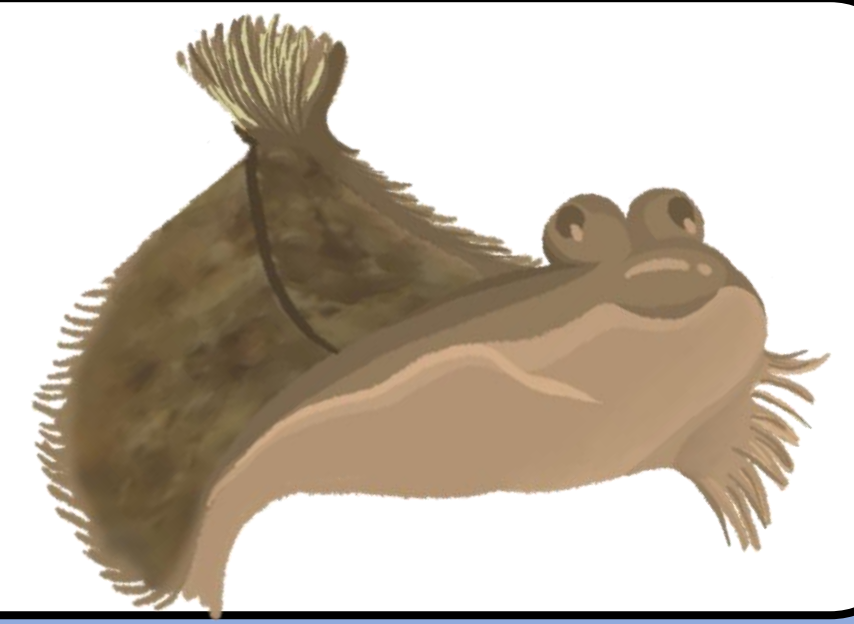


Caracterización de las regiones centroméricas del lenguado senegalés (*Solea senegalensis*, Kaup 1858)

Aaron Gálvez Salido, Claudio Jimenez, Ismael Manzanares García, Mateo Zapata, Roberto de la Herrán, Francisca Robles, Carmelo Ruiz Rejón, Rafael Navajas-Pérez



Introducción

Los centrómeros juegan un papel fundamental en el ciclo celular e intervienen en el proceso de unión entre cromátidas hermanas o cromosomas homólogos y su posterior segregación durante la mitosis y la meiosis, respectivamente. Las secuencias repetidas en tándem, como las secuencias satélite, son las componentes principales de estas regiones heterocromáticas. A pesar del alto número de especies que componen el grupo de los peces planos, únicamente se han caracterizado secuencias centroméricas de un reducido número de especies que son: el lenguado chino, *Cynoglossus semilaevis* (Ji et al., 2009), el lenguado *Achirus lineatus* (Carvalho de Azevedo et al. 2005), la acedia, *Dicologlossa cuneata* (de la Herrán et al., 2008) y el halibut, *Hippoglossus hippoglossus* (Robles et al., 2017). En cuanto al lenguado senegalés, se ha caracterizado una secuencia de ADN satélite, la familia *PvuII*, que se acumula específicamente en las regiones centroméricas de 19 parejas de cromosomas (Robles et al., 2017). Se ha encontrado la presencia bioinformática de *PvuII* en la mayoría de los cromosomas, con algunas excepciones debidas a problemas del ensamblaje y además las regiones centroméricas de los cromosomas 7 y 8 están sin caracterizar.

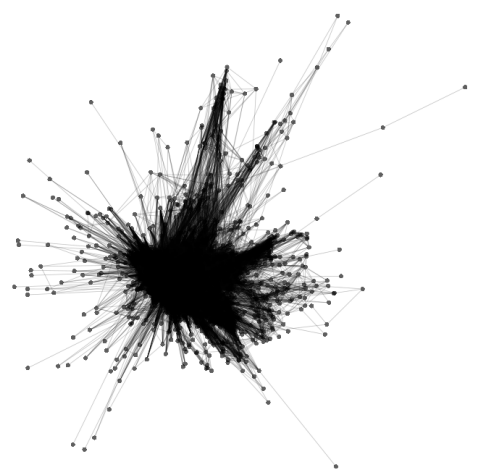
Material y Métodos

Se realizó un análisis bioinformático desde dos aproximaciones diferentes:

- Estudio de clusterización mediante **RepeatExplorer2** de lecturas de Illumina para detección de secuencias repetidas en tándem
- Estudio de la composición en elementos móviles con **RepeatMasker** usando el ensamblaje del genoma

