Efficacy of a non-invasive sampling method for genomic analysis of the critically endangered limpet *Patella candei candei* (d'Orbigny, 1840)

J. Quinteiro[1], G. González-Lorenzo[2], D. Hernández-Reyes[3], L. Quinteiro[1], R. Herrera[4], J. Martínez[4], J.A. González[3], M. Rey-Méndez[1] and N. González-Henríquez[3]

[1] Molecular Systematics Laboratory, Department of Biochemistry and Molecular Biology, Aquaculture Institute; University of Santiago de Compostela, A Coruña, Galicia, Spain.
[2] Canary Islands Oceanographic Center, Spanish Institute of Oceanography (IEO, CSIC), Tenerife, Spain.
[3] Departament of Biology, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain.
[4] Directorate General for the Fight against Climate Change and the Environment. Council for the Ecological Transition, Fight against Climate Change and Territorial Planning. Canary Islands Government.

ABSTRACT

In this study we tested a non-invasive, swab-based DNA sampling method that can be used *in situ* for the monitoring of *Patella candei candei* endangered populations. These DNA samples, obtained from the limpet mucus, must provide DNA for the implementation of genome reduced-representation methodological approaches (ddRAD) to obtain thousands of markers dispersed trough the genome in a cost-effective way. The use of relevant, but non-ideally collected and/or preserved DNA samples, including those from non-invasive mucosal swabs sampling provided acceptable genotyping results. The generated data allowed the correct use of SNPs identification pipelines and population genetic structure analysis. The achieved results are relevant for the design and monitoring of conservation strategies for the endangered majorera limpet in Fuerteventura Island.

KEYWORDS: DNA sampling method, swab-based, methodological ddRAD

RESUMEN

En este estudio probamos un método de muestreo de ADN no invasivo basado en hisopos que se puede utilizar *in situ* para el seguimiento de poblaciones en peligro de *Patella candei candei*. Estas muestras de ADN, obtenidas del mucus de la lapa, deben proporcionar ADN para la implementación de enfoques metodológicos de representación reducida del genoma (ddRAD), para obtener miles de marcadores dispersos a través del genoma de una manera efectiva. El uso de muestras de ADN relevantes, pero no idealmente recolectadas y / o preservadas, incluyendo las muestras de mucosas no invasivas, proporcionó resultados aceptables para la genotipificación. Los datos generados permitieron el uso correcto de sistemas de identificación de SNP y el análisis de la estructura genética de la población. Los resultados obtenidos son relevantes para el diseño y seguimiento de estrategias de conservación de la lapa majorera en peligro de extinción en la isla de Fuerteventura.

PALABRAS CLAVE: muestreo de ADN, ddRAD

INTRODUCTION

The development and testing of non-invasive sampling methods is crucial for the collection of samples for molecular taxonomy studies and analysis of the population genetic structure of critically endangered organisms, avoiding a negative impact on small and / or endangered populations.

Non-invasive samples have been proposed and evaluated for genetic analysis in order to minimize the impact on the studied populations using different DNA sources (Balázs *at al.*, Domingues *at al.*, Morinha *et al*., 2014). Of particular interest are those works that have successfully demonstrated the feasibility of analysis from fish and mollusc mucus (Domingues *at al.*, 2019; Kim and Roe, 2021; Morinha *at al.*, 2014). Although the non-invasive method that has the least effect on the environment and organisms is based on environmental DNA (eDNA), it does not allow obtaining individualized genetic information.

The limpet *P. candei candei* d'Orbigny, 1840 is endemic to Fuerteventura and Selvagens islands. The critical state of the "majorera limpet", from the unique Fuerteventura population in the Canary Islands, provoked its inclusion in the Spanish Catalogue of Endangered Species under the category 'in danger of extinction' (González-Lorenzo *et al.*, 2015). In contrast, the population of the remote Selvagens islands is diagnosed as healthy

(González-Lorenzo *et al.*, 2015; Hernández-Dorta, 1992).

