



**UNIVERSIDADE FEDERAL DO PIAUÍ
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA**

MICHELLI FERREIRA DOS SANTOS

**SEQUENCIAMENTO DO DNA, DESENVOLVIMENTO DE MARCADORES
MOLECULARES E MONTAGEM DO GENOMA MITOCONDRIAL DA LAGOSTA
Panulirus echinatus (Smith, 1869)**

Teresina

2016

MICHELLI FERREIRA DOS SANTOS

Sequenciamento do DNA, desenvolvimento de marcadores moleculares e montagem do genoma mitocondrial da lagosta *Panulirus echinatus* (Smith, 1869)

Tese apresentada ao Programa de Pós-Graduação da Rede Nordeste de Biotecnologia – Embrapa Meio-Norte - Universidade Federal do Piauí, como requisito parcial para a obtenção do grau de doutor em Biotecnologia.

Área de Concentração: Biotecnologia em Agropecuária

Orientador:

Prof. Dr. Fábio Mendonça Diniz

Teresina

2016

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Comunitária Jornalista Carlos Castello Branco
Serviço de Processamento Técnico

S237 Santos, Michelli Ferreira dos.
Sequenciamento do DNA, desenvolvimento de marcadores
moleculares e montagem do genoma mitocondrial da lagosta
Pamulirus echinatus (Smith, 1869) / Michelli Ferreira dos Santos. --
2016.
107 f.

Tese (Doutorado) – Biotecnologia. Rede Nordeste de
Biotecnologia (RENORBIO), Universidade Federal do Piauí,
Teresina, 2016.

Orientação: Prof. Dr. Fábio Mendonça Diniz.

1. Diversidade genética. 2. Microssatélites. 3. *Pamulirus echinatus*
- DNA. 4. Miseq. I. Título.

CDD 572.86



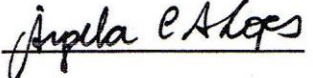


FOLHA DE APROVAÇÃO – DEFESA DE TESE

ALUNO: MICHELLI FERREIRA DOS SANTOS

TÍTULO DO PROJETO: “Sequenciamento do DNA, desenvolvimento de marcadores moleculares e montagem do genoma mitocondrial da lagosta *Panulirus echinatus* (Smith, 1869).”

PROFESSOR ORIENTADOR: Prof. Dr. Fábio Mendonça Diniz

BANCA EXAMINADORA:

	CONCEITO	ASSINATURA
Prof. Dr. Fábio Mendonça Diniz - EMBRAPA (Presidente)	<u>SATISFATORIO</u>	
Prof. Dr. Adalberto Socorro da Silva, UFPI (Examinador)	<u>SATISFATORIO</u>	
Profa. Dra. Ângela Celis de Almeida Lopes - UFPI (Examinadora)	<u>Satisfatório</u>	
Prof. Dr. Vladimir Costa Silva – UFPI (Examinador)	<u>Satisfatório</u>	
Prof. Dr. Fábio Barros Britto - UFPI (Examinador)	<u>SATISFATORIO</u>	

DATA DA AVALIAÇÃO: 27 de junho de 2016.

HORÁRIO: 08h30min

LOCAL: Auditório do NUPCelt/CCA/UFPI.

Dedicatória

À minha mãe Maria da Conceição Ferreira Lima

AGRADECIMENTOS

Agradeço a todos que me apoiaram nessa etapa. Minha gratidão especial:

À Universidade Federal do Piauí, pela formação;

Ao Programa de Pós-Graduação da Rede Nordeste de Biotecnologia – RENORBIO, pela oportunidade de cursar o doutorado;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo período de concessão da bolsa de estudos;

Aos Professores do Programa de Pós-Graduação da Rede Nordeste de Biotecnologia e da Universidade Federal do Piauí, pelos ensinamentos transmitidos;

Ao Dr. Fábio Mendonça Diniz, pela competência, responsabilidade e comprometimento na orientação desse trabalho;

À Empresa Brasileira de Pesquisa Agropecuária, Embrapa Meio-Norte, em especial ao Laboratório de Biologia Molecular e Biotecnologia pela oportunidade da execução do trabalho;

À Dalhousie University, pela parceria na realização do sequenciamento do genoma da lagosta *Panulirus echinatus*;

Ao Professor Dr. Francisco Prosdocimi, da Universidade Federal do Rio de Janeiro, pela contribuição na montagem do DNA mitocondrial da espécie;

Aos amigos do Laboratório de Biologia Molecular e Biotecnologia da Embrapa Meio-Norte, pelas experiências vividas e compartilhadas que só contribuíram para o meu crescimento;

Aos meus familiares e amigos pela força e incentivo;

À minha Mãe Maria da Conceição Ferreira Lima e as minhas irmãs Gisele Ferreira dos Santos e Isália Ferreira dos Santos por sempre acreditarem na minha capacidade.

RESUMO

SANTOS, M. F. **Sequenciamento do DNA, desenvolvimento de marcadores moleculares e montagem do genoma mitocondrial da lagosta *Panulirus echinatus* (Smith, 1869)**. 2016. 107f. Tese (Doutorado) – RENORBIO, Universidade Federal do Piauí, Teresina, 2016.

A lagosta espinhosa *Panulirus echinatus* (Smith, 1869) é amplamente distribuída nos Atlânticos Central, Ocidental e Oriental. Representa um dos mais importantes recursos pesqueiros do litoral das regiões Norte e Nordeste do Brasil e são intensamente exploradas por possuírem alto valor comercial, com isso, estratégias para o manejo sustentável devem compreender também estudos de diversidade genética da espécie. Nosso trabalho teve como objetivo o sequenciamento do genoma da lagosta, utilizando a tecnologia de sequenciamento de nova geração, visando a descrição de marcadores microssatélites, além da montagem do genoma mitocondrial da espécie. O DNA alvo utilizado para o sequenciamento foi extraído do tecido muscular da lagosta. Um nanograma desse DNA foi utilizado para obtenção dos *reads* que foram submetidos a posterior sequenciamento. Para tanto, uma biblioteca de extremidades emparelhadas da Illumina foi criada conforme especificações do kit de preparação da Illumina Nextera (Illumina Inc.). O sequenciamento foi conduzido usando o sequenciador Miseq Benchtop (Illumina Inc.). As sequências contíguas (*contigs*) foram criadas a partir dos dados das sequências emparelhadas pelas extremidades obtidas por meio do CLC Genomics Workbench 7.0.4 (Qiagen). Os *motifs* de microssatélites foram identificados e localizados utilizando o software MSDB, foram realizadas buscas por regiões microssatélites e 27 pares de locos SSRs tri e tetranucleotídicos foram selecionados. O genoma mitocondrial foi completamente montado por interações no *software* MITObim. O resultado obtido, após o sequenciamento foi uma plataforma com 7196 sequências contíguas, onde 2959 SSR foram identificados, entre seis classes de marcadores SSRs. Dos 27 *primers* previamente testados, 17 apresentarem polimorfismo e 3 foram monomórficos. O genoma mitocondrial completo da lagosta é de 15.808 pb, contendo 37 genes (13 genes codificadores de proteínas, 22 genes para RNA transportador e 2 genes para RNA ribossomais), a ordem dos genes é exatamente a mesma observada para outros mitogenomas do gênero *Panulirus*. Os dados fornecidos podem ajudar a esclarecer as relações evolutivas dentro da família Palinuridae, e ser usados para estudos de genética de populações e identificação de organismos em crustáceos.

Palavras-chave: Diversidade genética, microssatélites, Miseq, sequências.

ABSTRACT

SANTOS, M. F. **DNA sequencing, development of molecular markers and assembly of mitochondrial genome *Panulirus echinatus* lobster (Smith, 1869)**. 2016. 107f. Thesis (Doctor) – RENORBIO, Universidade Federal do Piauí, Teresina, 2016.

The spiny *Panulirus echinatus* lobster (Smith, 1869) is widely distributed in the Central Atlantic, Western and Eastern. It is one of the most important fishing grounds of coastal resources in North and Northeast Brazil and are heavily exploited by having high commercial value, therefore, strategies for sustainable management must also understand genetic diversity studies of the species. Our study aimed to sequence the genome of the lobster, using the new generation of sequencing technology, aiming at the description of microsatellite markers, in addition to the assembly of the mitochondrial genome of the species. The target DNA used for sequencing was extracted from muscle lobster. One nanogram of this DNA was used to obtain the reads that have undergone subsequent sequencing. To do so, a paired end library Illumina was created as preparation kit specifications of Illumina Nextera (Illumina Inc.). The sequencing was carried out using the Sequencer Miseq Benchtop (Illumina Inc.). Contiguous sequences (contigs) were created from the data of the sequences matched by the edges obtained through the CLC Genomics Workbench 7.0.4 (Qiagen). The microsatellite motifs were identified and located using the MSDB software, searches for microsatellite regions and 27 pairs of loci SSRs and tri tetranucleotídicos were selected were performed. The mitochondrial genome has been completely mounted interactions in MITObim software. The result obtained after the sequencing was a platform with 7196 contigs where SSR 2959 were identified among six classes of SSRs markers. Of the 27 primers previously tested, 17 polymorphics and 3 were monomorphic. The complete mitochondrial genome of the lobster is 15,808 bp, containing 37 genes (13 protein-coding genes, 22 genes for tRNA and 2 genes for ribosomal RNA), the order of genes is exactly the same observed for other mitogenomas of *Panulirus* genre. The data provided may help clarify the evolutionary relationships within the Palinuridae family, and be used for genetic studies of populations and identification of organisms in crustaceans.

Keywords: Genetic diversity, microsatellites, Miseq, sequences.

LISTA DE FIGURAS

CAPÍTULO 1

- Figura 1 - Distribuição geográfica da espécie *Panulirus echinatus*. 18
- Figura 2 - Desenho esquemático da lagosta *Panulirus echinatus*, destacando as características morfológicas (adaptado de HOLTHUIS, 1991). 20

CAPÍTULO 2

- Figura 1 -Contagem total e número de repetições de diferentes tipos de SSRs no genoma da lagosta *Panulirus echinatus*. 44
- Figura 2 - Repetições tetranucleotídicas mais frequentes em *P. echinatus*. 45

CAPÍTULO 3

- Figura 1 -Identificação dos locais amostrados da espécie *P. echinatus*. 53
- Figura 2 - Análise de coordenadas principais da variação de microssatélites em *P. echinatus*. 60
- Figura 3 - Estrutura da população obtida pelo software STRUCTURE para K = 2 a K = 6, permutado por CLUMPP e visualizados por DISTRUCT. Cada indivíduo é representado por uma barra vertical em segmentos coloridos K, o comprimento de cada cor é proporcional ao coeficiente de adesão estimada. Linhas pretas separam indivíduos de diferentes populações pré-definidos. 62

CAPÍTULO 4

- Figura 1 - Mapa do genoma mitocondrial da lagosta *Panulirus echinatus*. 73
- Figura 2 - Genomas mitocondriais completos de sete lagostas do gênero *Panulirus*, incluindo *Panulirus echinatus*. 75

LISTA DE TABELAS

CAPÍTULO 2

- Tabela 1 - Contagem, comprimento total, comprimento médio, frequência, densidade e percentual de seis tipos de microssatélites perfeitos na sequência do genoma da lagosta *P. echinatus*, Teresina, PI, 2016.....40
- Tabela 2 - *Motif*, contagem total, comprimento total, comprimento médio, frequência e densidade dos *motifs* de maior frequência em *P. echinatus*, Teresina, PI, 2016.43
- Tabela 3 - Características dos tipos de microssatélites em *P.echinatus*, Teresina, PI, 2016....45

CAPÍTULO 3

- Tabela 1 - Sequência dos iniciadores desenhados, estrutura de repetição, temperatura de anelamento (Ta), número de alelos (A), tamanho do alelo (bp) e número de acesso no GenBank, Teresina, PI, 2016..... 58
- Tabela 2 - Estatísticas de diversidade genética para lagosta espinhosa *Panulirus echinatus*, Teresina, PI, 2016..... 60
- Tabela 3 - Variabilidade de 17 locos de microssatélites e F-estatísticas em populações *Panulirus echinatus*, Teresina, PI, 2016..... 61
- Tabela 4 - Proporção de adesão das populações pré-definidos em clusters inferidos, Teresina, PI, 2016..... 63

CAPÍTULO 4

- Tabela 1 - Características do genoma mitocondrial da lagosta *P. echinatus*, Teresina, PI, 2016. 76

LISTA DE ABREVIATURAS E SÍMBOLOS

AFLP	Polimorfismo de comprimento de fragmento amplificado
ATP6	Trifosfato de adenosina sintetase subunidades 6
ATP8	Trifosfato de adenosina sintetase subunidades 8
COX1	<i>Citocromo c oxidase</i> subunidade I
COX2	<i>Citocromo c oxidase</i> subunidade II
COX3	<i>Citocromo c oxidase</i> subunidade III
CYTB	Citocromo b
DNAmt	DNA mitochondrial
F _{ST}	Índice de fixação
HWE	Equilíbrio de Hardy-Weinberg
NAD1	Desidrogenase subunidade 1
NAD2	Desidrogenase subunidade 2
NAD3	Desidrogenase subunidade 3
NAD4	Desidrogenase subunidade 4
NAD4-L	Desidrogenase subunidade 4L
NAD5	Desidrogenase subunidade 5
NAD6	Desidrogenase subunidade 6
PCR	Reação em Cadeia de Polimerase
RAPD	Polimorfismo de DNA amplificado ao acaso
RFLP	Polimorfismo no comprimento de fragmentos de restrição
SDS	Dodecil-sulfato de sódio
SNPs	Polimorfismo de nucleotídeo único
SSR	Sequência simples repetida
°C	Graus Celsius
mM	Milimolar
μL	Microlitros

SUMÁRIO

Introdução Geral	14
Capítulo 1	17
Revisão de Literatura.....	17
1.1 Distribuição geográfica e habitat	18
1.2 Morfologia e reprodução da lagosta <i>Panulirus echinatus</i>	19
1.3 Importância econômica.....	21
1.4 O papel da genética na gestão das pescas	22
1.5 Marcadores Moleculares na família Palinuridae.....	23
1.6 Referências.....	30
Capítulo 2	34
Abundância e análise de marcadores microssatélites em lagostas <i>Panulirus echinatus</i> (Smith, 1869) (Crustacea: Decapoda: Palinuridae).....	34
2.1 Resumo	35
2.2 Abstract.....	36
2.3 Introdução	37
2.4 Material e Métodos	38
2.5 Resultados e Discussão	39
2.6 Conclusão.....	46
2.7 Referências.....	47
Capítulo 3	49
Desenvolvimento de marcadores de sequências simples repetidas no genoma da lagosta <i>Panulirus echinatus</i> por meio do sequenciamento de nova geração e uma análise preliminar da diversidade genética da espécie	49
3.1 Resumo	50
3.2 Abstract.....	51
3.3 Introdução	52
3.4 Material e Métodos	53
3.5 Resultados e Discussão	56
3.6 Conclusão.....	63

3.7 Referências.....	64
Capítulo 4	67
Genoma mitocondrial completo da lagosta <i>Panulirus echinatus</i> (Smith, 1869) (Crustacea: Decapoda: Palinuridae).....	67
4.1 Resumo	68
4.2 Abstract.....	69
4.3 Introdução	70
4.4 Material e Métodos	71
4.5 Resultados e Discussão	73
4.6 Conclusão.....	78
4.7 Referências.....	79
APÊNDICES	82
APÊNDICE A.....	83
APÊNDICE B.....	86
APÊNDICE C.....	91
APÊNDICE D.....	108

Introdução Geral

A lagosta espinhosa *Panulirus echinatus* (Smith, 1869) é amplamente distribuída nos Atlânticos Central, Ocidental e Oriental. No Brasil está presente em toda a costa do litoral e em algumas ilhas oceânicas (HOLTHUIS, 1991). Atualmente a lagosta espinhosa representa um dos mais importantes recursos pesqueiros do litoral das regiões Norte e Nordeste do Brasil, sendo a única lagosta do gênero *Panulirus* comercialmente capturada, que não possui uma pesca legalizada (GÓES; LINS-OLIVEIRA; JE, 2009).

Apesar de não ser a espécie mais capturada na costa do litoral brasileiro, a lagosta espinhosa já tem representatividade nas capturas e já se encontra inserida no mercado de exportação (BRASIL, 2008). As espécies de lagostas espinhosas são intensamente exploradas por possuírem alto valor comercial, com isso, estratégias para o manejo sustentável devem compreender também estudos de diversidade genética da espécie (DINIZ et al., 2005).

A técnica inovadora de sequenciamento de nova geração *Next Generation Sequencing*, aliada a recursos computacionais da bioinformática vem sendo aplicada para gerar dados de sequências, na identificação de *motifs* de microssatélites, principalmente de espécies pouco estudadas ou sem nenhuma informação genética, mas de grande importância econômica (ABDELKRIM et al., 2009). Essa técnica foi utilizada nesse estudo para sequenciar o genoma da lagosta *P.echinatus* e com isso gerar informações que serão utilizadas em estudos de identificação de espécies, desenvolvimento de marcadores e estudos populacionais.

A diversidade genética das espécies vem sendo observada por meio da utilização de ferramentas moleculares. Os marcadores microssatélites ou SSR (*Simple Sequence Repeats*) são os mais utilizados em estudo de diversidade genética e populacional, possibilitando a caracterização genética e subsidiando estratégias de conservação. A identificação de marcadores específicos para a *P.echinatus* possibilitará a caracterização da estrutura populacional dessa lagosta de ocorrência em toda a costa do litoral brasileiro e também ajudará a entender a densidade populacional, principalmente pelo fato dessa espécie sofrer pressão das correntes marítimas. Além disso, a informação genética poderá ser incluída em planos de gestão de conservação e manejo da lagosta.

A técnica de sequenciamento também tem possibilitado a identificação de genomas mitocondriais de várias espécies. O DNA mitocondrial (DNAm) é muito utilizado como marcador molecular em estudos que visam inferir a relação evolutiva entre as espécies, pois

têm mostrado que as divergências nas sequências acumulam mais rapidamente do que em DNA nuclear. Isso ocorre devido a taxa de mutação ser mais rápida e também pela falta de recombinação durante a replicação (BROWN, 1985; WILSON et al., 1985). Devido a essa variabilidade no seu genoma, o DNAm_t é considerado um bom marcador para a detecção de possível diferenciação genética (AVISE, 1994, MORITZ et al., 1987, BIRKY et al., 1989). Nesse estudo a montagem completa do DNAm_t da lagosta possibilitará a utilização do genoma como um potencial marcador, podendo ser utilizado por exemplo em estudos de filogenia.

Desta forma, o objetivo do nosso trabalho foi o sequenciamento do genoma da lagosta *P. echinatus*, e a partir dele o desenvolvimento de marcadores microssatélites, inexistentes até o momento para espécie, visando utiliza-los em estudos de diversidade genética e estrutura populacional, além da montagem do genoma mitocondrial da espécie, que irá ajudar a esclarecer as relações evolutivas dentro da família Palinuridae, e ser usados para estudos de genética de populações e identificação de organismos em crustáceos.

Referências

ABDELKRIM, J. et al. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. **BioTechniques**, v. 46, n. 3, p. 185–192, 2009.

AVISE, J. C. *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York, 1994.

BIRKY, C. W.; FUERST, P.; MARUYAMA, T. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. **Genetics**, 121(3): 613-627, 1989.

BRASIL. Instituto Brasileiro do Meio Ambiente e dos Recursos Renováveis. Plano de gestão para o uso sustentável de lagostas no Brasil: *Panulirus argus* (Latreille, 1804) e *Panulirus laevicauda* (Latreille, 1817). DIAS-NETO, J. (org.). Brasília: **IBAMA**, 2008, 121p.

BROWN, W. M. The mitochondrial genome of animals. In: MacIntyre, R.J. (Ed.), **Molecular Evolutionary Genetics**. Plenum, New York, NY, pp: 95–130, 1985.

DINIZ, F. M. et al. Microsatellites in the overexploited spiny lobster, *Panulirus argus*: Isolation, characterization of loci and potential for intraspecific variability studies. **Conservation Genetics**, v. 6, n. 4, p. 637–641, 2005.

GÓES, C.; LINS-OLIVEIRA, J.E. Natural diet of the spiny lobster, *Panulirus echinatus* Smith, 1869 (Crustacea: Decapoda: Palinuridae), from São Pedro and São Paulo Archipelago, Brazil. **Brazilian Journal of Biology**, v. 69, n. 691, p. 143–148, 2009.

HOLTHUIS, L.B. FAO Species Catalogue. Marine Lobster of The World. An annotated and illustrated catalogue of species of interest to fisheries known to date. **FAOFish. Syn.**, Rome, v.13, n. 125, p.1-292, 1991.

MORITZ, C.; DOWLING, T. E.; BROWN, W. M. Evolution of animal mitochondrial DNA: relevance for population biology and systematics. **Annual Review of Ecology and Systematics**, 18: 269-292, 1987.

WILSON, A. C.; CANN, R. L.; CARR, S. M.; George, M.; GYLLENSTEN, U. B.; HELM-BYEHOWSKI, K. M.; HIGUCHI, R. G.; PALUMBI, S. R.; PRAGER, E. M.; SAGE, R. D.; STONEKING, M. Mitochondrial DNA and two perspectives on evolutionary genetics. **Biological Journal of the Linnean Society**, 26:375-400, 1985.

Capítulo 1

Revisão de Literatura

1.1 Distribuição geográfica e habitat

A lagosta *Panulirus echinatus* (Smith, 1869), também conhecida como lagosta espinhosa é uma espécie pertencente ao filo Arthropoda, ao subfilo Crustacea, a classe Malacostraca, a ordem Decapoda, a família Palinuridae e ao gênero *Panulirus*. As lagostas do gênero *Panulirus* possuem espinhos no cefalotórax e nas antenas, características que as diferenciam das outras espécies, sendo assim, conhecidas como lagosta espinhosa ou “*spiny lobster*” (HOLTHUIS, 1991).

Lagostas desse gênero são encontradas em regiões tropicais e subtropicais. A espécie *P. echinatus* distribui-se nos Atlânticos Oriental, Central e Ocidental (Figura 1). No Atlântico Oriental e Central, estão presente nas Ilhas Canárias e de Cabo Verde, e nas Ilhas Ascensão e Santa Helena. No Brasil, ela encontrada nos arquipélagos de Fernando de Noronha, São Pedro e São Paulo, Atol das Rocas e Ilha Trindade, e também está presente em boa parte da costa do litoral brasileiro, do Ceará ao Rio de Janeiro (HOLTHUIS, 1991).

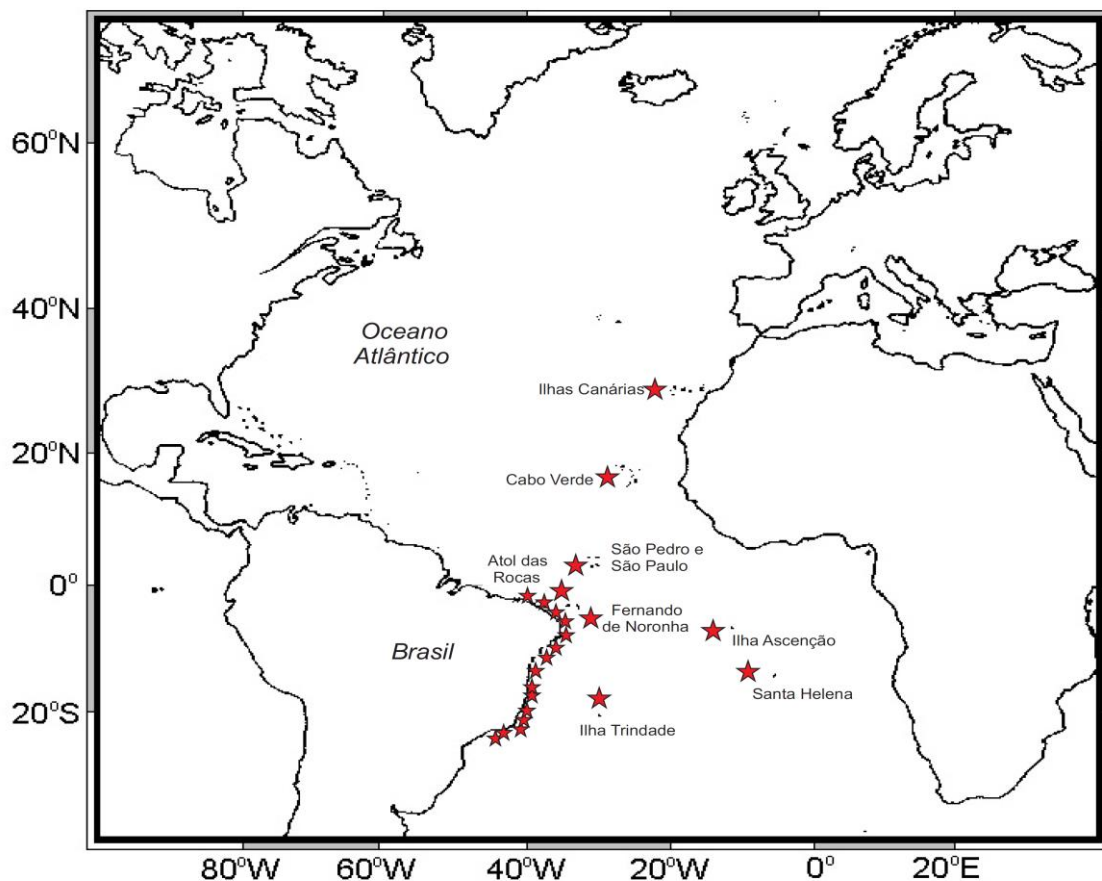


Figura 1 - Distribuição geográfica da espécie *Panulirus echinatus*.

As lagostas habitam substratos profundos, rochosos e emaranhados de algas calcárias bentônicas, constituídas por algas vermelhas e algas verdes. Essas algas são importantes fornecedores de carbonato de cálcio, magnésio e sedimentos essenciais no processo de formação do exoesqueleto durante o ciclo de vida, nas várias mudas ou ecdise que realiza da fase larval até a adulta (FONTELES-FILHO, 1992).

Possuem hábitos alimentares noturnos, vivem em ambientes em que possam se proteger, com ocorrência em profundidades menores que 25 metros (MELO, 1999). São animais onívoros e oportunistas, alimentando-se de presas que ficam nos substratos e de diversos organismos presente nos diferentes níveis tróficos (GÓES; LINS-OLIVEIRA, 2009). Possuem características predatórias, permitindo assim o controle na frequência de alguns organismos marinhos (TEGNER; LEVIN, 1983). Têm grande importância ecológica na cadeia alimentar marinha, pois, as várias fases larvais que possuem em seu ciclo de vida servem de alimento a vários outros organismos aquáticos (FONTELES-FILHO, 1994; PINHEIRO; LINS-OLIVEIRA, 2006)

1.2 Morfologia e reprodução da lagosta *Panulirus echinatus*

A lagosta *P. echinatus* tem uma carapaça espinhosa e corpo dividido em cefalotórax (cabeça e tórax) e abdome. Possui um par de longas antenas espinhosa, dois espinhos supra-oculares voltados para cima e para frente, cinco pares de pereiópodes, sendo o primeiro o mais curto e o terceiro mais longo. Todos os pares possuem espinhos terminais, no abdome possui somitos abdominais com pequenas manchas claras, presença de um leque caudal, com telso retangular longo e largo, exópodo e endópodo (Figura 2) (HOLTHUIS et al., 1980, MELLO, 1999).

O ciclo de vida da lagosta é longo e dividido em fases larvais. A fecundação e desenvolvimento embrionário das lagostas são externos e o ciclo de vida está dividido em cinco fases: adulto, ovos, filosoma, puerulus e juvenis. As fêmeas produzem uma grande quantidade de ovos, que logo chegam à fase larval, período em que sofrem bastante predação, dificultando na sua maioria atingirem a fase adulta (OLIVEIRA et al, 2014).

A fase larval é chamada de filosoma (estágios larvais), compreendendo 11 estágios com duração de aproximadamente 10 meses. Nessa fase as larvas possuem pouca capacidade natatória e são carregadas pelas correntes marítimas. O corpo é comprimido dorso-ventralmente e são totalmente transparentes. Depois desse período transformam-se em

puerulus (pós-larvais) com 12 meses apresentando uma forma já definida, com capacidade natatória. A fase seguinte é a pós-puerulus com hábitos bentônicos, torna-se juvenil com aproximadamente 36 meses, nessa fase adquire coloração e é morfologicamente semelhante ao indivíduo adulto. As lagostas tornam-se adultas somente quando atingem a maturidade sexual, onde têm a capacidade de se reproduzir com aproximadamente 48 meses, dando início a um novo ciclo (IGARASHI; MAGALHÃES NETO, 1999; PHILLIPS; KITAKA, 2000; IGARASHI, 2007).

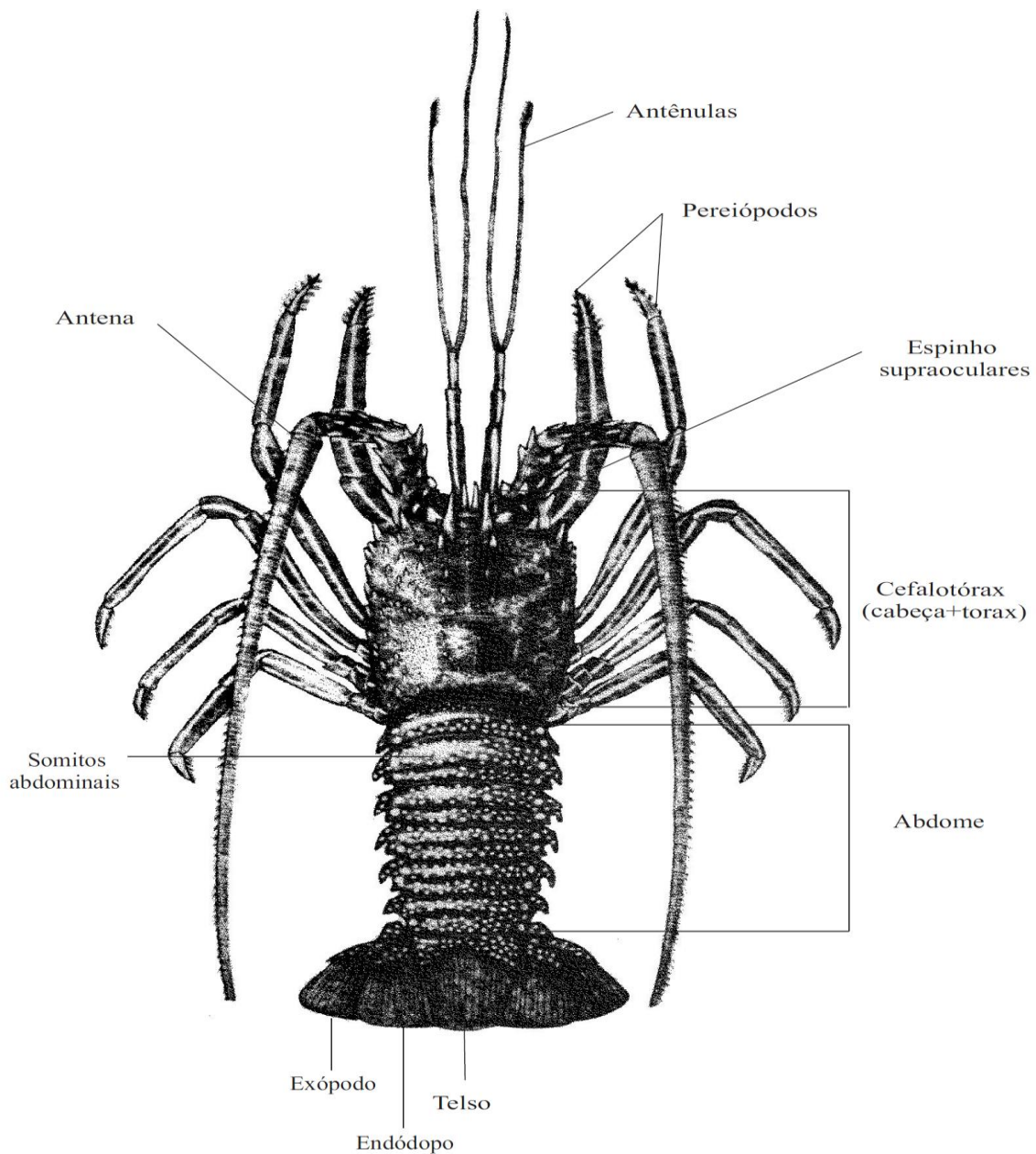


Figura 2 - Desenho esquemático da lagosta *Panulirus echinatus*, destacando as características morfológicas (adaptado de HOLTHUIS, 1991).

