



UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA  
FACULTAD DE CIENCIAS MARINAS  
INSTITUTO DE INVESTIGACIONES OCEANOLÓGICAS



**“ORGANIZACIÓN Y EVENTOS DE DUPLICACIÓN EN EL ADN  
MITOCONDRIAL EN EL GÉNERO *Calidris* (AVES, SCOLOPACIDAE)”**

T E S I S

QUE PARA CUBRIR PARCIALMENTE LOS REQUISITOS NECESARIOS PARA  
OBTENER EL GRADO DE MAESTRO EN CIENCIAS EN ECOLOGÍA  
MOLECULAR Y BIOTECNOLOGÍA

PRESENTA

**JAVIER ROBLES FLORES**

**ENSENADA, BAJA CALIFORNIA, OCTUBRE DEL 2016**

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA  
FACULTAD DE CIENCIAS MARINAS  
INSTITUTO DE INVESTIGACIONES OCEANOLÓGICAS

**“ORGANIZACIÓN Y EVENTOS DE DUPLICACIÓN EN EL ADN MITOCONDRIAL  
EN EL GÉNERO *Calidris* (AVES, SCOLOPACIDAE)”**


T E S I S

QUE PARA CUBRIR PARCIALMENTE LOS REQUISITOS NECESARIOS PARA  
OBTENER EL GRADO DE MAESTRO EN CIENCIAS EN ECOLOGÍA MOLECULAR Y  
BIOTECNOLOGÍA

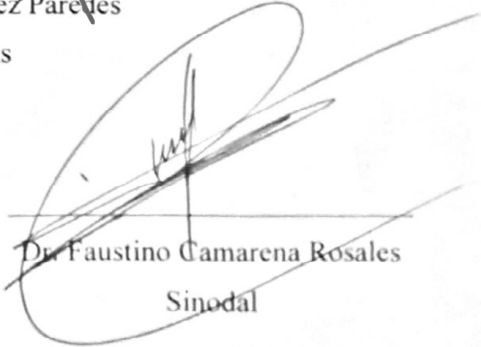
PRESENTA

**JAVIER ROBLES FLORES**

Aprobada por:

  
Dr. Luis Manuel Enriquez Paredes  
Director de tesis

  
Dr. Guillermo Juan Fernández Aceves  
Sinodal

  
Dr. Faustino Camarena Rosales  
Sinodal

## **Dedicatoria**

### ***A mi esposa***

Elizabeth Jauregui García, quien ha sido un apoyo incondicional en todo momento, quien ha estado a mi lado en momentos cruciales de mi vida, quien me alienta a enriquecer mi carrera profesional y quien siempre muestra su apoyo constante para poder lograr mis metas; pero sobre todo porque es ella quien me ha servido de inspiración.

Es una persona ejemplar que no se rinde nunca, que a pesar de que el camino sea largo no se detiene para tomar un respiro; que aunque la fatiga invada su cuerpo no busca pretextos para no completar sus metas. Es por eso, y por muchas razones más que me ha servido como un modelo a seguir, a nunca desistir, a nunca renunciar, a siempre seguir adelante con los propósitos que fije en la vida.

Le doy gracias a Dios y a la vida porque te puso en mi camino, porque gracias a ti he logrado llegar hasta aquí, este momento que considero un logro importantísimo en mi carrera.

Gracias Ely por todo lo que haces, hecho y harás por mí.

Te amo hermosa

.

## **Agradecimientos:**

Al Consejo Nacional de Ciencias y Tecnología (CONACYT), por el apoyo económico que me brindó a través de la beca # 360810, la cual fue indispensable para poder cumplir en tiempo y forma el periodo de estudios del posgrado.

Las muestras se obtuvieron mediante el apoyo de CONACYT a través de los Proyectos Investigación Científica Básica CB2007-82671 y CB2010-155353; de los Programa de Apoyo a Proyectos de Investigación e innovación Tecnológica de la Universidad Nacional Autónoma de México (Proyecto PAPIIT IN212809-3), del International Program del U.S. Forest Service a través de Copper River International Migratory Bird Initiative (CRIMBI) e Instituto de Ciencias del Mar y Limnología - UNAM (Proyecto Institucional)

Al pueblo de México porque con su valiosa contribución se apoya a estudiantes como yo, para aportar información científica, independientemente del impacto que puedan provocar los resultados generados, toda publicación es valiosa y puede contribuir al desarrollo de nuestro País.