The use of genome reduced-representation methodological approaches (RAD) allows to obtain thousands of markers dispersed trough the genome in a cost-effective way (Baird *at al.*, 2008; Etter *at al.*, 2011; Kess *at al.*, 2016; Peterson *at al.*, 2012).

In this study we tested a non-invasive, swab-based DNA sampling method that can be used in situ during the sampling and monitoring of *Patella candei candei* endangered populations. These DNA samples, obtained from the limpet mucus, must provide DNA that generates results, in terms of the number of reads and ddRAD loci, similar to those obtained from standard samples.

MATERIALS AND METHODS

The *Patella candei candei* samples were obtained from the abundant and healthy population of the Selvagens Islands (N = 30). From the Fuerteventura critically endangered population they were obtained samples from archived collections (1995) and from the seizure of illegal catches (N = 3). Additionally, they were collected by non-invasive mucosal swabs sampling (N = 2, Jandia, Fuerteventura). Samples were immersed in tubes with RNALater (Thermofisher) for their conservation, being temporarily 4°C and later at -20°C. Total DNA was isolated with the E.Z.N.A.® Mollusc DNA kit (Omega Bio-Tek). DNA was then ultrafiltrated with Amicon-M100K (Millipore). Quantity and integrity of purified DNA was estimated with a Biophotometer Eppendorf D30, Qubit fluorometer and visualized electrophoretically in 1% agarose gel.

The genomic libraries were elaborated by ddRAD methodology using the isolated DNA from the diverse sample typology and sequenced in Miseq (Illumina) (Baird *et al.*, 2008; Etter *et al.*, 2011; Kess *et al.*, 2016; Peterson *et al.*, 2012). SNPs identification and analysis were performed with Stacks and dDocent pipelines (Catchen *at al.*, 2013; Puritz *at al.,*2014).

RESULTS AND DISCUSSION

The integrity of the DNA isolated from the limpet samples *Patella candei candei* from the population of Fuerteventura was similar for all the samples, without important differences between the three types of analysed samples: archived samples, recent seized samples and non-invasive samples. Along with high molecular weight DNA (HMW DNA), DNA with

a certain degree of degradation is observed (Fig. 1A). This electrophoretic profile was modified after ultrafiltration in the Amicon-100K device, being removed most of the degradation smear, highlighting the HMW DNA fraction. After ultrafiltration, a lower amount of HMW DNA is observed in the samples from non-invasive sampling compared to the rest of samples (Fig. 1B).

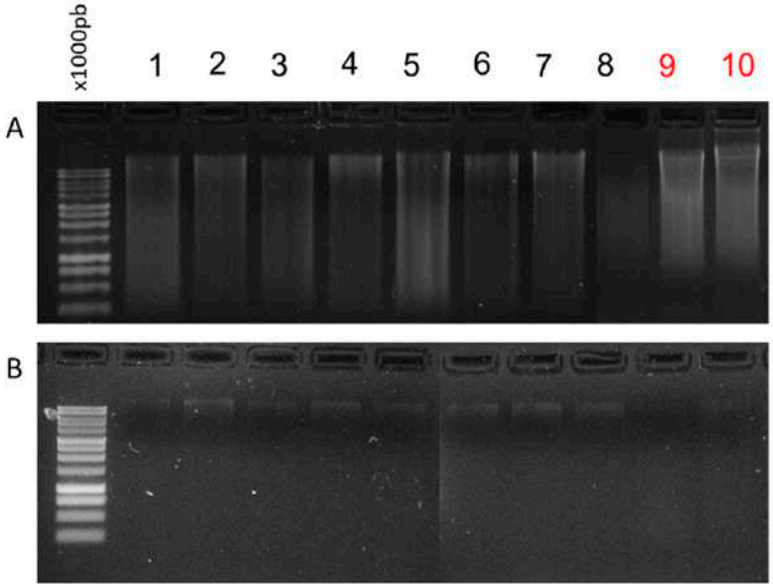

Figure 1.- Electrophoresis (1% agarose) of (A) total DNA isolated from diverse samples (N=10) of the limpet, *P. candei candei*, sampled in Fuerteventura Island. The total DNA after ultrafiltration (B) in Amicon-M100K (Millipore). The samples 9 and 10 were obtained by non-invasive mucosal swabs sampling.