1.3 Importância econômica

As lagostas são importantes recursos pesqueiros em muitas partes do mundo e em alguns países são a base da economia local. A captura mundial de lagostas registrados em 2010 ultrapassou 279.000 t (toneladas). As lagostas mais capturadas são as das família Nephropidae, conhecidas como lagostas verdadeiras, da família Palinuridae, lagostas espinhosas e da família Scyllaridae, lagostas sapateiras. O maior número de espécies é observado em regiões tropicais, mas é possível encontrar uma grande quantidade de lagostas em regiões frias. Elas possuem alto valor na gastronomia e são a base comercial de importantes setores industriais (PHILLIPS, 2006; FAO, 2010).

As lagostas são capturadas e comercializadas em alguns países como: Brasil, Cuba, Austrália, EUA e México, que são responsáveis pelas grandes pescarias comerciais. A pesca da lagosta também é base econômica e sustentável de pequenas populações que vivem nas localizações costeiras de ocorrência da espécie e dependem dessa atividade para a sobrevivência (PHILLIPS; KITAKA, 2000).

No Brasil a pesca da lagosta ocorre principalmente na costa do litoral brasileiro, proporcionando o desenvolvimento da atividade pesqueira, que é de grande importância social e econômica. A pesca da lagosta é voltada principalmente para o mercado internacional, em virtude da considerável aceitação, destacando-se os Estados Unidos, Japão e países da Europa como principais compradores (GALDINO, 1995).

No litoral brasileiro, três lagostas são bastante consumidas, todas pertencentes a família Palinuridae, são elas: a lagosta vermelha *Panulirus argus* (Latreille), a lagosta verde *Panulirus laevicauda* (Latreille) e a lagosta espinho *Panulirus echinatus* (Smith, 1869) (SILVA; FONTELES-FILHO, 2011). *P. argus* e *P. laevicauda* são as duas lagostas que atualmente movimentam o mercado pesqueiro de lagostas no Brasil, e a maior parte da sua captura é destinada ao mercado externo, que em 2011 movimentou cerca de US\$ 69.163,97 (BRASIL, 2010).

A *P. echinatus* vem aos poucos ganhando espaço na captura, devido principalmente a sua ocorrência em toda a costa do litoral brasileiro. Outro motivo é a não existência de uma regulamentação para a pesca, o que a torna alvo fácil da pesca ilegal.

O estado do Ceará tem um histórico na produção da lagosta para o mercado externo, desde o início dessa atividade pesqueira nos anos 50, mas nos últimos anos tem-se notado uma significativa queda nas exportações. Uma avaliação feita no período de 2000 a 2012

revela que elevadas exportações foram realizadas no ano de 2007 e a partir desse período vem sofrendo significativas quedas de produção. Essa redução está diretamente relacionada à exigência do mercado externo quanto à qualidade do produto exportado (ARAGÃO, 2013).

Devido a essa intensa exploração na pesca das lagostas e a notável queda que vem sendo observada ano a ano nas exportações tornam-se necessárias medidas que visem à conservação dos seus estoques, para que possam ser melhor monitorados.

1.4 O papel da genética na gestão das pescas

A pesca da lagosta começou no Brasil pelo estado do Ceará na década de 50, e permaneceu na costa do litoral cearense até 1960. Desde o início a pesca sempre teve uma característica artesanal. Somente a partir da década de 60, com a expansão para o mercado externo, que ela se tornou industrial, e foi se estendendo a toda costa do litoral brasileiro. As capturas comerciais de lagosta se concentram no estado do Ceará, mas ocorrem em toda a extensão da costa do litoral dos estados do Amapá até o Espírito Santo (PAIVA; COSTA 1963).

Por apresentar uma característica bastante artesanal em sua captura, tornam-se necessárias informações, que vão desde o conhecimento das espécies capturadas até o manejo adequado. Em 2008 foi lançado o primeiro Plano de Gestão do Uso sustentável de Recursos Pesqueiros (lagostas), elaborado pelo IBAMA, com o objetivo de adotar medidas e informações para o uso desse recurso (BRASIL, 2008).

A gestão da pesca da lagosta em várias partes do mundo onde há a sua ocorrência e também no Brasil, é sempre direcionada a várias técnicas de manejo, principalmente no que diz respeito à temporada da pesca e ao tamanho mínimo de captura. Além dessas duas medidas outras são observadas, como a proibição da pesca no período de reprodução, proibição de apetrechos de pesca, captura máxima permitida e proibição da captura de fêmeas ovadas, esses são alguns pontos que visam a gestão da pesca sustentável da lagosta (HILSDORF et al., 2006).

Todas essas medidas são importantes, mas a adoção de informações genéticas na atualidade torna-se indispensável para traçar um manejo adequado dos estoques. Apesar dessa necessidade, raramente aspectos da informação genética são discutidos em programas de conservação e manutenção de estoques pesqueiros.

Hoje a obtenção de dados genéticos ocorre principalmente através do uso de ferramentas como os marcadores moleculares na caracterização de diversidade genética das populações e seus estoques pesqueiros. O objetivo principal do uso dessas ferramentas moleculares é entender como essas populações estão distribuídas e como estão ligadas geneticamente, se existe ou não fluxo gênico entre elas, e se essas mudanças, em nível de DNA, podem estar influenciando na variabilidade genética das espécies.

O desafio maior é fazer com que essas informações genético-moleculares das espécies de lagostas exploradas comercialmente cheguem de fato a quem trabalha com essa captura, pois é um recurso natural que precisa ser conservado de forma sustentável, e a inserção de informações genéticas nos planos de gestão ajudará a desenvolver metodologias de previsões de captura, manejo e conservação.

1.5 Marcadores Moleculares na família Palinuridade

A infraordem Achelata contém três famílias principais, a família Palinuridae (lagostas espinhosas), Scyllaridae (lagostas sapatas) e Synaxidae (lagostas de coral), essas famílias compartilham uma característica: uma fase larval única chamada de filosoma (PALERO et al., 2009). As lagostas da família Palinuridae contêm dez gêneros, são eles: *Jasus*, *Justitia*, *Linuparus*, *Nupalirus*, *Palinurus*, *Palinustus*, *Panulirus*, *Projasus*, *Puerulus* e *Sagmariasus*, todos com grande interesse comercial (DALL'OCCHO, 2010). As lagostas, devido a sua importância econômica e apreciável fonte de alimento, são bastante capturadas, com isso o conhecimento da estruturação genética dos seus estoques é necessário para traçar planos de conservação e manejo sustentável das espécies, atualmente os marcadores moleculares são os mais utilizados para gerar informações de diversidade genética e estrutura populacional de várias espécies, incluindo estudos com lagostas.

Os marcadores moleculares podem ser classificados em dois tipos: (1) marcadores que estão associados a genes com função conhecida, e (2) marcadores associados a regiões genômicas anônimas. A utilização de marcadores associados a genes com função conhecida está se tornando extremamente importante para aquicultura genética, pois a conservação dentro dos genes é elevada, permitindo que esses marcadores sirvam como pontos na identificação de segmentos genômicos possíveis de serem comparados entre as espécies (HILSDORF, 2013). Vários tipos de marcadores são altamente populares em estudos de organismos aquáticos. Estudos mais atuais para a escolha de dados revelantes da pesca e

aquicultura são provenientes de avaliações diretas do DNA nuclear e variação de sequências (BROWN; EPIFANIO, 2003).

As principais ferramentas moleculares de DNA nuclear incluem *Random amplified polymorphic DNA* (RAPD), *Restriction fragment length polymorphism* (RFLP), *Amplified Fragment Length Polymorphism* (AFLP), *Simple Sequence Repeats* (SSR), *Individual Single Nucleotide Polymorphism* (SNP) e marcadores de DNA ribossomal nuclear. Essas ferramentas têm uso generalizado em estudos genéticos de populações para caracterizar divergência genética dentro e entre as populações ou espécies. A escolha dos marcadores moleculares pode ser feita com base no valor que eles têm em detectar polimorfismo numa população, geralmente observado pelo conteúdo de informação polimórfica (BOTSTEIN et al., 1980).

O RAPD é uma técnica que já foi bastante utilizada para avaliar a diversidade genética de espécies, populações de uma grande variedade de crustáceos decápodes e de tanto outros organismos. Atualmente ela vem sendo substituída por ferramentas moleculares mais específicas, como microssatélites e DNA mitocondrial. Poucas são as investigações realizadas com essa técnica em lagostas.

Um dos trabalhos utilizando a técnica RAPD foi à investigação preliminar das relações genéticas entre lagostas americanas (*Homarus americanus*) de regiões ecologicamente distintas. Os resultados obtidos a partir dos valores F_{ST} variaram de $<0,000-0,073$, indicando que as lagostas nos locais estudados não são geneticamente isoladas (HARDING et al., 1997). Outro estudo desenvolvido por meio de análise RAPD foi a observação do grau de variação genética da lagosta europeia (*Homarus gammarus*), onde as populações estudadas apresentaram altos valores de distância genética, indicando que a troca de alelos entre elas pode ser restrita (ULRICH et al., 2001).

A técnica RAPD também foi utilizada para discriminação de tecidos da lagosta americana e europeia (HUGHES; BEAUMONT, 2004), e também em espécies de lagosta da Coreia do Sul e do Oceano Índico *Ibacus ciliatus* e *Puerulus sewelli*, respectivamente, os resultados obtidos indicaram que a complexidade dos padrões de bandas variou drasticamente entre os iniciadores e locais amostrados (PARK et al., 2005).

O RAPD, por apresentar características dominantes, é comum em estudos de diversidade genética ele vir associado a outro marcador mais específico. Foi o que aconteceu em estudos desenvolvidos com uma importante lagosta explorada comercialmente na Índia. Para verificar a estruturação genética da lagosta *Thenus unimaculatus*, marcadores RAPD e

um conjunto de dados de genes de DNA mitocondrial (sequências parcial de citocromo-oxidase I e citocromo b) foram usados para detectar a diferenciação populacional, os dados obtidos de diversidade para os dois marcadores foram semelhantes (JEENA et al., 2015), estes são os poucos trabalhos sobre lagostas com o uso dessa técnica.

Outra ferramenta molecular utilizada, mas não tão sensível quanto os marcadores de DNA, são as isoenzimas, que identifica variações genéticas em nível de enzimas, que são diretamente codificadas pelo DNA. As isoenzimas podem ser utilizadas para quantificar a variação genética e distinguir níveis de populações, espécies e grupos taxonômicos superiores. Desvantagens associadas a esta ferramenta, incluem, deficiências heterozigotas ocasionais devido a alelos nulos (enzimaticamente inativos) e é sensível à quantidade bem como a qualidade de amostras de tecido, além disso, algumas alterações na sequência de DNA podem está mascaradas ao nível da proteína, reduzindo o nível de variação detectável (LIU; CORDES, 2004).

Isoenzimas foram utilizadas como parte de uma investigação genética abrangente de lagosta Europeia (*Homarus gammarus*) onde 1514 lagostas de 14 locais diferentes foram analisadas para seis locos isoenzimático polimórficos. Apesar dos baixos níveis de diferenciação genética ($F_{ST} = 0,016$), os testes para diferenciação populacional revelaram valores altamente significativos para todos os locos investigados (JORSTAD et al., 2005).

A estrutura genética populacional da lagosta *Nephrops norvegicus* foi estudada por meio de eletroforese de isoenzimas. Populações do Atlântico e de oito localidades do Mediterrâneo foram amostradas e dos 15 sistemas enzimáticos analisados 11 foram polimórficos. A análise das frequências genotípicas e alélicas revelou níveis de polimorfismo genéticos comparáveis com os decápodos já estudado, observando uma diferenciação genética moderada entre regiões geográficas (MALTAGLIATI et al., 1998).

Estudos populacionais em lagostas também são realizados com outra ferramenta molecular, o RFLP. Geralmente os estudos realizados através dessa técnica vem associado a um marcador molecular, como o DNA mitocondrial (DNAm). A técnica RFLP foi base de um estudo para determinar a identidade de 54 larvas (filosoma) de duas espécies de lagostas, *Panulirus inflantus* e *Panulirus gracilis*, para estudos de diferenciação morfológica e morfométrico (MUÑOZ-GARCÍA et al., 2014)

Já Chow et al., (2006) para identificar larvas de lagostas do gênero *Panulirus* com ocorrência nas águas do Japão e determinar a variação genética dentro e entre as 10 espécies

de ocorrência, utilizou o RFLP, e a partir de duas endonucleases (AluI e TaqI) foi possível a discriminação de todas as espécies avaliadas.

A técnica RFLP geralmente vem associada a outros marcadores como DNAMt e microssatélites, como foi observado em um estudo desenvolvido por Jones e colaboradores (2003), para lagosta americana *Homarus americanus*, onde foi desenvolvido e caracterizado locos de microssatélites tri e tetranucleotídios, e verificado os padrões de herança dentro da espécie, tanto por meio de marcadores microssatélites como DNAMt (PCR-RFLP).

Estudos de diversidade genética e estrutura populacional são comumente realizados com uma classe de marcador chamada de microssatélites (SSRs) Os microssatélites são muito frequentes e apresentam ampla cobertura no genoma, podendo estar associados a sequências expressas. É uma classe de marcador com o maior conteúdo de informação polimórfica (PIC). Em lagostas a maioria dos estudos de estruturação populacional, divergência genética utilizam-se dos SSRs, como podemos observar nos trabalhos a seguir.

O declínio observado nas populações de lagosta da espécie *Panulirus homarus* tem levado a alguns estudos sobre o conhecimento da base genética e estrutura populacional dessa espécie, como o estudo realizado por Delghandi et al. (2015), onde utilizaram 13 novos marcadores microssatélites di-, tri- e tetranucleotídeos polimórficos para *P. homarus*, a partir de uma estratégia de iniciadores para espécies cruzadas, baseado em dados de sequenciamento de outra espécie estreitamente relacionada a lagosta *Panulirus ornatus*. Os marcadores foram avaliados em 96 indivíduos de *P. homarus*, as amplificações cruzadas bem sucedidas destes microssatélite nessa espécie revelaram como eles podem ser utilizados para a pesquisa genética futura em outras populações de diferentes gêneros de *Panulirus*.

Truelove et al. (2015) utilizaram desse marcador para investigar a conectividade de populações de lagosta *Panulirus argus*, entre duas áreas marinhas protegidas na costa caribenha do México, os resultados obtidos foram a não evidência de diferenciação populacional entre as populações estudadas.

Trinta marcadores microssatélites foram utilizados para *Panulirus ornatus*, dos quais 15 locos foram consistentemente amplificados e exibiram níveis adequados de polimorfismo de parentesco nas populações estudadas. Esses marcadores também foram testados em mais três espécies do gênero *Panulirus*, *P. homarus*, *P. versicolor*, e *P. argus*, com diferentes níveis de amplificação e polimorfismo exibido (DAO; TODD; JERRY, 2013).

Dez loci microssatélites tetranucleotídeo foram descritos para a lagosta do caribe *Panulirus argus*, onde desmostraram elevados níveis de heterozigosidade observada e

esperada a partir de amostras capturadas em Belize e Porto Rico. Todos os loci microssatélites foram úteis para avaliar a discriminação das populações de lagosta que atualmente é submetido à pesca excessiva (DINIZ et al., 2004).

Estudos estão sendo realizados em conjunto com os dois marcadores mais informativos, microssatélites e DNAm, Dao et al. (2015), utilizando a região controle do DNA mitocondrial e 10 microssatélites polimórficos, não detectaram estrutura genética em 216 indivíduos de *P. ornatus* amostrados da Austrália, Indonésia e Vietnã. Nenhuma evidência foi observada para a diferenciação genética entre as populações (sequências da região controle do DNAm $\Phi_{ST} = -0,008$; locos de microssatélites $F_{ST} = 0,003$), revelando um padrão de alto fluxo gênico.

A classe de marcador molecular que vem sendo bastante utilizada é o DNA mitocondrial. Uma pequena porção do DNA das células eucarióticas é não-nuclear, e está localizada dentro de organelas no citoplasma chamadas mitocôndrias. Em crustáceos decápodes, o tamanho usual é de 14 a 18 kb, o DNAm, normalmente contém 36 ou 37 genes (sendo dois genes de RNAs ribossomais, 16SrRNA e 12SrRNA, 22 genes de tRNAs e 12 ou 13 genes codificadores de proteínas). Além disso, há geralmente pelo menos uma sequência de comprimento variável que não codifica nenhum gene, chamada de região controle.

Vários são os estudos que utilizam o DNAm em lagostas, gerando uma gama de informações que serão úteis em estudos principalmente de filogenia, estrutura populacional e divergência genética.

Diniz et al. (2005) por exemplo, exploraram o potencial do domínio hipervariável da região controle do DNA mitocondrial como um marcador genético para caracterizar subdivisão em população de lagostas, usando a lagosta *Panulirus argus* como referência. Já Garcia et al. (2008) investigaram as variações de nucleótidos do DNAm em lagostas adultas para obter marcadores genéticos úteis na identificação de *Panulirus interruptus*, *Panulirus inflatus*, e *Panulirus gracilis*, através de PCR-RFLP que amplificam tamanho diferentes de fragmentos em reação multiplex.

Farhadi et al. (2013) utilizaram sequências de DNA mitocondrial da região controle para investigar a estrutura genética populacional da espécie *Panulirus homarus* (Linnaeus, 1758), no Oceano Índico Ocidental, utilizando 137 indivíduos coletados de quatro pontos diferentes.

Análises filogenéticas utilizando 21 táxons do gênero *Panulirus*, foram examinados com dados de sequência de nucleotídicas parciais de porções de dois genes mitocondriais,

RNA ribossomal grande subunidade (16S rRNA) e citocromo-oxidase subunidade I (COI), procurando fornecer evidências moleculares e apoiar as relações inter-específicas (NAYAK; UMADEVI, 2012).

Dharani et al. (2009) em estudos para a identificação de larvas puerulus de *Panulirus homarus* da costa indiana utilizaram a reação em cadeia de polimerase com polimorfismo de fragmento de restrição (PCR-RFLP), e a análise do DNA mitocondrial (DNAMt) do gene citocromo-oxidase (COI). Os resultados obtidos para espécie foi uma amplificação padrão que se manteve em todas as fases do ciclo de vida da lagosta, permitindo assim a sua identificação nas diferentes fases da sua história de vida, mostrando assim ser ferramentas moleculares eficiente.

Grande parte de estudos com DNAMt visam esclarecer as relações evolutivas das espécies, e em lagosta principalmente do gênero *Panulirus* vem sendo bastante empregado, devido a dispersão das larvas através das correntes marítimas. Como exemplo podemos citar o estudo desenvolvido com a lagosta *Panulirus homarus* que é uma espécie difundida ao longo da costa dos oceanos Índico e Pacífico, mas possui uma taxonomia incerta, com três ou quatro subespécies geográficas tendo sido descritas. Lavery et al. (2014) utilizaram DNA mitocondrial (16S, COI e controle região) e nuclear (18S, ITS-1) para examinar espécimes de *P.homarus*, a fim de determinar a sua validade genética, e entender a história evolutiva desta espécie. Os resultados obtidos foram um alto grau de divergência evolutiva das populações, sugerindo que a especiação alopátrica tem sido um fator importante na evolução do gênero.

Para a espécie *Panulirus echinatus* um dos poucos estudos desenvolvidos com DNA mitocondrial foi a análise de larvas do gênero *Panulirus* coletadas no Atlântico Central, através de sequências nucleotídicas da região 16S do DNAMt. Os resultados obtidos identificaram as larva como sendo *P. echinatus*, juntamente com observação morfológicas, e de acordo com os resultados da árvore filogenética molecular com outras espécies do mesmo gênero a colocou estreitamente relacionada com a lagosta *Panulirus penicillatus*. O estudo mostrou que análises moleculares juntamente com investigações morfológicas podem contribuir significativamente para aperfeiçoar a identificação da espécie na forma de larva no gênero *Panulirus* (KONISHI; SUZUKI; CHOW, 2006).

A disponibilidade de vários mitogenomas completos de lagostas depositados no GenBank tem auxiliado estudos de filogenia, para o gênero *Panulirus*, podemos citar os mitogenomas completos das lagostas *Panulirus cygnus* (KIM et al., 2015), *Panulirus ornatus* (LIANG, 2012), *Panulirus versicolor* (SHEN; BRABAND; SCHOLTZ, 2013), *Panulirus*

japonicus (YAMAUCHI; MIYA; NISHIDA, 2002), *Panulirus homarus* (NC016015) e *Panulirus stimpsoni* (LIU; CUI, 2011).

As informações aqui descritas mostram uma visão geral sobre o uso das principais ferramentas moleculares em lagostas da família Palinuridae. O objetivo principal desse levantamento é relatar a importância das ferramentas na caracterização genética das populações, e com isso auxiliar em estudos de conservação dessas espécies de grande importância econômica.

1.6 Referências

- ARAGÃO, J. N. Boletim Técnico Científico. CEPENE, Tamandaré - PE - v. 19, n. 1, p. 103-106, 2013
- BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, 32:314- 331, 1980.
- BRASIL. Instituto Brasileiro do Meio Ambiente e dos Recursos Renováveis. Plano de gestão para o uso sustentável de lagostas no Brasil: *Panulirus argus* (Latreille, 1804) e *Panulirus laevicauda* (Latreille, 1817). DIAS-NETO, J. (org.). Brasília: IBAMA, 2008, 121p.
- BROWN, B.; EPIFANIO, J. Nuclear DNA. Population genetics: principles and applications for fisheries scientists. **American Fisheries Society**, Bethesda, Maryland, 101-123, 2003.
- CHOW, S. et al. Molecular species identification of spiny lobster phyllosoma larvae of the genus *Panulirus* from the northwestern Pacific. **Marine Biotechnology**, v. 8, n. 3, p. 260–267, 2006.
- DAO, H. T. et al. Oceanographic Currents and Local Ecological Knowledge Indicate , and Genetics Does Not Refute , a Contemporary Pattern of Larval Dispersal for The Ornate Spiny Lobster , *Panulirus ornatus* in the South-East Asian Archipelago. p. 1–19, 2015.
- DAO, H. T.; TODD, E. V.; JERRY, D. R. Characterization of polymorphic microsatellite loci for the spiny lobster *Panulirus* spp. and their utility to be applied to other *Panulirus* lobsters. **Conservation Genetics Resources**, v. 5, n. 1, p. 43–46, 2013.
- DALL'OCCHO, P. L. Taxonomia e Distribuição das Lagostas (Crustacea: Decapoda: Achelata e Polychelida) no Oceano Atlântico. Rio Claro: [s.n.], 432f, 2010.
- DELGHANDI, M. et al. Isolation, characterization, and multiplexing of novel microsatellite markers for the tropical scalloped spiny lobster (*Panulirus homarus*). **Genetics and Molecular Research**, v. 14, n. 4, p. 19066–19070, 2015.
- DHARANI G.; MAITRAYEE, G.A; KARTHIKAYALU, S.; KUMAR, T. S.; ANBARASU, M.; VIJAYAKUMARAN, M. Identification of *Panulirus homarus* puerulus larvae by restriction fragment length polymorphism of mitochondrial cytochrome oxidase I gene.. Pak. J. Biol. Sci., 12: 281-285, 2009.
- DINIZ, F. M. et al. Polymorphic tetranucleotide microsatellite markers in the Caribbean spiny lobster, *Panulirus argus*. **Molecular Ecology Notes**, v. 4, n. 3, p. 327–329, 2004.
- DINIZ, F. M. et al. The hypervariable domain of the mitochondrial control region in Atlantic spiny lobsters and its potential as a marker for investigating phylogeographic structuring. **Marine Biotechnology**, v. 7, n. 5, p. 462–473, 2005.
- FAO. Fishery statistical collections: global production. **Food and Agriculture**

Organization (FAO) of the UN, 2010.

FARHADI, A. et al. Mitochondrial DNA population structure of the scalloped *Panulirus homarus* (Linnaeus 1758) from the West Indian Ocean. v. 70, p. 1491–1498, 2013.

FONTELES-FILHO, A.A. Population dynamics of spiny lobsters (Crustacea: Palinuridae) in Northeast Brazil. **Ciência e Cultura**, 44:192-196, 1992.

FONTELES-FILHO, A.A. State of the lobster fishery in North-east Brazil, p. 108-118, in Phillips, B.F.; Cobb, J.S. & Kittaka, J. (eds.), Spiny lobster management. Fishing News Books, 550 p., Oxford, 1994.

GARCÍA-RODRÍGUEZ, F. J. et al. Mitochondrial DNA markers to identify commercial spiny lobster species (*Panulirus* spp.) from the Pacific coast of Mexico: An application on phyllosoma larvae. **Fishery Bulletin**, v. 106, n. 2, p. 204–212, 2008.

GALDINO, J. W. A intermediação e os problemas socio-econômicos no defeso da pesca de lagosta em Redonda, Icapuí (CE). Dissertação (Mestrado em Economia Rural). Departamento de Economia Agrícola, Universidade Federal do Ceará, UFC, Fortaleza, 1995. 134p

GÓES, C. A.; LINS-OLIVEIRA, J. E. Natural diet of the spiny lobster, *Panulirus echinatus* Smith, 1869 (Crustacea: Decapoda: Palinuridae), from São Pedro and São Paulo Archipelago, Brazil. *Braz. J. Biol.*, 69(1): 143-148, 2009.

HARDING, G. C.; KENCHINGTON, E. L.; BIRD, C. J.; PEZZACK, D. S.; LANDRY, D. C. Genetic relationships among subpopulations of the American lobster (*Homarus americanus*) as revealed by random amplified polymorphic DNA. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(8): 1762-1771, 1997.

HILSDORF, A. W. S. Marcadores moleculares e a caracterização dos recursos genéticos de peixes : desenvolvimento sustentável da aquicultura e da pesca de espécies nativas de água doce no Brasil. p. 159, 2013.

HILSDORF, A.W.S.; RESENDE, E.K.; Marques D. K. S. Genética e Conservação de Estoques Pesqueiros de Águas Continentais no Brasil: Situação Atual e Perspectivas. Embrapa. 2006: 44p.

HOLTHUIS, L.B. FAO Species Catalogue. Marine Lobster of The World. An annotated and illustrated catalogue of species of interest to fisheries known to date. **FAOFish**. Syn., Rome, v.13, n. 125, p.1-292, 1991.

HOLTHUIS, L.B., EDWARDS, A.J.; LUBBCKOCK, H.R. The decapoda and stomatopod Crustacea of St. Paul's Rocks. *Zool. Meded. Rijksmus. Natuur. Hist. Leiden*, v. 56, n. 3, p. 27-51, 1980.

HUGHES, G.; BEAUMONT, A. R. A potential method for discriminating between tissue from the European Lobster (*Homarus gammarus*) and the American Lobster (*H. americanus*). **Crustaceana**, 77(3): 371-376, 2004.

- IGARASHI, M. A.; MAGALHÃES NETO, E. O. O ciclo de desenvolvimento e a situação atual do cultivo de lagostas no Brasil e no mundo. *Revista Econômica do Nordeste*, Fortaleza, v. 30, n. 1, p. 94-99, 1999.
- IGARASHI, M. A. Sinopse da situação atual, perspectivas e condições de cultivo para lagostas Palinuridae. *Ciência Animal Brasileira*, Goiânia, v. 8, n. 2, p. 155-166, 2007.
- JEENA, N. S. et al. Population genetic structure of the shovel-nosed lobster *Thenus unimaculatus* (Decapoda, Scyllaridae) in Indian waters based on RAPD and mitochondrial gene sequences. *Hydrobiologia*, v. 766, n. 1, p. 225–236, 2015.
- JONES, M. W. et al. Development, characterisation, inheritance, and cross-species utility of American lobster (*Homarus americanus*) microsatellite and mtDNA PCR-RFLP markers. **Genome / National Research Council Canada = Genome / Conseil national de recherches Canada**, v. 46, n. 1, p. 59–69, 2003.
- JORSTAD, K. E. et al. Allozyme variation in European lobster (*Homarus gammarus*) throughout its distribution range. **New Zealand Journal of Marine and Freshwater Research**, v. 39, n. April 2004, p. 515–526, 2005.
- KONISHI, K.; SUZUKI, N.; CHOW, S. A late-stage phyllosoma larva of the spiny lobster *Panulirus echinatus* Smith, 1869 (Crustacea: Palinuridae) identified by DNA analysis. **Journal of Plankton Research**, v. 28, n. 9, p. 841–845, 2006.
- KIM, G. et al. Complete mitochondrial genome of Australian spiny lobster, *Panulirus cygnus* (George, 1962) (Crustacea: Decapoda: Palinuridae) from coast of Australia. **Mitochondrial DNA**, v. 1736, n. January 2016, p. 1–2, 2015.
- LAVERY, S. D. et al. Evolutionary divergence of geographic subspecies within the scalloped spiny lobster *Panulirus homarus* (Linnaeus 1758). **PLoS ONE**, v. 9, n. 6, 2014.
- LIANG, H. Complete mitochondrial genome of the ornate rock lobster *Panulirus ornatus* (Crustacea: Decapoda). **African Journal of Biotechnology**, v. 11, n. 80, p. 14519–14528, 2012.
- LIU, Z. J.; CORDES, J. F. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture*, 238(1): 1-37, 2004.
- LIU, Y.; CUI, Z. Complete mitochondrial genome of the Chinese spiny lobster *Panulirus stimpsoni* (Crustacea: Decapoda): Genome characterization and phylogenetic considerations. **Molecular Biology Reports**, v. 38, n. 1, p. 403–410, 2011.
- MALTAGLIATI, F. et al. Genetic structure of Norway lobster, *Nephrops norvegicus* (Crustacea: Nephropidae), from the Mediterranean Sea. **Scientia Marina**, v. 62, n. Suppl. 1, p. 91–99, 1998.
- MELO, G.A.S. Manual de identificação dos crustáceos Decapoda do litoral brasileiro: Anomura, Thalassinidea, Palinuridea, Astacidea. Plêiade/ FAPESP, 551 p., São Paulo, 1999.
- MUÑOZ-GARCÍA, I. et al. Taxonomy of the phyllosoma of *Panulirus inflatus* (Bouvier, 1895) and *P. gracilis* Streets, 1871, based on morphometry and molecular analysis. **Nauplius**, v. 22, n. 1, p. 41–51, 2014.

