53°9′29.736″N, 9°5′27.384″W) on the west coast of Ireland. Blackhead and Flaggy Shore are exposed and sheltered shores, respectively. Fifty specimens were taken from the upper and lower intertidal zones of each site ($n = 200$) using haphazard quadrat sampling. Samples were taken in February 2017. Individuals with damaged or eroded shells were discarded, and only adult specimens were used. Since shell growth in this species is indeterminate, adulthood was assessed by shell height, according to the method used by Williams (1964), Saier (2000) and De Wolf, Blust and Backeljau (2001). Individuals with shell height ≥ 12 mm were considered adults, as that is the length at which sexual maturity is generally reached (Williams, 1964; Yamada, 1987). The specimens were killed by freezing (-20 °C) in order to preserve the reproductive organs and to avoid any damage to the shells. Soft parts were removed from

shells by detaching the columellar muscle. Once removed, the specimens were sexed based on the presence/absence of a penis.

Landmark collection

Shells were digitized in 2D using a Canon EOS 1200D SLR camera mounted on a tripod. Each shell was photographed with the aperture facing directly upwards and with the columella along the vertical axis. Shells were placed on a bed of white cement powder, which provided support and contrast. In order to compensate for a lack of type I (Bookstein, 1997) biologically homologous landmarks, a virtual grid (constructed using Adobe Illustrator) was superimposed over each image to aid in the identification of landmarks. This was adapted from the method used by Maddux & Franciscus (2009) and Vaux *et al.* (2017). The position of the grid was defined by biologically homologous structures on each individual, i.e. the final suture on the right side of the body and the most extreme point of the lower basal lip. Anchoring the grid at these homologous points and rescaling the grid from sample to sample provided consistent locations for digitization along the outline of each individual. Landmarks were applied to the points where the grid lines intersected the shell. In addition, the apex of the shell (the protoconch is rarely preserved in this species) and the penultimate suture both on the right and left side of the body were also landmarked. The shells were landmarked using TPSDIG2 software (Rohlf, 2010). This provided a set of X and Y Cartesian coordinates that contained the size and shape information for each specimen. To remove variation due to size, position and orientation, the coordinates were subjected to generalized least-squares Procrustes superimposition (Rohlf & Slice, 1990). This removed all confounding information that was not directly related to shape and provided a set of Procrustes residuals for use in the analysis.

Outline digitization

Shells were digitized using the same method as above, with one exception. In order to extract just the outline, the shells were secured to a glass panel in a constant position and lit from beneath using an LED spotlight. This provided a silhouette of each shell. The chain code was extracted by binarizing the images and automatically tracing the curve of each specimen. This chain code was used to compute normalized elliptic Fourier descriptors based on the first harmonic. This provided a set of 20 harmonics, each containing four coefficients ($n = 80$). By visual inspection, it was determined that the first ten harmonics were sufficient to capture accurately the relatively simple shell shape. The first harmonic (which contains size and rotation information) was removed. This left 36 Fourier coefficients (harmonics 2–10) which could be analysed using conventional multivariate methods. Removal of noninformative harmonics (noise) also greatly reduced the size of the overall dataset. This was important as the data were later analysed using LDA, which requires a matrix inversion of the pooled covariance matrix. Data reduction is particularly important for EFA, as the Fourier coefficients are composed of the trigonometric amplitudes of the X and Y increments, which generally results in large numbers of variables. Shell outlines were extracted and normalized elliptic Fourier descriptors were calculated using various packages within the software suite SHAPE v. 1.3 (Iwata & Ukai, 2002; Tracey *et al.*, 2006).

Statistical analysis

To classify individuals into groups, a number of different methods were employed. We refer to the four populations sampled from different exposures and zonations as: exposed–lower (EL), exposed–upper (EU), sheltered–lower (SL) and sheltered–upper

(SU). Individuals were also assessed to determine if sexual shape dimorphism was present.

Multivariate statistical analyses were carried out using the Procrustes residuals for the GPA data and Fourier coefficients for the EFA data. To ordinate and visually explore the data, PCA was used, as it is simply an ordination method that makes no assumptions about the data. Also, it subjects the data to a rigid transformation so that no information is lost. A broken-stick test (Jackson, 1993) was used to determine which principal components (PCs) were statistically significant for each dataset. Warped outline-deformation grids along PC1 and PC2 were generated for the GPA data, while contour deformations were generated for the EFA data in order to visualize morphological changes along the axes of greatest variation. Warped outline-deformation grids were constructed in TPSRELW v. 1.67 (Rohlf, 2007) and contour deformations in SHAPE v. 1.3 – PrinComp (Iwata & Ukai, 2002).

Similarity between the two different methods was assessed through the correlation of pairwise distances. Euclidean distance matrices were constructed from both the GPA residuals and the EFA coefficients. Both distance matrices were assessed for similarity through the use of a Mantel test. This test randomly permutes columns and rows to provide matrix correlations for unrelated matrices (Smouse, Long & Sokal, 1986). This correlation was used to infer how similar or dissimilar the two different methods were, based on the r -test statistic. The Mantel test was carried out with 9,999 permutations. The data were assessed for multivariate normality by computing Mardia's skewness and kurtosis (Mardia, 1970), in addition to a Doornik and Hansen omnibus test (Doornik & Hansen, 2008). Orthogonal PCs were extracted from the Procrustes residuals and the Fourier coefficients.