When compared with the conspecific samples, captured in the Selvagens islands population, the Fuerteventura sample showed a low DNA concentration. The total DNA isolated from Selvagens population (Mean=30.57 ng; SD=65.29) was clearly higher than values from Fuerteventura sample (mean=2.17 ng; SD=1.89) (Figure 2A). Consequently, the amount of loaded DNA for the ddRAD libraries construction (Mean = 471.54 ng; SD = 302.75) was clearly higher in the case of the Selvagens libraries (Mean = 39.02ng; SD = 33.96), with respect to those of Fuerteventura (Figure 2B).

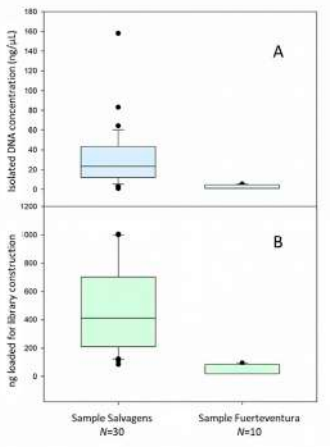

Figure 2.- DNA concentration (ng/uL) of isolated DNA (40µL) (A) and quantity finally loaded for ddRAD library construction (B).

The final results obtained for each of the libraries, samples from Fuerteventura and Selvagens, did not reflect the large initial discrepancy in the amount of initial DNA. Even so, the number of reads obtained from the individual libraries of the Selvagens sample was slightly higher than that of the Fuerteventura libraries (Mean= 162369,6; SD= 59109,05) (Figure 3A). Similarly, when depth coverage was considered, Selvagens libraries showed higher values (Mean= 9,48; SD= 1,59) than those from Fuerteventura (Mean=11,11; SD=2,44) (Figure 3B). In both cases the mean number of reads and depth coverage was slightly lower than that of other libraries built in the same experiment from *Patella crenata* samples.

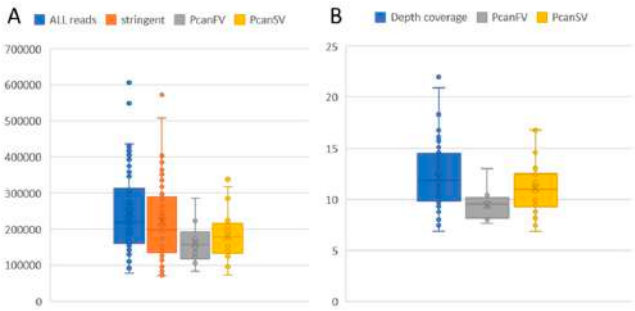

Figure 3.- Results from the NGS sequencing of the ddRAD libraries. The total number of obtained reads (A) and the depth coverage for *P. candei* candei libraries sampled in Selvagens (PcanSV) and Fuerteventura Island (Pcan FV) (B).

In contrast, regardless of the pipeline used (Stacksand dDocent), the number of loci detected was slightly higher in the case of the Fuerteventura libraries (Mean=1300 -Stacks-, 3481 -dDocent-) than those of Selvagens (Mean=1074 -Stacks-, 3199 -dDocent-).

The initial concentration of DNA to be incorporated in the ddRAD library construction methodology is an important factor in generating the final number of reads. However, the variation of the PCR cycles used to perform the barcodes incorporation, makes it possible to slightly harmonize the number of fragments available in each library. Thus, a difference of about an order of magnitude in the initial DNA concentration can generate a more approximate number of reads, if increasing by a number of 2 to 3 PCR cycles for that library with less DNA. However, the number of cycles to be increased must be kept to the minimum necessary to avoid the generation of excessive PCR replicates.