- NAYAK, S.; UMADEVI, K. In silico comparative molecular phylogeny of mitochondrial 16s rRNA and COI genes of the spiny lobster genus *Panulirus* (Decapoda: Palinuridae). v. 3, n. 3, p. 364–373, 2012.
- OLIVEIRA, P. A.; CRISPIM, VENDEL, A.L. Caracterização populacional das lagostas do gênero *Panulirus* nos ambientes recifais da Praia do Seixas e da Penha-PB. **Gaia Scientia**, v.8, n.1, p. 365-383, 2014.
- PAIVA, M. P.; COSTA, R. S. Tamanho de fêmeas de lagosta em reprodução nas águas costeiras do Ceará. **Arq. Est. Biol. Mar. Univ. Fed. Ceará**, Fortaleza, v. 3, p. 53-56, 1963.
- PALERO, F.; CRANDALL, K. A.; ABELLÓ, P.; MACPHERSON, E.; PASCUAL, M. Phylogenetic relationships between spiny, slipper and coral lobsters (Crustacea, Decapoda, Achelata). **Molecular Phylogenetics and Evolution**, 50(1): 152-162, 2009.
- PARK, S. Y.; YOON, J. M. Genetic differences and variations in slipper lobster (*Ibacus ciliatus*) and deep sea lobster (*Puerulus sewelli*) determined by RAPD analysis. *Gene and Genomics*, 27(4): 307-317. 2005.
- PINHEIRO, A. P.; LINS-OLIVEIRA, J. E. Reproductive Biology of *Panulirus echinatus* (Crustacea: Palinuridae) from São Pedro and São Paulo Archipelago, Brazil. **Nauplius** 14(2): 89-97, 2006.
- PHILLIPS, B. F. Lobsters: biology, management, aquaculture and fisheries. Blackwell Publishing Limited, 2006.
- PHILLIPS, B. F.; KITAKA, J. Spiny lobsters: fisheries and culture (p.679). Oxford: Fishing News Books, 2000.
- SILVA, A.C.; FONTELES-FILHO, A.A. Avaliação do Defeso Aplicado à Pesca da Lagosta no Nordeste do Brasil. Fortaleza, Editora Expressão Gráfica. 110p, 2011.
- SHEN, H.; BRABAND, A.; SCHOLTZ, G. Mitogenomic analysis of decapod crustacean phylogeny corroborates traditional views on their relationships. **Molecular Phylogenetics and Evolution**, v. 66, n. 3, p. 776–789, 2013.
- ULRICH, I.; MULLER, J.; SCHUTT, C.; BUCHHOLZ, F. A study of population genetics in the European lobster, *Homarus gammarus* (Decapoda, Nephropidae). *Crustaceana*, 74(9): 825-837, 2001.
- TEGNER, M. J.; LEVIN, L. A. Spiny lobsters and sea urchins: analysis of a predator-prey interaction. **Journal of Experimental Marine Biological and Ecology**, Amsterdam, 1983.v. 73, p. 50-125,
- TRUELOVE, N. K. et al. Genetic analysis reveals temporal population structure in Caribbean spiny lobster (*Panulirus argus*) within marine protected areas in Mexico. **Fisheries Research**, v. 172, n. April, p. 44–49, 2015.
- YAMAUCHI, M. M.; MIYA, M. U.; NISHIDA, M. Complete mitochondrial DNA sequence of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Decapoda). **Gene**, v. 295, n. 1, p. 89–96, 2002.

Capítulo 2

Abundância e análise de marcadores microsatélites em lagostas *Panulirus echinatus* (Smith, 1869) (Crustacea: Decapoda: Palinuridae)

*Artigo “**Abundance and analysis of microsatellite markers in lobsters *Panulirus echinatus* (Smith, 1869) (Crustacea: Decapoda: Palinuridae)**”

2.1 Resumo

A lagosta *Panulirus echinatus* (Smith, 1869), também conhecida como lagosta espinhosa, tem uma grande importância econômica por ser uma apreciável fonte de alimento. As espécies de lagosta espinhosa possuem um alto valor comercial e, devido a uma intensa exploração pesqueira, podem ter suas populações reduzidas em diversidade, culminando com o declínio da pesca. Com o objetivo de estudar a distribuição de SSRs no genoma da lagosta, e facilitar a compreensão da sua estrutura, tornando assim base para o isolamento de marcadores SSRs específicos, foi realizada uma busca em todo o seu genoma sequenciado. A partir do sequenciamento do genoma da lagosta, obteve-se uma plataforma com 7196 *contigs*, os *motifs* de microssatélites foram identificados e localizados utilizando o software MSDB (*Microsatellite Search Building Database*), um total de 2959 SSRs divididos em seis classes, foram identificados. Os comprimentos totais e médios foram de 67.049 pb e 125,39 pb, respectivamente. Os mononucleotídeos foi o tipo mais abundante, com a mais alta frequência relativa (399,1 loci/Mb), correspondendo a 43,60% de todos os SSR, seguido de dinucleotídeos (43,26%), trinucleotídeos (7,47%), tetranucleotídeos (3,61%), pentanucleotídeos (1,49%) e hexanucleotídeos (0,57%). As sequências completas obtidas a partir desse sequenciamento proporcionou a identificação de tipos de SSRs, que irá possibilitar a análise e desenvolvimento de microssatélites específicos para *P.echinatus*.

Palavras-chave: Genoma, estrutura, sequenciamento, SSRs

2.2 Abstract

The lobster *Panulirus echinatus* (Smith, 1869), also known as spiny lobster or brown, have a great economic importance to be a significant source of food. The spiny lobster species have a high commercial value, and due to intense fishing operation, may have its population reduced by diversity, culminating with the decline of fishing. With the aim of studying the distribution SSRs in the genome of lobster, and facilitate the understanding of its structure, thus making basis for the isolation of specific SSRs markers, a search was performed in all their genome sequenced. From the lobster genome sequencing was obtained a platform 7196 contigs motifs microsatellite was identified and located using the MSDB software (Microsatellite search Building Database), a total of 2959 SSRs and six classes of SSRs was identified the genome of the lobster, the total length and average were 67049 bp and 125.39 bp, respectively. The mononucleotide was the most abundant type with the highest relative frequency (399.1 loci / Mb), corresponding to 43.60% of all SSR, followed by dinucleotides (43.26%), trinucleotide (7.47%) tetranucleotideos (3.61%), pentanucleotide (1.49%) and hexanucleotideos (0.57%). The complete sequence derived from this sequence yielded identifying types of SSRs, which will enable the analysis and development of specific microsatellite for *P.echinatus*.

Keywords: Genome, structure, sequencing, SSRs

2.3 Introdução

A lagosta marrom *Panulirus echinatus* (Smith, 1869), também conhecida como lagosta espinhosa, tem uma grande importância econômica por ser um apreciável fonte de alimento. Possui grande abundância em áreas costeiras do litoral brasileiro e em ilhas oceânicas (BARRETO; KATSURAGAWA, 2008).

As espécies de lagosta espinhosa possuem um alto valor comercial, e devido a uma intensa exploração pesqueira, podem ter suas populações reduzidas em diversidade, culminando com o declínio da pesca. O conhecimento de como essas populações encontra-se distribuídas podem ser mensurados através de estudos com ferramentas moleculares, como os marcadores do tipo microssatélites ou SSR (*Simple Sequence Repeats*).

Os microssatélites são marcadores de caráter codominante e multialélicos, tendo maior conteúdo informativo por loco (CAIXETA et al., 2009). As sequências repetidas, comuns ao longo de todo o genoma eucarioto, caracterizam locos microssatélites, que consistem de dois a seis nucleotídeos repetidos em *tandem* (JARNE; LAGODA, 1996). As sequências de DNA que flanqueiam os microssatélites geralmente são conservadas entre os indivíduos de uma mesma espécie, permitindo a amplificação do DNA repetitivo, por meio de *primers* específicos. Essas sequências formam locos genéticos altamente polimórficos, tornando o marcador microssatélite um dos mais utilizados atualmente.

Em contraste com o DNA não repetitivo, o polimorfismo nos microssatélites ocorre principalmente devido à variação do número de motivos repetidos em vez de substituições. Os microssatélites surgiram como um dos marcadores neutros mais populares e versáteis utilizados em estudos genéticos, incluindo o mapeamento de genoma, estrutura populacional, filogenia e relações de parentesco (SELKOE; TOONEN, 2006).

O procedimento mais comum para o isolamento de marcadores microssatélites, é o enriquecimento do DNA genômico e o rastreamento da biblioteca resultante, no entanto, esse processo é desafiador, caro e demorado, e estão sendo substituídos rapidamente pela mineração *in silico* de sequências de SSRs, *in silico* é uma expressão usada no âmbito da simulação computacional e áreas correlatas para indicar algo ocorrido "em ou através de uma simulação computacional" (JAYASHREE et al., 2006; ZANE; BARGELLONI; PATARNELLO, 2002).

A grande disponibilidade de sequências de genomas de vários organismos depositados em bancos de dados e os novos desenvolvimentos metodológicos de mineração *in silico* de

microsatélites, aumentaram as pesquisas e caracterização desse marcador, buscando principalmente compreender a origem, a distribuição e a função dos SSRs no genoma dos eucariotos (SHARMA; GROVER; KAHL, 2007).

As novas técnicas de sequenciamento de nova geração NSG (*Next Generation Sequencing*), aliadas a ferramentas computacionais permite um conhecimento do genoma das espécies estudadas e possibilita a identificação de marcadores microsatélites, para futuros estudos de diversidade genética.

Com o objetivo de estudar a frequência de SSRs no genoma da lagosta *P. echinatus*, e descrever a sua estrutura, tornando assim base para o isolamento de marcadores SSRs específicos, foi realizada uma busca em todo o seu genoma.

2.4 Material e Métodos

O sequenciamento do genoma da *P. echinatus* foi realizado em parceria com a Dalhousie University (Canadá), Department of Biology, Marine Gene Probe Laboratory.

O DNA alvo utilizado para o sequenciamento foi extraído do tecido muscular da lagosta, utilizando-se 20 mg de material macerado, digerido em tampão contendo, 50 mM Tris-HCl pH 8.0, 2% de SDS, 0,75M NaCl, 10 mM EDTA e 100 µg/mL de proteinase K, incubado em banho térmico a 55°C *overnight*, com agitações brandas. A seguir, a amostra foi purificada utilizando o protocolo padrão PCI (SAMBROOK; RUSSELL, 2001), com Fenol-Clorofórmio/Álcool isoamílico.

Para a obtenção dos *reads* que foram submetidos a posterior sequenciamento, foi utilizado 1 ng de DNA genômico. Uma biblioteca de extremidades emparelhadas da Illumina foram produzidas conforme especificações do kit de preparação da Illumina Nextera (Illumina Inc.). O sequenciamento foi conduzido usando o sequenciador Miseq Benchtop (Illumina Inc.). As sequências contíguas (*contigs*) foram criadas a partir dos dados das sequências emparelhados pelas extremidades obtidas por meio do CLC Genomics Workbench 7.0.4 (Qiagen).

Os *motifs* de microsatélites foram identificados e localizados utilizando o software MSDB (*Microsatellite Search Building Database*) (<http://msdb.biosv.com/>) (DU et al., 2013), a partir dos *contigs* obtidos no formato FASTA.

O software MSDB tem dois modos de pesquisa: modo de pesquisa perfeito usado para procurar microsatélites puros e o modo de pesquisa imperfeito usados para pesquisar todos

os tipos de microssatélites, que permite a análise de microssatélites composto e complexos (DU et al., 2013). Os microssatélites podem ser agrupados em seis categorias: microssatélites puros ou perfeitos; microssatélites puros interrompidos; microssatélites compostos; microssatélites compostos interrompidos; microssatélites complexos e microssatélites complexos interrompidos (BACHMANN; BAREI; TOMIUK, 2004; CHAMBERS; MACAVOY, 2000).

Com a finalidade de procurar SSRs perfeitos, parâmetros mínimos de repetição foram definidos em 14, 7, 5, 4, 3 e 3, para análises de mononucleotídeos, dinucleotídeos, trinucleotídeos, tetranucleotídeos, pentanucleotídeos e hexanucleotídeos respectivamente. Os resultados geraram informações sobre o número de *motifs*, o comprimento total, média de cada *motif*, abundância relativa (Nº de *motifs*/ tamanho total das sequências analisadas em Mb) e densidade (comprimento total do motif em bp/ tamanho total das sequências analisadas em Mb) do genoma da lagosta *P. echinatus*.

2.5 Resultados e Discussão

Frequência e densidade de seis classes de microssatélites

Após o sequenciamento, obteve-se uma plataforma com 7196 *contigs* e um total de 2959 SSR foi identificado no genoma da lagosta (Tabela 1). Os comprimentos totais e médios foram 67049 bp e 125,39 bp, respectivamente. A frequência relativa e densidade foram de 915,45 loci/Mb e 20743,696 bp/Mb, respectivamente. A contagem, comprimento total, comprimento médio, frequência, densidade e o percentual das seis classes de SSR perfeitas estão resumidos na Tabela 1.

Os mononucleotídeos foi o tipo mais abundante, com a mais alta frequência relativa (399,1 loci/Mb), correspondendo a 43,60% de todos os SSR, seguido de dinucleotídeos (43,26%), trinucleotídeos (7,47%), tetranucleotídeos (3,61%) e pentanucleotídeos (1,49%). Hexanucleotídeos foram os menos abundantes, representando apenas 0,57% de todos os SSRs (Tabela 1).

A prevalência de repetições mononucleotídicas é comum em genomas eucarióticos, sobre as outras classes de nucleotídeos (SHARMA et al., 2007). Já a segunda classe de repetições mais abundante nesse estudo, os dinucleotídeos, com uma porcentagem bem

próximas dos mono- e densidade maior (Tabela 1), foi relatada como a classe de maior frequência em caragueijo da espécie *Ucides cordatus* (ARAÚJO et al., 2015)

Tabela 1 - Contagem, comprimento total, comprimento médio, frequência, densidade e percentual de seis tipos de microssatélites perfeitos na sequência do genoma da lagosta *P. echinatus*, Teresina, PI, 2016.

Tipo	Contagem Total (N)	Comprimento Total (pb)	Comprimento Médio (pb)	Frequência (loci/Mb)	Densidade (pb/Mb)	(%)
Mononucleotídeos	1290	21440	16,62	399,1	6633,132	43,60
Dinucleotídeos	1280	36978	28,89	396,0	11440,296	43,26
Trinucleotídeos	221	5313	24,04	68,3	1643,742	7,47
Tetranucleotídeos	107	2320	21,68	33,1	717,764	3,61
Pentanucleotídeos	44	680	15,45	13,6	210,379	1,49
Hexanucleotídeos	17	318	18,71	5,26	98,383	0,57
Total	2959	67049	125,39	915,3	20743,696	

Abundância e números de repetição para as diferentes categorias de microssatélites

Mononucleotídeos

A categoria de repetição mononucleotídica mais abundante foi o *motif* A (n) com 788 loci respondendo a 61,08 % dos SSRs mononucleotídicos. O comprimento total, a frequência e a densidade do *motif* A (n) foi 13064 bp, 243,79 loci / Mb e 4041,76 bp / Mb, respectivamente, e o comprimento médio foi de 16,58 bp (Tabela 2). No entanto, o *motif* G (n) foi o menos abundante, representando apenas 0,85 % do total.

Os números de repetições de mononucleotídicos variaram de 14 a 67 vezes, sendo que as repetições mais predominantes variaram de 14 a 30 vezes, contando com 1275 SSRs, correspondendo a 98,68% da contagem total de SSRs mononucleotídicos (Fig. 1A). De modo geral o tipo de repetição (A/T)n foi predominante, essa repetição também é comum em alguns fungos como *Agaricus bisporus*, *Volvariella volvacea* e *Coprinus cinereus* (WANG et al., 2014).

Dinucleotídeos

As repetições dinucleotídicas que mais predominaram, incluem AT, AC, CT e categoria GT dos SSRs, repetições semelhantes foram observadas em outros invertebrados, como em caranguejo da espécies *Ucides cordatus* e camarões *peneídeos* (ARAÚJO et al., 2015, DE FREITAS et al., 2007). Os resultados mostraram que as frequências de AT, AC e

CT foram as mais elevadas (103,02 loci/Mb, 88,17 loci/Mb e 80,44 loci/Mb, respectivamente), GT teve a menor frequência de 69,3 loci / Mb (Tabela 2).

Essas quatro categorias de SSR equivalem a 1102 e representaram 86,09% do número total de repetições dinucleotídicas. As repetições dinucleotídicas variaram de 7 a 66 vezes, no entanto, as repetições predominantes variaram de 7 a 30 vezes, somando 1226 SSRs, correspondendo a 95,78% do total de dinucleotídeos (Fig. 1B)

Trinucleotídeos

Os *motifs* de maiores repetição trinucleotídica foram AAT, ATT, CCT e CTT. Os resultados mostraram que as frequências de AAT e ATT foram as mais elevadas (11,76 loci/Mb, 11,14 loci/Mb, respectivamente), CTT teve a menor frequência de 4,64 loci / Mb (Tabela 2), frequências de AAT foram observadas em nematóides *Meloidogyne hapla*, e *Pristionchus pacificus* (CASTAGNONE-SERENO et al., 2010), e ATT foi uma das repetições mais frequentes no camarão chinês (*Fenneropenaeus chinensis*) (WANG; LI; XIANG, 2005)

Essas quatro categorias de SSR equivalem a 110 e representaram 49,77 % do número total de repetições trinucleotídica. Os números de repetição dos SSRs trinucleotídicos variou entre 5 e 37 vezes, as repetições predominantes variou de 5 a 18 vezes o que equivale a 211 SSRs e representam 95,47% da contagem total de trinucleotídeos SSR (Fig. 1C).

Tetranucleotídeos

Um total de 47 categorias de repetições tetranucleotídicas foi obtido neste estudo. Análise de frequências e de densidades de cada categoria tetramérica repetidas revelou que ATCT, AGAT, AATA e CTGT foram predominantes em todo a genoma, com frequência de 7,43 loci / Mb, 4,64 loci / Mb, 3,4 loci / Mb, 3,09 loci / Mb (Tabela 2), respectivamente.

Essas quatro categorias de repetições tetramérica correspondem a 60 SSRs do total de tetranucleotídios. As repetições de SSRs tetranucleotídicos variou entre 4 e 17 vezes, no entanto as repetições mais predominantes variaram de 4 a 9 vezes, o que representa 93,45% da contagem total de tetranucleotídeos SSR (Fig. 1D). O número e distribuição dos 13 *motifs* mais frequentes das repetições tetranucleotídicas estão resumidos na Figura 2.

Entre as seis classes de marcadores microssatélites, os tetranucleotídios são os marcadores mais selecionados em estudos populacionais, pois geram melhor visualização nas imagens e análise do tamanho dos fragmentos.

Pentanucleotídeos

As categorias de repetição pentanucleotídica que mais se repetiram foram ACAAC e TGTTT e apresentaram uma maior frequência de 1,86 loci / Mb e 1,55 loci / Mb, e densidade de 27,84 pb / Mb e 23,2 bp / Mb (Tabela 2), respectivamente.

Os números de repetição de SSRs pentanucleotídeos variou entre 4 e 6 vezes, correspondendo a 44 SSRs, representando assim 100 % das repetições pentanucleotídica (Fig. 1E).

Hexanucleotídeos

As frequências de todas as categorias de repetição hexanucleotídica foram menores do que os cinco tipos de repetição acima, com 0,94 a 0,31 loci / Mb. Os tempos de repetição dos SSRs hexanucleotídicos variou entre 3 a 4 vezes. No entanto, os tempos de repetição predominante variou entre 3 a 4, correspondendo a 17 SSRs, representando 100% (Fig. 1F). O número, comprimento total, frequência e densidade de todos os *motifs* encontrados em *Panulirus echinatus*, consta no APÊNDICE A.

Tabela 2 - *Motif*, contagem total, comprimento total, comprimento médio, frequência e densidade dos *motifs* de maior frequência em *P. echinatus*, Teresina, PI, 2016.

<i>Motif</i>	Contagem Total	Comprimento Total (pb)	Comprimento Médio(pb)	Frequência (loci/Mb)	Densidade (pb/Mb)
Mononucleotídeos					
A	788	13064	16,58	243,79	4041,76
T	463	7585	16,38	143,24	2346,66
C	28	604	21,57	8,66	186,87
G	11	187	17	3,4	57,85
Dinucleotídeos					
AT	333	10584	31,78	103,02	3274,49
AC	285	7108	24,94	88,17	2199,08
CT	260	8112	31,2	80,44	2509,7
GT	224	5364	23,95	69,3	1659,52
Trinucleotídeos					
AAT	38	1236	32,53	11,76	382,4
ATT	36	1068	29,67	11,14	330,42
CCT	21	381	18,14	6,5	117,87
CTT	15	411	27,4	4,64	127,16
Tetranucleotídeos					
ATCT	24	700	29,17	7,43	216,57
AGAT	15	328	21,87	4,64	101,48
AATA	11	236	21,45	3,4	73,01
CTGT	10	184	18,4	3,09	56,93
Pentanucleotídeos					
ACAAC	6	90	15	1,86	27,84
TGTTT	5	75	15	1,55	23,2
AGGTT	4	65	16,25	1,24	20,11
ACACC	3	45	15	0,93	13,92
Hexanucleotídeos					
ATAATG	3	60	20	0,93	18,56
ACCCAC	2	36	18	0,62	11,14
ATGTGT	2	36	18	0,62	11,14
ACCACA	1	18	18	0,31	5,57

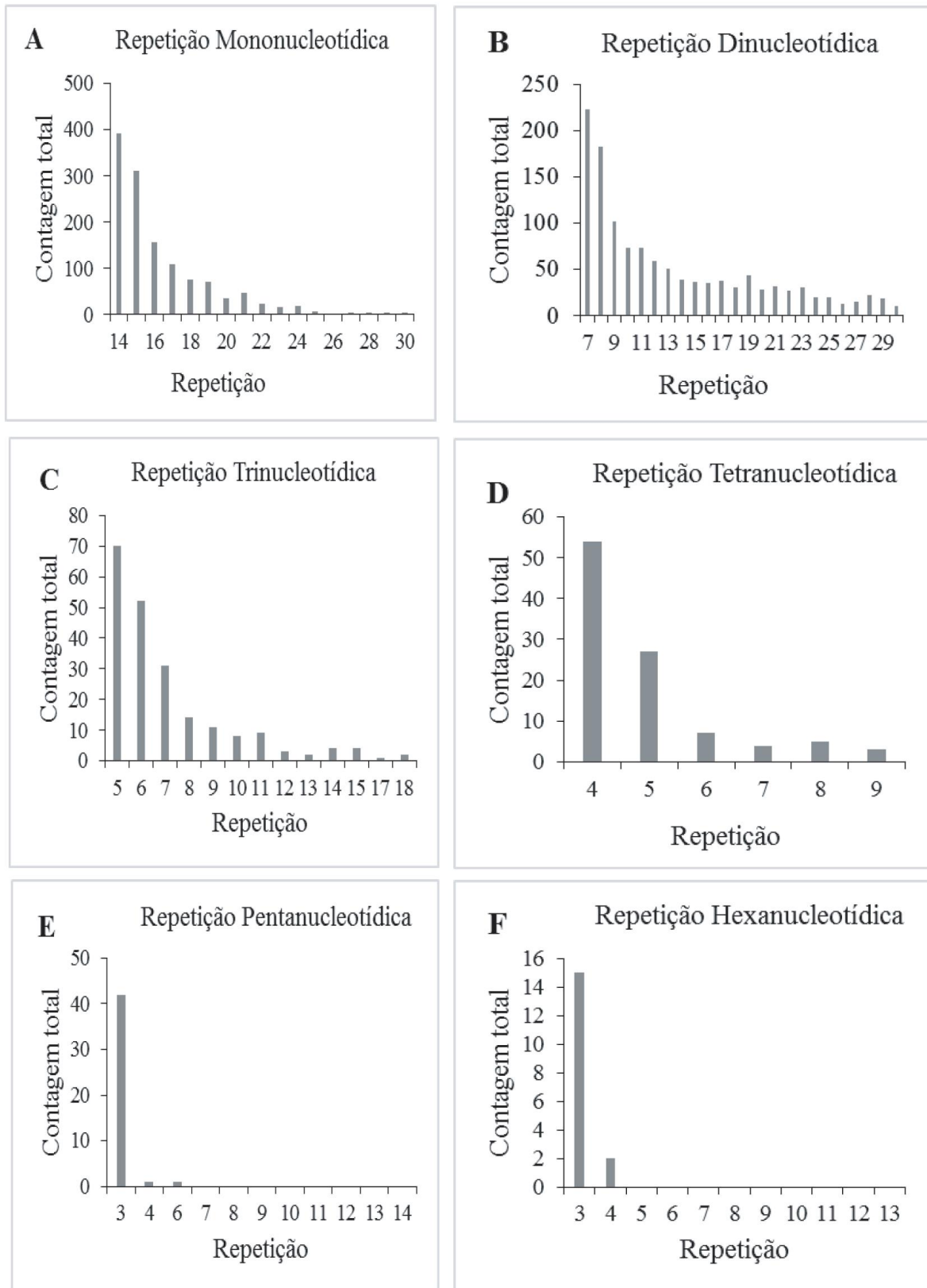


Figura 1 -Contagem total e número de repetições de diferentes tipos de SSRs no genoma da lagosta *Panulirus echinatus*.

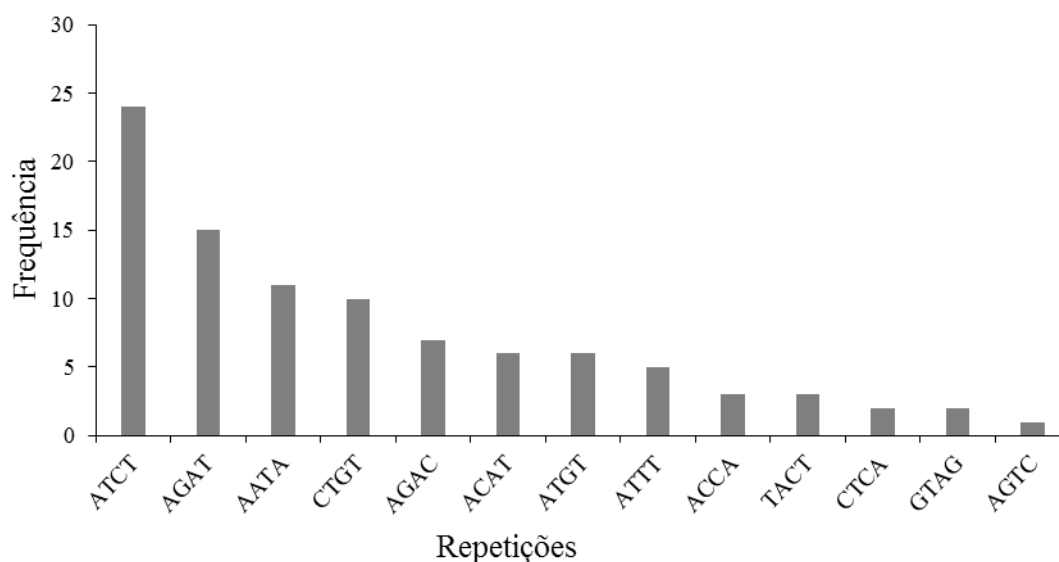


Figura 2 - Repetições tetranucleotídicas mais frequentes em *P. echinatus*.

Os microssatélites formam grupos em seis classes quanto à natureza de sua repetição: composto, complexo, composto interrompido, complexo interrompido, puro interrompido e puros. Em *P.echinatus* os microssatélites puros foram os mais abundantes (2460), seguidos por composto interrompido (88), puro interrompido (67), composto (43), complexo interrompido (29) e complexo (2), o comprimento total, frequência e densidade foi de 52846, 761,08 e 16349,6 respectivamente (Tabela 3).

Tabela 3 - Características dos tipos de microssatélites em *P.echinatus*, Teresina, PI, 2016.

Tipo	Composto	Complexo	Composto Interrompido	Complexo Interrompido	Puro Interrompido	Puro
Contagem Total	43	2	88	29	67	2460
Comprimento Total (pb)	2679	169	5123	2952	4068	52846
Comprimento médio (pb)	62,3	84,5	58,22	101,79	60,72	21,48
Frequência (loci/Mb)	13,3	0,62	27,23	8,97	20,73	761,08
Densidade (pb/Mb)	828,83	52,29	1584,96	913,29	1258,56	16349,6

Os marcadores microssatélites tem sido utilizados com sucesso em alguns estudos populacionais com crustáceos (BEACHAM et al, 2008;. BALL et al, 2003; KENCHINGTON et al, 2009), mas um número crescente de relatórios sobre as dificuldades no isolamento, otimização e seleção para amplificação vem sendo relatado com frequência (BROOKER et al, 2000;. MOORE et al, 1999; ROBAINAS et al, 2002; WEETMAN et al., 2007).

As dificuldades com o desenvolvimento de microssatélites é predominantemente associada à sua baixa abundância no genoma de determinadas espécies (FLETCHER; BAILIE; PRODÖHL, 2010), por tanto conhecer a distribuição e os de tipos microssatélites existentes no genoma facilita o seu isolamento.

2.6 Conclusão

Com a abordagem realizada nesse estudo foi possível o sequenciamento do genoma da lagosta *P. echinatus*, utilizando a técnica de sequenciamento de nova geração, NSG, ou seja, o conhecimento da disposição dos nucleotídios ao longo do seu genoma.