LDA was carried out on the raw EFA coefficients/GPA residuals, maximum PCs and on a variable number of PCs. To prevent overfitting, the number of PCs was incrementally reduced until the highest jack-knifed cross-validation group assignment percentage was achieved for each group. This method was used based on the findings of Sheets *et al.* (2006). All multivariate statistics were computed using PAST v. 3.15 (Hammer, Harper & Ryan, 2008).

RF was adopted, because the method is extremely resistant to overfitting of the data and it requires very little tuning to produce the optimal classification algorithm, compared with methods such as Support Vector Machines which, despite their incorporation into a number of recent morphometric studies (Santos, Guyomarc'h & Bruzek, 2014), are regarded as more difficult to train to produce reliable classification. RF classification rate also improves with larger datasets, which are often found in morphometric-based studies dealing with large numbers of variables per sample (Díaz-Uriarte & De Andres, 2006). RF was carried out on the raw EFA coefficients and GPA residuals, the maximum number of PCs for both methods, and also a variable number of PCs for both methods. The number of PCs used as input was again incrementally reduced until the highest cross-validated classification percentage was achieved. Here, tenfold cross-validation was employed. This method splits the data into training and testing sets and then provides an average of the results for each split in order to give an indication of the effectiveness of a model. The model was run with 100 iterations (trees). RF analysis was conducted using WEKA v. 3.8 (Hall *et al.*, 2009).

RESULTS

Visual inspection of the PCA scatter plots for both GPA (Fig. 1) and EFA (Fig. 2) indicated that both methods recorded similar levels of variation in the dataset. The Mantel test revealed that the distance matrices of the Procrustes residuals and the Fourier coefficients were positively correlated with the r statistic, indicating weak positive correlation ($r = 0.327$, $P = 0.0001$). This suggests that the two datasets are relatively similar. The broken-stick test