The use of relevant, but non-ideally collected and/or preserved DNA samples, including those from non-invasive mucosal swabs sampling still provide acceptable genotyping results. Particularly, for the monitoring of critically endangered limpet species, *Patella candei candei*, the generated data allowed the correct use of SNPs identification pipelines and population genetic structure analysis (Figure 4). The achieved results are relevant for the design and monitoring of conservation strategies for the endangered majorera limpet in Fuerteventura Island.

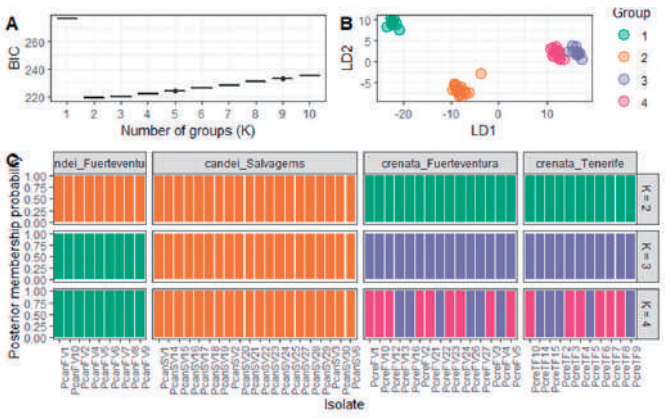

Figure 4.- Examples of analysis performed with SNP data attained from ddRAD libraries; including (A) estimation of the number of groups for structure type analysis; (B) PCA analysis and (C) Structure estimation of the membership probability for the individual libraries.

CONCLUSIONS

- Archived samples, non-ideally collected and non-invasive mucosal swabs sampling are adequate sources of DNA for genome reduced-representation methodological approaches (RAD) in limpets.
- Number of reads, depth coverage and loci number are similar in the diverse sample typology tested.

ACKNOWLEDGEMENTS

This research was performed within the SNPs4candei Project (2020-2021) funded by the Ministry for the Ecological Transition and the Demographic Challenge (MITECO), via the Biodiversity Foundation. Diverse bioinformatics workflows and jobs were executed at the Centre of Supercomputation of Galicia (CESGA). We are grateful to Ricardo Araújo (Natural History Museum of Funchal, Funchal, Madeira). Material from Cape Verde Islands was provided by BIOVAL (MAC/3/C216) and BIOTECMAR (MAC/3/C156) both projects framed on the EU ERDF Programme INTERREG (MAC) 2007–2013.

REFERENCES

Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A. and Johnson E.A. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE,* 3: 1–7. doi:10.1371/journal.pone.0003376

Balázs G., Vörös J., Lewarne B. and Herczeg G. 2020. A new non-invasive in situ underwater DNA sampling method for estimating genetic diversity. *Evolutionary Ecology*, 34(4): 633–644. doi:10.1007/s10682-020-10053-1

Catchen J., Hohenlohe P.A., Bassham S., Amores A. and Cresko W.A. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22(11): 3124–3140. doi:10.1111/mec.12354

Domingues R.R., Garrone-Neto D., Hilsdorf A.W.S. and Gadig O.B.F. 2019. Use of mucus as a non-invasive sampling method for DNA barcoding of stingrays and skates (batoid elasmobranchs). *Journal of Fish Biology*, 94(3): 512–516. doi:10.1111/jfb.13919

Etter P.D., Bassham S., Hohenlohe P.A., Johnson E.A. and Cresko W.A. 2011. SNP Discovery and Genotyping for Evolutionary Genetics Using RAD Sequencing. *Methods in Molecular Biology*, 772: 157–178. doi:10.1007/978-1-61779-228-1_9

González-Lorenzo G., Mesa Hernández E., Pérez-Dionis G., Brito Hernández A., Galván Santos B. and Barquín Diez, J. 2015. Ineffective conservation threatens *Patella candei*, an endangered limpet endemic to the Macaronesian islands. *Biological Conservation*,

192: 428–435. doi: https://doi.org/10.1016/j.biocon.2015.11.005

Hernández-Dorta F.J. 1992. Revisión del género *Patella* Linné, 1758 (Mollusca , Gastropoda) en las Islas Canarias. *Vierae*a, 21: 109–135.