Foi possível também, realizar a partir do genoma sequenciado, busca por microssatélites perfeitos e imperfeitos, onde encontramos as seis classes de microssatélites: mono-, di-, tri-, tetra-, pent- e hexanucleotídios, com suas abundâncias e repetições para as diferentes categorias, sendo os mono- e dinucleotídios os mais abundantes.

Com os dados obtidos a partir dessa busca de distribuição de SSRs no genoma da lagosta *P.echinatus* será possível o desenvolvimento dos primeiros microssatélites para essa lagosta, que são inexistentes até o momento e que será utilizado em estudos populacionais da espécie.

2.7 Referências

ARAÚJO, E. S.; BRITTO, F. B.; CARVALHO, A. M. F.; ORASMO, G. R.; DINIZ, F. M. Isolation and identification of a new set of microsatellite loci from *Ucides cordatus* genome. **Archives of Biological Sciences**, 67(4), 1369-1376, 2015.

BARRETO, A. DO V; KATSURAGAWA, M. Estádios de desenvolvimento dos órgãos reprodutivos dos machos de *Panulirus echinatus* Smith (Decapoda: Palinuridae). **Revista Brasileira de Zoologia**, v. 25, n. 1, p. 74–82, 2008.

BACHMANN, L.; BAREI, P.; TOMIUK, J. Allelic variation , fragment length analyses and population genetic models : a case study on *Drosophila* microsatellites *. v. 42, n. May, p. 215–222, 2004.

BEACHAM, T. D.; SUPRNAULT. J.; MILLER, K. M. Population structure of Dungeness crab (*Cancer magister*) in British Colombia. **Journal of Shellfish Research**, 27: 901-906, 2008.

BALL, A. O.; CHAPMAN R. W. Population genetic analysis of white shrimp, *Litopenaeus setiferus*, using microsatellite genetic markers. **Molecular Ecology**, 12: 2319-2330, 2003.

BROOKER, A. L.; BENZIE, J. A. H.; BLAIR, D.; VERSINI, J-J. Population structure of the giant tiger prawn *Penaeus monodon* in Australian waters, determined using the microsatellite markers. **Marine Biology**, 136:149-157, 2000.

CASTAGNONE-SERENO, P. et al. Genome-wide survey and analysis of microsatellites in nematodes, with a focus on the plant-parasitic species *Meloidogyne incognita*. **BMC genomics**, v. 11, n. 1, p. 598, 2010.

CAIXETA, E.T.; OLIVEIRA, A.C.B.; BRITO, G.G.; SAKIYAMA, N.S. Tipos de Marcadores Moleculares. In: BORÉM, A.; CAIXETA, E.T. (Org.). Marcadores moleculares. 2. ed. Viçosa: Editora Jard., v. 1. p.11-101, 2009.

CHAMBERS, G. K.; MACAVOY, E. S. Microsatellites: Consensus and controversy. **Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology**, v. 126, n. 4, p. 455–476, 2000.

DE FREITAS, P. D. et al. In silico analysis of polymorphic microsatellites in penaeid shrimp and construction of a free-access database. **Genetics and Molecular Biology**, v. 30, n. 4, p. 1194–1197, 2007.

DU, L. et al. MSDB: A user-friendly program for reporting distribution and building databases of microsatellites from genome sequences. **Journal of Heredity**, v. 104, n. 1, p. 154–157, 2013.

FLETCHER, H.; BAILIE, D. A.; PRODÖHL, P. A. High Incidence of Cryptic Repeated Elements in Microsatellite Flanking Regions of Galatheid Genomes and Its Practical Implications for Molecular Marker Development. **Journal of Crustacean Biology**, v. 30, n. 4, p. 664–672, 2010.

JARNE, P.; LAGODA, P. J. L. Microsatellites, from molecules to populations and back. **Trends in Ecology and Evolution**, v. 11, n. 10, p. 424–429, 1996.

- JAYASHREE, B. et al. A database of simple sequence repeats from cereal and legume expressed sequence tags mined in silico: survey and evaluation. **In silico biology**, v. 6, n. November, p. 607–620, 2006.
- KENCHINGTON, E. L.; Harding, G. C.; Jones, M. W.; PRODOHL, P. A. 2009. Pleistocene glaciation events shape genetic structure across the range of the American lobster, *Homarus americanus*. **Molecular Ecology**, 18: 1654-1667, 2009.
- MOORE, S. S.; WHAN, V.; DAVIS, G. P.; BYRNE, K.; HETZEL, D. J. Z.; PRESTON, N. The development and application of genetic markers for the Kumura prawn *Penaeus japonicus*. **Aquaculture**, 173: 19-32, 1999.
- ROBAINAS, A.; MONNEROT, M.; SOLIGNAC, M.; DENNENBOUY, N.; ESPINOSA, G.; MACHADO, E. G. Microsatellite loci from the pink shrimp *Farfantepenaeus notialis* (Crustacea, Decapoda). **Molecular Ecology**, Notes 2: 344-345, 2002.
- SAMBROOK, J.; RUSSELL, D.W. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 2001.
- SELKOE, K. A.; TOONEN, R. J. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. **Ecology Letters**, v. 9, n. 5, p. 615–629, 2006.
- SHARMA, P. C.; GROVER, A.; KAHL, G. Mining microsatellites in eukaryotic genomes. **Trends in Biotechnology**, v. 25, n. 11, p. 490–498, 2007.
- WANG, Y. et al. Microsatellites in the genome of the edible mushroom, *Volvariella volvacea*. **BioMed Research International**, v. 2014, 2014.
- WANG, H.; LI, F.; XIANG, J. Polymorphic EST-SSR markers and their mode of inheritance in *Fenneropenaeus chinensis*. **Aquaculture**, v. 249, n. 1-4, p. 107–114, 2005.
- WEETMAN, D.; RUGGIERO, A.; MARIANI, S.; SHAW, P. W.; LAWLER, A. R.; HAUSER, L. Hierarchical population genetic structure in the commercially exploited shrimp *Crangon crangon* identified by AFLP analysis. **Marine Biology**, 151: 565-575, 2007.
- ZANE, L.; BARGELLONI, L.; PATARNELLO, T. Strategies for microsatellite isolation: A review. **Molecular Ecology**, v. 11, n. 1, p. 1–16, 2002.

Capítulo 3

Desenvolvimento de marcadores de sequências simples repetidas no genoma da lagosta *Panulirus echinatus* por meio do sequenciamento de nova geração e uma análise preliminar da diversidade genética da espécie

***Artigo “Development of genome-wide informative simple sequence repeat markers and a snapshot of the genetic diversity of the spiny lobster, *Panulirus echinatus* using next generation sequencing”, submetido ao periódico: *Molecular Genetics and Genomics*, ISSN: 1617-4615 (Print) 1617-4623 (Online), Fator de Impacto: 2.728.**

3.1 Resumo

Marcadores de sequências simples repetidas (SSR) de tri- e tetranucleotídeos foram desenvolvidos para a lagosta espinhosa *Panulirus echinatus* usando a tecnologia de sequenciamento de nova geração (NGS). Baseado em 7.196 *contigs* gerados por NGS, 27 loci SSR foram inicialmente selecionadas para o desenho dos iniciadores, dos quais 20 consistentemente amplificaram bandas específicas entre os indivíduos. Polimorfismo foi identificado em 17 locos de três populações insulares no Oceano Atlântico, São Pedro e São Paulo, Atol das Rocas, e Cabo Verde, fornecendo uma visão geral da variação genética na espécie. O número de alelos por loco variou de 2 a 8 e a média da heterozigosidade observada e esperada de 0,418 e 0,587, respectivamente. O conteúdo de informação polimórfica (PIC) variou entre 0,173-0,811. Desvios significativos do equilíbrio de Hardy-Weinberg foram encontrados em oito loci (*Pech-3*, *Pech-4*, *Pech-6*, *Pech-8*, *Pech-11*, *Pech-14*, *Pech-15*, *Pech-16*), mas não foi detectada nenhuma evidência de desequilíbrio de ligação. Frequência de alelos nulos foi $< 0,2$ para a maioria dos locos, exceto para *Pech-11* e *Pech-15*. Uma análise de diversidade genética preliminar revelou baixos valores de estimativas de diferenciação da população ($F_{ST} = 0,039$, $R_{ST} = 0,036$, e $D_{EST} = 0,089$). PCoA e Bayesian clustering não conseguiram distinguir as populações de lagosta uma das outras evitando separação das espécies em grupos distintos. Estes marcadores moleculares são o primeiro relato de SSR para a espécie e possivelmente será usado no futuro para pesquisas de conservação genética de *P. echinatus*.

Palavras-chave: Genética de populações, microsatélites, conectividade, lagosta espinhosa, conservação.

3.2 Abstract

Simple sequence repeat markers (SSR) of tri- and tetranucleotide motifs were developed for the brown spiny lobster *Panulirus echinatus* using next-generation sequencing (NGS). Based on 7,196 contigs generated by NGS, 27 SSR loci were initially selected for *primer* designing, of which 20 consistently amplified specific bands across individuals. Polymorphism was identified in 17 loci for three insular populations in the Atlantic Ocean, Saint Peter and Saint Paul Rocks, The Rocas Atoll, and Cape Verde, providing an overview of the genetic variation in the species. The number of alleles per locus ranged from 2 to 8 and observed and expected heterozygosities averaged 0.418 and 0.587, respectively. Polymorphic information content (PIC) ranged from 0.173 to 0.811. Significant deviations from Hardy–Weinberg equilibrium were found in eight loci (*Pech-3*, *Pech-4*, *Pech-6*, *Pech-8*, *Pech-11*, *Pech-14*, *Pech-15*, *Pech-16*), but no evidence of linkage disequilibrium was detected. Frequency of null alleles was < 0.2 for most loci, except *Pech-11* and *Pech-15*. A preliminary genetic diversity analysis revealed low values for population differentiation estimates ($F_{ST} = 0.039$, $R_{ST} = 0.036$, and $D_{EST} = 0.089$). PCoA and Bayesian clustering could not distinguish lobster populations from each other avoiding separation of the species into distinct clusters. These molecular markers are the first SSR reported for the species and maybe be used for future conservation genetics research of *P. echinatus*.

Keywords: Population genetics, microsatellites, connectivity, brown spiny lobster, conservation.

3.3 Introdução

A pesca da lagosta é, sem dúvida, uma das pesca de mariscos financeiramente mais valiosas em todo o mundo, especialmente para as lagosta espinhosas. A lagosta *Panulirus echinatus* (SMITH, 1869), ocorre em recifes rochosos do Oceano Atlântico tropical e ocidental e em ilhas do Oceano Atlântico Central (HOLTHUIS, 1991). A espécie é pescada em toda a sua área de distribuição. Contudo, é no Brasil e nas ilhas de Santa Helena e Cabo Verde que registra as maiores atividades de pesca comercial.

Pesca em determinados habitats pode comprometer as populações localmente adaptadas (AGARDY, 2000), resultando em perda de diversidade genética. Não há dados disponíveis sobre as espécies de lagostas que poderia ser usado para indicar se um impacto sobre a população mundial já existe. Portanto, é de suma importância estudos de variabilidade genética das espécies visando uma estratégia de gestão para garantir que as populações continuam a um nível sustentável. Até agora, não existe trabalhos na literatura contendo registros de distribuição dessa espécie, por conseguinte, pouco se sabe sobre a diversidade de populações de *P.echinatus*.

Marcadores de sequências simples repetidas (SSR), também conhecidos como microssatélites de DNA, são marcadores moleculares úteis em estudos populacionais de organismos aquáticos e em estudos de conservação genética, pois apresenta altos níveis de variabilidade interespecífica, em consequência da elevada taxa de mutação destas regiões não codificadoras (WRIGHT; BENTZEN, 1994).

Como não existem marcadores moleculares identificados para *P. echinatus* até o momento e a espécie tem limitada informação genômica, uma alternativa é a descoberta de microssatélites no seu genoma e consequentemente o desenho de *primers* desenvolvidos a partir da tecnologia de sequenciamento de nova geração.

Nesse estudo, foram isolados e caracterizados um conjunto de 20 marcadores de sequência simples repetidas de tri- e tetranucleotídeos para *P.echinatus* usando o sequenciamento de nova geração. Além disso, discutimos brevemente o potencial da utilização desses marcadores desenvolvidos para avaliar a variabilidade genética da espécie.

3.4 Material e Métodos

Amostragem e extração de DNA

Amostras de *P. echinatus* foram coletados em três localidades diferentes em ilhas centrais do oceano Atlântico: São Pedro e São Paulo ($0^{\circ} 55' 2''$ N, $29^{\circ} 20' 44''$ W); Atol das Rocas ($3^{\circ} 51' 50''$ S, $33^{\circ} 48' 48''$ W) e Cabo Verde ($14^{\circ} 55' 0''$ N, $23^{\circ} 31' 0''$ W) (Figura 1)

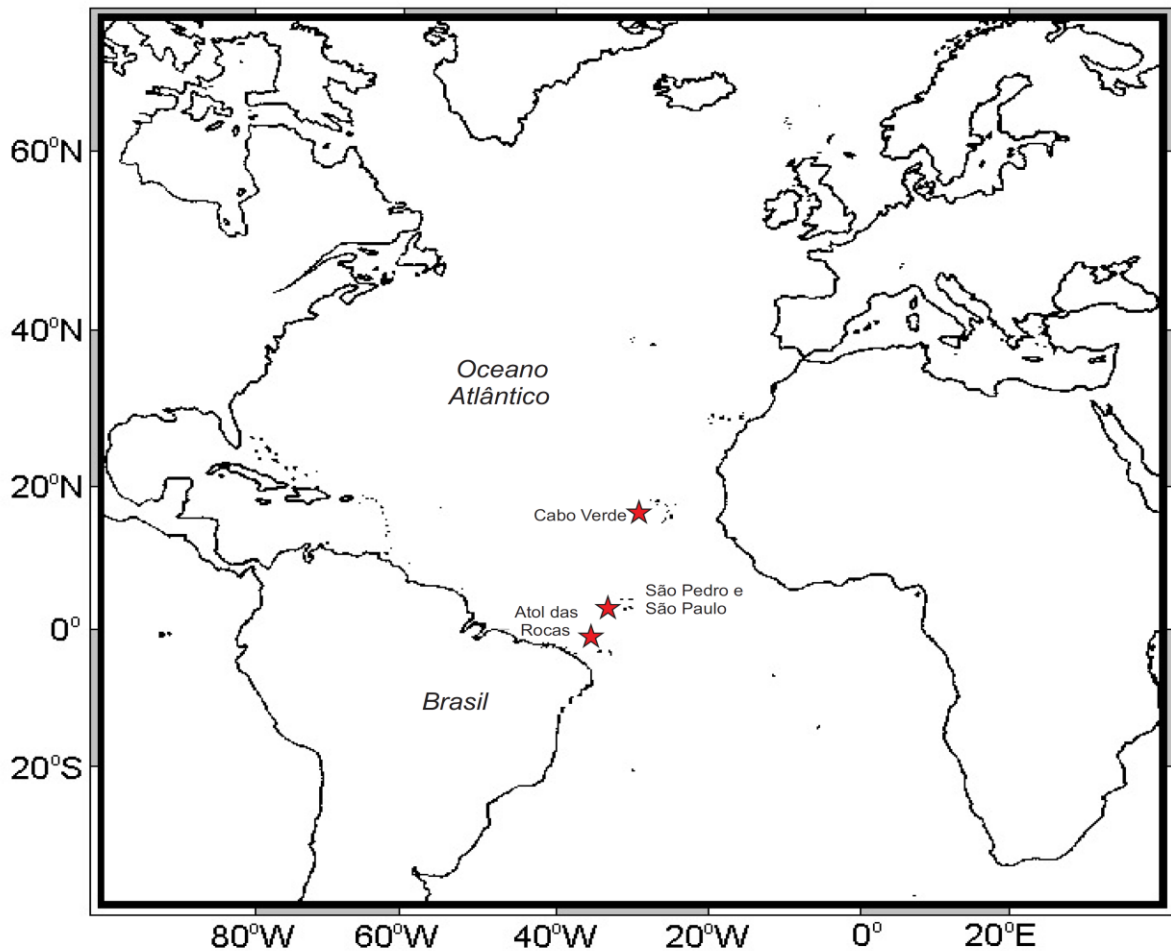


Figura 1 - Identificação dos locais amostrados da espécie *P. echinatus*.

Amostras de tecido muscular das patas (pereiópodos) e das antenas foram coletadas e imediatamente preservadas em etanol a 95%, e armazenadas a -20°C até à sua utilização. O DNA foi extraído do tecido do músculo utilizando o protocolo *HotShot* (*Hot Sodium Hydroxide and Tris*) como descrito por Truett et al. (2000). DNA de alto peso molecular foi isolado por precipitação em etanol e submetido à eletroforese em gel de agarose.

Sequenciamento do DNA da lagosta

O sequenciamento do genoma da *P. echinatus* foi realizado em parceria com a Dalhousie University (Canadá), Department of Biology, Marine Gene Probe Laboratory. Uma biblioteca Illumina *paired-end* foi criada utilizando 1ng de DNA genômico da *P. echinatus*, seguindo o protocolo padrão da Illumina Nextera DNA Sample Preparation Kit (Illumina Inc., San Diego, CA). O sequenciamento do DNA foi conduzido usando o sequenciador Miseq Benchtop (Illumina Inc., San Diego, CA). Os *contigs* foram criados a partir dos dados das sequências emparelhadas pelas extremidades obtidas por meio do CLC Genomics Workbench 7.0.4 (Qiagen).

Desenho dos primers

Todos os *contigs* foram analisados utilizando o software MSDB (Microsatellite Search Building Database) (<http://msdb.biosv.com/>) (DU et al., 2013), para a detecção de possíveis locos microssatélites com pelo menos quatro repetições de *motifs* tri- e tetranucleotídicos para desenhar os pares de *primers* para cada loco detectado nas suas regiões flanqueadoras. Longos trechos de repetição mononucleotídicas foram ignorados para o desenvolvimento do marcador. Os desenhos de *primers* foram realizados com o Primer3 (ROZEN; SKALETSKY, 2000),

Otimização dos SSR, amplificação via PCR e genotipagem

A etapa de otimização, amplificação e genotipagem foi realizada no Laboratório de Biologia Molecular e Biotecnologia da Embrapa Meio Norte. O DNA genômico de quatro indivíduos foram inicialmente utilizados para validar todos os pares de *primers* desenhados usando reação em cadeia da polimerase (PCR). As reações de PCR foram realizadas para um volume final de 10µL, contendo aproximadamente 20ng de DNA genômico, 1,5µL de buffer 10X (Thermo Fisher Scientific Inc), 2,5mM de MgCl₂ (Thermo Fisher Scientific Inc), 200 mM dNTPs (New England BioLabs Inc.), 0,75 µM de *primer*, 2,0µL de Q-solution (Qiagen), 1U de Taq DNA polimerase (Thermo Fisher Scientific Inc).

As reações foram conduzidas em termociclador Veriti 96 Well Thermal Cycler (Applied Biosystems), utilizando as seguintes condições para amplificação: 95°C por 1 minuto, 30 ciclos de 95°C por 1 minuto, a temperatura de anelamento variou conforme cada *primer* específico (50 – 57°C) por 1 minuto e 72°C por 1 minuto, seguidos de um ciclo de extensão

final de 72°C por 7 minutos. Todas as amplificações foram realizadas com um controle negativo no qual foram adicionados todos os componentes, exceto o DNA molde.

Os produtos da amplificação foram analisados em gel de poliacrilamida a 6% e corados com nitrato de prata. Amostras adicionais de lagosta (N = 42) foram genotipadas das três localizações diferentes descritas acima, para obter informações de frequência dos alelos.

Análise dos dados

Os dados obtidos foram inicialmente analisados no MICRO-CHECKER 2.2.3 (VAN OOSTERHOUT et al., 2004) para testar a presença de alelos nulos. O número de alelos (A), a heterozigosidade observada e esperada (H_O e H_E), e o conteúdo de informação polimórfica (PIC) foram determinados utilizando o CERVUS 3.0 (KALINOWSKI; TAPER; MARSHALL, 2007). A riqueza alélica (AR) uma medida do número de alelos por loco independentes do tamanho da população foi calculado pela *FSTAT* versão 2.9.3.2 (GOUDET, 1995). Os testes de equilíbrio de Hardy-Weinberg (HWE) foram realizados utilizando o software GENEPOP (RAYMOND; ROUSSET, 1995). A correção de Bonferroni (RICE, 1989) foi aplicada para avaliar a significância ($P < 0,05$).

A diversidade genética de cada loco foi avaliada no ARLEQUIN ver. 3.5 (EXCOFFIER; LISCHER, 2010), para determinar o valor $\theta (F_{ST})$ para o conjunto da amostra.

Um modelo de mistura “*admixture model*” de análise Bayesiana de agrupamento foi utilizado para inferir possível estruturação populacional utilizando o programa STRUCTURE v.2.3.2 (PRITCHARD; STEPHENS; DONNELLY, 2000). O modelo foi testado usando 1.000.000 interações na cadeia de Monte Carlo Markov (MCMC) após 500.000 interações *burn-in period*. A estimativa do melhor K foi calculada com base em cinco repetições para cada K (1-6) como descrito por Evanno et al. (2005) usando o STRUCTURE HARVESTER v.0.6.92 (<http://taylor0.biology.ucla.edu/structureHarvester/>) (EARL; VONHOLDT, 2012). O programa CLUMPP v.1.1.2 (JAKOBSSON; ROSENBERG, 2007) foi usada para alinhar as cinco repetições do melhor K . O programa DISTRUCT v.1.1 (ROSENBERG, 2004) foi usado para visualizar graficamente os resultados produzidos pelo CLUMPP. A estrutura da população também foi analisada utilizando análise de coordenadas principais (PCoA), o R_{ST} (SLATKIN, 1995), uma medida de diferenciação genética análoga à F_{ST} , e o D_{est} (JOST, 2008) valor de diferenciação real, foi calculado utilizando o GenAIEx v.6.5 (PEAKALL; SMOUSE, 2012).

3.5 Resultados e Discussão

Montagem da sequência, identificação de SSR e desenho de primers

Uma biblioteca genômica, que foi previamente carregada com 16% do MiSeq Reagente Kit v2 com 300 ciclos de execução de sequenciamento, produziu 3.232.259 *reads*, que foram reunidos em 7.196 *contigs*. O software MSDB identificou 2.959 locos de microssatélites, sendo a maioria das repetições mononucleotídeos (1.290; 43,60%) e dinucleotídeos (1280; 43,26%). Para facilidade de imagem e de *scoring*, escolheu examinar apenas locos tri- e tetranucleotídeos, que apresentaram 221 (7,47%) e 107 (3,61%) repetições, respectivamente.

Destes, 27 locos foram selecionados para validação, dos quais 20 consistentemente amplificaram bandas específicas entre os indivíduos. Polimorfismo foi identificado em 17 locos em todo o conjunto de dados, sendo três locos monomórficos (*Pech-18*, *Pech-19* e *Pech-20*) (Tabela 1). No entanto, é esperado que estes locos possam tornar-se polimórficos quando mais indivíduos forem amostrados. Se o monomorfismo for confirmado, estes locos ainda podem ser úteis com uma nova fonte de dados e regiões flaqueadoras, podendo ser usado como ferramenta para novos estudos genéticos (NAZARENO; DOS REIS, 2011; CHATROU et al., 2009). Todas as sequências microssatélites foram depositados no banco de dados NCBI sob os números de acesso GenBank KX237634 a KX237653.

Polimorfismo e validação de marcadores microssatélites

A análise com o programa Micro-Checker identificou frequências de alelos nulos de baixa a moderada em alguns locos, o que não é surpreendente, pois isso é frequentemente observado numa variedade de espécies de invertebrados marinhos (KAUKINEN et al., 2004). Frequências de alelos nulos abaixo de 0,2 são aceitáveis na maioria dos conjuntos de dados de microssatélites (DAKIN; AVISE, 2004). Apenas em dois loci (*Pech-11* e *Pech-15*) esta estimativa foi maior do que 0,2. A razão mais provável para a ocorrência deste fenômeno são os erros de genotipagem, que ocorre durante os *scoring* dos géis (Tabela 2).

A genotipagem de todo o conjunto de dados (n = 42) também revelou 91 alelos em todos os 17 locos polimórficos, variando de 2 (*Pech-10*) a 8 (*Pech-4*), com uma média de $5,4 \pm 1,4$ alelos por loco (Tabela 1). O número de alelos em microssatélites tri- e tetranucleotídeos

em lagostas foi observado por variar de 5,3 para lagosta brasileira *P. Argus* (DINIZ et al., 2005) e 11,9 para *P. cygnus* (KENNINGTON et al. 2010).

O tamanho do alelo no loco mais polimórfico, *Pech-14*, com base em A_R , H_e e PIC, variou 160-178 pb. Para o loco menos polimórfico, *Pech-10*, o tamanho alelo variou 78-86 pb, um loco tri- e tetranucleotídeo, respectivamente (Tabela 2).

O nível de polimorfismo de cada loco também foi avaliado pela riqueza alélica (A_R), heterozigosidades (H) e o conteúdo de informação polimórfica (PIC) (Tabela 2). Os valores de riqueza alélica variou 1,8 a 6,2 (média de 3,8), enquanto que os valores de PIC variaram entre 0,173 e 0,811. A média do PIC (0,537) caracteriza estes conjuntos de locos microssatélites como marcadores altamente informativos, como descrito por Botstein et al., (1980). Em geral a média da heterozigosidade observada e esperada foi estimada em 0,418 e 0,587, respectivamente. Esses valores foram superiores ao observado em *Panulirus homarus* em que a heterozigosidade observada (H_o) e esperada (H_E) variou de 0,265 e 0,387, respectivamente (DELGHANDI et al., 2015). No entanto, as heterozigosidades foram inferiores quando comparadas com outras lagostas, tais como *P. argus* (DINIZ et al 2004; DINIZ et al 2005; TRINGALI et al 2008), *P. ornatus* (DAO et al., 2013), *P. cygnus* (KENNINGTON et al. 2010), *P guttatus* (TRUELOVE; BEHRINGER; PREZIOSI, 2016), *P. interruptus* (BEN-HORIN et al. 2009).

Oito locos de microssatélites apresentaram probabilidades significativas de equilíbrio de Hardy-Weinberg após as correções de Bonferroni, P-valor ajustado <0,0029. Nestes casos, o desvio ocorreu sob a forma de deficiência heterozigótica, talvez resultante do pequeno tamanho da amostra ou considerando o grande potencial de fluxo gênico nesta espécie, o efeito Wahlund (JOHNSON; BLACK, 1984). Além disso, alelos nulos, mesmo em baixas frequências podem ter desempenhado um papel importante no equilíbrio de HWE. Alelos nulos é comum em marcadores microssatélites de muitos invertebrados marinhos (DAILIANIS et al., 2011).

Tabela 1 - Sequência dos iniciadores desenhados, estrutura de repetição, temperatura de anelamento (Ta), número de alelos (A), tamanho do alelo (bp) e número de acesso no GenBank, Teresina, PI, 2016.

Loco	Sequências dos <i>primer</i> (5' - 3')	Estrutura de repetição	Ta (°C)	A	Tamanho (bp)	Acesso no GenBank
Pech-1	F: AAGCAGAACGAACTCACCGTA R: TTTCTTCATGGTCAGAAGGGT	(CTA) ₁₁	55	5	119-146	KX237634
Pech-2	F: TCCTGGTTACGGAATAATGACC R: CTGTTGTTGTTTCCCATACACG	(CTG) ₉	55	6	109-115	KX237635
Pech-3	F: ACATAGATGGATGGGTAGATAGAC R: CTAAGTCTTTCTTAACTGCCT	(GATA) ₁₂	57	6	118-134	KX237636
Pech-4	F: GGGCCTTTCTTCATCTATTCC R: GTCCAGGGCAAGATTGTGATA	(ATA) ₁₂	57	8	183-196	KX237637
Pech-5	F: TGTGTGTGTGTGTTTGTGTGTA R: CACACATATACGAGTCAAAACGAA	(TCTG) ₈	55	5	96-114	KX237638
Pech-6	F: CAATCTTAAAAGCAGACCAACG R: GAGGAAGAATAGGTAGAAATGCAGA	(ATA) ₁₅	55	6	134-158	KX237639
Pech-7	F: AGAGGCAAAGGGGTTTCATCTA R: TATACCACCATCACCAGCCATA	(ATCT) ₆	55	5	118-150	KX237640
Pech-8	F: TATTCTTGTTCCTTTCGCCC R: GAGTAGGAGGAGGAGGAAGAAGA	(TCT) ₁₄	55	3	117-123	KX237641
Pech-9	F: TCTATCTATCTATCTTTCATTAACGC R: GTTTTCGTCTCCTTTCCTGTGA	(TATA) ₇ (TA) ₄	55	6	72-86	KX237642
Pech-10	F: TGTGTGGTGAGAGGAGAGTTGA R: CTGGAAAGGATATGTTCAAAGGAG	(GATA) ₈	56	2	78-86	KX237643
Pech-11	F: TCAGGTGGACTGTGACCAAG R: CATATTATCATTATTATTATAATTATTATTAT	(TAA) ₄ (TAA) ₁₀	52	5	190-211	KX237644
Pech-12	F: CCTCCACTTTGAAGCTCTG R: TTATCATTAGTATCATTATGATCAT	(ATA) ₁₇	52	6	99-117	KX237645
Pech-13	F: CATAACGTGATGATAATGATGGTG R: GAGAAGAGGTAATAAATACTACTGGC	(ATA) ₁₀	57	5	101-107	KX237646
Pech-14	F: TTTTGCTTGTGTGGCAAGATAC R: ACTCCGCCCTTGTTTAAGTTT	(ATA) ₁₃	57	7	160-178	KX237647
Pech-15	F: GTATTGGTGCCTGTTTATGC R: TGATGATGATGATAATGATAATAATGA	(TTA) ₁₅	55	6	162-194	KX237648
Pech-16	F: GCTTTCAGTATATGTCTATCTACCTG R: GATAGATAAACAGAGAAATAGATAAA	(TATC) ₁₁	55	6	87-107	KX237649
Pech-17	F: GCTGCCAAGCATTACGTTTT R: CTGAATGAGTACCTGCGTGC	(TATT) ₉	55	4	190-214	KX237650
Pech-18	F: CTCTCTCTATCTGTCTATCTATCTATC R: TGGCATCAGAGGTATAGAAAGACA	(TATC) ₁₁ (TCTA) ₁₇	56	1	175	KX237651
Pech-19	F: CACACACACATCATTGTATTATCATC R: TGATAACAACAACAACACCAACA	(TAA) ₁₄	52	1	136	KX237652
Pech-20	F: CGGGAAACAGCGATCATGTA R: GGGGAAGATGTGTGTAGTGCAA	(TAA) ₂₇	57	1	194	KX237653

Varição genética em Panulirus echinatus

Um nível de polimorfismo ligeiramente maior foi revelado para população de São Pedro e São Paulo, com base nas heterozigosidades (H) e no conteúdo de informação polimórfica (PIC) (Tabela 3). Essa alteração é provável devido ao fato desse arquipélago está situado entre Atol das Rocas e Cabo Verde, e potenciais movimentos larvais ocorram de ambos os locais no sentido São Pedro e São Paulo, aqui favorecido pelas correntes e marítimas na região.

