

IDENTIFICAÇÃO DE LOCI MICROSSATÉLITES EM PEIXES UTILIZANDO NGS

HUDSON LIMA DIAS¹; NATÁLIA CARRILHO BARRETO²; ADRIANA PEREIRA DA FRANCA³; VITÓRIA CARRILHO BARRETO⁴; RAFAEL ALDRIGHI TAVARES⁵

¹Universidade Federal de Pelotas – Pelotas, RS – huddias96@hotmail.com

²Universidade Federal de Pelotas – Pelotas, RS – nataliacbrt@gmail.com

³Universidade Federal de Pelotas – Pelotas, RS – vcarrilho05@gmail.com

⁴Universidade Federal de Pelotas – Pelotas, RS – drikafranca13@gmail.com

⁵Universidade federal de Pelotas, Dep. de Zootecnia – Pelotas, RS – r.tavares@ufpel.edu.br

1. INTRODUÇÃO

O uso de ferramentas moleculares tem impulsionado os estudos genéticos em organismos aquáticos, especialmente por meio dos marcadores microssatélites (SSRs), que se destacam pela alta variabilidade e ampla aplicação em conservação e melhoramento (SELKOE; TOONEN, 2006). A redução da diversidade genética causada pela pesca e pela ausência de manejo reforça a importância de monitorar espécies nativas de interesse produtivo.

O objetivo do trabalho foi utilizar o sequenciamento de nova geração como ferramenta para descoberta de *loci* potencialmente amplificáveis de microssatélites em diferentes espécies de peixes.

2. METODOLOGIA

Para a identificação de microssatélites em sete espécies de peixes (*Brycon orbignyanus*, *Salminus brasiliensis*, *Pimelodus maculatus*, *Hemisorubim platyrhynchos*, *Leiarius marmoratus*, *Odontesthes humensis* e *Loricariichthys anus*) utilizando o sequenciamento de nova geração, inicialmente a extração de DNA total foi realizada a partir da nadadeira caudal das diferentes espécies de peixes, utilizando o protocolo de cloreto de sódio (TAVARES, 2011), conhecido por sua simplicidade, rapidez e facilidade. O DNA obtido foi mantido a -20°C. A qualidade e integridade do DNA extraído foram avaliados por meio de eletroforese em gel de agarose a 1%, corado com GELRED® e visualizado sob luz ultravioleta (UV) em transiluminador (Clinx Science®). Após a quantificação e verificação da pureza, o DNA foi fragmentado em tamanhos adequados e enriquecido com adaptadores necessários para o sequenciamento *Illumina paired-end* (GOODWIN; MCPHERSON; MCCOMBIE, 2016 e GRIFFITHS, et al. 2016). O sequenciamento desta única biblioteca genômica foi conduzido em um sequenciador MiSeq (Illumina), San Diego, (USA) com 150 pares de bases. As leituras obtidas foram analisadas no programa PAL_FINDER_V0.02.03 para identificar as leituras que continham microssatélites com dinucleotídeos (2mer), trinucleotídeos (3mer), tetranucleotídeos (4mer), pentanucleotídeos (5mer) e hexanucleotídeos (6mer). Estas sequências, para serem consideradas SSRs, deveriam ter pelo menos 12 pb de comprimento para 2mer, 3mer e 4mer; e de 15pb a 18pb para 5mer e 6mers, respectivamente.