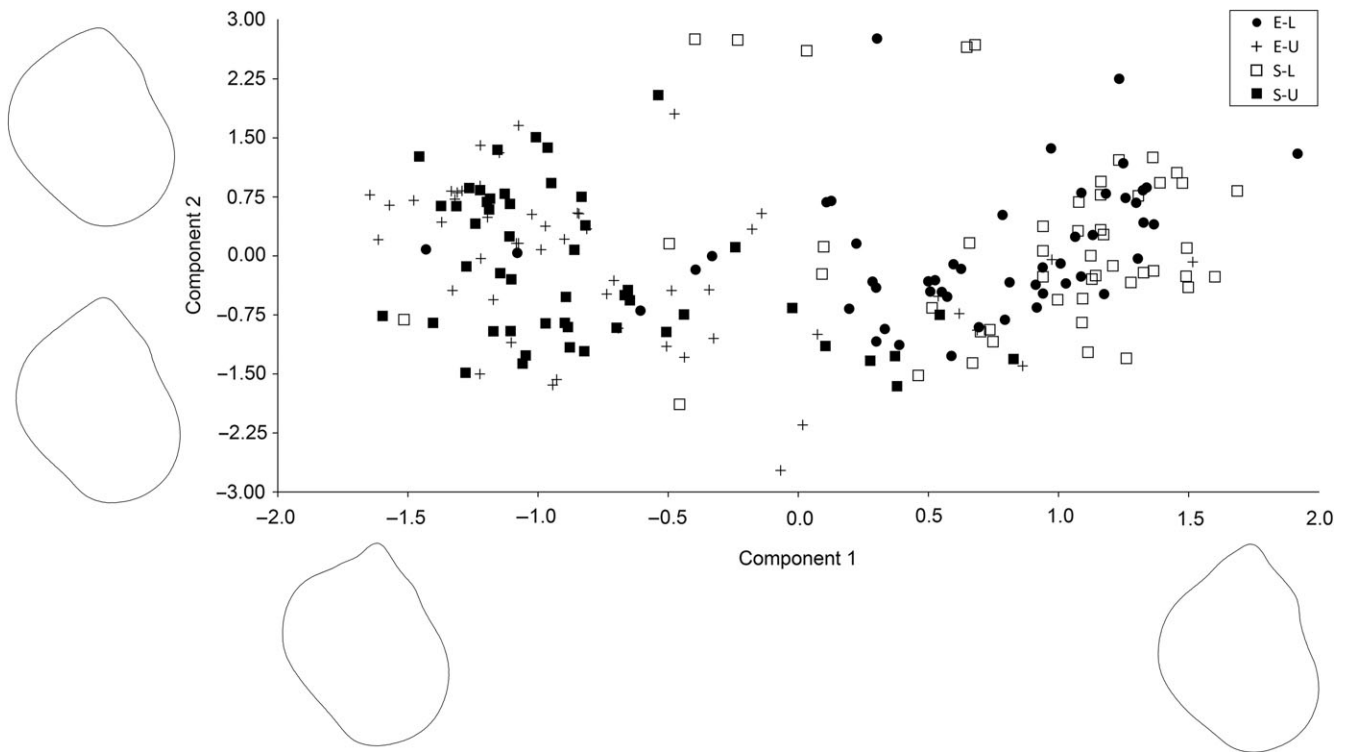


Figure 1. Principal components analysis scatterplot of Procrustes residuals showing morphological variation of shells of *Littorina littorea* individuals based on zonation and shore exposure. Shape deformations are included to show morphological change along principal component 1 and principal component 2. Abbreviations: E-L, exposed lower shore; E-U, exposed upper shore; S-L, sheltered lower shore; S-U, sheltered upper shore.

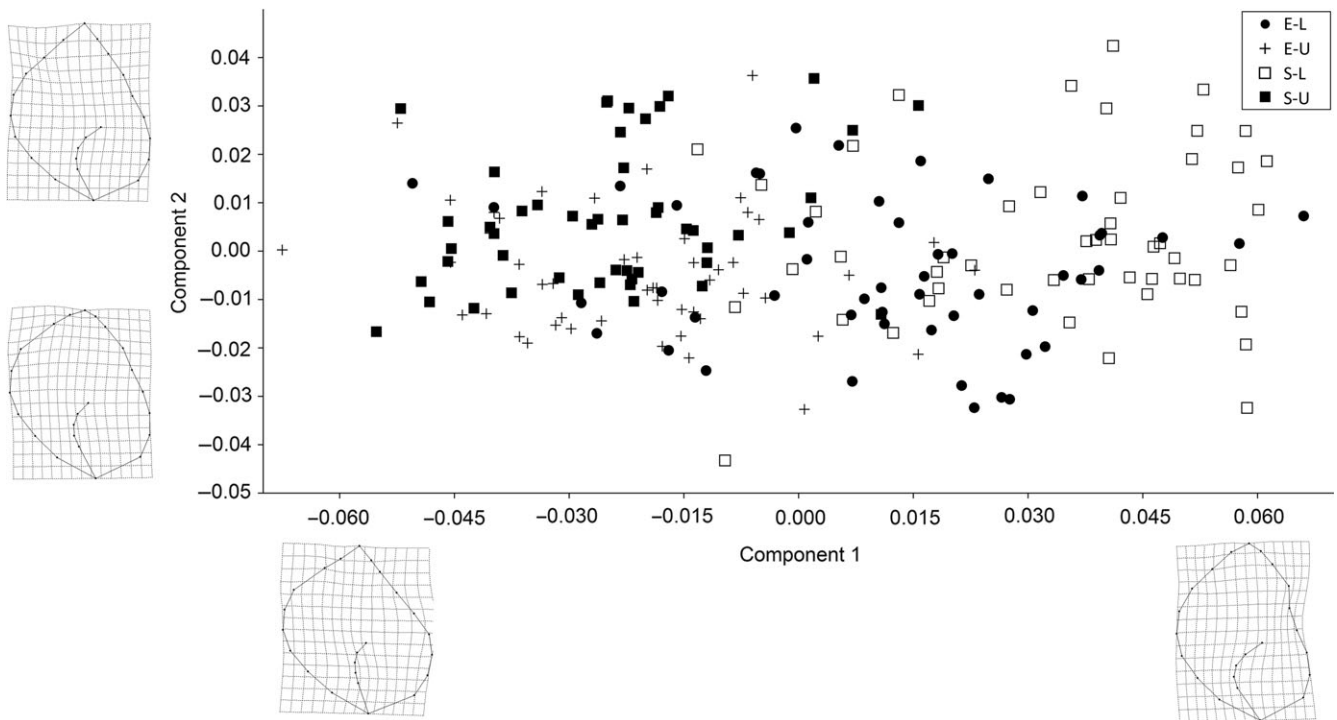


Figure 2. Principal components analysis scatter plot of Elliptic Fourier coefficients showing morphological variation of shells of *Littorina littorea* individuals based on zonation and shore exposure. Closed-contour shape deformations are included to show morphological change along principal component 1 and principal component 2. Abbreviations: E-L, exposed lower shore; E-U, exposed upper shore; S-L, sheltered lower shore; S-U, sheltered upper shore.

revealed the first four principal components to be statistically significant for both EFA and GPA. Individuals appeared to cluster most clearly based on zonation, rather than shore type. SU and

EU individuals clustered together, as did SL and EL individuals. The clustering based on zonation was most apparent for the PCA based on EFA coefficients (Fig. 2). Shape deformations generated

for GPA revealed the morphological variation along the first two principal components. Upper-shore specimens displayed a broader shell with more pointed apex, as opposed to a narrower shape with a flatter apex for lower-shore specimens. As the aperture was recorded using GPA, deformations in aperture shape could also be visualized. This was not possible for EFA, in which structures inside the shell contour (i.e. the aperture) were not recorded. For the EFA contour deformations, individuals displayed a slightly narrower shell with a taller spire, in moving from negative to positive along PC1. No evidence of sexual dimorphism was found through PCA group separation or LDA (results not shown).