As baixas estimativas para F_{ST} , R_{ST} e D_{est} encontradas em *P.echinatus*, 0,039, 0,036 e 0,089, respectivamente, são um indicativo de pouca (baixo nível) diferenciação entre as populações pré-definidas (Tabela 3). No entanto, medidas de diferenciação genética podem ser superestimadas por alelos nulos, mesmo sem modificar o significado do resultado geral (CHAPUIS; ESTOUP 2007; CARLSSON, 2008). Portanto, uma investigação mais extensa deve ser realizada para confirmar esse padrão observado.

Acredita-se que a troca de larvas entre as populações de *P.echinatus* possa ser grande na região oceânica estudada, as correntes oceânicas desempenham um importante papel no transporte de larvas para longas distâncias das áreas de desovas originais e, conseqüentemente, os padrões de diferenciação genética são suscetíveis de ser sutis ou inexistentes. Altas taxas de dispersão também pode ter contribuído para os níveis de diferenciação populacional. Além disso, a espécie tem uma fase larval planctônica longa com duração de cerca de 12 meses (GÓES; LINS-OLIVEIRA; JE, 2009).

De fato, muitos estudos genéticos com lagostas mostraram pouca evidência de subdivisão entre as populações geograficamente separadas. Baixas estimativas de F_{ST} foram relatadas em outras espécies de lagosta, como *P. ornatus* (DAO et al., 2015), *P. argus* (TRUELOVE et al., 2015) e *P. interruptus* (BEN-HORIN et al., 2009).

A análise de coordenadas principais (PCoA) não revelou uma clara separação das espécies em grupos distintos (Figura 2), e todos os indivíduos das populações apresentaram-se misturados, corroborando com os resultados obtidos na análise Bayesiana de agrupamento (Figura 3).

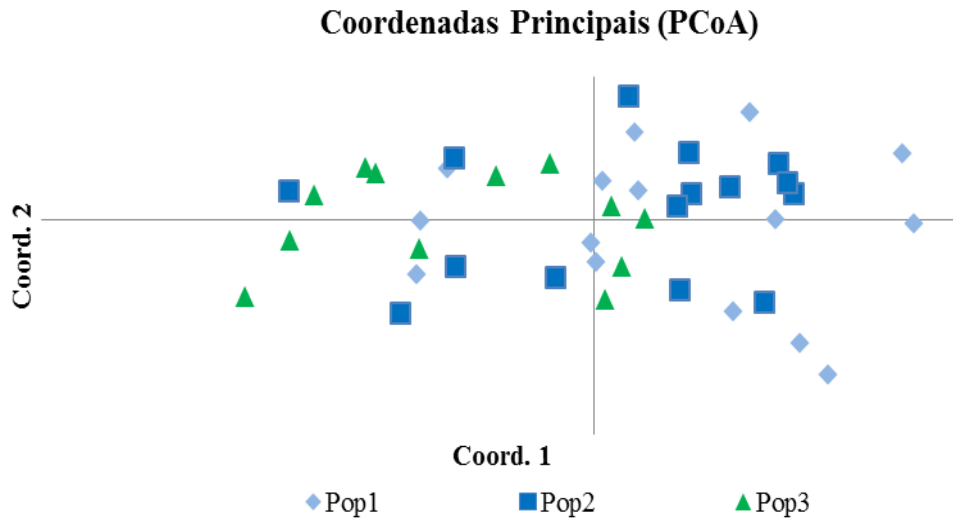


Figura 2 - Análise de coordenadas principais da variação de microsatélites em *P. echinatus*.

Tabela 2 - Estatísticas de diversidade genética para lagosta espinhosa *Panulirus echinatus*, Teresina, PI, 2016

Locos	Todas as populações (N = 42)						
	A_R	H_o	H_E	PIC	P_{HWE}	F_{IS}	F_{Nulls}
<i>Pech-1</i>	3,6	0,667	0,656	0,583	0,2629	-0,0718	-0,0112
<i>Pech-2</i>	4,0	0,595	0,660	0,589	0,1778	0,0806	0,0345
<i>Pech-3</i>	3,9	0,395	0,654	0,583	0,0006*	0,3998	0,1522
<i>Pech-4</i>	4,5	0,472	0,698	0,644	0,0004*	0,3155	0,1278
<i>Pech-5</i>	4,2	0,564	0,683	0,628	0,0616	0,1261	0,0656
<i>Pech-6</i>	4,7	0,415	0,755	0,704	0,0000*	0,4051	0,1899
<i>Pech-7</i>	3,3	0,424	0,493	0,439	0,3548	0,0820	0,0414
<i>Pech-8</i>	2,4	0,095	0,293	0,266	0,0000*	0,8609	0,1504
<i>Pech-9</i>	4,0	0,524	0,618	0,568	0,2030	0,1431	0,0542
<i>Pech-10</i>	1,8	0,214	0,194	0,173	1,0000	-0,0943	-0,0193
<i>Pech-11</i>	4,4	0,281	0,730	0,672	0,0000*	0,5892	0,2543
<i>Pech-12</i>	3,8	0,469	0,515	0,470	0,1240	0,0639	0,0256
<i>Pech-13</i>	3,5	0,524	0,616	0,538	0,1258	0,1517	0,0527
<i>Pech-14</i>	6,2	0,595	0,842	0,811	0,0005*	0,2696	0,1293
<i>Pech-15</i>	3,9	0,118	0,549	0,506	0,0000*	0,7747	0,2746
<i>Pech-16</i>	4,8	0,405	0,696	0,652	0,0010*	0,4045	0,1666
<i>Pech-17</i>	2,6	0,357	0,343	0,306	0,6560	-0,0275	-0,0140
Média	3,8	0,418	0,587	0,537	-	0,2528	-

Riqueza alélica (A_R), Heterozigosidade observada (H_o), Heterozigosidade esperada (H_E), Conteúdo de informação polimórfica (PIC), Probabilidade de equilíbrio de Hardy-Weinberg (p_{HWE}), *Locus que se desviaram significativamente de HWE após a correção de Bonferroni (P crítica ajustado <0,0029). F_{IS} : coeficiente de endogamia, Frequência de alelos nulos (F_{Nulls}).

Tabela 3 -Variabilidade de 17 locos de microssatélites e F-estatísticas em populações *Panulirus echinatus*, Teresina, PI, 2016.

Locos	Atol das Rocas			São Pedro e São Paulo			Cabo Verde			F_{ST}	R_{ST}	D_{est}
	H_O	H_E	PIC	H_O	H_E	PIC	H_O	H_E	PIC			
<i>Pech-1</i>	0,467	0,531	0,475	0,933	0,674	0,589	0,467	0,531	0,475	0,07429	0,070	0,125
<i>Pech-2</i>	0,467	0,549	0,421	0,667	0,756	0,682	0,467	0,549	0,421	0,02968	-0,013	0,052
<i>Pech-3</i>	0,357	0,632	0,528	0,308	0,588	0,496	0,357	0,632	0,528	-0,00100	-0,065	-0,003
<i>Pech-4</i>	0,429	0,638	0,541	0,571	0,717	0,649	0,429	0,638	0,541	0,02325	0,123	0,082
<i>Pech-5</i>	0,500	0,518	0,443	0,600	0,655	0,583	0,500	0,518	0,443	0,08107	0,080	0,169
<i>Pech-6</i>	0,467	0,701	0,619	0,467	0,726	0,652	0,467	0,701	0,619	0,11736	0,104	0,312
<i>Pech-7</i>	0,333	0,373	0,338	0,500	0,598	0,483	0,333	0,373	0,338	0,09271	-0,068	0,081
<i>Pech-8</i>	0,067	0,186	0,164	0,000	0,349	0,309	0,067	0,186	0,164	-0,04239	-0,046	-0,018
<i>Pech-9</i>	0,467	0,540	0,492	0,667	0,644	0,566	0,467	0,540	0,492	0,01974	0,096	0,033
<i>Pech-10</i>	0,133	0,129	0,117	0,267	0,239	0,204	0,133	0,129	0,117	-0,01867	-0,019	-0,005
<i>Pech-11</i>	0,100	0,511	0,406	0,143	0,772	0,704	0,100	0,511	0,406	0,10222	0,070	0,202
<i>Pech-12</i>	0,500	0,416	0,347	0,500	0,753	0,677	0,500	0,416	0,347	0,04287	0,120	0,046
<i>Pech-13</i>	0,600	0,602	0,519	0,467	0,572	0,460	0,600	0,602	0,519	-0,00106	0,012	0,001
<i>Pech-14</i>	0,667	0,766	0,709	0,400	0,841	0,786	0,667	0,766	0,709	0,05190	0,096	0,252
<i>Pech-15</i>	0,000	0,304	0,272	0,154	0,545	0,449	0,000	0,304	0,272	0,08538	-0,043	0,113
<i>Pech-16</i>	0,462	0,726	0,656	0,267	0,671	0,608	0,462	0,726	0,656	0,03976	0,119	0,098
<i>Pech-17</i>	0,333	0,297	0,260	0,333	0,402	0,361	0,333	0,297	0,260	-0,02275	-0,024	-0,012
Média	0,373	0,495	0,429	0,426	0,617	0,544	0,373	0,495	0,429	0,0396	0,036	0,089

H_O : heterozigidade observada; H_E : heterozigidade esperada; PIC : conteúdo de informação polimórfica; F_{ST} : Diversidade genética; R_{ST} : Medida de diferenciação genética; D_{est} : Diferenciação Real.

A análise Bayesiana de agrupamento foi incapaz de identificar qualquer estrutura populacional resultante do conjunto de dados e, portanto, as três populações pré-definidas não são geograficamente claras. Quando o K foi selecionado entre 2 a 6, a proporção de indivíduos atribuídos a cada grupo assumiu uma forma quase uniforme, indicando que a maioria das adesões são indivíduos bastante misturados, com alguns originados principalmente de uma população (Figura 3).

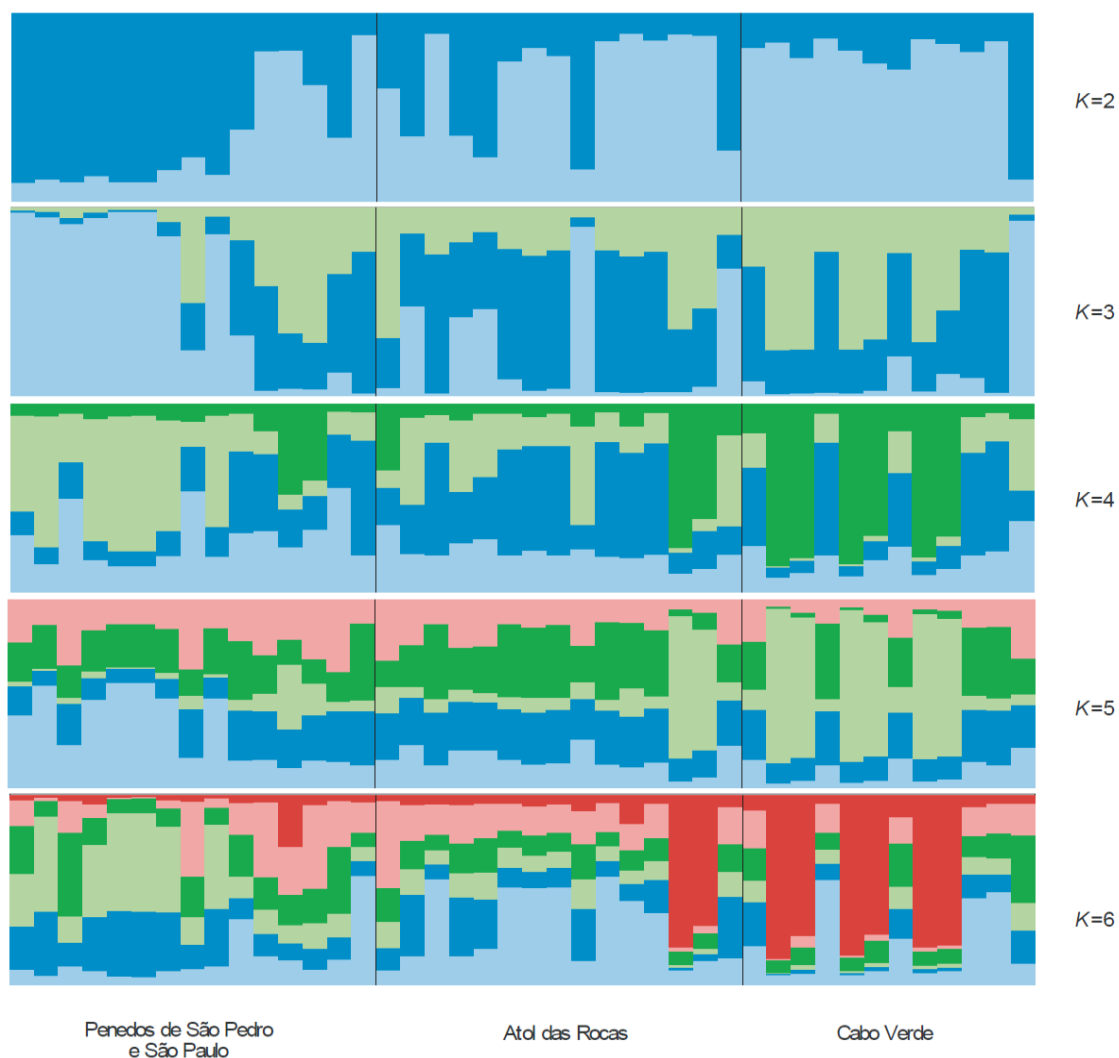


Figura 3 - Estrutura da população obtida pelo software STRUCTURE para $K = 2$ a $K = 6$, permutado por CLUMPP e visualizados por DISTRUCT. Cada indivíduo é representado por uma barra vertical em segmentos coloridos K , o comprimento de cada cor é proporcional ao coeficiente de adesão estimada. Linhas pretas separam indivíduos de diferentes populações pré-definidas.

Todas as amostras individuais foram dispersas em todas as ilhas do Atlântico com magnitudes semelhantes de proporções de indivíduos (Tabela 4). A co-ocorrência de indivíduos nos seis grupos indica mistura dentro das populações. Essa mistura indica que os indivíduos podem ter ascendência mista e ter herdado alguma fração do genoma de ancestrais em conjunto K.

Tabela 4 - Proporção de adesão das populações pré-definidas em clusters inferidos, Teresina, PI, 2016.

Populações Pre-definidas	Clusters inferidos						Número de indivíduos
	1	2	3	4	5	6	
Atol das Rocas	0,196	0,156	0,192	0,196	0,126	0,136	15
São Pedro e São Paulo	0,127	0,179	0,187	0,184	0,158	0,166	15
Cabo Verde	0,198	0,163	0,131	0,132	0,200	0,175	12

3.6 Conclusão

Neste estudo, relatamos o isolamento, utilizando a tecnologia NGS, e caracterização de 17 locos microssatélites polimórficos para a lagosta espinhosa *P. echinatus*. Estes microssatélites demonstraram potencial para estudos genéticos em nível de população e podem ser usados eficazmente como ferramenta molecular para ajudar na conservação e manejo dessa espécie da pesca altamente valiosa.

As conclusões deste estudo genético exploratório oferecem também um olhar sobre a variação genética da espécie entre as ilhas oceânicas de São Pedro e São Paulo, Atol das Rocas e Cabo Verde. Um padrão sutil de diferenciação genética foi detectado, sem revelar, porém, uma separação geográfica bem definida.

Um estudo mais detalhado da conectividade entre as populações poderá ser investigado utilizando o mesmo conjunto de marcadores microssatélites, em uma amostragem maior e em toda faixa de distribuição da espécie.

3.7 Referências

AGARDY, T. Effects of fisheries on marine ecosystems: a conservationist's perspective. **ICES Journal of Marine Science**, 57: 761-765, 2000.

BEN-HORIN, T. et al. Characterization of eight polymorphic microsatellite loci for the California spiny lobster, *panulirus interruptus* and cross-amplification in other achelate lobsters. **Conservation Genetics Resources**, v. 1, n. 1, p. 193–197, 2009.

BOTSTEIN, D.; WHITE, R. L.; SKOLNICK, M.; DAVIS, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**. 32:314- 331, 1980.

CARLSSON, J. Effects of microsatellite null alleles on assignment testing. **Journal of Heredity**. 99:616-623, 2008.

CHAPUIS, M.P.; ESTOUP, A. Microsatellite null alleles and estimation of population differentiation. **Molecular Biology and Evolution**. 24:621-631, 2007.

CHATROU, L.W.; ESCRIBANO, M.P.; VIRUEL, M.A.; MAAS, J.W.; RICHARDSON, J.E.; HORMAZA, J.I. Microsatellite flanking regions provide a new source of data for plant species-level phylogenetics. **Molecular Phylogenetics and Evolution**, Philadelphia, v.53, p.726-733, 2009.

DAILIANIS, T.; TSIGENOPOULOS, C. S.; DOUNAS, C.; VOULTSIADOU, E. Genetic diversity of the imperilled bath sponge *Spongia officinalis* Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation. **Molecular Ecology**, 20:3757-3772, 2011.

DAKIN, E. E.; AVISE, J. C. Microsatellite null alleles in parentage analysis. **Heredity**, v. 93, n. 5, p. 504–9, 2004.

DAO, H. T. et al. Oceanographic Currents and Local Ecological Knowledge Indicate , and Genetics Does Not Refute , a Contemporary Pattern of Larval Dispersal for The Ornate Spiny Lobster , *Panulirus ornatus* in the South-East Asian Archipelago. **PLoS ONE**. p. 1–19, 2015.

DAO, H. T.; TODD, E. V.; JERRY, D. R. Characterization of polymorphic microsatellite loci for the spiny lobster *Panulirus* spp. and their utility to be applied to other *Panulirus* lobsters. **Conservation Genetics Resources**, v. 5, n. 1, p. 43–46, 2013.

DELGHANDI, M. et al. Isolation, characterization, and multiplexing of novel microsatellite markers for the tropical scalloped spiny lobster (*Panulirus homarus*). **Genetics and Molecular Research**, v. 14, n. 4, p. 19066–19070, 2015.

DINIZ, F. M. et al. Polymorphic tetranucleotide microsatellite markers in the Caribbean spiny lobster, *Panulirus argus*. **Molecular Ecology Notes**, v. 4, n. 3, p. 327–329, 2004.

DINIZ, F. M. et al. Microsatellites in the overexploited spiny lobster, *Panulirus argus*: Isolation, characterization of loci and potential for intraspecific variability studies.

Conservation Genetics, v. 6, n. 4, p. 637–641, 2005.

DU, L. et al. MSDB: A user-friendly program for reporting distribution and building databases of microsatellites from genome sequences. **Journal of Heredity**, v. 104, n. 1, p. 154–157, 2013.

EARL, D.A.; VONHOLDT, B.M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.

Conservation Genetics Resources. 4:359–361, 2012

EVANNO, G.; REGNANT, S.; GOUDET, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. **Molecular Ecology**, v. 14, p. 2611–2620, 2005.

EXCOFFIER, L.; LISCHER, H. E. L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. **Molecular Ecology Resources**, v. 10, n. 3, p. 564–567, 2010.

GÓES, C.; LINS-OLIVEIRA, J.E. Natural diet of the spiny lobster, *Panulirus echinatus* Smith, 1869 (Crustacea: Decapoda: Palinuridae), from São Pedro and São Paulo Archipelago, Brazil. **Brazilian Journal of Biology**, v. 69, n. 691, p. 143–148, 2009.

GOUDET, J. FSTAT Version 2932: a computer program to calculate F-statistics. *J Hered* 86: 485–486, 1995.

HOLTHUIS, L.B. FAO Species Catalogue. Marine Lobster of The World. An annotated and illustrated catalogue of species of interest to fisheries known to date. **FAO Fish Syn.**, Rome, v.13, n. 125, p.1–292, 1991.

JOHNSON, M.S.; BLACK, R. The Wahlund effect and the geographical scale of variation in the intertidal limpet *Siphonaria* sp. **Marine Biology**. 79:295–302, 1984.

JAKOBSSON, M.; ROSENBERG, N. A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. **Bioinformatics**, v. 23, n. 14, p. 1801–1806, 2007.

JOST, L. GST and its relatives do not measure differentiation. **Molecular Ecology**, v. 17, n. 18, p. 4015–4026, 2008.

KAUKINEN, K.H.; SUPERNAL, K.J.; MILLER, K.M. Enrichment of tetranucleotide microsatellite loci from invertebrate species. **Journal of Shellfish Research**. 23:621–626, 2004.

KALINOWSKI, S. T.; TAPER, M. L.; MARSHALL, T. C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. **Molecular Ecology**, v. 16, n. 5, p. 1099–1106, 2007.

KENNINGTON, W. J. et al. Characterization of 18 polymorphic microsatellite loci for the western rock lobster *panulirus cygnus*. **Conservation Genetics Resources**, v. 2, n.

SUPPL.1, p. 389–391, 2010.

NAZARENO, A. G. ; REIS, M. S. d. ; REIS, Maurício Sedrez dos . The same but different: Monomorphic microsatellite markers as a new tool for genetic analysis. **American Journal of Botany**, v. 98, p. e265-e267, 2011.

PEAKALL,R.; SMOUSE, P.E. GenAIEx 65: Genetic analysis in Excel Population genetic software for teaching and research: An update. **Bioinformatics**, 28:2537–2539, 2012.

PRITCHARD, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. **Genetics**, v. 155, n. 2, p. 945–959, 2000.

RAYMOND, M.; ROUSSET, F. An exact test for population differentiation. **Evolution**, v. 49, n. April, p. 1280–1283, 1995.

RICE, W. R. **Analyzing Tables of Statistical Tests**.**Evolution**, 1989.

ROSENBERG, N. A. DISTRUCT: A program for the graphical display of population structure. **Molecular Ecology Notes**, v. 4, n. 1, p. 137–138, 2004.

ROZEN, S.; SKALETSKY, H.J. Primer3 on the WWW for general users and for biologist programmers In: Krawetz, S, Misener, S (Eds), **Bioinformatics Methods and Protocols: Methods in Molecular Biology**. **Humana Press**, Totowa, NJ, pp 365–386, 2000.

SLATKIN, M. A measure of population subdivision based on microsatellite allele frequencies. **Genetics**, v. 139, n. 1, p. 457–462, 1995.

TRINGALI, M. D.; SEYOUM, S.; SCHMITT, S. L. Ten di- and trinucleotide microsatellite loci in the Caribbean spiny lobster, *Panulirus argus*, for studies of regional population connectivity. **Molecular Ecology Resources**, v. 8, n. 3, p. 650–652, 2008.

TRUELOVE, N.; BEHRINGER, D. C.; PREZIOSI, R. F. Isolation and characterization of eight polymorphic microsatellites for the spotted spiny lobster, *Panulirus guttatus*. **PeerJ**, p. 4–11, 2016.

TRUELOVE, N. K. et al. Genetic analysis reveals temporal population structure in Caribbean spiny lobster (*Panulirus argus*) within marine protected areas in Mexico. **Fisheries Research**, v. 172, n. April, p. 44–49, 2015.

TRUETT, G. E. et al. Preparation of PCR-quality mouse genomic dna with hot sodium hydroxide and tris (HotSHOT). **BioTechniques**, v. 29, n. 1, p. 52–54, 2000.

VAN OOSTERHOUT, C. et al. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. **Molecular Ecology Notes**, v. 4, n. 3, p. 535–538, 2004.

WRIGHT, J.M., BENTZEN, P. Microsatellites – Genetic Markers For The Future. **Reviews in Fish Biology and Fisheries**, 4: 384-388, 1994.

Capítulo 4

**Genoma mitocondrial completo da lagosta *Panulirus echinatus* (Smith, 1869)
(Crustacea: Decapoda: Palinuridae)**

***Artigo “Complete mitochondrial genome of lobster *Panulirus echinatus* (Smith, 1869) (Crustacea: Decapoda: Palinuridae)”**

4.1 Resumo

A lagosta espinhosa *Panulirus echinatus* (Crustacea: Decapoda: Palinuridae) habita substratos profundos, rochosos e emaranhados de algas de regiões tropicais e subtropicais. Essa espécie é a terceira mais capturada no litoral do Nordeste brasileiro, tendo uma significativa participação no comércio interno e externo. Nosso estudo teve por objetivo sequenciar, determinar a estrutura, a ordem dos genes e composição das bases do genoma mitocondrial da lagosta *Panulirus echinatus*. O DNA da lagosta foi extraído e seu genoma foi parcialmente sequenciado, utilizando a tecnologia do sequenciador Illumina MiSeq. Um total de 423Mb foi obtido e estes dados foram usados para a montagem do genoma como base de referência. Usando o software SOAPdenovo-Trans, MIRA e MITObim foi possível descrever toda a versão circularizada do genoma mitocondrial. O genoma mitocondrial completo da lagosta é de 15.808 pb, contendo 37 genes (13 genes codificadores de proteínas, 22 genes para RNA transportador e 2 genes para RNA ribossomais), a ordem dos genes é exatamente a mesma observada para outros mitogenomas do gênero *Panulirus*. Os dados fornecidos podem ajudar a esclarecer as relações evolutivas dentro da família Palinuridae, e ser usados para estudos de genética de populações e identificação de organismos em crustáceos.

Palavras-chave: crustáceos, mitogenoma, relações evolutivas, MiSeq.

4.2 Abstract

The spiny lobster *Panulirus echinatus* (Crustacea: Decapoda: Palinuridae) dwells deep substrates, rocky and tangles of seaweed tropical and subtropical regions. This species is the third most captured in the Brazilian Northeast coast, with a significant share in domestic and foreign trade. Our study aimed to sequence, determine the structure, the order of genes and composition of bases of the mitochondrial genome of *echinatus* *Panulirus* lobster. The DNA was extracted from lobster and its genome has been partially sequenced using the technology of the sequencer Illumina MiSeq. A total of 423Mb was obtained and these data were used for assembling of the genome as the basis. Using SOAPdenovo-Trans software, MIRA and MITObim was possible to describe all the circularized version of the mitochondrial genome. The complete mitochondrial genome of the lobster is 15,808 bp, containing 37 genes (13 protein-coding genes, 22 genes for tRNA and 2 genes for ribosomal RNA), the order of genes is exactly the same observed for other mitogenomas of *Panulirus* genre. The data provided may help clarify the evolutionary relationships within the Palinuridae family, and be used for genetic studies of populations and identification of organisms in crustaceans.

Keywords: crustacean, mitogenoma, evolutionary relationships, MiSeq.

4.3 Introdução

A lagosta espinhosa *Panulirus echinatus* (Crustacea: Decapoda: Palinuridae) habita substratos profundos, rochosos e emaranhados de algas de regiões tropicais e subtropicais, sendo facilmente encontradas em arquipélagos de ilhas oceânicas como: Penedos de São Pedro e São Paulo, Atol das Rocas, Ilha Trindade, Cabo Verde, Ilhas Canárias, Ilhas Ascensão e Santa Helena, e uma boa parte da costa do litoral brasileiro (HOLTHUIS, 1991; MELO, 1999).

Essa espécie é a terceira mais capturada no litoral do Nordeste brasileiro, tendo uma significativa participação no comércio interno e externo (GÓES; LINS-OLIVEIRA; JE, 2009). Devido a sua importância econômica, alguns estudos têm sido realizados, como: informações sobre a biologia reprodutiva (PINHEIRO; LINS-OLIVEIRA, 2006), estudos sobre a dieta alimentar (GÓES; LINS-OLIVEIRA; JE, 2009), desenvolvimento dos ovários e órgãos reprodutivos em machos (BARRETO; KATSURAGAWA, 2008; BARRETO et al., 2008) e estágio larval (KONISHI; SUZUKI; CHOW, 2006), mas informações genéticas sobre a espécie que visem esclarecer as relações evolutivas e estudos populacionais, que possam contribuir para identificação, conservação e manejo são escassas.

O DNA mitocondrial (DNAm_t) vem tornando-se cada vez mais popular em estudos de relações filogenéticas de crustáceos decápodes que utilizam genes e regiões específicas, devido às suas características estruturais e funcionais. É um genoma curto com conteúdo gênico bastante conservado, possui poucos genes 37 no total (13 genes codificadores de proteínas, 22 para RNA transportadores e 2 para as subunidades ribossômicas). Possui uma região não codificadora com uma taxa evolutiva alta (BOORE, 1999; KIM et al., 2012; SHEN; BRABAND; SCHOLTZ, 2013). Com o advento do sequenciamento de nova geração uma ampla produção de genomas parciais de animais, plantas e microorganismos, marcadores genéticos e mitogenomas completos vêm sendo obtidos (MARDIS, 2008).

Esse estudo teve por objetivo sequenciar, determinar a estrutura, a ordem dos genes e composição das bases do genoma mitocondrial da lagosta *Panulirus echinatus*.

4.4 Material e Métodos

Extração de DNA e Sequenciamento

A extração de DNA foi realizada a partir do tecido muscular da lagosta. Utilizou-se 20mg de tecido macerado digerido em tampão contendo, 50 mM Tris-HCl pH 8,0, 2% de SDS, 0,75M NaCl, 10 mM EDTA e 100 µg/mL de proteinase K, incubado em banho térmico a 55°C *overnight*, com intervalo de agitações, em seguida a amostra foi purificada utilizando o protocolo padrão PCI, com Fenol-Clorofórmio/Álcool isoamílico (SAMBROOK; RUSSELL, 2001).

A montagem do genoma mitocondrial da espécie *P. echinatus* foi realizada em parceria com a Universidade Federal do Rio de Janeiro, Instituto de Bioquímica Médica, Laboratório Multidisciplinar para Análise de Dados.

Uma biblioteca Illumina *paired-end* foi preparada seguindo o protocolo padrão da Illumina Nextera DNA Kit Preparação de Amostras (Illumina Inc., San Diego, Califórnia). Um único DNA de maior peso molecular foi sequenciado utilizando uma plataforma MiSeq Benchtop, produzindo 2 fastq files de 846,501 *reads*. A quantidade de DNA total sequenciada foi calculada em 423,25 Mb e usada como entrada de dados para um alinhamento de nucleotídeos contra um conjunto de dados que contém os mitogenomas de duas lagostas já conhecidas, como, *Panulirus versicolor* (NC_017868) (SHEN; BRABAND; SCHOLTZ, 2013) e *Panulirus ornatus* (NC_014854) (QIAN et al., 2011).