Uma vez identificadas as leituras, com *loci* SSRs, estas foram agrupadas para um subdiretório local do programa Primer 3 v.2 (ROZEN & SKALETSKY, 2000) para o desenho dos primers e identificação de *Loci Potencialmente Amplificáveis* (PAL). Foram utilizados alguns critérios para o desenho dos primers, os mesmos utilizados por CASTOE et al. (2102): 1 – Conteúdo de GC maior que 30%; 2 - temperaturas de melting entre 58°C e 65°C, com um máximo de 2°C de diferença

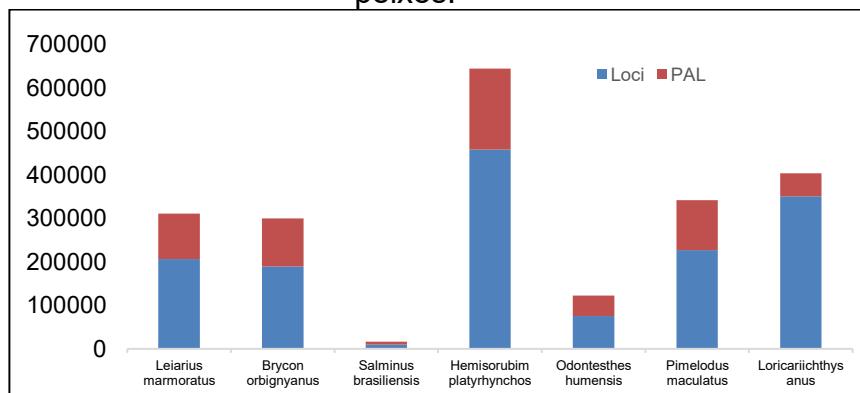
entre os primers; 3 - os últimos dois nucleotídeos na extremidade 3', sendo G ou C, 4 - máximo poli-N de 4 nucleotídeos. Se todos os outros critérios forem atingidos, um único par de primer é escolhido, apresentando o maior escore assinalado pelo Primer 3, além do maior tamanho da região de amplificação da sequência repetida.

3. RESULTADOS E DISCUSSÃO

A utilização do sequenciamento *Illumina paired-end* para descoberta de microssatélites em peixes representa um avanço tecnológico expressivo. A geração massiva de dados, aliada à precisão das leituras, possibilita não apenas a identificação de milhares de marcadores em curto período, mas também o acesso a regiões genômicas pouco exploradas (EKBLOM; GALINDO, 2011).

A espécie *Hemisorubim platyrhynchos* apresentou o maior número de leituras (1.816.200) e o maior número de SSRs identificados (271.144) e a espécie *Loricariichthys anus* o menor número de leituras (1.103.580), porém a espécie *Salminus brasiliensis* o menor número de SSRs identificados (8.770). Os SSRs identificados variam em número de *loci* e número de *loci* potencialmente identificados (PAL), com destaque para a espécie *Hemisorubim platyrhynchos* com maior número de *loci* e PAL e a espécie *Salminus brasiliensis* com menores números (Figura 1).

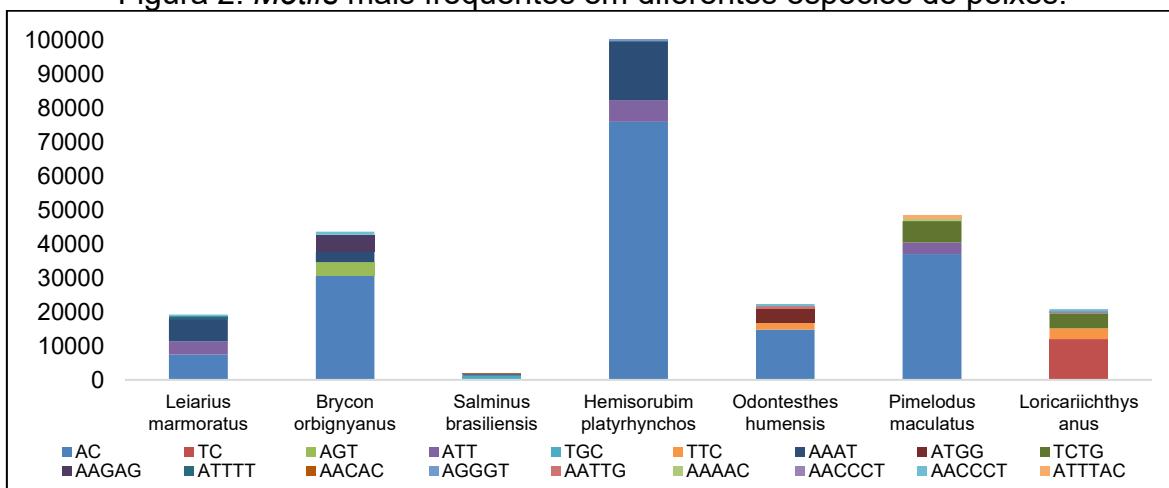
Figura 1: *Loci* e *Loci* Potencialmente amplificáveis em diferentes espécies de peixes.