A mis padres, quienes desde siempre se esmeraron por mi desarrollo y bienestar, que están siempre atentos a mis pasos, quienes me han aconsejado y apoyado en las decisiones que tomo, a ellos que me alientan transformando el trayecto de mis proyectos en algo más sencillo.

A mi comité de tesis, al Dr. Faustino Camarena, y al Dr. Guillermo Fernández por dedicarle una parte de su tiempo y su esfuerzo en revisar y estar pendiente de la realización de esta tesis; a la Dra. Enda Sánchez del CICESE por la ayuda en la etapa experimental y en la revisión de este trabajo. Un agradecimiento especial al Dr. Luis M. Enríquez, por ser un gran apoyo, brindarme su amistad, por dedicarme paciencia y

siempre estar atento en mi aprendizaje, gracias a él yo pude desarrollar mis habilidades teóricas y técnicas más de lo pensado, gracias “profe” por ser tan dedicado y atento.

A mis amigos, todos los que estuvieron siempre pendientes a la culminación de mi maestría, quienes en cualquier oportunidad me preguntaban con real interés cual era el estado de mi tesis, a ellos quienes con humor me metían presión para finalizarla más pronto. Gracias, Berenice Silva, Kee Whun Dan, Yesenia Balderas, Sara Isabel, Lluvia Dorantes, Omar Hernández, Joel Suárez, Obed Álvarez, Rodrigo Jauregui, Rosario Hernández, Virginia García, Astrid Luna, Paola Batta, Anaid Saavedra y a ti, que estás leyendo esto, gracias por el interés y atención.

## Resumen

La región control (RC) del ADN mitocondrial es importante en el proceso de replicación y transcripción, pero además es un marcador genético muy usado en estudios poblacionales y filogenéticos. La duplicación de genes es frecuente en el genoma mitocondrial y se han descrito varios arreglos de genes que involucran la duplicación de la RC o remanentes de ésta. Las diferencias en el orden de los genes han sido útiles para la resolución filogenética de algunos grupos de especies. Las Aves (Clase: Aves) también tienen distinto orden de genes comparado con otros vertebrados, además se han descrito varios acomodos alternativos al ancestral y se pueden presentar uno o varios acomodos distintos incluso dentro de un mismo Orden (ej. Psittaciformes, Passeriformes). El Orden *Charadriiformes* es un clado de Aves muy diverso que alberga a especies con una amplia variabilidad morfológica, conductual, crianza y otros caracteres ecológicos. Dado que existe evidencia de rearrreglos similares que han aparecido independientemente en diferentes momentos e incluso en un mismo Orden, el objetivo de este estudio fue la descripción de estas regiones en *Calidris mauri* y otras especies cercanas. (Charadriiformes: Scolopacidae). Dentro de la Familia Scolopacidae, *Arenaria interpres* presenta una RC sencilla, lo que se considera un ordenamiento de tipo ancestral, mientras que *Philomachus pugnax* posee RC duplicada y un reacomodo de los genes adyacentes involucrando la sección *Thr-RC* (*Cytb/Thr/Pro/ND6/Glu/RC*); siendo este un ordenamiento distinto a los reportados para Charadriiformes hasta el momento. *Calidris mauri* presenta duplicación de la sección *Thr-RC* separada por una copia degenerada *Cytb* que difieren en el extremo 3' de  $RC^2$  por varias repeticiones CAAA. Como se espera de la evolución concertada, las copias de  $RC$ 's son casi idénticas en cada individuo, mientras que existen variaciones entre individuos concentradas en los extremos de los dominios I y III de la RC. El tamaño de los amplicones y sus secuencias parciales sugieren el mismo arreglo en *C. minutilla*, *C. alpina*, *C. canutus*, *C. himantopus*, *Tringa semipalmata* y *Tringa flavipes*. La baja tasa de éxito en la amplificación de estos genes o regiones en *Limnodromus* spp. y *Limosa fedoa* spp., son congruentes con las relaciones filogenéticas reportadas y no permitió confirmar si este arreglo de genes es el común denominador en escolopácidos. Ya que la tasa de sustitución de la RC duplicada será más lenta que la deducida en una RC no duplicada, es importante evaluar si existe esta duplicación para estudios filogenéticos. Algunos autores han planteado que al existir regiones duplicadas las homologas están sujetas a la recombinación esto puede homogenizar las secuencias parálogas retardando su divergencia evolutiva, por lo tanto esto confunde la determinación de su antigüedad. El género *Calidris* ha sido ampliamente estudiado utilizando la RC como marcador molecular. Es imperativo utilizar con exactitud las secuencias ortólogas al realizar análisis filogenéticos; para identificar duplicaciones de genes en Aves recomendamos utilizar el amplicón  $RC^1-Pro^2$  ya que nos permite discriminar entre las  $RC$ 's seleccionando específicamente la  $RC^1$  y así tomar las medidas necesarias para análisis interindividuales..