Para las anotaciones de las secuencias de forma manual se utilizó una **base de datos secuencias repetidas de peces planos**, proporcionada por el grupo de Microbiología Aplicada y Genética Molecular del departamento de Biomedicina, Biotecnología y Salud Pública de la Universidad de Cádiz. Las secuencias detectadas mediante los análisis anteriores se mapearon sobre los cromosomas del ensamblaje con **BLASTn** para obtener las coordenadas y se representaron con un script en Python para realizar ideogramas. Para comprobar la estructura secundaria de las secuencias de ADN satélite centromérico se utilizó **RNAfold**. También se hicieron análisis más precisos mediante dotplots y **Tandem Repeat Finder** en los cromosomas 7 y 8 para caracterizar la región centromérica

Cluster 3 RepeatExplorer2



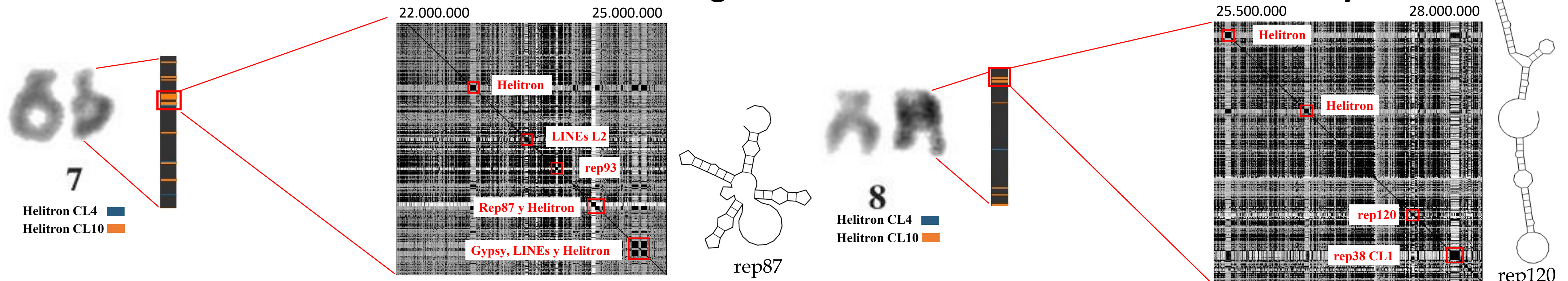
Identidad del 91,5 % a la familia de ADN satélite *PvuII* previamente caracterizada en esta especie como una secuencia localizada en la región centromérica

Alineamiento consenso cluster 3 y elemento tipo LINE tipo L1

LINE/L1	1	AGTGTTTGATTTTGCNANGATTTTCANTGTTNGGATGNTTTTTCTATTGTTGTCGTGTCGTTTTCNCAATACGATAAAGGTGTTTTCGTTGTTGTTTTTTTTAAATTTTTCTAGCGTCCGTAATAT	128
Tetrámero pvuII	1	-----GTGAC-----TTTTTAAATTTTTCTAGCTGTCGTTACAT	35
LINE/L1	126	AAATCTCACAAATGAAGTTAACATCTTAAATAATTTACTTCTGTTCCTCAGTGTGATTCTGCAATGATTTCACTGTTGGGATGTTTTCTATTAGTGTCTCTTTGTGGCCATTCAAGCTGA	256
Tetrámero pvuII	36	AAAACCTCACAAATGAAGTTAACATCTTAAAGTTACTTCTGTTCCTCAGTGTGATTCTGCAATGATTTCACTGTTGGGATGTTTTCTATTAGTGTCTCTTTGTGGCCATTCAAGCTGA	102
LINE/L1	257	GATAAAGGCTGTTTASTGACTTTTTCT-----AGTGTTTGATACGTAAATGATTTCACTGTT	310
Tetrámero pvuII	103	GATAAAGGCTGTTTCTGACTTTTTTAATTTTTCTACTGTCTCGTTACATAAAACTCACAAATGAAGTTAACATCTTAAAGTTTACTTCTGTACCCCTCAGTGTGATTTGCAATGATTTCACTGTT	290
LINE/L1	311	TGGGATGTTTTCTATTAGTGTCTCTTTGTGGCCATTCAAGCTGAGATAAGGCTGTTTCGTTGACTTTTTAAATTTTTCTACTGTCTCGTTACATAAAACTCACAAATGAAGTTAACATCTTAA	438
Tetrámero pvuII	291	TGGGATGTTTTCTATTAGTGTCTCTTTGTGGCCATTCAAGCTGAGATAAGGCTGTTTCGTTGACTTTTTAAATTTTTCTACTGTCTCGTTACATAAAACTCACAAATGAAGTTAACATCTTAA	418
LINE/L1	439	GTATACTTCTGTACCCCTCAGTGTGATTGCAATGATTTCACTGTTGGGATGTTTTCTATTAGTGTCTCTTTGTGGCCATTCAAGCTGAGATAAGGCTGTTTCGTTGACTTTTTAAATTTTT	506
Tetrámero pvuII	419	GTATACTTCTGTACCCCTCAGTGTGATTGCAATGATTTCACTGTTGGGATGTTTTCTATTAGTGTCTCTTTGTGGCCATTCAAGCTGAGATAAGGCTGTTTCGTTGACTTTTTAAATTTTT	540
LINE/L1	507	CTACTGTCCGTTNACATAAANCTCACAAATGAAGTTAACATCTTAAAGTTTAACTTCTGTACCCCTCAGTGTGATTGCAATGATTTCACTGTTGGGATGTTTTCTATTAGTGTCTCTTTGT	604
Tetrámero pvuII	547	CTACTGTCTCGTTACATAAANCTCACAAATGAAGTTAACATCTTAAAGTTTAACTTCTGTACCCCTCAGTGTGATTGCAATGATTTCACTGTTGGGATGTTTTCTATTAGTGTCTCTTTGT	674
LINE/L1	605	GTGGCCATTCAAGCTGAGATAAGGCTGTTTAGTACTTTTTNGTTTTTAAACATCTTAAAGATATACTTCTGTACCCCTGACATAAAATACCATCACTGTCACTACAGTACAATTTAACTAGA	822
Tetrámero pvuII	675	GTGGCCATTCAAGCTGAGATAAGGCTGTTT-----	704