Os *motifs* mais frequentes, entre as espécies de peixes, dos tipos de dinucleotídeos, trinucleotídeos, tetranucleotídeos e hexanucleotídeos foram AC, ATT, AAAT e AACCT respectivamente (Figura 2). As repetições de pentanucleotídeos foram muito variáveis entre as espécies, não apresentando um *motif* com maior frequencia, mas o *motif* AAGAG apresentou um número bem elevado em relação aos outros, porém somente na espécie *Brycon orbignyanus* (Figura 2).

Em peixes, a caracterização genômica de SSRs tem sido realizada em diversos genomas e transcriptomas, demonstrando variabilidade na distribuição e abundância desses marcadores em regiões codificantes e não codificantes (ZHANG et al., 2021). O SSRs comumente encontrado na literatura em peixes, segundo os padrões de distribuição são: di- > mono- > tetra- > tri- > penta- > hexanucleotídeos para espécies como *Danio rerio*, *Takifugu rubripes*, *Ictalurus punctatus*, *Esox lucius*, *Oncorhynchus mykiss* e o *Astyanax mexicanus* da Ordem Characiformes (LEI et al., 2021).

Figura 2: *Motifs* mais frequentes em diferentes espécies de peixes.



Em *M. anguillicaudatus*, foram identificados 8.194 loci SSRs (FENG et al., 2018) e em *Dissostichus eleginoides*, a frequência de SSR foi estimada como aproximadamente um SSR por ~1 kB de transcrito, com predominância de *motif* dinucleotídeo, refletindo padrões observados em teleósteos (SOTO et al., 2019).

Os *motifs* TC e AC estão entre os mais frequentes, como em *Raja pulchra* (KANG et al. 2012) e *Schizothorax biddulphi* (LUO et al. 2012), indicando que mesmo com estas diferenças entre as espécies, o valor encontrado vai de encontro com MEGLÉCZ et al. (2012), de que os *motifs* mais comuns de dinucleotídeos em Chordata, dependendo da espécie, são AC e TC. Novamente segundo MEGLÉCZ et al. (2012) em geral os motivos mais encontrados para trinucleotídeos e tetranucleotídeos para o grupo Chordata são AAT e AGAT.

Diferenças entre o número de SSRs entre espécie pode estar relacionado com a fragmentação do genoma para a criação da biblioteca, por ser aleatória, não existindo garantia de que o genoma inteiro siga o mesmo padrão do encontrado durante a criação da biblioteca (CALABUIG et al. 2012), ainda mais que se obteve uma pequena cobertura do genoma neste trabalho. Mas ainda assim deve-se levar em conta a origem da diversidade da composição dos microssatélites entre os diferentes grupos filogenéticos, que pode ser resultado dos diferentes mecanismos de mutação, que alcançam processos evolutivos e a seleção natural (MEGLÉCZ et al. 2012), sugerindo então que as frequências e motivos dos microssatélites são específicos e fazem parte do processo evolutivo.

4. CONCLUSÕES

A tecnologia NSG consolidou-se como ferramenta essencial para a descoberta e aplicação de microssatélites em peixes, integrando avanços em conservação e melhoramento genético.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- CALABUIG, C. M. D. N. et al. Genome-wide identification and characterization of microsatellite loci in coscoroba swan (*Coscoroba coscoroba*). **Genomics and Quantitative Genetics**, v.5, p.14-19, 2012. ISSN 2157-9903.
- EKBLOM, R.; GALINDO, J. Applications of next generation sequencing in molecular ecology of non-model organisms. **Heredity**, London, v.107, p.1-15, 2011.