## Abstract

Mitochondrial control region (CR) plays an important role in replication and transcription, but it is also a widely used genetic marker in population genetics and phylogeny reconstruction. Duplication events leading to gene rearrangements are common in the mitochondrial genome of birds, most of which involve complete or partial copies of CR and their adjacent regions. Differences in gene order have been useful for the phylogenetic resolution of some vertebrate taxa. Within vertebrates, birds (Class: Aves) have also a distinct gene order reinforcing its monophyly, but several rearrangements to the typical mtDNA avian pattern had been described so far, both within and among several avian Orders (i.e. Psittaciformes, Passeriformes). The Order Charadriiformes is a very diverse clade of birds including species with a wide morphological variability, behavioral, breeding and other ecological characters. Since there is strong evidence for similar mitochondrial gene arrangements arising independently and repeatedly between different and even within some avian Orders, the goal of the present study was to describe if *Calidris mauri* other related scolopacid species (Charadriiformes: Scolopacidae). Within Scolopacidae (Charadriiformes), *Arenaria interpres* appears to have a single RC, which has been considered as the ancestral gene arrangement, while *Philomachus pugnax* exhibits a rearrangement of RC and some adjacent genes has a duplication involving *Thr-CR* (*Cytb/Thr/Pro/ND6/Glu/CR*), this being a different arrangement from those reported so far within Charadriiformes. *Calidris mauri* also presented duplication *Thr-RC* section which are separated by a degenerate copy *Cytb* and differ only at the 3'-end of *CR*<sup>2</sup> by several tandem CAAA repeats. As expected in concerted evolution we found almost identical duplicated CRs within individuals, while variations among individuals were mainly concentrated at the end of CR domains I and III. Based on amplicon size and partial sequence analysis the same arrangement is present in *C. minutilla*, *C. alpina*, *C. canutus*, *C. himantopus*, *Tringa semipalmata* and *Tringa flavipes*. Poor sequence quality in more divergent relatives as *Limnodromus spp.* and *Limosa spp.*, preclude us from confirm if *Thr-CR* duplication is common among Scolopacidae. As the substitution rate of duplicated CR will be slower than the deduced from non-duplicated CR, it is important to evaluate whether duplication is present or not in further phylogenetic studies some authors have suggested that to be duplicated regions homologous are subject to recombination it can homogenize the paralogous sequences delaying their evolutionary divergence, so this confuses the determination of its age. The genus *Calidris* has been extensively studied using the RC as molecular marker; it is imperative to use exactly the orthologous to perform phylogenetic analyzes because the evolutionary rate between orthologous and paralogous sequences may differ. To identify gene duplications in Birds we recommend using the amplicon *RC*<sup>1</sup>-*Pro*<sup>2</sup> because it allows us to discriminate between the RC's specifically selecting the *RC*<sup>1</sup>, and thus take the necessary measures to interindividual analyses.