Kess T., Gross J., Harper F. and Boulding, E. G. (2016). Low-cost ddRAD method of SNP discovery and genotyping applied to the periwinkle *Littorina saxatilis*. *Journal of Molluscan Studies,* 82: 104–109. doi:10.1093/mollus/eyv042

Kim K.S. and Roe K.J. 2021. Genome-wide SNPs redefines species boundaries and conservation units in the freshwater mussel genus *Cyprogenia* of North America. *Scientific Reports*, 11(1): 10752. doi:10.1038/s41598-021-90325-0

Morinha F., Travassos P., Carvalho D., Magalhães P., Cabral J.A. and Bastos, E. 2014. DNA sampling from body swabs of terrestrial slugs (Gastropoda: Pulmonata): a simple and non-invasive method for molecular genetics approaches. *Journal of Molluscan Studies*, 80(1): 99–101. doi:10.1093/mollus/eyt045

Peterson B.K., Weber J.N., Kay E.H., Fisher, H.S. and Hoekstra H.E. 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7. doi:10.1371/journal.pone.0037135

Puritz J.B., Hollenbeck C.M. and Gold J.R. 2014. dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. PeerJ, 2, e431. doi:10.7717/peerj.431

Efficacy of a non-invasive sampling method for genomic analysis of the critically endangered limpet *Patella candei candei* d'Orbigny, 1840

Quinteiro, J. [1]; González-Lorenzo, G. [2]; Hernández-Reyes, D. [3]; Quinteiro, L. [1]; Herrera, R. [4]; Martínez, J. [4]; González, J.A. [3]; Rey-Méndez, M. [1]; González-Henríquez, N. [3]

1 Molecular Systematics Laboratory, Department of Biochemistry and Molecular Biology, Aquaculture Institute; University of Santiago de Compostela, A Coruña, Galicia, Spain.
2 Canary Islands Oceanographic Center, Spanish Institute of Oceanography (IEO, CSIC), Tenerife, Spain.
3 Departament of Biology, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain.
4 Directorate General for the Fight against Climate Change and the Environment. Council for the Ecological Transition, Fight against Climate Change and Territorial Planning. Canary Islands Government

Abstract

In this study we tested a non-invasive, swab-based DNA sampling method that can be used in situ for the monitoring of *Patella candei candei* endangered populations. These DNA samples, obtained from the limpet mucus, must provide DNA for the implementation of genome reduced-representation methodological approaches (ddRAD) to obtain thousands of markers dispersed trough the genome in a cost-effective way. The use of relevant, but non-ideally collected and/or preserved DNA samples, including those from non-invasive mucosal swabs sampling provided acceptable genotyping results. The generated data allowed the correct use of SNPs identification pipelines and population genetic structure analysis. The achieved results are relevant for the design and monitoring of conservation strategies for the endangered majorera limpet in Fuerteventura Island.

Introduction

The development and testing of non-invasive sampling methods is crucial for the collection of samples for molecular taxonomy studies and analysis of the population genetic structure of critically endangered organisms, avoiding a negative impact on small and / or endangered populations.

Non-invasive samples have been proposed and evaluated for genetic analysis in order to minimize the impact on the studied populations using different DNA sources (Balázs, Vörös, Lewarne, & Herczeg, 2020; Domingues, Barrone-Neto, Hilsdorf, & Gadig, 2019; Morinha et al., 2014). Of particular interest are those works that have successfully demonstrated the feasibility of analysis from fish and mollusc mucus (Domingues et al., 2019; Kim & Ree, 2021; Morinha et al., 2014). Although the non-invasive method that has the least effect on the environment and organisms is based on environmental DNA (eDNA), it does not allow obtaining individualized genetic information.