O software GMAP (WU; WATANABE, 2005) foi utilizado para mapear o conjunto completo de *reads* para esses dois mitogenomas do mesmo gênero, 567 sequências *paired-end* emparelhadas foram encontrados para mapear os genomas e ainda utilizadas para uma montagem de base de referência, utilizando o software MIRA (CHEVREUX; WETTER; SUHAI, 1999). A referência utilizada foi baseada numa versão quase completa do mitogenoma da *P. echinatus*, gerado pelo novo montador SOAPdenovo-Trans (LUO et al., 2012).

Análise da sequência

A sequência montada foi então usada como entrada para o montador interativo MITObim (HAHN; BACHMANN; CHEVREUX, 2013) para fechar pequenas aberturas, agora com todo o conjunto de dados de leituras como entrada. Após 10 interações, MITObim chegaram a uma série de leitura fixa de 834 *reads* montado, produzindo uma versão completa e circularizada do mitogenoma da lagosta do *P. echinatus*. Um rendimento médio de 13,5 *reads* por posição de nucleotídeo foi observado e uma sobreposição geral de leitura foi verificada manualmente usando software Tablet (MILNE et al., 2013). A anotação do genoma foi realizada utilizando um servidor MITOS WebServer (BERNT et al., 2013), mitoMaker e curação manual utilizando o BLAST (ALTSCHUL et al., 1997).

Previsões de RNA de transferência foram realizadas no servidor web tRNAscan-SE (LOWE; EDDY, 1997). As posições de início e fim de cada gene foram manualmente localizadas por meio do software Artemis (CARVER et al., 2012).

4.5 Resultados e Discussão

Organização do Genoma Mitocondrial da Panulirus echinatus

O genoma circularizado completo da lagosta *Panulirus echinatus* contém 15.808 pb de comprimento (Figura 1). Possui um dos mais longos genomas mitocondriais do seu gênero, comparados com *Panulirus versicolor* (15.700 pb) (SHEN; BRABAND; SCHOLTZ, 2013), *Panulirus cygnus* (15.724 pb) (KIM et al., 2015), *Panulirus japonicus* (15.717 pb) (YAMAUCHI; MIYA; NISHIDA, 2002), *Panulirus stimpsoni* (15.677 pb) (LIU; CUI, 2011) e *Panulirus homarus* (15.665 pb) (Acesso GenBank, NC_016015.1), menor somente que o mitogenoma da lagosta *Panulirus ornatus* (16.105 pb) (LIANG, 2012).

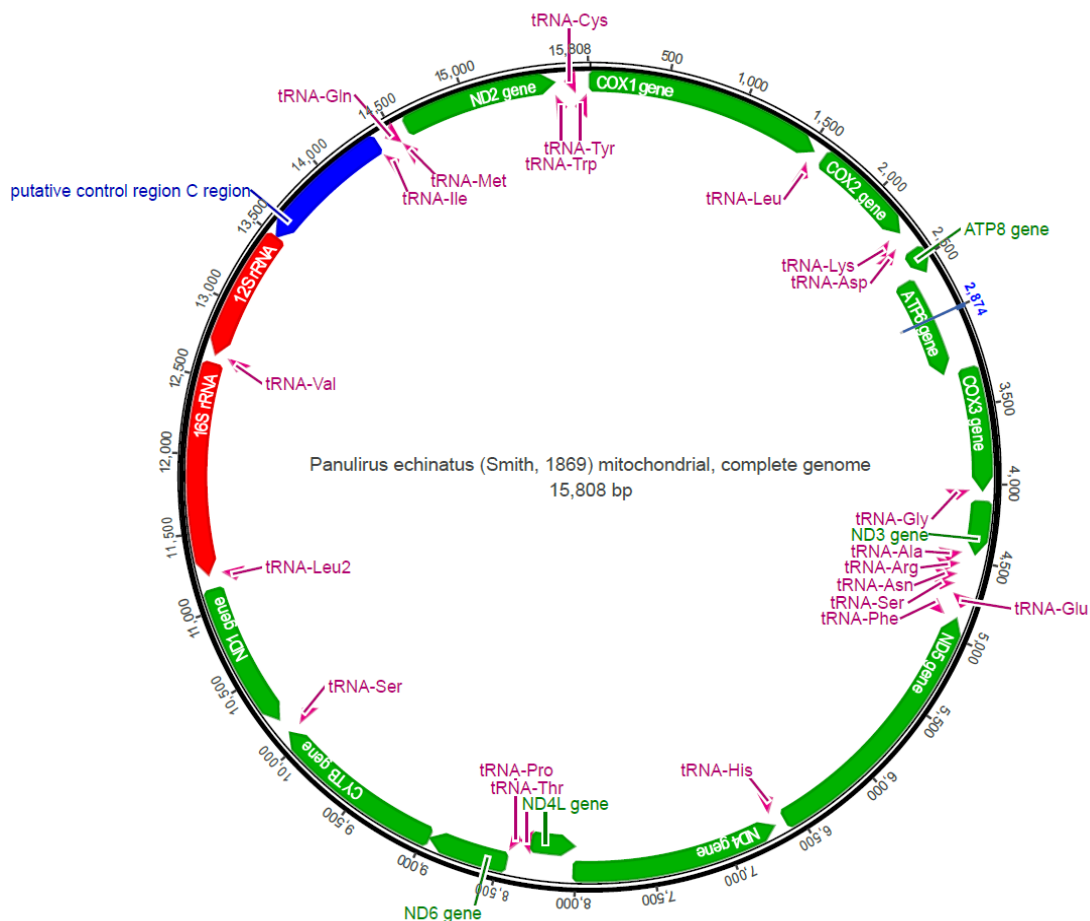


Figura 1 - Mapa do genoma mitocondrial da lagosta *Panulirus echinatus*.

O genoma mitocondrial da lagosta contém 13 genes codificadores de proteínas, 22 tRNAs necessários para a tradução das proteínas codificadas pelo mtDNA, duas subunidades do ribossomo mitocondrial, e uma região controle (Tabela 1). A quantidade de genes e seu arranjo foram conservados comparados com outros mitogenomas do gênero *Panulirus*, de COX1 a tRNA(Tyr) (Figura 2).

Composição das bases

O conteúdo A + T do genoma completo da lagosta *P. echinatus* é 60,6% (A = 31,0%, C = 23,5%, G = 15,8%, T = 29,5%), que é próximo ao da lagosta *P. japonicus* (64,5%), *P. stimpsoni* (65,6%) (LIU; CUI, 2011), *P. ornatus* (66,7%) (LIANG, 2012), *P. versicolor* (66,8%), *P. cygnus* (63,5%) (KIM et al., 2015) e *P. homarus* (67,1 %) (Acesso GenBank, NC_016015.1). O teor de A + T dos genes codificadores de proteínas é 58,66% (A = 23,64%, C = 21,92%, G = 19,4%, T = 35,01%), inferior ao conteúdo de A+T das lagostas do gênero *Panulirus*.

O comprimento das regiões não codificantes do genoma de *P. echinatus* é 979 pb, correspondendo a 6% do genoma total, distribuídos em dez regiões intergênicas que variam de 4 a 33pb, a maior região não codificadora (884pb) é a região controle (RC), com um elevado teor de A + T (59,8%) localizada entre os genes 12S-rRNA e o gene tRNAIle.



Figura 2 - Genomas mitocondriais completos de sete lagostas do gênero *Panulirus*, incluindo *Panulirus echinatus*.

Tabela 1 - Características do genoma mitocondrial da lagosta *P. echinatus*, Teresina, PI, 2016.

Gene	Posição		Comprimento		Códona		Nucleotídios Intergênicos	Cadeia ^b
	Início	Fim	Nucleotídio	Aminoácido	Início	Parada		
COX1	1	1534	1534	511	ACG	T*	0	H
tRNA(Leu)	1535	1598	64				6	H
COX2	1605	2292	688	229	ATG	T*	0	H
tRNA(Lys)	2293	2357	65				6	H
tRNA(Asp)	2364	2428	65				0	H
ATP8	2429	2587	159	52	ATG	TAA	-7	H
ATP6	2581	3257	677	225	ATG	TA*	0	H
COX3	3258	4047	790	263	ATG	T*	0	H
tRNA(Gly)	4048	4112	65				0	H
NAD3	4113	4464	353	117	ATT	T*	-1	H
tRNA(Ala)	4463	4530	66				3	H
tRNA(Arg)	4534	4596	63				9	H
tRNA(Asn)	4606	4670	65				0	H
tRNA(Ser)	4671	4738	68				-1	H
tRNA(Glu)	4738	4808	71				0	H
tRNA(Phe)	4809	4877	69				1	H
NAD5	4879	6606	1728	576	ATG	T*	0	L
tRNA(His)	6607	6674	68				0	H
NAD4	6675	8013	1339	446	ATG	T*	-4	L
NAD4-L	8010	8309	300	100	ATG	T*	2	L
tRNA(Thr)	8312	8379	68				0	H
tRNA(Pro)	8380	8445	66				2	L
NAD6	8448	8962	515	171	ATC	TA*	0	H
CYTB	8963	10097	1135	378	ATG	T*	0	H
tRNA(Ser)	10098	10164	67				33	H
NAD1	10198	11140	943	314	GTG	T*	15	L
tRNA(Leu)	11156	11225	70				4	L
16S-rRNA	11230	12599	1370				-17	L
tRNA(Val)	12583	12653	71				0	L
12S-rRNA	12654	13500	847				-1	L
Control Region	13502	14385	884					
tRNA(Ile)	14385	14451	67				-3	H
tRNA(Gln)	14449	14517	69				8	L
tRNA(Met)	14526	14594	69				0	H
NAD2	14595	15596	1002	333	GTG	TAA	-2	H
tRNA(Trp)	15595	15662	68				-1	H
tRNA(Cys)	15662	15727	66				0	L
tRNA(Tyr)	15728	15793	66				15	L

^a Segmento do DNA que codifica proteínas.

^bFita: H = *heavy* (+); L = *light* (-)

T* ou TA* indicam códon de finalização incompleto.

Genes codificadores de proteínas

O genoma mitocondrial da lagosta *P. echinatus* contém 13 genes codificadores de proteínas, localizados nas duas fitas do mitogenoma, designadas como H (*heavy*) e L (*light*), sendo nove genes (COX1, COX2, ATP8, ATP6, COX3, NAD3, CYTB, NAD6 e NAD2) localizados na fita H do mitogenoma e os outros quatro (NAD5, NAD4, NAD4-L, e NAD1) na fita L (Tabela 1). Desses 13 genes, dois pares se sobrepõem na mesma cadeia, o ATP8 e ATP6 e o NAD4 e NAD4-L compartilhando sete nucleotídeos.

O ATG é o códon de iniciação de 8 genes codificadores de proteínas (COX2, ATP8, ATP6, COX3, NAD5, NAD4, NAD4-L, CYTB), o GTG é o códon de iniciação dos genes (NAD1 e NAD2), o ATT do NAD3, o ATC do NAD6, e o códon de iniciação do gene COX1 é ACG. O códon de parada dos genes ATP8 e NAD2 é TAA, os demais apresentaram códon de parada incompletos, TA* (AT6 e NAD6) e T* (COX1, COX2, COX3, NAD3, NAD5, NAD4, NAD4-L, CYTB, NAD1). Códon de parada incompletos são comumente relatado em outros invertebrados (CROZIER; CROZIER, 1993, MASTA; BOORE, 2004). Esses códon de terminação não são totalmente codificados no genoma, mas, podem ser completados através da poliadenilação pós-transcricional na extremidade 3' (CROSBY et al., 2015, WOLSTENHOLME, 1992) (Tabela 1).

Genes de RNA ribossomal e de transferência

Os genes de RNA de subunidades ribossomais pequeno e grande (12S-rRNA e 16S-rRNA) de *P. echinatus* apresentaram um tamanho de 847 e 1370 pb, respectivamente (Tabela 1), semelhante ao observado em *P. cygnus* 852 a 1359 pb (KIM et al., 2015) em *P. japonicus* 855-1355 (YAMAUCHI; MIYA; NISHIDA, 2002) e *P. versicolor* 895-1358 (SHEN; BRABAND; SCHOLTZ, 2013). O teor de A + T no gene 12S-rRNA é de 63,6% e está situado entre os genes tRNA (leu) e tRNA (Val). O gene 16S-rRNA possui um teor de A + T de 67,6% e situa-se entre o gene tRNA (Val) e a região controle. Tanto o tamanho das subunidades ribossomais quanto a posição são as mesmas observada em outros mitogenomas do mesmo gênero.

O genoma mitocondrial da *P. echinatus* codifica 22 tRNAs com comprimentos de 63 a 71 pb (Tabela 1), muito semelhante a outras espécies do mesmo gênero, como:

Panulirus cygnus (63 a 73)(KIM et al., 2015), *Panlirus ornatus* (62 a 72) (LIANG, 2012) e *P. stimpsoni* (63 a 72) (LIU; CUI, 2011).

Como observados na maioria dos invertebrados, dentre os genes tRNA que compõem o mitogenoma de *P. echinatus*, foram identificadas as duas serinas e as duas leucinas. Com exceção do gene tRNA(Arg) e tRNA(Ser 1), todos os outros dobraram-se nas típicas estruturas secundárias de folha de trevo, que foram obtidas pelo *software* tRNAScan (APÊNDICE B), essa ausência na forma de trevo observada nesses dois genes de tRNA, está relacionado com a falta da alça DHU (dihidrouridina) na estrutura secundária do tRNA, impedindo assim o seu dobramento (WILSON et al., 2000; YAMAUCHI; MIYA; NISHIDA, 2002)

4.6 Conclusão

Nesse estudo a combinação da técnica de sequenciamento de nova geração com ferramentas da bioinformática permitiu montar e anotar o DNA mitocondrial completo da lagosta *P. echinatus*, sendo a ordem dos genes exatamente a mesma observada nas outras espécies do gênero *Panulirus*.

A montagem do DNA mitocondrial da lagosta irá fornecer um recurso valioso de informação genética sobre a espécie, além de fornecer dados que irão ajudar a esclarecer as relações evolutivas dentro da família Palinuridae, e ser usados para estudos de genética de populações e identificação de organismos em crustáceos.

4.7 Referências

- ALTSCHUL, S. F. et al. Gapped BLAST and PS I-BLAST: a new generation of protein database search programs. **Nucleic Acids Res**, v. 25, n. 17, p. 3389–3402, 1997.
- BARRETO, A. V et al. Desenvolvimento dos ovários lagosta *Panulirus echinatus* (Crustacea: Palinuridae) baseados nas análises macroscópica, microscópica e relação gonadossomática (RGS). **Revista Brasileira de Zoologia**, v. 25, n. 4, p. 689–695, 2008.
- BARRETO, A. DO V; KATSURAGAWA, M. Estádios de desenvolvimento dos órgãos reprodutivos dos machos de *Panulirus echinatus* Smith (Decapoda: Palinuridae). **Revista Brasileira de Zoologia**, v. 25, n. 1, p. 74–82, 2008.
- BERNT, M. et al. MITOS: Improved de novo metazoan mitochondrial genome annotation. **Molecular Phylogenetics and Evolution**, v. 69, n. 2, p. 313–319, 2013.
- BOORE, J. L. Do animal mitochondrial genomes. v. 5347, n. 99, p. 209–210, 1999.
- CARVER, T. et al. Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. **Bioinformatics**, v. 28, n. 4, p. 464–469, 2012.
- CROZIER R. H.; CROZIER Y. C. The mitochondrial genome of the honeybee *Apis mellifera*: Complete sequence and genome organization. **Genetics**, 133:97–117, 1993.
- CROSBY, M. A. et al. Gene Model Annotations for *Drosophila melanogaster*: The Rule-Benders. **G3: Genes|Genomes|Genetics**, v. 5, n. August, p. 1737–1749, 2015.
- CHEVREUX, B.; WETTER, T.; SUHAI, S. Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. **Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB)**, n. 1995, p. 45–56., 1999.
- GÓES, C.; LINS-OLIVEIRA; JE. Natural diet of the spiny lobster, *Panulirus echinatus* Smith, 1869 (Crustacea: Decapoda: Palinuridae), from São Pedro and São Paulo Archipelago, Brazil. **Braz. J. Biol. J. Biol**, v. 69, n. 691, p. 143–148, 2009.
- HAHN, C.; BACHMANN, L.; CHEVREUX, B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - A baiting and iterative mapping approach. **Nucleic Acids Research**, v. 41, n. 13, p. 1–9, 2013.
- HOLTHUIS, L.B. FAO Species Catalogue. Marine Lobster of The World. An annotated and illustrated catalogue of species of interest to fisheries known to date. **FAOFish. Syn.**, Rome, v.13, n. 125, p.1-292, 1991.
- KIM, G. et al. Complete mitochondrial genome of Australian spiny lobster, *Panulirus cygnus* (George, 1962) (Crustacea: Decapoda: Palinuridae) from coast of Australia. **Mitochondrial DNA**, v. 1736, n. January 2016, p. 1–2, 2015.
- KIM, S. et al. The mitochondrial genomes of *Cambaroides similis* and *Procambarus clarkii* (Decapoda: Astacidea: Cambaridae): The phylogenetic implications for

- Reptantia. **Zoologica Scripta**, v. 41, n. 3, p. 281–292, 2012.
- KONISHI, K.; SUZUKI, N.; CHOW, S. A late-stage phyllosoma larva of the spiny lobster *Panulirus echinatus* Smith, 1869 (Crustacea: Palinuridae) identified by DNA analysis. **Journal of Plankton Research**, v. 28, n. 9, p. 841–845, 2006.
- LIANG, H. Complete mitochondrial genome of the ornate rock lobster *Panulirus ornatus* (Crustacea: Decapoda). **African Journal of Biotechnology**, v. 11, n. 80, p. 14519–14528, 2012.
- LIU, Y.; CUI, Z. Complete mitochondrial genome of the Chinese spiny lobster *Panulirus stimpsoni* (Crustacea: Decapoda): Genome characterization and phylogenetic considerations. **Molecular Biology Reports**, v. 38, n. 1, p. 403–410, 2011.
- LOWE, T. M.; EDDY, S. R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. **Nucleic Acids Research**, v. 25, n. 5, p. 955–964, 1997.
- LUO, R. et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. **GigaScience**, v. 1, n. 1, p. 18, 2012.
- MASTA, S. E.; BOORE, J. L. The Complete Mitochondrial Genome Sequence of the Spider *Habronattus oregonensis* Reveals Rearranged and Extremely Truncated tRNAs. **Molecular Biology and Evolution**, v. 21, n. 5, p. 893–902, 2004.
- MARDIS, E. R. The impact of next-generation sequencing technology on genetics. **Trends in Genetics**, v. 24, n. 3, p. 133–141, 2008.
- MELO, G.A.S. Manual de identificação dos crustáceos Decapoda do litoral brasileiro: Anomura, Thalassinidea, Palinuridea, Astacidea. Plêiade/ FAPESP, 551 p., São Paulo, 1999.
- MILNE, I. et al. Using tablet for visual exploration of second-generation sequencing data. **Briefings in Bioinformatics**, v. 14, n. 2, p. 193–202, 2013.
- PINHEIRO, A. P.; LINS-OLIVEIRA, J. E. Reproductive Biology of *Panulirus echinatus* (Crustacea: Palinuridae) from São Pedro and São Paulo Archipelago, Brazil. **Nauplius**, v. 14, n. 2, p. 89–97, 2006.
- QIAN, G. et al. Two new decapod (Crustacea, Malacostraca) complete mitochondrial genomes: Bearings on the phylogenetic relationships within the Decapoda. **Zoological Journal of the Linnean Society**, v. 162, n. 3, p. 471–481, 2011.
- SAMBROOK, J.; RUSSELL, D.W. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. 2001.
- SHEN, H.; BRABAND, A.; SCHOLTZ, G. Mitogenomic analysis of decapod crustacean phylogeny corroborates traditional views on their relationships. **Molecular Phylogenetics and Evolution**, v. 66, n. 3, p. 776–789, 2013.
- WILSON, K. et al. The complete sequence of the mitochondrial genome of the

crustacean *Penaeus monodon*: are malacostracan crustaceans more closely related to insects than to branchiopods? **Molecular biology and evolution**, v. 17, n. 6, p. 863–874, 2000.

WOLSTENHOLME, D. R. Animal mitochondrial DNA: Structure and evolution. **International Review of Cytology**, 141:173–216, 1992.

WU, T. D.; WATANABE, C. K. GMAP: A genomic mapping and alignment program for mRNA and EST sequences. **Bioinformatics**, v. 21, n. 9, p. 1859–1875, 2005.

YAMAUCHI, M. M.; MIYA, M. U.; NISHIDA, M. Complete mitochondrial DNA sequence of the Japanese spiny lobster, *Panulirus japonicus* (Crustacea: Decapoda). **Gene**, v. 295, n. 1, p. 89–96, 2002.

APÊNDICES

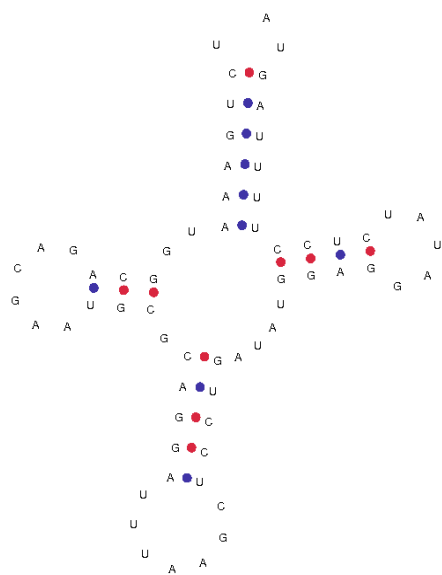
APÊNDICE A - Número, comprimento total, frequência e densidade de todos os *motifs* encontrados em *Panulirus echinatus*.

Motif	Motif Length	Contagem Total	Comprimento Total (pb)	Comprimento Médio(pb)	Frequência (loci/Mb)	Densidade (bp/Mb)
Mononucleotídios						
A	1	788	13064	16,58	243,79	4041,76
T	1	463	7585	16,38	143,24	2346,66
C	1	28	604	21,57	8,66	186,87
G	1	11	187	17	3,4	57,85
Dinucleotídios						
AT	2	333	10584	31,78	103,02	3274,49
AC	2	285	7108	24,94	88,17	2199,08
CT	2	260	8112	31,2	80,44	2509,7
GT	2	224	5364	23,95	69,3	1659,52
AG	2	178	5810	32,64	55,07	1797,5
Trinucleotídios						
AAT	3	38	1236	32,53	11,76	382,4
ATT	3	36	1068	29,67	11,14	330,42
CCT	3	21	381	18,14	6,5	117,87
CTT	3	15	411	27,4	4,64	127,16
AAG	3	13	441	33,92	4,02	136,44
ACT	3	12	249	20,75	3,71	77,04
GTT	3	12	198	16,5	3,71	61,26
AGC	3	11	204	18,55	3,4	63,11
AGT	3	11	219	19,91	3,4	67,75
AGG	3	10	180	18	3,09	55,69
CTG	3	9	174	19,33	2,78	53,83
ATC	3	8	129	16,13	2,48	39,91
ATG	3	8	150	18,75	2,48	46,41
AAC	3	7	111	15,86	2,17	34,34
GTG	3	6	93	15,5	1,86	28,77
ACC	3	2	30	15	0,62	9,28
GTC	3	2	39	19,5	0,62	12,07
Tetranucleotídios						
ATCT	4	24	700	29,17	7,43	216,57
AGAT	4	15	328	21,87	4,64	101,48
AATA	4	11	236	21,45	3,4	73,01
CTGT	4	10	184	18,4	3,09	56,93
AGAC	4	7	148	21,14	2,17	45,79
ACAT	4	6	116	19,33	1,86	35,89
ATGT	4	6	100	16,67	1,86	30,94
ATTT	4	5	112	22,4	1,55	34,65
ACCA	4	3	48	16	0,93	14,85
TACT	4	3	52	17,33	0,93	16,09
CTCA	4	2	36	18	0,62	11,14

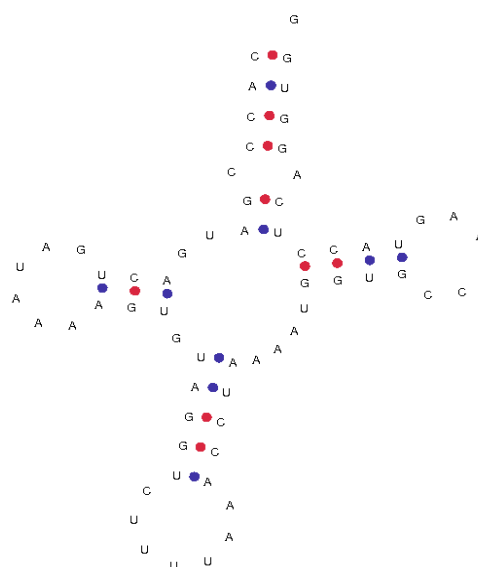
GTAG	4	2	36	18	0,62	11,14
AGTC	4	1	20	20	0,31	6,19
AGTG	4	1	20	20	0,31	6,19
CAAA	4	1	16	16	0,31	4,95
CCTC	4	1	16	16	0,31	4,95
CGTC	4	1	16	16	0,31	4,95
GAAA	4	1	16	16	0,31	4,95
GCCA	4	1	16	16	0,31	4,95
GCTT	4	1	20	20	0,31	6,19
GGAG	4	1	16	16	0,31	4,95
GGTG	4	1	16	16	0,31	4,95
GTTT	4	1	16	16	0,31	4,95
TCGG	4	1	16	16	0,31	4,95
TGAA	4	1	20	20	0,31	6,19
Pentanucleotídios						
ACAAC	5	6	90	15	1,86	27,84
TGTTT	5	5	75	15	1,55	23,2
AGGTT	5	4	65	16,25	1,24	20,11
ACACC	5	3	45	15	0,93	13,92
ACCTC	5	3	45	15	0,93	13,92
GTGTT	5	3	45	15	0,93	13,92
ACCTA	5	2	30	15	0,62	9,28
CTGGC	5	2	30	15	0,62	9,28
GAGGT	5	2	30	15	0,62	9,28
ACTAT	5	1	15	15	0,31	4,64
AGCAT	5	1	15	15	0,31	4,64
AGGTC	5	1	15	15	0,31	4,64
ATGTT	5	1	15	15	0,31	4,64
CCGTA	5	1	15	15	0,31	4,64
CTGAC	5	1	15	15	0,31	4,64
CTGTC	5	1	15	15	0,31	4,64
GAAAA	5	1	30	30	0,31	9,28
GAGGA	5	1	15	15	0,31	4,64
TAAAA	5	1	15	15	0,31	4,64
TCACA	5	1	15	15	0,31	4,64
TCCCC	5	1	15	15	0,31	4,64
TCTCT	5	1	15	15	0,31	4,64
TTATT	5	1	15	15	0,31	4,64
Hexanucleotídios						
ATAATG	6	3	60	20	0,93	18,56
ACCCAC	6	2	36	18	0,62	11,14
ATGTGT	6	2	36	18	0,62	11,14
ACCACA	6	1	18	18	0,31	5,57
ACCCCC	6	1	18	18	0,31	5,57
ATAGAT	6	1	18	18	0,31	5,57
ATATGT	6	1	18	18	0,31	5,57
ATCATT	6	1	24	24	0,31	7,43

CAAAAA	6	1	18	18	0,31	5,57
CTTCAC	6	1	18	18	0,31	5,57
GGAGGG	6	1	18	18	0,31	5,57
TGATGG	6	1	18	18	0,31	5,57
TGTTTT	6	1	18	18	0,31	5,57

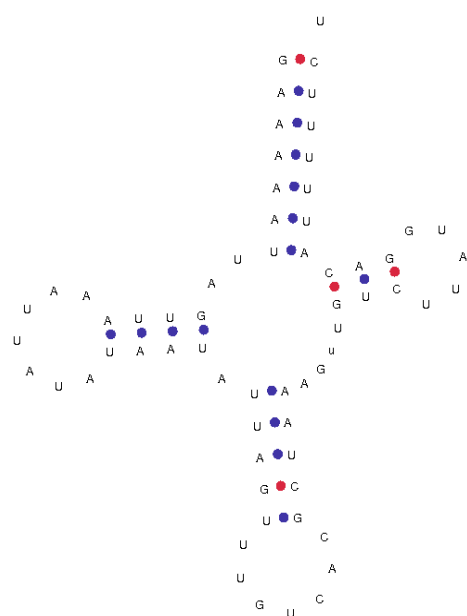
APÊNDICE B. Estruturas secundárias dos genes tRNAs do mitogenoma da *Panulirus echinatus* obtidas com o software tRNAScan-SE 1.21 (LOWE; EDDY, 1997).



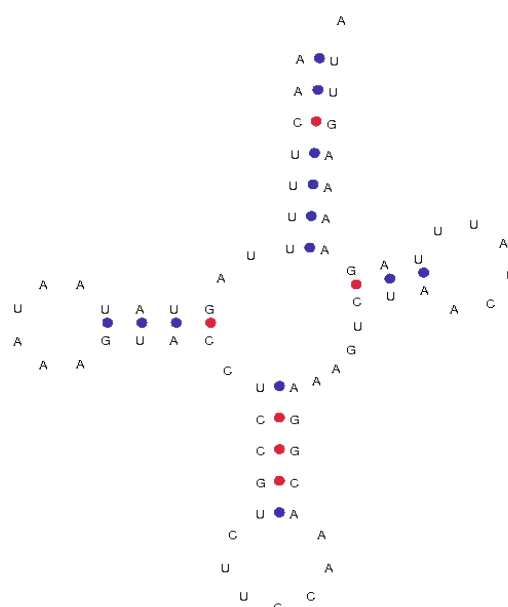
LEU 1, Anti-códon (TAA)



LYS, Anti-códon (TTT)

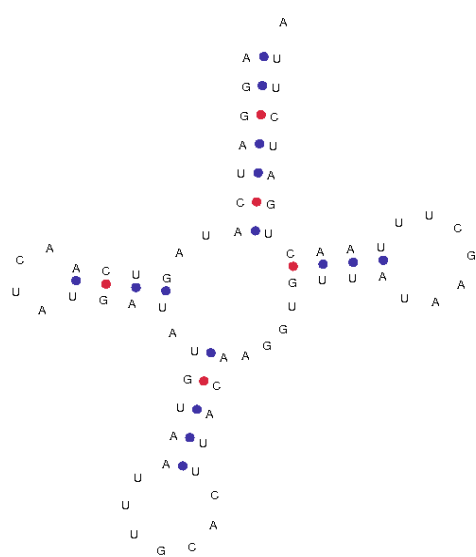


ASP, Anti-códon (GTC)

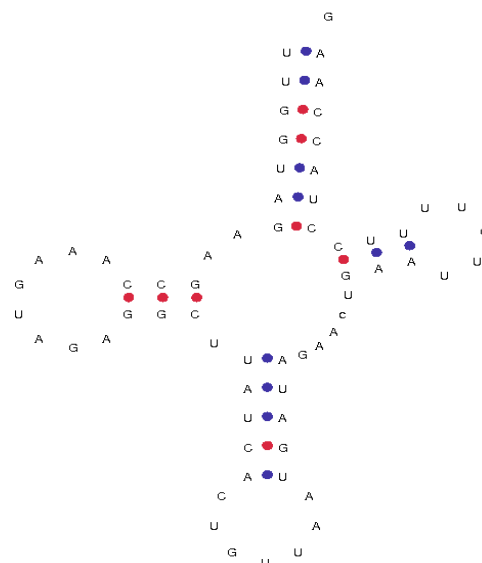


GLY, Anti-códon (TCC)

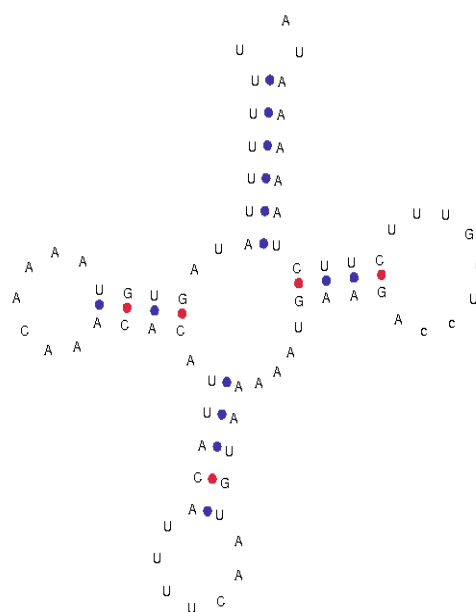
APÊNDICE B. Estruturas secundárias dos genes tRNAs do mitogenoma da *Panulirus echinatus* obtidas com o software tRNAscan-SE 1.21 (LOWE; EDDY, 1997).