## Tabla de contenido

Dedicatoria .....	II
Agradecimientos: .....	III
Resumen.....	V
Abstract .....	VI
Lista de Tablas .....	VIII
Lista de Figuras .....	IX
Introducción .....	13
Antecedentes .....	18
Objetivos .....	31
Objetivo general.....	31
Objetivos particulares .....	31
Justificación .....	32
Resultados .....	37
Amplicón RC <sup>1</sup> -Pro <sup>2</sup> :.....	37
Cobertura de la secuenciación de la región Cytb-12S.....	39
Comparación de las copias de la región control en <i>Calidris mauri</i> .....	48
Discusión.....	53
Aspectos metodológicos .....	53
Ordenamiento de genes en el ADNmt en el género <i>Calidris</i> : Duplicación de la región control y los genes adyacentes .....	56
Duplicación de la Región Control en <i>Calidris mauri</i> y otras especies del género.....	59
Mantenimiento y evolución de los genes duplicados.....	63
Conclusiones .....	66
Referencias.....	68
Lista de Anexos.....	72

## Lista de Tablas

<b>Tabla 1.</b> Resumen del control de calidad y evaluación de la amplificación y la secuenciación .....	40
<b>Tabla 2.</b> Sitio de los genes desde el Cytb hasta el ARNr 12S. Los genes se encontraron utilizando el programa DOGMA (Dual Organellar Genome Annotator) para identificar los límites de codificación de proteínas, ARNr, ARNt y genes utilizando la configuración por defecto de BLAST (Wyman et al., 2004) basados en datos de ADNmt para vertebrados .....	45

## Lista de Figuras

- Figura 1. Organización de los genes mitocondriales desde el gen ND5 hasta el 12S ribosomal en Vertebrados y aves. A) Muestra el acomodo de los genes que comparten los vertebrados. B) Acomodo de los genes de tipo ancestral en aves, se observa una diferencia en el orden del gen que codifica para la NADH deshidrogenasa 6 (ND6), ácido glutámico (E), citocromo b (Cytb), treonina (T) y prolina (P); los cuales, en aves, están invertidos ubicados adyacentemente al gen NADH deshidrogenasa 5 (ND5). La región control (CR), el gen que codifica para Fenilalanina (F) y el ARNr 12S no presentan ningún cambio en el orden. .... 18
- Figura 2: Formación de ortólogos y parálogos. El árbol muestra seis genes homólogos de tres especies designadas A, B y C. Los genes están representados por círculos y cada color representa una especie. Los genes parálogos se rodean con una línea más gruesa (Sólo el gen del linaje A no tiene un parálogos). Las cajas en los nodos representan eventos de duplicación. Duplicación 1 produce los parálogos  $\alpha$  y  $\beta$  en el ancestro de B y C; mientras que la duplicación 2 produce los parálogos  $\beta_1$  y  $\beta_2$  en el linaje C. Todos los genes de B y C son co-ortólogos del gen A. Los genes A y B son “in-paralogs” relativos a la especiación 1, pero son “out-paralogs” relativos a la especiación 2. Genes  $\beta_1$  y  $\beta_2$  están “in-paralogs” en relación con ambas especiaciones en el árbol. Genes  $B\alpha$  y  $C\alpha$  son “one-to-one orthologs” (Studer y Robinson-Rechavi, 2009; Gibson y Baker (2012)). ..... 20
- Figura 3: Diferentes eventos de Conversión de genes homogenizan al mínimo genes que se duplican y divergen en cada especie hija (A y B), de modo que, mientras que los genes parálogos son muy similares, los ortólogos divergen en el tiempo. (tomado de (Hurles, 2004))..... 21
- Figura 4: Distintos Acomodos en Aves alrededor de la Región Control (RC). En la parte superior se describe el acomodo de los genes que se comparte entre los Vertebrados. (A) Acomodo común reportado en Aves. (B) Acomodo encontrado en algunas Garzas (Ardeidae) (Boore (2000); Zhou et al., 2014) y en Albatros (*Phoebastria albatrus*) (Lounsberry et al., 2015). (C) Acomodo descrito para *Philomachus pugnax* (Verkuil et al., 2010). (D) Acomodo encontrado en Loros del Amazonas (Eberhard et al., 2001), (E) Acomodo encontrado en el lorito de Luzón (*Prioniturus luconensis*) (Eberhard y Wright, 2016). (F) Acomodo encontrado en *Ardea cinerea* y *Ardea purpurea* (Zhou et al., 2014). (G) Acomodo encontrado en el avetorillo canelo (*Ixobrychus cinnamomeus*) (Zhou et al., 2014). (H) Primer orden alternativo encontrado en Aves (Mindell et al., 1998; Bensch y Anna, 2000). (tomado de Eberhard y Wright (2016)). ..... 23