The limpet P. candei candei d'Orbigny, 1840 is endemic to Fuerteventura and Selvagens Islands. The critical state of the "majorera limpet", from the unique Fuerteventura population in the Canary Islands, provoked its inclusion in the Spanish Catalogue of Endangered Species under the category 'in danger of extinction' (González-Lorenzo et al., 2015). In contrast, the population of the remote Selvagens Islands is diagnosed as healthy (González-Lorenzo et al., 2015; Hernández-Dorta, 1992).

The use of genome reduced-representation methodological approaches (RAD) allows to obtain thousands of markers dispersed trough the genome in a cost-effective way (Baird et al., 2008; Etter, Bassham, Hohenlohe, Johnson, & Cresko, 2011; Kess, Gross, Harper, & Boulding, 2016; Peterson, Weber, Kay, Fisher, & Hoekstra, 2012).

In this study we tested a non-invasive, swab-based DNA sampling method that can be used in situ during the sampling and monitoring of *Patella candei candei* endangered populations. These DNA samples, obtained from the limpet mucus, must provide DNA that generates results, in terms of the number of reads and ddRAD loci, similar to those obtained from standard samples.

Materials and Methods

The *Patella candei candei* samples were obtained from the abundant and healthy population of the Salvagens Islands (N = 30). From the Fuerteventura critically endangered population they were obtained samples from archived collections (1995) and from the seizure of illegal catches (N = 3). Additionally, they were collected by non-invasive mucosal swabs sampling (N = 2, Jandía, Fuerteventura). Samples were immersed in tubes with RNALater (Thermofisher) for their conservation, being temporarily 4°C and later at -20°C. Total DNA was isolated with the E.Z.N.A.® Mollusc DNA kit (Omega Bio-Tek). DNA was then ultrafiltrated with Amicon-M100K (Millipore). Quantity and integrity of purified DNA was estimated with a Biophotometer Eppendorf D30, Qubit fluorometer and visualized electrophoretically in 1% agarose gel.

The genomic libraries were elaborated by ddRAD methodology using the isolated DNA from the diverse sample typology and sequenced in Miseq (Illumina) (Baird et al., 2008; Etter et al., 2011; Kess et al., 2016; Peterson et al., 2012). SNPs identification and analysis were performed with Stacks and dDocent pipelines (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013; Puritz, Hollenbeck, & Gold, 2014).

Figure 1. Electrophoresis (1% agarose) of (A) total DNA isolated from diverse samples (N=10) of the limpet, Patella candei candei, sampled in Fuerteventura Island. The total DNA after ultrafiltration (B) in Amicon-M100K (Millipore). The samples 9 and 10 were obtained by non-invasive mucosal swabs sampling.

Results and Discussion

The integrity of the DNA isolated from the limpet samples *Patella candei candei* from the population of Fuerteventura was similar for all the samples, without important differences between the three types of analysed samples: archived samples, recent seized samples and non- invasive samples. Along with high molecular weight DNA (HMW DNA), DNA with a certain degree of degradation is observed (Figure 1.A). The electrophoretic profile was modified after ultrafiltration in the Amicon-100K device, being removed most of the degradation smear, highlighting the HMW DNA fraction. After ultrafiltration, a lower amount of HMW DNA is observed in the samples from non-invasive sampling compared to the rest of samples (Figure 1.B).

When compared with the conspecific samples, captured in the Salvagens islands population, the Fuerteventura sample showed a low DNA concentration. The total DNA isolated from Salvagens population (Mean=30.57 ng; SD=65.29) was clearly higher than values from Fuerteventura sample (mean=2.17 ng; SD=1.89) (Figure 2.A). Consequently, the amount of loaded DNA for the ddRAD libraries construction (Mean = 471.54 ng; SD = 302.75) was clearly higher in the case of the Salvagens libraries (Mean = 39.02ng; SD = 33.96), with respect to those of Fuerteventura (Figure 2.B).