Based on the Mardia multivariate normality and Hansen and Doornik omnibus tests, the data were found not to adhere to a normal probability distribution function, indicating that the data were significantly nonnormal. As such, the assumptions of the LDA were deemed to be violated. However, the test is known to be robust to violations of both multivariate normality and equal covariance matrices for each class (Lachenbruch & Goldstein, 1979), performing well in both dimensionality reduction and classification despite these violations (Li *et al.*, 2006). For this reason, and in the interest of comparing the method with RF, LDA was still carried out.

The optimal number of PCs to use for the LDA was found by incrementally reducing the number of PCs until the jack-knifed cross-validated number of correctly assigned individuals peaked (Table 1). For EFA, the optimal number of PCs was found to be 17, while for GPA the optimal number was 28. Before reducing the PC numbers used, strong evidence of overfitting was found. The highest classification success was found by using the first 17 PCs from the EFA coefficients, which correctly assigned 78%, representing 156 individuals. LDA of the first 28 PCs from the GPA residuals produced a slightly lower classification rate of 75%, representing 150 individuals. The difference in classification rate was therefore 3%, or six individuals. GPA required a greater number of PCs than EFA in order to produce optimal results.

RF was used as a nonparametric means of classification. As with LDA, a variable number of PCs were used in order to find the highest classification percentage (Table 2). For EFA, the highest classification success was achieved when a variable number of

PCs were used. Reducing the data to the first 25 PCs produced a classification rate of 75.5% (Kappa = 0.6733, mean absolute error = 0.2597), while for GPA the optimal input was the first 20 PCs, which produced a classification rate of 61% (Kappa = 0.48, mean absolute error = 0.2867). For GPA, the greatest classification success was achieved with a RF analysis of the raw GPA residuals. RF of the raw GPA residuals produced a classification rate of 65% (Kappa = 0, mean absolute error = 0.375). However, a Kappa value of zero indicates a poor model, or a result which could be expected by chance (Carletta, 1996). Taking the optimal method for both EFA and GPA (variable PCs), RF of EFA was 14.5% more effective than that of GPA, equating to a difference of 29 correctly assigned individuals.

DISCUSSION

Two methods of data extraction were employed in this study: EFA and GPA. Visual inspection of the PCA scatter plots for both the EFA and GPA data indicate that both methods recorded similar variation in the dataset, both methods revealing broader shell shape in upper shore specimens. In addition, the Mantel test showed that the distance matrices for both datasets were positively correlated, indicating that the two data extraction methods, despite their theoretical differences, recorded similar morphological variation. The results of this study are consistent with the findings of Van Bocxlaer & Schultheiß (2010), who found that EFA performed better than semilandmark analysis in providing input for classification. These authors also found that the optimal method to employ for the analysis of gastropod shells depends on the level of shell ornamentation, with semilandmark analysis performing better with ornamented shells. However, for species with a relatively simple shape and low levels of sculpture such as *Littorina littorea*, EFA is the optimal method for obtaining morphological data, as evidenced by the current study. Sheets *et al.* (2006) also compared a number of morphometric methods, including semilandmark-based methods and EFA, and found comparable rates of classification success in all of them. However, these authors were unable to use automatic outline detection, due to the

Table 1. Classification percentages produced by linear discriminant analysis.

Linear discriminant analysis							
Data acquisition	EFA coefficients and GPA residuals		Maximum PCs		Variable no. of PCs		
	Observed	Cross-validated	Observed	Cross-validated	No. of PCs used	Observed	Cross validated
EFA	86.5	74	86.5	74	17	82	78
GPA	86.5	68	86.5	68	28	84	75

Raw coefficients from elliptic Fourier analysis (EFA) and residuals from generalized Procrustes analysis (GPA); maximum principal components (PCs) and variable numbers of PCs provided data input.

Table 2. Classification percentages produced by random forests.

Random forest analysis							
Data acquisition	EFA coefficients and GPA residuals		Maximum PCs		Variable no. of PCs		
	Training	Cross-validated	Training	Cross-validated	No. of PCs used	Training	Cross-validated
EFA	100	75	100	71	25	100	75.5
GPA	100	65	100	56.5	20	100	61

Raw coefficients from elliptic Fourier analysis (EFA) and residuals from generalized Procrustes analysis (GPA); maximum principal components (PCs) and variable numbers of PCs provided data input.

irregular shape of the material (feathers) and instead had to rely on manual tracing of the curves. This may have introduced a source of error that could explain the discrepancy in findings between their study and that of Van Bocxlaer & Schultheiß (2010).