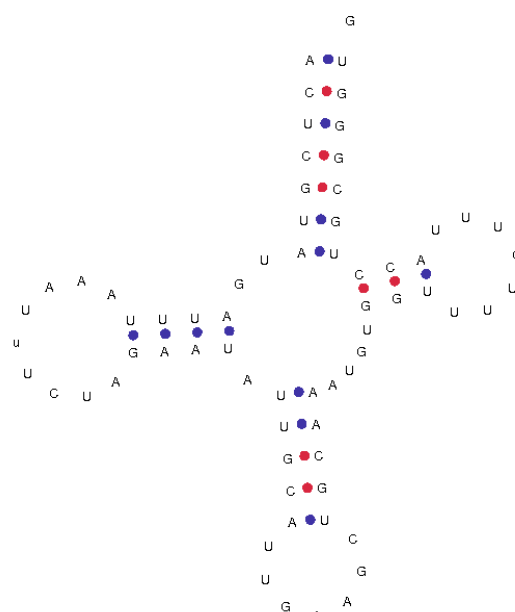
ALA, Anti-códon (TGC)



ASN, Anti-códon (GTT)

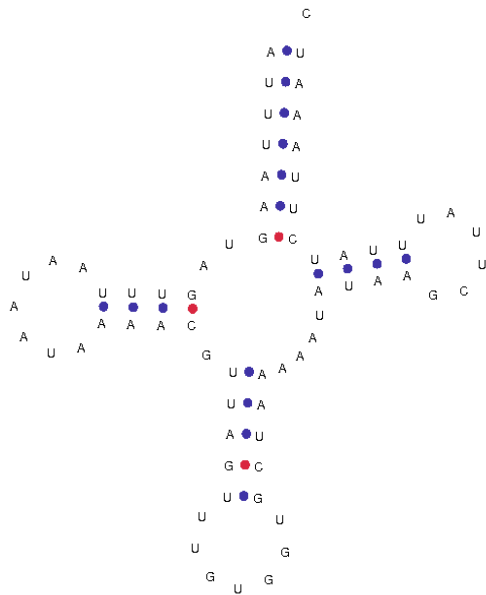


GLU, Anti-códon (TTC)

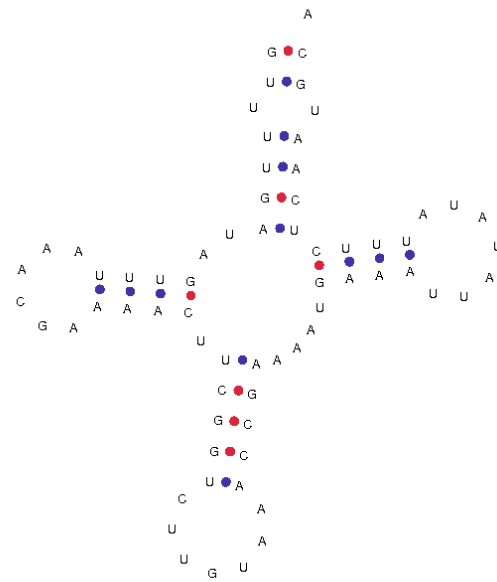


PHE, Anti-códon (GAA)

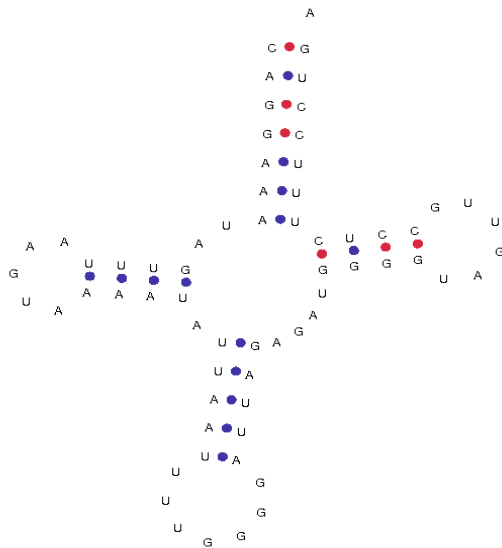
APÊNDICE B. Estruturas secundárias dos genes tRNAs do mitogenoma da *Panulirus echinatus* obtidas com o software tRNAScan-SE 1.21 (LOWE; EDDY, 1997).



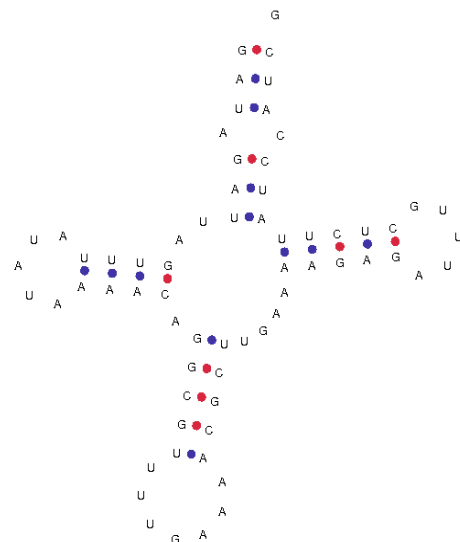
HIS, Anti-códon (GTG)



THR, Anti-códon (TGT)

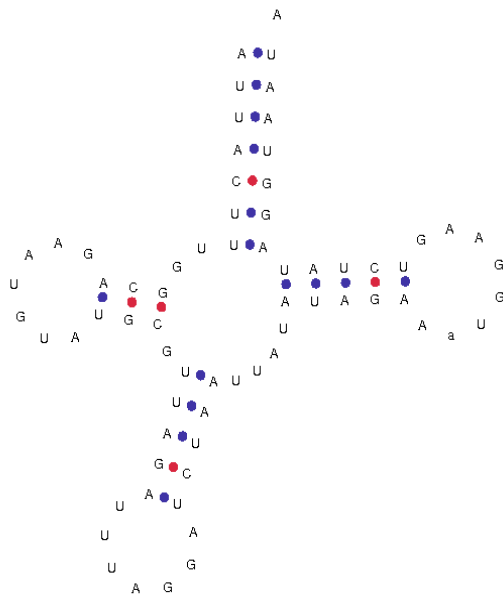


PRO, Anti-códon (TGG)

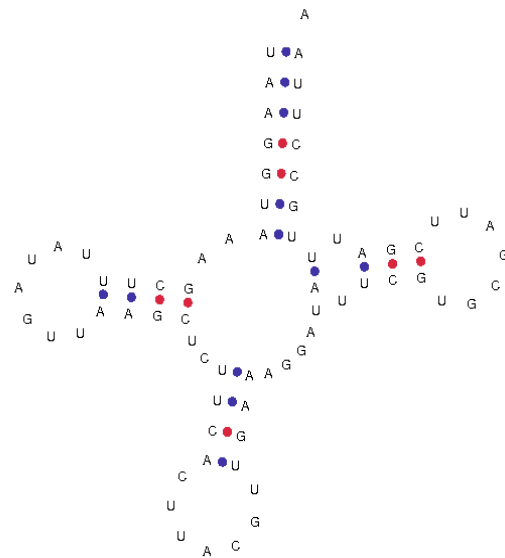


SER 2 , Anti-códon (TGA)

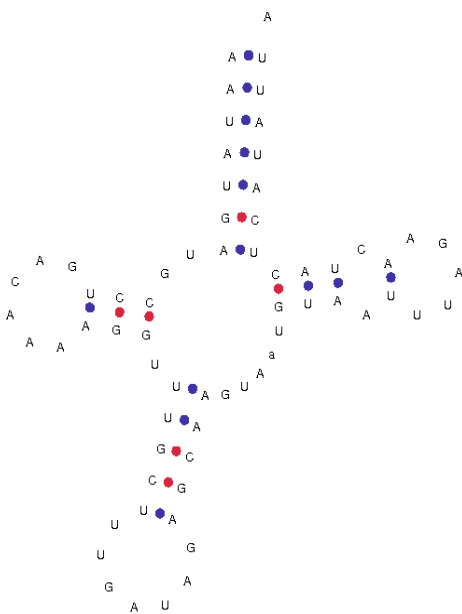
APÊNDICE B. Estruturas secundárias dos genes tRNAs do mitogenoma da *Panulirus echinatus* obtidas com o software tRNAscan-SE 1.21 (LOWE; EDDY, 1997).



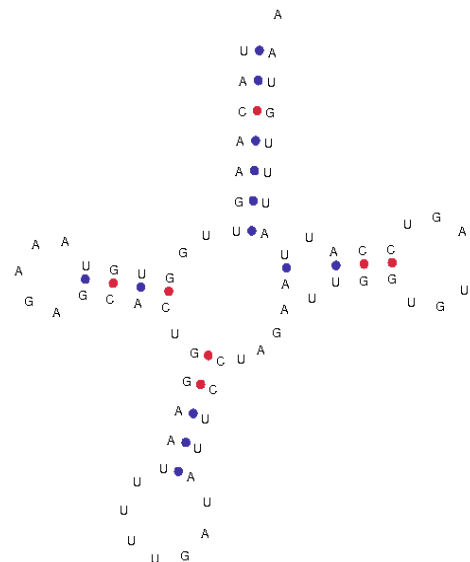
LEU 2, Anti-códon (TAG)



VAL, Anti-códon (TAC)

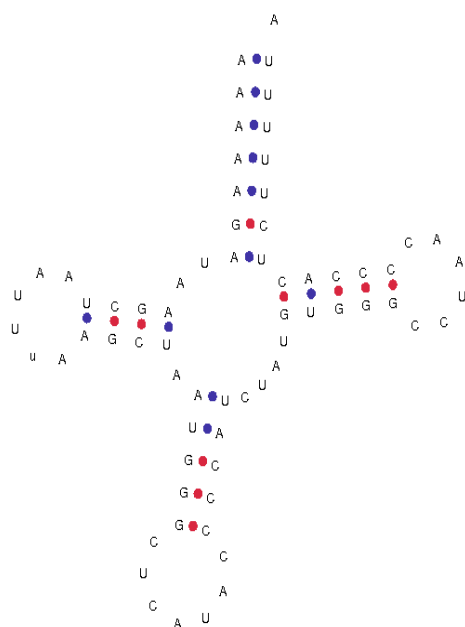


ILE, Anti-códon (GAT)

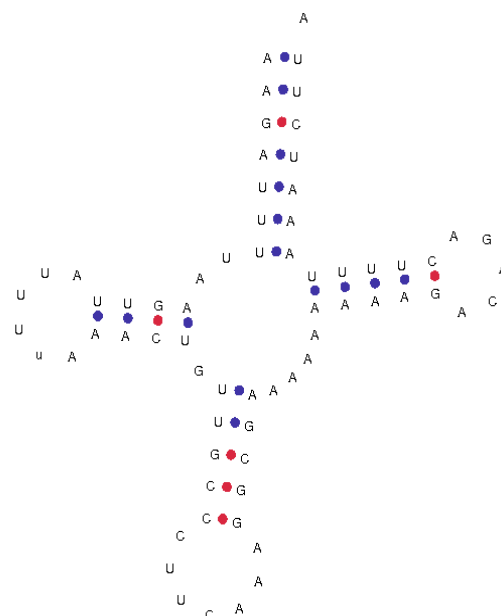


GLN, Anti-códon (TTG)

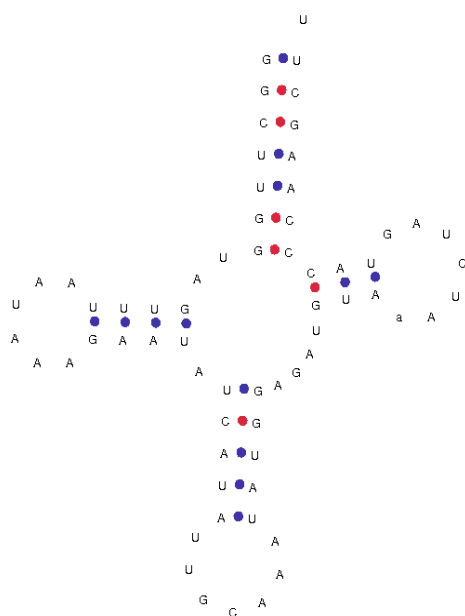
APÊNDICE B. Estruturas secundárias dos genes tRNAs do mitogenoma da *Panulirus echinatus* obtidas com o software tRNAScan-SE 1.21 (LOWE; EDDY, 1997)



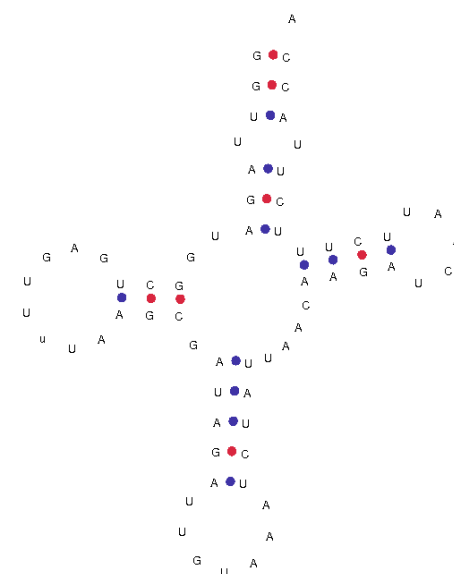
MET, Anti-códon (CAT)



TRP, Anti-códon (TCA)



CYS, Anti-códon (GCA)



TYR, Anti-códon (GTA)

APÊNDICE C

Molecular Genetics and Genomics

Development of genome-wide informative simple sequence repeat markers and a snapshot of the genetic diversity of the spiny lobster, *Panulirus echinatus* using next generation sequencing

--Manuscript Draft--

Manuscript Number:	MGAG-D-16-00274
Full Title:	Development of genome-wide informative simple sequence repeat markers and a snapshot of the genetic diversity of the spiny lobster, <i>Panulirus echinatus</i> using next generation sequencing
Article Type:	Original Article
Corresponding Author:	Fabio Mendonça Diniz, Ph.D. Embrapa Caprinos e Ovinos BRAZIL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Embrapa Caprinos e Ovinos
Corresponding Author's Secondary Institution:	
First Author:	Michelli Ferreira do Santos, DSc
First Author Secondary Information:	
Order of Authors:	Michelli Ferreira do Santos, DSc Isis Gomes de Brito Souza, DSc Sulimary Oliveira Gomes, DSc Geice Ribeiro da Silva Paul Bentzen, PhD Fabio Mendonça Diniz, Ph.D.
Order of Authors Secondary Information:	
Funding Information:	
Abstract:	Simple sequence repeat markers (SSR) of tri- and tetranucleotide motifs were developed for the brown spiny lobster <i>Panulirus echinatus</i> using next-generation sequencing (NGS). Based on 7,196 contigs generated by NGS, 27 SSR loci were initially selected for primer designing, of which 20 consistently amplified specific bands across individuals. Polymorphism was identified in 17 loci for three insular populations in the Atlantic Ocean, Saint Peter and Saint Paul Rocks, The Rocas Atoll, and Cape Verde, providing an overview of the genetic variation in the species. The number of alleles per locus ranged from 2 to 8 and observed and expected heterozygosities averaged 0.422 and 0.603, respectively. Polymorphic information content (PIC) ranged from 0.173 to 0.811. Significant deviations from Hardy-Weinberg equilibrium were found in six loci (Pech-1404, -1793, -3301, -4440, -7142, -7144), but no evidence of linkage disequilibrium was detected. Frequency of null alleles was < than 0.2 for most loci, except Pech3301 and Pech7142. A preliminary genetic diversity analysis revealed low values for population differentiation estimates ($F_{ST} = 0.039$, $R_{ST} = 0.036$, and $DEST = 0.089$). PCoA and Bayesian clustering could not distinguish lobster populations from each other avoiding separation of the species into distinct clusters. These molecular markers are the first SSR reported for the species and maybe be used for future conservation genetics research of <i>P. echinatus</i> .
Suggested Reviewers:	Luis Fernando Marins, DSc Professor, Universidade Federal do Rio Grande dqmluf@furg.br Molecular Genetics expert

Development of genome-wide informative simple sequence repeat markers and a snapshot of the genetic diversity of the spiny lobster, *Panulirus echinatus* using next generation sequencing

Michelli Ferreira do Santos^{a,b,1}, Isis Gomes de Brito Souza^{a,b,1}, Sulimary Oliveira Gomes^{a,b}, Geice Ribeiro da Silva^{a,b}, Paul Bentzen^c, Fabio Mendonça Diniz^{d,e}

^a Universidade Federal do Piauí, Campus Universitário Ministro Petrônio Portella, Ininga, Teresina, PI, 64049-550, Brazil

^b Universidade Federal do Piauí, Northeast Biotechnology Network RENORBIO, Teresina, PI, 64049-550, Brazil

^c Department of Biology, Dalhousie University, Halifax, NS, B3H 4J1, Canada

^d Empresa Brasileira de Pesquisa Agropecuária, Embrapa Meio-Norte, CP: 01, Teresina, PI, CEP: 64.006-220, Brazil

^e Empresa Brasileira de Pesquisa Agropecuária, Embrapa Caprinos e Ovinos, Caixa Postal: 145, Sobral, CE, CEP: 62010-970, Brazil

* Corresponding author. Tel.: +55 (88) 3112-7588; Fax: +55 (88) 3112-7455.

E-mail address: fabio.diniz@embrapa.br (F.M. Diniz).

¹ These authors contributed equally to this work.

Abstract: Simple sequence repeat markers (SSR) of tri- and tetranucleotide motifs were developed for the brown spiny lobster *Panulirus echinatus* using next-generation sequencing (NGS). Based on 7,196 *contigs* generated by NGS, 27 SSR loci were initially selected for primer designing, of which 20 consistently amplified specific bands across individuals. Polymorphism was identified in 17 loci for three insular populations in the Atlantic Ocean, Saint Peter and Saint Paul Rocks, The Rocas Atoll, and Cape Verde, providing an overview of the genetic variation in the species. The number of alleles per locus ranged from 2 to 8 and observed and expected heterozygosities averaged 0.418 and 0.587, respectively. Polymorphic information content (PIC) ranged from 0.173 to 0.811. Significant deviations from Hardy–Weinberg equilibrium were found in six loci (*Pech-6*, *Pech-8*, *Pech-11*, *Pech-14*, *Pech-15*, *Pech-16*), but no evidence of linkage disequilibrium was detected. Frequency of null alleles was < than 0.2 for most loci, except *Pech-11* and *Pech-15*. A preliminary genetic diversity analysis revealed low values for population differentiation estimates ($F_{ST} = 0.039$, $R_{ST} = 0.036$, and $D_{EST} = 0.089$). PCoA and Bayesian clustering could not distinguish lobster populations from each other avoiding separation of the species into distinct clusters. These molecular markers are the first SSR reported for the species and maybe be used for future conservation genetics research of *P. echinatus*.

Keywords: Population genetics, microsatellites, connectivity, brown spiny lobster, conservation.

1. Introduction

Lobster fisheries is, undoubtedly, one of the most financially valuable shellfish fisheries around the world, especially for spiny and clawed lobsters. The brown spiny lobster, *Panulirus echinatus* (**Smith, 1869**), occurs on rocky reefs in the tropical western Atlantic Ocean and central Atlantic Islands (**Holthuis, 1991**). The species is fished throughout its range, however, in Brazil and the islands of St. Helena (UK) and Cape Verde support commercial fishing activities.

Fishing in certain habitats may compromise distinct, locally adapted populations (**Agardy, 2000**) resulting in loss of genetic diversity. No fisheries data is available on the species that could be used to indicate whether an impact on the global population is existing. Therefore, it is of paramount importance to studying the genetic variability of the species throughout its distribution aiming at a management strategy to ensure that the populations remain at a sustainable level. So far, only very few papers can be found in the literature containing more than distribution records on this species, therefore, little is known about the population diversity of *P. echinatus*.

Polymorphic simple sequence repeat (SSR) markers, also known as DNA microsatellites, are useful molecular markers for fisheries population and conservation genetic studies due to their levels of intraspecific variability in many taxa, a consequence of the high mutation rate of these noncoding regions (**Wright and Bentzen, 1994**). As no molecular markers have been identified in *P. echinatus* until now and the species has limited genomic information, it seems, therefore, an efficient alternative for microsatellite discovery and subsequent PCR primer design to use genome assembled Illumina single-end DNA sequences via next generation sequencing technology (**Abdelkrim et al., 2009**).

In this study we have isolated and characterized a set of 20 simple sequence repeat markers of tri- and tetranucleotide motifs from *P. echinatus* using next-generation sequencing. Additionally, we briefly discuss the potential of using the developed markers to assess the genetic variability of the species.

2. Results and Discussion

2.1. Sequence assembly, SSR identification and primer design

The genomic library, which was previously loaded as 16 % of a MiSeq Reagent Kit v2 300 cycle sequencing run, produced 3,232,259 reads, which were assembled into 7,196 contigs. The software MSDB identified 2,959 microsatellite loci, being in the majority mononucleotide (1290; 43.60 %) and dinucleotide (1280; 43.26 %) repeats. For ease of imaging and scoring, we chose to examine only tri- and tetranucleotide loci, respectively showing 221 (7.47 %) and 107 (3.61 %) repeats. From these, 27 loci were chosen for primer designing and validation, of which 20 consistently amplified specific bands across individuals. Polymorphism was identified in 17 loci across the entire data set, the remaining three loci being monomorphic (*Pech-18*, *Pech-19* and *Pech-20*) (**Table 1**). Nevertheless, it is still expected that these loci may become polymorphic once again when additional individuals are sampled. If monomorphism is fully confirmed, these loci may also become useful as a new source of data as their flanking regions may also be used as a new tool for genetic studies (**Nazareno and dos Reis, 2011; Chatrou et al., 2009**). All microsatellite sequences were deposited in the NCBI database under GenBank accession numbers KX237634 to KX237653.

2.2. Polymorphism and validation of genomic SSR markers

The analysis with the program Micro-Checker showed in a pool of sampled individuals the presence of low to moderate frequencies of null alleles in these loci, which is not surprising as this is commonly observed in a variety of marine invertebrate species (**Kaukinen et al. 2004**). Null frequencies below 0.2 are acceptable in most microsatellite data sets (**Dakin and Avise, 2004**). Only in two loci (*Pech-11* and *Pech-15*) this estimate was higher than 0.2. The most probable reason for occurrence of this phenomenon being scoring errors due to stuttering or large allele drop out. No loci showed significant linkage disequilibrium between any pair of loci after corrections for multiple comparisons.

The genotyping of the entire dataset ($n = 42$) has also revealed 91 alleles at all 17 polymorphic loci, ranging from 2 (*Pech-10*) to 8 (*Pech-4*), with an average of 5.4 ± 1.4 alleles per locus (**Table 1**). The number of alleles at tri- and tetranucleotide

microsatellites in spiny lobsters has been reported to vary from 5.3 for Brazilian *P. argus* (Diniz et al., 2005) to 11.9 for *P. cygnus* (Kennington et al. 2010).

The size of alleles in the most polymorphic locus, *Pech-14*, based on A_R , H_E and PIC , varied from 160 to 178. For the less polymorphic locus, *Pech-10*, allele size ranged from 78 to 86 bp; a tri- and tetranucleotide motif loci, respectively.

The level of polymorphism of each locus was also evaluated by the allelic richness (A_R), heterozygosities (H) and the polymorphic information content (PIC) (Table 2). The values of allelic richness varied from 1.8 to 6.2 (average of 3.8 ± 1.0), while PIC values ranged between 0.173 and 0.811. Mean PIC (0.537 ± 0.165) characterize this set of microsatellite loci as highly informative markers, as defined by Botstein et al. (1980). Overall mean observed and expected heterozygosity was estimated to be 0.418 ± 0.165 and 0.587 ± 0.173 , respectively. These estimates were higher when compared to mean H_O and H_E of the scalloped spiny lobster (*P. homarus*) genotyped with microsatellite markers, 0.265 and 0.387, respectively (Delghandi et al., 2015). However, the heterozygocities were of lower magnitude when compared to other spiny lobsters, such as *P. argus* (Diniz et al. 2004; Diniz et al. 2005; Tringali et al. 2008), *P. ornatus* (Dao et al. 2013), *P. cygnus* (Kennington et al. 2010), *P. guttatus* (Truelove; Behringer; Preziosi et al. 2016), *P. interruptus* (Ben-Horin et al. 2009).

Eight microsatellite loci exhibited significant probabilities of departure from Hardy-Weinberg equilibrium expectations after sequential Bonferroni correction, adjusted critical $P < 0.0029$. In these cases the deviations occurred in the form of heterozygote deficiency, perhaps resulting from the small sample size or, considering the potential for extensive gene flow in this species, the Wahlund effect (Johnson and Black, 1984). Additionally, null alleles, even in low frequencies could have played an important part in the departures of HWE. Null allelism is common in microsatellite markers from many marine invertebrates (Dailianis et al., 2011).

Table 1 Characterization of 20 primer pairs that amplify short tandem repeats from *Panulirus echinatus*[†].

Locus	Primer sequence (5' - 3')	Core motif	Ta (°C)	A	Size range [§] (bp)	GenBank access
<i>Pech-1</i>	F: AAGCAGAACGAACTCACCGTA R: TTTCTTCATTGGTCAGAAGGGT	(CTA) ₁₁	55	5	119-146	KX237634
<i>Pech-2</i>	F: TCCTGGTTACGGAATAATGACC R: CTGTTGTTGTTTCCCATACACG	(CTG) ₉	55	6	109-115	KX237635
<i>Pech-3</i>	F: ACATAGATGGATGGGTAGATAGAC R: CTAAGTCTTTCTTAACTGCCT	(GATA) ₁₂	57	6	118-134	KX237636
<i>Pech-4</i>	F: GGGCCTTTCTTCATCTATTTCC R: GTCCAGGGCAAGATTGTGATA	(ATA) ₁₂	57	8	183-196	KX237637
<i>Pech-5</i>	F: TGTGTGTGTGTGTTTGTGTGTA R: CACACATATACGAGTCAAAACGAA	(TCTG) ₈	55	5	96-114	KX237638
<i>Pech-6</i>	F: CAATCTTAAAAGCAGACCAACG R: GAGGAAGAATAGGTAGAAATGCAGA	(ATA) ₁₅	55	6	134-158	KX237639
<i>Pech-7</i>	F: AGAGGCAAAGGGGTTTCATCTA R: TATACCACCATCACCAGCCATA	(ATCT) ₆	55	5	118-150	KX237640
<i>Pech-8</i>	F: TATTCTTGTTCTTTCTTCGCCC R: GAGTAGGAGGAGGAGGAAGAAGA	(TCT) ₁₄	55	3	117-123	KX237641
<i>Pech-9</i>	F: TCTATCTATCTATCTTTCATTAACGC R: GTTTTCGTCTCCTTTCCTGTGA	(TATA) ₇	55	6	72-86	KX237642
<i>Pech-10</i>	F: TGTGTGGTGAGAGGAGAGTTGA R: CTGGAAAGGATATGTTCAAAGGAG	(GATA) ₈	56	2	78-86	KX237643
<i>Pech-11</i>	F: TCAGGTGGACTGTGACCAAG R: CATATTATCATTATTATTATAATTATTATTAT	(TAA) ₄ N(TAA) ₁₀	52	5	190-211	KX237644
<i>Pech-12</i>	F: CCTCCCCTTTGAAGCTCTG R: TTATCATTAGTATCATTATGATCAT	(ATA) ₁₇	52	6	99-117	KX237645
<i>Pech-13</i>	F: CATAACGTGATGATAATGATGGTG R: GAGAAGAGGTACTAACTCACTGGC	(ATA) ₁₀	57	5	101-107	KX237646
<i>Pech-14</i>	F: TTTTGCTTGTGTGGCAAGATAC R: ACTCCGCCCTTGTTTTAAGTTT	(ATA) ₁₃	57	7	160-178	KX237647
<i>Pech-15</i>	F: GTATTGGTGCCTGGTTTATGC R: TGATGATGATGATAATGATAATAATGA	(TTA) ₁₅	55	6	162-194	KX237648
<i>Pech-16</i>	F: GCTTTCAGTATATGTCTATCTACCTG R: GATAGATAAACAGAGAAATAGATAAA	(TATC) ₁₁	55	6	87-107	KX237649
<i>Pech-17</i>	F: GCTGCCAAGCATTACGTTTT R: CTGAATGAGTACCTGCGTGC	(TATT) ₉	55	4	190-214	KX237650
<i>Pech-18</i>	F: CTCTCTCTATCTGTCTATCTATCTATC R: TGGCATCAGAGGTATAGAAAGACA	(TATC) ₁₁ (TCTA) ₁₇	56	1	175	KX237651
<i>Pech-19</i>	F: CACACACACATCATTGTATTATCATC R: TGATAACAACAACAACCAACA	(TAA) ₁₄	52	1	136	KX237652
<i>Pech-20</i>	F: CGGGAAACAGCGATCATGTA R: GGGGAAGATGTGTGTAGTGCAA	(TAA) ₂₇	57	1	194	KX237653

[†]Forty-two samples genotyped; [§]observed distribution of alleles at each locus; Ta, annealing temperature; A, number of alleles.