The final results obtained for each of the libraries, samples from Fuerteventura and Salvagens, did not reflect the large initial discrepancy in the amount of initial DNA. Even so, the number of reads obtained from the individual libraries of the Salvagens sample was slightly higher than that of the Fuerteventura libraries (Mean= 162369.6; SD= 59109.05) (Figure 3.A). Similarly, when depth coverage was considered, Salvagens libraries showed higher values (Mean= 9.48; SD= 1.59) than those from Fuerteventura (Mean=11.11; SD=2.44) (Figure 3.B). In both cases the mean number of reads and depth coverage was slightly lower than that of other libraries built in the same experiment from *Patella crenata* samples.

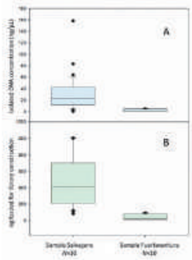

Figure 2. DNA concentration (ng/uL) of isolated DNA (40µL) (A) and quantity finally loaded for ddRAD library construction (B).

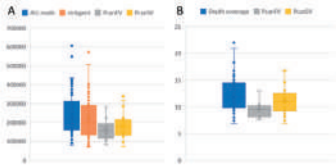

Figure 3. Results from the NGS sequencing of the ddRAD libraries. The total number of obtained reads (A) and the depth coverage for P. candei candei libraries sampled in Selvagens (PcanSV) and Fuerteventura Island (Pcan FV) (B).

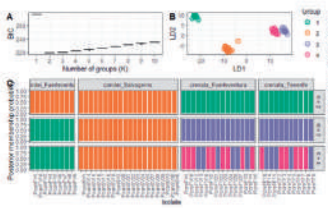

Figure 4. Examples of analysis performed with SNP data attained from ddRAD libraries; including (A) estimation of the number of groups for structure type analysis; (B) PCA analysis and (C) Structure estimation of the membership probability for the individual libraries.

In contrast, regardless of the pipeline used (Stacksand dDocent), the number of loci detected was slightly higher in the case of the Fuerteventura libraries (Mean=1300 -Stacks-, 3481 -dDocent-) than those of Selvagens (Mean=1074 -Stacks-, 3199 -dDocent-).

The initial concentration of DNA to be incorporated in the ddRAD library construction methodology is an important factor in generating the final number of reads. However, the variation of the PCR cycles used to perform the barcodes incorporation, makes it possible to slightly harmonize the number of fragments available in each library. Thus, a difference of about an order of magnitude in the initial DNA concentration can generate a more approximate number of reads, if increasing by a number of 2 to 3 PCR cycles for that library with less DNA. However, the number of cycles to be increased must be kept to the minimum necessary to avoid the generation of excessive PCR replicates.

The use of relevant, but non-ideally collected and/or preserved DNA samples, including those from non-invasive mucosal swabs sampling still provide acceptable genotyping results. Particularly, for the monitoring of critically endangered limpet species, *Patella candei candei* , the generated data allowed the correct use of SNPs identification pipelines and population genetic structure analysis (Figure 4). The achieved results are relevant for the design and monitoring of conservation strategies for the endangered majorera limpet in Fuerteventura Island.

Conclusions

- Archived samples, non-ideally collected and non-invasive mucosal swabs sampling are adequate sources of DNA for genome reduced-representation methodological approaches (RAD) in limpets.
- Number of reads, depth coverage and loci number are similar in the diverse sample typology tested.

Acknowledgements

This research was performed within the SNPsKandei Project (2020-2021) funded by the Ministry for the Ecological Transition and the Demographic Challenge (MITECO), via the Biodiversity Foundation. Diverse bioinformatics workflows and jobs were executed at the Centre of Supercomputation of Galicia (CESGA). We are grateful to Ricardo Araújo (Natural History Museum of Funchal, Funchal, Madeira). Material from Cape Verde Islands was provided by [illegible] (MAC/1/C216) and BIOTECMAR (MAC/1/C116) both projects framed on the EU ERDF Programme INTERREG (MAC) 2007–2013.

Referencias

[illegible]