In contrast to the comparative study by Sheets *et al.* (2006), the method of data extraction employed in the current study had a significant impact on classification. Data obtained through EFA performed consistently better at *a priori* classification in comparison with data obtained through GPA. This held true regardless of whether the data used were raw residuals/coefficients or PC scores, and regardless of the classification method used. This supports the findings of Van Bocxlaer & Schultheiß (2010), who found that outline data provided the optimal input for the classification of unornamented shells. Also, EFA was much quicker to carry out than GPA. For EFA, the method of extracting morphometric data from each of the shells was automated from binarized images, whereas for GPA each shell required time-consuming individual landmarking. This is not a major consideration for the relatively small number of specimens used here ($n = 200$) but, for morphometric studies with a greater number of samples and replicates, speed becomes more important. Regardless of data collection method or classification method, the highest classification percentages were obtained when using a variable number of PCs, as also found by Sheets *et al.* (2006).

As a classification method, LDA superficially performed better than RF analysis upon jack-knifed cross-validation. The highest classification percentages were obtained when using a variable number of PCs. The main purpose of carrying out LDA on the data here was to provide a comparison for the RF. That the LDA produced very similar classification rates to that of the RF suggests that LDA is in fact somewhat robust to certain violations of the test assumptions, as indicated by previous studies (Lachenbruch & Goldstein, 1979; Li *et al.*, 2006).

RF proved to be a relatively successful means of nonparametric classification, producing comparable rates to LDA for the EFA data. In order to produce the optimal results, a number of factors within the model needed to be fine-tuned. A major factor influencing classification rate was the number of iterations. RF proved less effective in the classification of GPA data, despite efforts to produce classification rates comparable to those of EFA data. Based on this, RF is recommended for outline data which do not adhere to multivariate normality or which otherwise do not meet the assumptions of LDA. For landmark data, there may be preferable machine-learning methods with the capability to produce better classification results—such as SVMs (Van Bocxlaer & Schultheiß, 2010) or logistic regression (Navega *et al.*, 2015). Our method did not achieve the extremely high rates of classification (>90%) commonly reported in the literature (e.g. Calle *et al.*, 2002; Huber *et al.*, 2011). However, many of the previous studies have explored interspecific variation (e.g. Jaramillo-O *et al.*, 2015), in which morphological differentiation is often more pronounced. Future morphometric studies exploring relatively slight morphological variation, as in the present case of *L. littorea*, will likely report lower rates of correct classification.

Visual inspection of PCA scatter plots for both EFA and GPA data revealed that individuals of *L. littorea* showed a high degree of morphological similarity. This is consistent with the low levels of genetic variation between populations of the species reported in previous studies (e.g. Fevolden & Garner, 1987; Hollander *et al.*, 2006). However, despite group separation being slight, differentiation could still be observed, especially based on tidal zonation. Upper-shore individuals displayed a broader shell with a more pointed apex. This contradicted expectations, because in *Littorina* species larger and broader individuals are usually more prevalent on the lower shore (Cummins *et al.*, 2002), where greater shell girth acts as a defence against crushing predators such as *Carcinus* (Johannesson, 1986). The protoconch and early whorls of the shell

of *L. littorea* are always sharply pointed (Reid, 1996), so the finding of a flatter apex low on the shore could imply a degree of erosion of the very tip, which is less pronounced at higher tidal levels. Only mature adults were examined here, and this suggestion would require testing with juvenile shells showing well-preserved apices. Morphological separation based on exposure was not clearly defined. This is consistent with the work of Janson (1987), who found virtually no morphological variation between *L. littorea* from exposed and sheltered shores. However, Cummins *et al.*, (2002) reported a significant correlation between shell width and exposure. No sexual dimorphism was detected in the present study. Saur (1990) and De Wolf *et al.* (2001) also reported no difference in shell height between male and female *L. littorea*. However, other authors (Moore, 1937; Van den Broeck *et al.*, 2007) have found evidence of sexual dimorphism. GPA warped outline grids provided detailed shape deformations that included shape information for structures within the shell contour, i.e. the aperture. This is perhaps the greatest advantage of landmark-based methods as opposed to outline-based methods. It is especially significant for caenogastropods, where aperture shape can be highly variable depending on environmental factors and often provides an important aid to classification. EFA deformations were also generated for the closed contours, but revealed less information of biological significance than those of the GPA.

Recommendations for future studies

Based on the results of this study, outline-based methods appear preferable to landmark-based methods for the extraction of morphological data when few unambiguously homologous points are present. The subsequent classification approach that can be used depends on whether the data meet or violate the assumptions of LDA. In either case, RF is recommended because the method makes no assumptions about the data, and is a straightforward and robust method for classification as compared with LDA. In addition, it may be more straightforward for most morphometric practitioners to tune a RF model than to determine whether the assumptions of LDA are violated (Karels, Bryant & Hik, 2004). It is also recommended that, whether using LDA or RF, cross-validation should be employed to ensure that statistical overfitting does not occur. However, RF is well known to be resistant to overfitting (Breiman, 2001), especially when higher iterations are used, and this is another reason to favour this method. Using an incrementally reduced variable number of PCs resulted in the highest classification rate regardless of the data extraction or classification method used. If the goal of the study is to observe the exact features which best distinguish between populations or groups of specimens, then GPA is preferable to EFA, because structures within the 2D shell contour can be digitized, landmarked and then visualized.