2.3. Genetic variation in *Panulirus echinatus*

A slightly higher level of polymorphism was revealed for the population of St. Peter & St. Paul Rocks, based on heterozygosities (H) and the polymorphic information content (PIC). The change is likely to be due to the fact that this archipelago is between Rocas and Cape Verde, and potential larval movements maybe occurring from both sites towards St. Peter & St. Paul Rocks, here favored by currents, counter-currents and gyres in the region.

The low magnitude for F_{ST} , R_{ST} and D_{est} estimates found in *P. echinatus*, respectively 0.039, 0.036 and 0.089, suggest some degree of genetic differentiation among the predefined populations (**Table 3**). However, measures of genetic differentiation may be overestimated by null alleles, even without modifying the overall significance of the result (**Chapuis and Estoup, 2007; Carlsson, 2008**). Therefore, more extensive research should be carried out to confirm this observed pattern.

Larval exchange among populations of *P. echinatus* is believed to be extensive within the studied oceanic region, as ocean currents are expected to play an important role on transporting larvae away from the original spawning areas, and consequently, patterns of genetic differentiation are likely to be subtle or inexistent. High rates of dispersion might have also contributed to the levels of population differentiation. Additionally, the species has a long-lived planktonic larval phase of about 12 months (**Góes et al., 2007**).

In fact, many genetic studies of spiny lobsters have shown little evidence of subdivision among geographically separated populations. Low F_{ST} estimates were reported in other spiny lobster species, such as *P. ornatus* (**Dao et al., 2015**), *P. argus* (**Truelove et al., 2015**) and *P. interruptus* (**Ben-Horin et al., 2009**).

The scatter-plot of the principal coordinates analysis (PCoA) failed to reveal a clear separation of the species in distinct clusters (**Figure 1**).

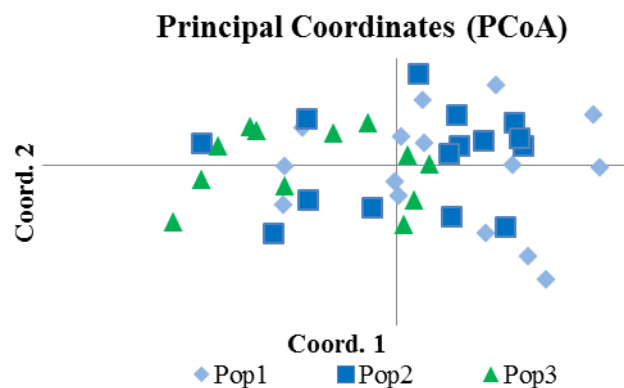


Figure 1. A. Principal coordinates analysis of microsatellite variation in *P. echinatus*.

Table 2. Genetic diversity statistics for the brown spiny lobster *Panulirus echinatus*.

Locus	All populations ($N = 42$)						
	A_R	H_O	H_E	PIC	P_{HWE}	F_{IS}	F_{Nulls}
<i>Pech-1</i>	3.6	0.667	0.656	0.583	0.2629	-0.0718	-0.0112
<i>Pech-2</i>	4.0	0.595	0.660	0.589	0.1778	0.0806	0.0345
<i>Pech-3</i>	3.9	0.395	0.654	0.583	0.0006*	0.3998	0.1522
<i>Pech-4</i>	4.5	0.472	0.698	0.644	0.0004*	0.3155	0.1278
<i>Pech-5</i>	4.2	0.564	0.683	0.628	0.0616	0.1261	0.0656
<i>Pech-6</i>	4.7	0.415	0.755	0.704	0.0000*	0.4051	0.1899
<i>Pech-7</i>	3.3	0.424	0.493	0.439	0.3548	0.0820	0.0414
<i>Pech-8</i>	2.4	0.095	0.293	0.266	0.0000*	0.8609	0.1504
<i>Pech-9</i>	4.0	0.524	0.618	0.568	0.2030	0.1431	0.0542
<i>Pech-10</i>	1.8	0.214	0.194	0.173	1.0000	-0.0943	-0.0193
<i>Pech-11</i>	4.4	0.281	0.730	0.672	0.0000*	0.5892	0.2543
<i>Pech-12</i>	3.8	0.469	0.515	0.470	0.1240	0.0639	0.0256
<i>Pech-13</i>	3.5	0.524	0.616	0.538	0.1258	0.1517	0.0527
<i>Pech-14</i>	6.2	0.595	0.842	0.811	0.0005*	0.2696	0.1293
<i>Pech-15</i>	3.9	0.118	0.549	0.506	0.0000*	0.7747	0.2746
<i>Pech-16</i>	4.8	0.405	0.696	0.652	0.0010*	0.4045	0.1666
<i>Pech-17</i>	2.6	0.357	0.343	0.306	0.6560	-0.0275	-0.0140
Mean	3.8	0.418	0.587	0.537	-	0.2528	-

A_R , allelic richness; H_O , observed heterozygosity; H_E , expected heterozygosity; PIC , polymorphic information content; P_{HWE} , probabilities of departure from Hardy–Weinberg Equilibrium (adjusted critical $P < 0.0029$); F_{Nulls} : null allele frequency.

Table 3. Variability of 17 microsatellite loci and F-statistics in *Panulirus echinatus* populations.

Locus	Rocas Atoll			St. Peter & St. Paul Rocks			Cape Verde			F_{ST}	R_{ST}	D_{est}
	H_O	H_E	PIC	H_O	H_E	PIC	H_O	H_E	PIC			
<i>Pech-1</i>	0.467	0.531	0.475	0.933	0.674	0.589	0.467	0.531	0.475	0.07429	0.070	0.125
<i>Pech-2</i>	0.467	0.549	0.421	0.667	0.756	0.682	0.467	0.549	0.421	0.02968	-0.013	0.052
<i>Pech-3</i>	0.357	0.632	0.528	0.308	0.588	0.496	0.357	0.632	0.528	-0.00100	-0.065	-0.003
<i>Pech-4</i>	0.429	0.638	0.541	0.571	0.717	0.649	0.429	0.638	0.541	0.02325	0.123	0.082
<i>Pech-5</i>	0.500	0.518	0.443	0.600	0.655	0.583	0.500	0.518	0.443	0.08107	0.080	0.169
<i>Pech-6</i>	0.467	0.701	0.619	0.467	0.726	0.652	0.467	0.701	0.619	0.11736	0.104	0.312
<i>Pech-7</i>	0.333	0.373	0.338	0.500	0.598	0.483	0.333	0.373	0.338	0.09271	-0.068	0.081
<i>Pech-8</i>	0.067	0.186	0.164	0.000	0.349	0.309	0.067	0.186	0.164	-0.04239	-0.046	-0.018
<i>Pech-9</i>	0.467	0.540	0.492	0.667	0.644	0.566	0.467	0.540	0.492	0.01974	0.096	0.033
<i>Pech-10</i>	0.133	0.129	0.117	0.267	0.239	0.204	0.133	0.129	0.117	-0.01867	-0.019	-0.005
<i>Pech-11</i>	0.100	0.511	0.406	0.143	0.772	0.704	0.100	0.511	0.406	0.10222	0.070	0.202
<i>Pech-12</i>	0.500	0.416	0.347	0.500	0.753	0.677	0.500	0.416	0.347	0.04287	0.120	0.046
<i>Pech-13</i>	0.600	0.602	0.519	0.467	0.572	0.460	0.600	0.602	0.519	-0.00106	0.012	0.001
<i>Pech-14</i>	0.667	0.766	0.709	0.400	0.841	0.786	0.667	0.766	0.709	0.05190	0.096	0.252
<i>Pech-15</i>	0.000	0.304	0.272	0.154	0.545	0.449	0.000	0.304	0.272	0.08538	-0.043	0.113
<i>Pech-16</i>	0.462	0.726	0.656	0.267	0.671	0.608	0.462	0.726	0.656	0.03976	0.119	0.098
<i>Pech-17</i>	0.333	0.297	0.260	0.333	0.402	0.361	0.333	0.297	0.260	-0.02275	-0.024	-0.012
Mean	0.373	0.495	0.429	0.426	0.617	0.544	0.373	0.495	0.429	0.0396	0.036	0.089

H_O , observed heterozygosity; H_E , expected heterozygosity; PIC , polymorphic information content.

The Bayesian cluster analysis was unable to identify any apparent population structure from the data set and, therefore, the three predefined populations were not geographically clear cut. When K was set between 2 to 6, the proportion of individuals assigned to each cluster assumed a near-uniform fashion, indicating that most accessions are fairly admixed individuals, with a few originated from mainly one population (see **Figure 2**).

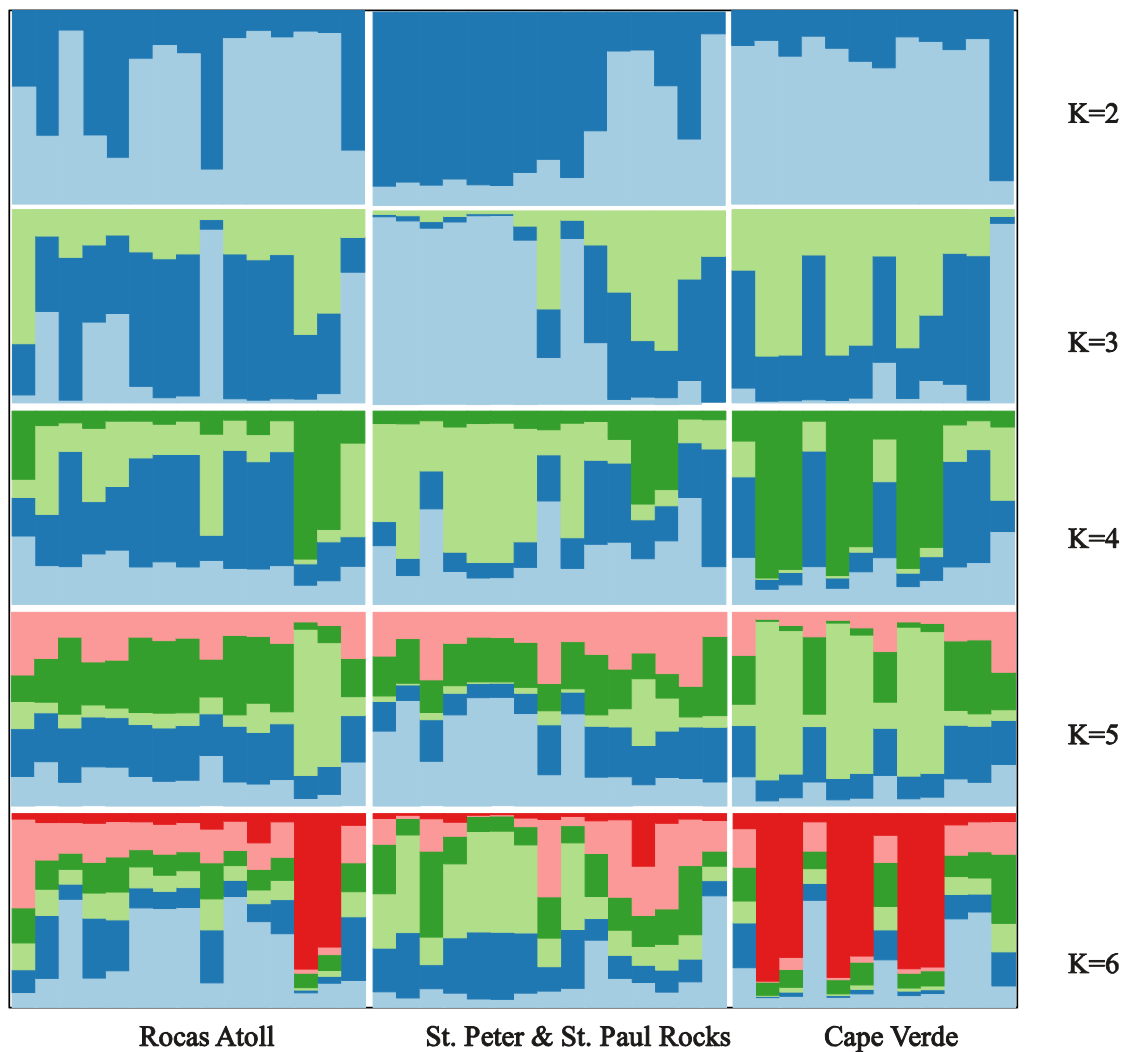


Figure 2. Estimated population structure in STRUCTURE for $K = 2$ to $K = 6$, permuted by CLUMPP and visualised by DISTRUCT. Each individual is represented by a vertical bar in K coloured segments, the length of each colour being proportional to the estimated membership coefficient. White lines separate individuals from different predefined populations.

All individual samples were dispersed majorly throughout the surveyed Atlantic islands with similar magnitudes of proportions of individuals (**Table 4**). The co-occurrence of individuals in the six clusters indicates admixture within the population. Admixture indicates that individuals may have mixed ancestry and have inherited some fraction of their genome from ancestors in cluster K.

Table 4. Proportion of membership of predefined populations in inferred clusters.

Predefined Populations	Inferred Clusters						Number of individuals
	1	2	3	4	5	6	
Rocas Atoll	0.196	0.156	0.192	0.196	0.126	0.136	15
St. Peter & St. Paul Rocks	0.127	0.179	0.187	0.184	0.158	0.166	15
Cape Verde	0.198	0.163	0.131	0.132	0.200	0.175	12

3. Experimental Section

3.1. Sampling and DNA extraction

Specimens of *P. echinatus* were collected from three different localities within the central Atlantic Islands of , Saint Peter and Saint Paul Rocks (0° 55' 2" N, 29° 20' 44" W), The Rocas Atoll (3° 51' 50" S, 33° 48' 48" W), and Cape Verde (14° 55' 0" N, 23° 31' 0" W). Tissue samples from walking legs (pereiopods) or antennae were immediately preserved either in 20% dimethylsulfoxide (DMSO) in a saturated sodium chloride solution (Amos and Hoelzel, 1991; Dawson et al., 1998) or in 95% ethanol, and stored at -20 °C until further use. PCR-ready DNA was extracted from the muscle tissue using a Sodium Hydroxide and Tris (HotShot) protocol as described previously by **Truett et al., 2000**. High molecular weight DNA was isolated by ethanol precipitation and visualized by gel electrophoresis.

3.2. NGSequencing of lobster DNA

At first, an Illumina paired-end library was created using 1 ng of *P. echinatus* genomic DNA, following the standard protocol of the Illumina Nextera XT Library Preparation kit (Illumina Inc., San Diego, CA). DNA was tagged and fragmented by the Nextera XT transposome, followed by limited-cycle PCR amplification, AMPure XP magnetic-bead purification (Agencourt Bioscience Corporation, Beverly, MA) and the Illumina Nextera XT bead-based normalization protocol. The DNA library was sequenced using a MiSeq Benchtop Sequencer (Illumina Inc., San Diego, CA). Contigs were created from the resulting paired-end sequence data (reads) using CLC Genomics Workbench 7.0.4 (Qiagen).

3.3. SSR mining and primer design

All these contigs were subsequently input into MSDB (Microsatellite Search Building Database) (<http://msdb.biosv.com/>; **Du et al., 2013**) for the detection of possible microsatellite loci with at least four repeats for tri- and tetranucleotide motifs and designing of primer pairs for each detected locus at their flanking regions. Long mononucleotide repeat stretches in sequences were ignored for marker development. Primer design was performed with the Primer3 (**Rozen and Skaletsky, 2000**)

3.4. SSR-PCR amplification for primer optimization and genotyping

Genomic DNA from 4 individuals were initially used to validate all designed primer pairs using polymerase chain reactions (PCRs). Reactions were performed in a 10 μ L total volume containing at least 20 ng of genomic DNA, with 1.5 μ L of 10 \times buffer (Thermo Fisher Scientific Inc), 2 to 2.5 mM MgCl₂ (Thermo Fisher Scientific Inc), 10 mM dNTP mix (New England BioLabs Inc.), 0.25 mM of each primer, 2.0 μ L de Q-solution (Qiagen), and 1 U of Taq DNA polymerase (Thermo Scientific Inc.) or HotStarTaq DNA Polymerase (Qiagen). All amplifications were run in a Veriti 96-well Thermal Cycler (Applied Biosystems) using the following PCR-cycle conditions: 95 °C for 10 minutes, 30 \times (1 minute at 95 °C, 1 minute at annealing temperature of 50 – 57 °C, depending on primer pair, 1 minute at 72 °C), and 7 minutes at 72 °C. All amplifications were run against a negative control in which all components except DNA template were added.

The amplification products were screened by silver nitrate detection on denatured 6% polyacrylamide gels. Additional lobster samples (N = 42) were genotyped from the three different locations described above to obtain baseline allele frequency information.

3.5. Data analysis

The genotyped data was initially analyzed using Micro-Checker 2.2.3 (**Van Oosterhout et al., 2004**) to test for the presence of null alleles, large alleles dropout and scoring errors by stuttering. Observed and expected heterozygosities (H_O and H_E), the number of alleles (N_a), and the polymorphic information content (PIC) were determined using CERVUS 3.0 (**Kalinowski; Taper Marshall, 2007**). Allelic richness (A_R) as a *measure* of the number of alleles per locus independent of the population size was calculated by *FSTAT* version 2.9.3.2 (**Goudet, 1995**). Deviations from Hardy–Weinberg equilibrium (HWE) and tests for linkage

disequilibrium were conducted using Genepop software (**Raymond and Rousset, 1995**). The Bonferroni correction (Rice, 1989) as applied when multiple pairwise tests were performed to assess the significance ($P < 0.05$).

The genetic diversity for each locus was evaluated by ARLEQUIN ver. 3.5 (**Excoffier and Lischer, 2010**), which determined the value of $\theta(F_{ST})$ for the whole sample set.

A Bayesian grouping admixture model was used to infer possible population structuring using the software STRUCTURE v2.3.3 (**Pritchard; Stephens; Donnelly, 2000**). The program was set up for 1,000,000 Markov chain Monte Carlo repetitions after an initial burn-in of 500,000 steps. The estimate of the best K was calculated based on 5 replications for each K (from 1 to 6) as described by **Evanno et al. (2005)** using Structure Harvester v.0.6.92 (<http://taylor0.biology.ucla.edu/structureHarvester/>) (**Earl and Vonholdt, 2012**). The program CLUMPP v.1.1.2 (**Jakobsson and Rosenberg, 2007**) was used to align the five repetitions of the best K. The program DISTRUCT v.1.1.1 (**Rosenberg, 2004**) was used to graphically display the results produced by CLUMPP. Population structure was also analyzed using principal coordinate analysis (PCoA), R_{ST} (**Slatkin, 1995**), a measure of genetic differentiation analogous to F_{ST} , and the D_{est} estimator of actual differentiation (**Jost, 2008**) as implemented in GenAlEx v.6.5 (**Peakall and Smouse, 2012**).

4. Conclusions

In this study, we report the isolation, using the NGS technology, and characterization of 17 polymorphic microsatellite loci for the brown spiny lobster *P. echinatus*. These microsatellites have demonstrated potential for population-level genetic studies and can be used effectively as molecular tool to aid in the conservation and management of this highly valuable fishery species.

The findings of this exploratory genetic study provide also a glimpse into the genetic variation in the species among the oceanic islands of Saint Peter and Saint Paul Rocks, The Rocas Atoll, and Cape Verde. A subtle pattern of genetic differentiation was detected, without revealing, however, a clear-cut geographic separation. A more detailed picture of population connectivity within the species' range is being further investigated using the same set of microsatellites markers, but larger sample size and more widespread sampling throughout the distribution of the species.

Acknowledgments

Sequencing was conducted on a MiSeq DNA sequencer purchased with a bequest from Elizabeth Ann Nielsen to the Marine Gene Probe Laboratory. The authors acknowledge financial support from the Brazilian Federal Government (Science without Borders program) scholarship. The authors MFS, IGBS, SOG and GRS were supported by scholarships from the *National Council for Scientific and Technological Development (CNPq)*.

Author Contributions

FMD, MFS and PB participated in the design of the study and coordinated the experiment, the analysis on the genetic diversity, and the draft of the manuscript; IGBS and PB performed NGSequencing; MFS, IGBS, SOG and GRS participated in the design of the study and executed the experiment, the analysis on the genetic diversity, and collected samples.

Conflicts of Interest

The authors declare no conflict of interest.

References

- Abdelkrim J, Robertson BC, Stanton JAL, Gemmell NJ. Fast, cost effective development of species specific microsatellite markers by genomic sequencing. *Biotechniques*. 2009; 46: 185-192
- Agardy, T. (2000). Effects of fisheries on marine ecosystems: a conservationist's perspective. *ICES J. Mar. Sci.*, 57: 761-765.
- Amos, B. and A. R. Holzel. 1991. Long term preservation of whale skin for DNA analysis. Rept. Intl. Whaling Commission. Special Issue No. 13:99-103.
- Ben-Horin, Matthew Iacchei, Kim A. Selkoe, Thien T. Mai, Rob J. Toonen. Characterization of eight polymorphic microsatellite loci for the California spiny lobster, *panulirus interruptus* and cross-amplification in other achelate lobsters. *Conservation Genetics Resources*, v. 1, n. 1, p. 193–197, 2009.
- Botstein D, White RL, Skolnick M (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 32:314–331.
- Carlsson J. 2008. Effects of microsatellite null alleles on assignment testing. *Journal of Heredity* 99:616-623.
- Chapuis MP, Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621_631.
- Chatrou, L.W.; Escribano, M.P.; Viruel, M.A.; Maas, J.W.; Richardson, J.E.; Hormaza, J.I. Microsatellite flanking regions provide a new source of data for plant species-level phylogenetics. *Molecular Phylogenetics and Evolution*, Philadelphia, v.53, p.726-733, 2009.
- Dakin, E. E.; Avise, J. C. Microsatellite null alleles in parentage analysis. *Heredity*, v. 93, n. 5, p. 504–9, 2004.

- Dailianis T, Tsigenopoulos CS, Dounas C, Voultziadou E. 2011. Genetic diversity of the imperilled bath sponge *Spongia officinalis* Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation. *Molecular Ecology* 20:3757_3772
- Dao, H. T.; Todd, E. V.; Jerry, D. R. Characterization of polymorphic microsatellite loci for the spiny lobster *Panulirus* spp. and their utility to be applied to other *Panulirus* lobsters. *Conservation Genetics Resources*, v. 5, n. 1, p. 43–46, 2013.
- Dao, H. T. et al. Oceanographic Currents and Local Ecological Knowledge Indicate , and Genetics Does Not Refute , a Contemporary Pattern of Larval Dispersal for The Ornate Spiny Lobster , *Panulirus ornatus* in the South-East Asian Archipelago. p. 1–19, 2015.
- Dawson MN, Raskoff KA, Jacobs DK (1998) Field preservation of marine invertebrate tissue for DNA analyses. *Mol Mar Biol Biotechnol* 7, 145–152.
- Delghandi, M. Goddard, S., Jerry, D.R., Dao, H.T., Afzal H. and Al-Jardani S.S. Isolation, characterization, and multiplexing of novel microsatellite markers for the tropical scalloped spiny lobster (*Panulirus homarus*). *Genetics and Molecular Research*, v. 14, n. 4, p. 19066–19070, 2015.
- Diniz, FM, Maclean, N, Ogawa, M, Paterson, IG and Bentzen, P. Microsatellites in the overexploited spiny lobster, *Panulirus argus*: Isolation, characterization of loci and potential for intraspecific variability studies. *Conservation Genetics*, v. 6, n. 4, p. 637–641, 2005.
- Diniz, FM, Maclean, N, Paterson, IG and Bentzen, P. Polymorphic tetranucleotide microsatellite markers in the Caribbean spiny lobster, *Panulirus argus*. *Molecular Ecology Notes*, v. 4, n. 3, p. 327–329, 2004.
- Du, L., Li Y, Zhang X, Yue B. MSDB: A user-friendly program for reporting distribution and building databases of microsatellites from genome sequences. *Journal of Heredity*, v. 104, n. 1, p. 154–157, 2013.
- Earl DA, VonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361.
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14:2611–2620.
- Excoffier, L.; Lischer, H. E. L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, v. 10, n. 3, p. 564–567, 2010.
- Peakall R, Smouse PE (2012) GenAIEx 65: Genetic analysis in Excel Population genetic software for teaching and research: An update. *Bioinformatics* 28:2537–2539.
- Góes, C.; Lins-Oliveira; JE. Natural diet of the spiny lobster, *Panulirus echinatus* Smith, 1869 (Crustacea: Decapoda: Palinuridae), from São Pedro and São Paulo Archipelago, Brazil. *Braz. J. Biol. J. Biol*, v. 69, n. 691, p. 143–148, 2009.
- Goudet J (1995) FSTAT Version 2932: a computer program to calculate F-statistics. *J Hered* 86: 485–486.
- Holthuis, L.B. (1991) FAO species catalogue. Vol. 13. Marine Lobsters of the World. An Annotated and Illustrated Catalogue of Species of Interest to Fisheries Known to Date. *FAO Fisheries Synopsis* 125:1-292.
- Johnson MS, Black R. 1984. The Wahlund effect and the geographical scale of variation in the intertidal limpet *Siphonaria* sp. *Marine Biology* 79:295-302
- Jakobsson, M.; Rosenberg, N. A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, v. 23, n. 14, p. 1801–1806, 2007.

- Jost, L. GST and its relatives do not measure differentiation. *Molecular Ecology*, v. 17, n. 18, p. 4015–4026, 2008.
- Kalinowski, S. T.; Taper, M. L.; Marshall, T. C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, v. 16, n. 5, p. 1099–1106, 2007.
- Kaukinen KH, Supernault KJ, Miller KM (2004) Enrichment of tetranucleotide microsatellite loci from invertebrate species. *J Shel Res* 23:621–626.
- Kennington, W. J., Levy E., Berry O., Groth D. M., Waite A. M., Johnson M. S., Melville-Smith R. Characterization of 18 polymorphic microsatellite loci for the western rock lobster *Panulirus cygnus*. *Conservation Genetics Resources*, v. 2, n. SUPPL.1, p. 389–391, 2010.
- Nazareno, A. G. ; Reis, M. S. d. ; REIS, Maurício Sedrez dos . The same but different: Monomorphic microsatellite markers as a new tool for genetic analysis. *American Journal of Botany* , v. 98, p. e265-e267, 2011.
- Pritchard, J. K.; STEPHENS, M.; DONNELLY, P. Inference of population structure using multilocus genotype data. *Genetics*, v. 155, n. 2, p. 945–959, 2000.
- Raymond, M.; Rousset, F. An exact test for population differentiation. *Evolution*, v. 49, n. April, p. 1280–1283, 1995.
- Rice, W. R. *Analyzing Tables of Statistical Tests*. *Evolution*, 1989.
- Rosenberg, N. A. DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, v. 4, n. 1, p. 137–138, 2004.
- Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers In: Krawetz, S, Misener, S (Eds), *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365–386.
- Slatkin, M. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, v. 139, n. 1, p. 457–462, 1995.
- Tringali, M. D.; Seyoum, S.; Schmitt, S. L. Ten di- and trinucleotide microsatellite loci in the Caribbean spiny lobster, *Panulirus argus*, for studies of regional population connectivity. *Molecular Ecology Resources*, v. 8, n. 3, p. 650–652, 2008.
- Truelove, N.; Behringer, D. C.; Preziosi, R. F. Isolation and characterization of eight polymorphic microsatellites for the spotted spiny lobster, *Panulirus guttatus*. p. 4–11, 2016.
- Truelove, N. K., Behringer D.C., Butler IV, MJ and Preziosi, RF. Genetic analysis reveals temporal population structure in Caribbean spiny lobster (*Panulirus argus*) within marine protected areas in Mexico. *Fisheries Research*, v. 172, n. April, p. 44–49, 2015.
- Truett GE, Heeger P, Mynatt RL, TrueWalker JA, Warman ML (2000) Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques* 29:52–54
- Van Oosterhout C, Hutchinson WF, Wills DPM, *et al.* (2004) Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4, 535-538.
- Wright, J.M., and Bentzen, P. (1994). Microsatellites – Genetic Markers For The Future. *Rev. Fish Biol. Fish.*, 4: 384-388.

APÊNDICE D



International Journal of Biological Sciences

Impact factor **4.509**  **RSS FEED**

Manuscript Status/Login Contact

ISSN 1449-2288 24 May 2016

[Home](#)[Index & Ranking](#)[Current Issue](#)[Archive](#)[Cover Images](#)[Editorial Board](#)[Author Info](#)[Submission](#)[Special Issues](#)[Contact](#)

Journal of Genomics now in PubMed/PubMed Central. Submit manuscript...

International Journal of Medical Sciences

Journal of Cancer

Theranostics

Journal of Genomics

Journal of Bone and Joint Infection (JBJI)

**Manuscript ID: 16278c**

Your manuscript has been successfully received! Thank you. A confirmation email has also been sent to fabio.diniz@embrapa.br.

Manuscript file received: Article Genetic Differentiation_Suli et al_24 05 2016_IJBS.doc.docx
File size: 0.2124 MB

Covering letter file received: Cover letter IJBS Suli et al 24 05 2016.pdf
File size: 0.1969 MB

Submission type: new

Manuscript title: Development and characterization of tetranucleotide microsatellite markers for the blue land crab *Cardisoma guanhumi* using Illumina paired-end sequencing data and a preliminary evidence of genetic differentiation in the Caribbean

Manuscript author(s): Sulimary Oliveira Gomes, Isis Gomes de Brito Souza, Michelli Ferreira do Santos, Geice Ribeiro da Silva, Marc Albrecht, Alan Shane McKinley, Paul Bentzen, Fábio Mendonça Diniz

Submitting author name: Dr. Fábio Mendonça Diniz

Contact email: fabio.diniz@embrapa.br

Additional email or contact info (optional): fmdiniz1@yahoo.co.uk

Manuscript category: Short Research Communication|Ecology

Suggested reviewers: Rodrigo Maggioni, PhD
Professor, Universidade Federal do Ceará
rodrigo.maggioni@gmail.com
Population Genetics expert

Luis Fernando Marins, D.Sc.
Professor, Universidade Federal do Rio Grande
dqmluf@furg.br
Molecular Genetics expert

Arati Iyengar, PhD
Senior Lecturer , University of Central Lancashire
aiyengar@uclan.ac.uk
Population genetics expert

Declaration made by the submitting author:

On behalf of all co-authors, I/we declare:

that I/we agree to all the terms of the [Copyright and Publishing Agreement](#) and declare that the article is original, has not been previously published and is not under consideration for publication elsewhere;

that I/we understand there is a [Publication Fee](#) charged by the Journal and agree to pay the fee if and when the article is accepted for publication.

IP address recorded: 187.60.32.115