Figura 5: Duplicación en tándem mediada por errores de replicación y pérdida al azar (modelo TDRL; Boore, 2000).....	24
Figura 6: Estructura general de la región control mitocondrial de vertebrados. Las flechas indican la ubicación del origen de replicación para la cadena pesada (H-strand) y el promotor bidireccional de transcripción para la cadena ligera y pesada (L y H-strand). TAS, secuencias de terminación asociado; F-B, cajas de secuencias conservadas en el dominio central; CSBs, bloques de secuencia conservados. (Tomado deRuokonen y Kvist (2002)) .....	25
Figura 7: Filogenia basada en secuencias de cinco genes (RAG1, CYT B, 12S, ND2 y COI) de algunas especies de la Familia Scolopacidae. Tomado deGibson y Baker (2012).....	28
Figura 8. Amplicón RC <sup>1</sup> -Pro <sup>2</sup> (L141-PropR). (A) Fotografía del gel de agarosa al 1.5%; las bandas indican la efectividad de amplificación de los cebadores, se indica en cada una de las bandas el peso aproximado en pares de bases (pb) con respecto a la escalera (500pb) el sentido de la corrida va de + a – en sentido vertical (de arriba abajo), las bandas más ligeras se sitúan en la parte superior. En la parte inferior se muestra el nombre de las aves en el carril sembrado; C. alpina (DUNL), C. minutilla (LESA), C. canutus (REKN), C. himantopus (STSA) y C. mauri (WESA). (B) Se muestra el esquema del orden de los genes de P. pugnax y el sitio de anclaje de los cebadores; directo (L141) e inverso (PropR).....	38
Figura 9. Muestra la cobertura de amplificaciones. En la parte superior se muestra el esquema del acomodo de los genes observado en Calidris mauri. Debajo del esquema se muestran los nombres de las aves playeras disponibles y debajo de estos los cuadros grises representan la cobertura de los amplicones obtenidos para cada una de ellas. ....	42
<b>Figura 10</b> Secuencia del ADN mitocondrial de Calidris mauri que comprende desde el Citocromo b hasta el ARNr 12S.....	44
<b>Figura 11.</b> Ubicación de los genes desde el Cytb hasta el ARNr 12S. La línea delgada muestra con números el sitio de las bases en esta sección del ADNmt. Las flechas de color verde indican el nombre de los genes (Cytb, ND6). Las flechas rosas indican el nombre de los ARNt (Thr, Pro, Glu y Phe). Las flechas de color negro indican las regiones no codificantes (RC <sup>1</sup> y RC <sup>2</sup> ). La flecha en rojo indica el ARNr (12S). Las anotaciones se realizaron manualmente en el programa Geneious V4.7 (Drummond et al., 2009).....	46
<b>Figura 12:</b> Orden de los genes en Calidris mauri; se muestra la región que comprende desde el citocromo b (Cytb B) hasta el ARN ribosomal 12S (12S).....	47