Our conclusions and recommendations are based on two data extraction methods and two classifications, and aim to provide a general guide for the analysis of gastropods with few clearly homologous landmarks. However, a wealth of data-extraction methods exist in addition to the methods used here, each with variants of their own. The same is true for classification methods—ensemble learning is simply one suite within a broad range of machine-learning methods. Future comparative studies to investigate the effectiveness of various machine-learning methods are recommended as this means of classification becomes increasingly popular in morphometric studies.

ACKNOWLEDGEMENTS

We thank Galway–Mayo Institute of Technology and the Marine and Freshwater Research Centre for the use of their facilities. We thank Otto Storan, Steve Barrett and Mary Veldon for their

technical assistance with the field and laboratory work. We also thank Associate Editor John Grahame for his careful reading of the manuscript and for his many insightful comments and suggestions.

REFERENCES

- ADAMS, D.C., ROHLF, F.J. & SLICE, D.E. 2004. Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology*, **71**: 5–16.
- BOOKSTEIN, F.L. 1997. *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press, Cambridge, UK.
- BREIMAN, L. 2001. Random forests. *Machine Learning*, **45**: 5–32.
- CALLE, L., DAVID, A., QUIÑONES, M.L., ERAZO, H.F. & JARAMILLO, O. 2002. Morphometric discrimination of females of five species of *Anopheles* of the subgenus *Nyssorhynchus* from southern and northwest Colombia. *Memórias do Instituto Oswaldo Cruz*, **97**: 1191–1195.
- CAPLE, J., BYRD, J. & STEPHAN, C.N. 2017. Elliptical Fourier analysis: fundamentals, applications, and value for forensic anthropology. *International Journal of Legal Medicine*, **131**: 1–16.
- CARLETTA, J. 1996. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics*, **22**: 249–254.
- CONDE-PADÍN, P., GRAHAME, J.W. & ROLÁN-ALVAREZ, E. 2007. Detecting shape differences in species of the *Littorina saxatilis* complex by morphometric analysis. *Journal of Molluscan Studies*, **73**: 147–154.
- COTTON, P.A., RUNDLE, S.D. & SMITH, K.E. 2004. Trait compensation in marine gastropods: shell shape, avoidance behavior, and susceptibility to predation. *Ecology*, **85**: 1581–1584.
- CRAMPTON, J.S. 1995. Elliptical Fourier shape analysis of fossil bivalves: some practical considerations. *Lethaia*, **28**: 179–186.
- CUMMINS, V., COUGHLAN, S., MCCLEAN, O., CONNOLLY, N., MERCER, J. & BURNELL, G. 2002. An assessment of the potential for the sustainable development of the edible periwinkle, *Littorina littorea*, industry in Ireland. *Marine Resource Series*, **22**: 64–154.
- DE WOLF, H., BLUST, R. & BACKELJAU, T. 2001. Shell size variation in *Littorina littorea* in the western Scheldt estuary. *Journal of Shellfish Research*, **20**: 427–430.
- DÍAZ-URIARTE, R. & DE ANDRES, S.A. 2006. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics*, **7**: 3.
- DOORNIK, J.A. & HANSEN, H. 2008. An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, **70**: 927–939.
- DUB, J.D., REDMAN, R.A., WAHL, D.H. & CZESNY, S.J. 2013. Utilizing random forest analysis with otolith mass and total fish length to obtain rapid and objective estimates of fish age. *Canadian Journal of Fisheries and Aquatic Sciences*, **70**: 1396–1401.
- FEVOLDEN, S.E. & GARNER, S.P. 1987. Environmental stress and allozyme variation in *Littorina littorea* (Prosobranchia). *Marine Ecology: Progress Series*, **39**: 129–136.
- FISHER, R.A. 1936. The use of multiple measurements in taxonomic problems. *Annals of Human Genetics*, **7**: 179–188.
- GLASBY, C.J. & GLASBY, T.M. 2006. Two types of uncini in *Polycirrus* (Polychaeta: Terebellidae: Polycirrinae) revealed using geometric morphometrics. *Journal of Natural History*, **40**: 237–253.
- HAINES, A.J. & CRAMPTON, J.S. 2000. Improvements to the method of Fourier shape analysis as applied in morphometric studies. *Palaentology*, **43**: 765–783.
- HALL, M., FRANK, E., HOLMES, G., PFAHRINGER, B., REUTEMANN, P. & WITTEN, I.H. 2009. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, **11**: 10–18.
- HAMMER, Ø., HARPER, D.A.T. & RYAN, P.D. 2008. PAST—paleontological statistics, ver. 1.89. Paleontological Museum, University of Oslo. Available at: <http://folk.uio.no/ohammer/past/index.html>.