FENG, B. et al. Development of novel EST-SSR markers for ploidy identification based on de novo transcriptome assembly for *Misgurnus anguillicaudatus*. **PLOS One**, San Francisco, v.13, n.4, e0195829, 2018.

GOODWIN, S.; MCPHERSON, J. D.; MCCOMBIE, W. R. Coming of age: ten years of next-generation sequencing technologies. **Nature Reviews Genetics**, London, v.17, p.333–351, 2016.

GRIFFITHS, S. M. et al. A Galaxy-based bioinformatics pipeline for optimised, streamlined microsatellite development from Illumina next-generation sequencing data. **Conservation Genetics Resources**, Dordrecht, v.8, n.4, p.481-486, 2016.

KANG, J. H.; PARK, J. Y.; JO, H. S. Rapid development of microsatellite markers with 454 pyrosequencing in a vulnerable fish, the mottled skate, *Raja pulchra*. **International Journal of Molecular Sciences**, Basel, v.13, p.7199-7211, 2012. Disponível em: <https://doi.org/10.3390/ijms13067199>. Acesso em: 1 jul. 2024. doi: 10.3390/ijm1307199.

LEI, Y. et al. Genome-wide characterization of microsatellite DNA in fishes: survey and analysis of their abundance and frequency in genome-specific regions. **BMC Genomics**, London, v.22, n.1, p.1-22, 2021. Disponível em: <https://doi.org/10.1186/s12864-021-07752-6>. Acesso em: 28 jun. 2024. doi: 10.1186/s12864-021-07752-6.

LUO, W. et al. Rapid development of microsatellite markers for the endangered fish *Schizothorax biddulphi* (Günther) using next generation sequencing and cross-species amplification. **International Journal of Molecular Sciences**, Basel, v.13, n.11, p.14946-14955, 2012. Disponível em: <https://doi.org/10.3390/ijms131114946>. Acesso em: 1 jul. 2024. doi: 10.3390/ijms131114946.

MEGLÉCZ, E. et al. Breakdown of phylogenetic signal: a survey of microsatellite densities in 454 shotgun sequences from 154 non model eukaryote species. **PLOS One**, San Francisco, v.7, n.7, e40861, 2012. Disponível em: <https://doi.org/10.1371/journal.pone.0040861>. Acesso em: 1 jul. 2024. doi: 10.1371/journal.pone.0040861.

ROZEN, S.; SKALETSKY, H. Primer3 sobre a WWW para usuários gerais e para programadores biólogos. In: MISENER, S.; KRAWETZ, S. A. (Eds.). **Métodos e protocolos de bioinformática. Métodos em Biologia Molecular™**, v.132. Totowa: Humana Press, 2000. p.365-386. doi: <https://doi.org/10.1385/1-59259-192-2:365>.

SELKOE, K. A.; TOONEN, R. J. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. **Ecology Letters**, Oxford, v.9, p.615–629, 2006.

SOTO, L. et al. De novo assembly and characterization of Patagonian toothfish transcriptome and development of EST-SSR markers for population genetics. **Frontiers in Marine Science**, Lausanne, 2019.

TAVARES, R. A. et al. Utilization of microsatellite markers to form families of *Odontesthes bonariensis* in a genetic breeding program. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v.63, p.1263-1267, 2011. Disponível em: <https://dx.doi.org/10.1590/S0102-09352011000500034>. Acesso em: 2 jul. 2024. doi: 10.1590/S0102-09352011000500034.

ZHANG, X. et al. Dynamic transcriptome sequencing and analysis during early development in the bighead carp. **BMC Genomics**, London, 2019.