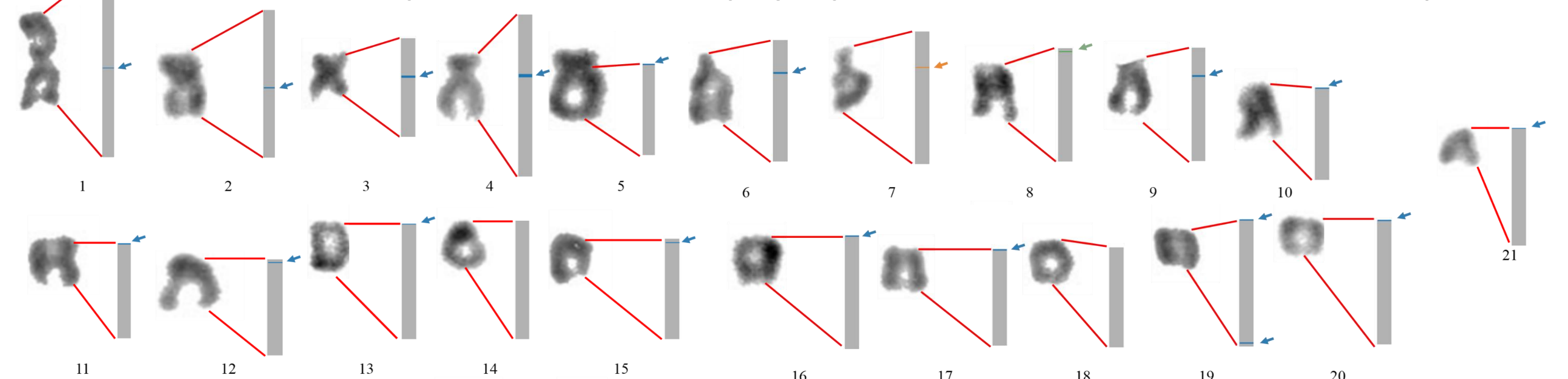
Se descubrió mediante BLASTn que esta secuencia satélite se asocia con un elemento repetido encontrado en la base de datos de peces planos, clasificado como LINE de tipo L1. El alineamiento de la secuencia consenso del cluster 3 frente a la secuencia de la LINE L1 verificó su homología. El alineamiento mostró que dentro de la LINE L1 se encontraban varias repeticiones (4) en tándem de la secuencia consenso

Resultados: Caracterización de las regiones centroméricas de los cromosomas 7 y 8



Exploración de la región centromérica candidata mediante dotplots, Tandem Repeats Finder, BLASTn y RNAfold para estimar la estructura secundaria en los cromosomas 7 y 8

Resultados: Comparación con el cariotipo y la posición centromérica sobre el ensamblaje



Cromosomas modificados del cariotipo de Vega et al., 2002. Los cromosomas del cariotipo están a escala entre sí y los ideogramas de cada uno también. Las flechas azules indican las posiciones donde se ha detectado *PvuII*, la flecha naranja indica la posición de la secuencia rep87 y la flecha verde indica la posición de la secuencia rep120

Conclusión: La secuencia *PvuII*, la cual corresponde a al consenso del cluster 3, presenta homología con un elemento móvil de tipo LINE L1. Posiblemente el origen de este satélite sea la homogeneización de una parte del elemento móvil. Los centrómeros de *S. senegalensis* están ocupados principalmente por secuencias de ADN satélite: las regiones centroméricas de los cromosomas del 1 al 6, del 9 al 13, del 15 al 17, 20 y 21 estarían ocupadas por la familia de ADN satélite *PvuII*, las del cromosoma 7 por la familia rep87 y la del cromosoma 8 por la familia rep120. Además, las posiciones sobre el ensamblaje se corresponden salvo algunas excepciones con el la morfología cromosómica que se observa en el cariotipo