- HOLLANDER, J., COLLYER, M.L., ADAMS, D.C. & JOHANNESSEN, K. 2006. Phenotypic plasticity in two marine snails: constraints superseding life history. *Journal of Evolutionary Biology*, **19**: 1861–1872.
- HUBER, H.R., JORGENSEN, J.C., BUTLER, V.L., BAKER, G. & STEVENS, R. 2011. Can salmonids (*Oncorhynchus* spp.) be identified to species using vertebral morphometrics? *Journal of Archaeological Science*, **38**: 136–146.
- IWATA, H. & UKAI, Y. 2002. SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *Journal of Heredity*, **93**: 384–385.
- JACKSON, D.A. 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. *Ecology*, **74**: 2204–2214.
- JANSON, K. 1987. Allozyme and shell variation in two marine snails (*Littorina*, Prosobranchia) with different dispersal abilities. *Biological Journal of the Linnean Society*, **30**: 245–256.
- JARAMILLO-O, N., DUJARDIN, J.P., CALLE-LONDOÑO, D. & FONSECA-GONZÁLEZ, I. 2015. Geometric morphometrics for the taxonomy of 11 species of *Anopheles* (*Nyssorhynchus*) mosquitoes. *Medical and Veterinary Entomology*, **29**: 26–36.
- JOHANNESSEN, B. 1986. Shell morphology of *Littorina saxatilis* Olivi: the relative importance of physical factors and predation. *Journal of Experimental Marine Biology and Ecology*, **102**: 183–195.
- JOHANNESSEN, K. 1988. The paradox of Rockall: why is a brooding gastropod (*Littorina saxatilis*) more widespread than one having a planktonic larval dispersal stage (*L. littorea*)? *Marine Biology*, **99**: 507–513.
- JOHANNESSEN, K. 1992. Genetic variability and large scale differentiation in two species of littorinid gastropods with planktotrophic development, *Littorina littorea* (L.) and *Melarhaphe* (*Littorina*) *neritoides* (L.) (Prosobranchia: Littorinacea), with notes on a mass occurrence of *M. neritoides* in Sweden. *Biological Journal of the Linnean Society*, **47**: 285–299.
- JOHNSTON, M.R., TABACHNICK, R.E. & BOOKSTEIN, F.L. 1991. Landmark-based morphometrics of spiral accretionary growth. *Paleobiology*, **17**: 19–36.
- KARELS, T.J., BRYANT, A.A. & HIK, D.S. 2004. Comparison of discriminant function and classification tree analyses for age classification of marmots. *Oikos*, **105**: 575–587.
- KEMP, P. & BERTNESS, M.D. 1984. Snail shape and growth rates: evidence for plastic shell allometry in *Littorina littorea*. *Proceedings of the National Academy of Sciences of the USA*, **81**: 811–813.
- KUHL, F.P. & GIARDINA, C.R. 1982. Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, **18**: 236–258.
- LACHENBRUCH, P.A. & GOLDSTEIN, M. 1979. Discriminant analysis. *Biometrics*, **35**: 69–85.
- LI, T., ZHU, S. & OGIHARA, M. 2006. Using discriminant analysis for multi-class classification: an experimental investigation. *Knowledge and Information Systems*, **10**: 453–472.
- LIU, A. & WIENER, M. 2002. Classification and regression by random Forest. *R News*, **2**: 18–22.
- LUBCHENCO, J. 1983. *Littorina* and *Fucus*: effects of herbivores, substratum heterogeneity, and plant escapes during succession. *Ecology*, **64**: 1116–1123.
- MADDUX, S.D. & FRANCISCUS, R.G. 2009. Allometric scaling of infraorbital surface topography in *Homo*. *Journal of Human Evolution*, **56**: 161–174.
- MARDIA, K.V. 1970. Measures of multivariate skewness and kurtosis with applications. *Biometrika*, **57**: 519–530.
- MITTEROECKER, P. & GUNZ, P. 2009. Advances in geometric morphometrics. *Evolutionary Biology*, **36**: 235–247.
- MONNET, C., ZOLLIKOFER, C., BUCHER, H. & GOUEMAND, N. 2009. Three-dimensional morphometric ontogeny of mollusc shells by micro-computed tomography and geometric analysis. *Paleontologia Electronica*, **12**: 1–13.
- MOORE, H.B. 1937. The biology of *Littorina littorea*. Part I. Growth of the shell and tissues, spawning, length of life and mortality. *Journal of the Marine Biological Association of the United Kingdom*, **21**: 721–742.
- NAVEGA, D., VICENTE, R., VIEIRA, D.N., ROSS, A.H. & CUNHA, E. 2015. Sex estimation from the tarsal bones in a Portuguese sample: a machine learning approach. *International Journal of Legal Medicine*, **129**: 651–659.
- PRIMOST, M.A., BIGATTI, G. & MÁRQUEZ, F. 2016. Shell shape as indicator of pollution in marine gastropods affected by imposex. *Marine and Freshwater Research*, **67**: 1948–1954.

- REID, D.G. 1996. *Systematics and evolution of Littorina*. Ray Society, London.
- REXSTAD, E.A., MILLER, D.D., FLATHER, C.H., ANDERSON, E.M., HUPP, J.W. & ANDERSON, D.R. 1990. Questionable multivariate statistical inference in wildlife habitat and community studies: a reply. *Journal of Wildlife Management*, **52**: 189–193.
- ROHLF, F.J. 2007. *tpsRelw version 1.45*. Department of Ecology and Evolution, State University of New York, Stony Brook, NY.
- ROHLF, F.J. 2010. *TPSDig2, version 2.16*. Department of Ecology and Evolution, State University of New York, Stony Brook, NY.
- ROHLF, F.J. & ARCHIE, J.W. 1984. A comparison of Fourier methods for the description of wing shape in mosquitoes (Diptera: Culicidae). *Systematic Zoology*, **33**: 302–317.
- ROHLF, F.J. & MARCUS, L.F. 1993. A revolution morphometrics. *Trends in Ecology & Evolution*, **8**: 129–132.
- ROHLF, F.J. & SLICE, D. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Biology*, **39**: 40–59.
- SAIER, B. 2000. Age-dependent zonation of the periwinkle *Littorina littorea* (L.) in the Wadden Sea. *Helgoland Marine Research*, **54**: 224–229.
- SANTOS, F., GUYOMARCH, P. & BRUZEK, J. 2014. Statistical sex determination from craniometrics: comparison of linear discriminant analysis, logistic regression, and support vector machines. *Forensic Science International*, **245**: 204 e1–e8.
- SAUR, M. 1990. Mate discrimination in *Littorina littorea* (L.) and *L. saxatilis* (Oliv.) (Mollusca: Prosobranchia). In: *Progress in littorinid and muricid biology* (K. Johannesson, D.G. Raffaelli & C.J. Hannaford-Ellis, eds), pp. 261–270. Springer, Dordrecht.
- SCHMITTBUHL, M., ALLENBACH, B., LE MINOR, J.M. & SCHAAF, A. 2003. Elliptical descriptors: some simplified morphometric parameters for the quantification of complex outlines. *Mathematical Geology*, **35**: 853–871.
- SHEETS, H.D., COVINO, K.M., PANASIEWICZ, J.M. & MORRIS, S.R. 2006. Comparison of geometric morphometric outline methods in the discrimination of age-related differences in feather shape. *Frontiers in Zoology*, **3**: 15–27.
- SHERRATT, E., ALEJANDRINO, A., KRAEMER, A.C., SERB, J.M. & ADAMS, D.C. 2016. Trends in the sand: directional evolution in the shell shape of recessing scallops (Bivalvia: Pectinidae). *Evolution*, **70**: 2061–2073.
- SHERRATT, E., SERB, J.M. & ADAMS, D.C. 2017. Rates of morphological evolution, asymmetry and morphological integration of shell shape in scallops. *BMC Evolutionary Biology*, **17**: 248.
- SMOUSE, P.E., LONG, J.C. & SOKAL, R.R. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology*, **35**: 627–632.
- STROBL, C., MALLEY, J. & GERHARD, T. 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging and random forests. *Psychological Methods*, **14**: 323–348.
- SWETS, D.L. & WENG, J.J. 1996. Using discriminant eigenfeatures for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**: 831–836.
- TRACEY, S.R., LYLE, J.M. & DUHAMEL, G. 2006. Application of elliptical Fourier analysis of otolith form as a tool for stock identification. *Fisheries Research*, **77**: 138–147.
- URRA, A., OLIVA, D. & SEPÚLVEDA, M. 2007. Use of a morphometric analysis to differentiate *Adelomelon ancilla* and *Odontocymbiola magellanica* (Caenogastropoda: Volutidae) of southern Chile. *Zoological Studies, Taipei*, **46**: 253–261.
- VALENZUELA, N., ADAMS, D.C., BOWDEN, R.M. & GAUGER, A.C. 2004. Geometric morphometric sex estimation for hatchling turtles: a powerful alternative for detecting subtle sexual shape dimorphism. *Copeia*, **2004**: 735–742.
- VAN BOCXLAER, B. & SCHULTHEIB, R. 2010. Comparison of morphometric techniques for shapes with few homologous landmarks based on machine-learning approaches to biological discrimination. *Paleobiology*, **36**: 497–515.
- VAN DEN BROECK, H., DE WOLF, H., BACKELJAU, T. & BLUST, R. 2007. Effects of environmental stress on the condition of *Littorina littorea* along the Scheldt estuary (The Netherlands). *Science of the Total Environment*, **376**: 346–358.
- VAUX, F., CRAMPTON, J.S., MARSHALL, B.A., TREWICK, S.A. & MORGAN-RICHARDS, M. 2017. Geometric morphometric analysis reveals that the shells of male and female siphon whelks *Penion chathamensis* are the same size and shape. *Molluscan Research*, **37**: 194–201.
- WEBSTER, M. & SHEETS, H.D. 2010. A practical introduction to landmark-based geometric morphometrics. *Quantitative Methods in Paleobiology*, **16**: 168–188.
- WILLIAMS, E.E. 1964. The growth and distribution of *Littorina littorea* (L.) on a rocky shore in Wales. *Journal of Animal Ecology*, **33**: 413–432.
- YAMADA, S.B. 1987. Geographic variation in the growth rates of *Littorina littorea* and *L. saxatilis*. *Marine Biology*, **96**: 529–534.