- Figura 13:** Alineamiento de las regiones control de *Calidris mauri*, los puntos indican nucleótidos idénticos a la secuencia de referencia, y los guiones indican huecos (gaps). Las secuencias que contienen los sitios característicos de la RC como la poli C Stretch, TAS motif, F box, E box, D box, C box D box BSB y el CSB1 están resaltados en color azul..... 49
- Figura 14:** Árbol de distancias filogenéticas entre las RC's. El árbol muestra las secuencias de la RC<sup>1</sup> y la RC<sup>2</sup> de *C. mauri* (WESA), *C. alpina* (DUNL), *C. minutilla* (LESA), *C. himantopus* (STSA), *Tringa semipalmata* (WILL), *Tringa flavipes* (LEYE), *A. interpres* (RUTU) y *P. pugnax* (RUFF). Los números en los nodos indican el porcentaje de identidad entre las secuencias agrupadas. RUTU, al poseer sólo una RC, funcionó de grupo externo. .... 50
- Figura 15** La figura muestra el orden de los genes desde el gen *Cytb* hasta el 12S para *Calidris mauri*. Los rectángulos en blanco tienen en el interior los códigos de los cebadores (JCytbF, Cytbend, b71, etcétera). Las líneas gruesas indican la longitud de los fragmentos y el par de cebadores que las generan. Las flechas indican amplificaciones utilizando la técnica de “primer walking” utilizando los amplicones largos como template. .... 51
- Figura 16:** Amplicón ND6L3-H451. En la figura superior se muestra el orden de los genes mitocondriales de *C. mauri* en la región *Cytb*-12S. Los cuadros en blanco tienen en el interior el nombre de los cebadores. Se muestran dos sitios distintos de anclajes de los cebadores originando un par de amplicones de distinta longitud, generando los productos 1 y 2. En la figura de la derecha se muestra un gel de agarosa; se corrió el producto de PCR del amplicón obtenido. En el carril de la izquierda se muestra la escalera (esc) con los pesos de las bandas marcados; en el carril de la derecha se muestra la banda del producto de PCR..... 52
- Figura 17** Filogenia basada en secuencias de cinco genes (RAG1, CYT B, 12S, ND2 y COI) estimados con el análisis bayesiano con particiones de 84 especies del suborden Scolopaci (Charadriiformes). Tomado de Gibson y Baker (2012)..... 54
- Figura 18** El árbol filogenético muestra la agrupación del género *Calidris* junto con las especies más cercanas a este género en la Familia Scolopacidae (tomado de Gibson y Baker (2012) ). Se indica la longitud aproximada del amplicón RC<sup>1</sup>-Pro<sup>2</sup> para las cinco especies de *Calidridos* de las que se obtuvo esta secuencia .. 57
- Figura 19** Acomodo de los genes en la region *Cytb*-12S. La figura **A** muestra el acomodo de los genes en *P. pugnax*. La figura **B** muestra el acomodo encontrado en *Calidridos*..... 58
- Figura 20** Duplicación en tándem mediada por errores de replicación y pérdida aleatoria (modelo TDRL; (Boore, 2000)(Tomada de Kurabayashi et al. (2008)). El

esquema superior muestra de manera simplificada el mecanismo por el cual se duplican varios genes mediante el modelo de TDRL, en este caso por el deslizamiento y mal complementación de las cadenas (slipped-strand mispairing). En el esquema inferior se muestra el acomodo final del segmento duplicado Cytb--CR en *Calidris mauri* conforme al modelo TDRL..... 60

## Introducción

El ADN mitocondrial (ADNmt) es una molécula de doble cadena circular, de 37 genes que codifican para 13 subunidades del sistema de fosforilación oxidativa, 2 ARN ribosómicos (ARNr), y 22 ARN de transferencia (ARNt). Está presente en cientos de miles de copias en cada célula, dentro de orgánulos que generan energía de la célula, las mitocondrias (Pakendorf y Stoneking, 2005).

Consiste predominantemente por ADN que codifica, a excepción de un fragmento largo de ~1100 pb de longitud que está involucrado en funciones regulatorias en la replicación del ADNmt; la Región Control (RC) (Pakendorf y Stoneking, 2005). Debido a su elevada tasa de evolución, de heredabilidad de modo estrictamente maternal, y la aparente falta de recombinación, el ADN mitocondrial (ADNmt) es la molécula de elección tanto para los estudios de la secuencia de divergencia intraespecífica (Gyllensten et al., 1991) y para la construcción de árboles de genes dentro de especies y entre especies relacionadas (Avise et al., 1987; Avise, 1989).

La elección de los marcadores moleculares se vuelve crítica cuando se estudian divergencias muy antiguas (Ericson et al., 2003). Las diferentes regiones de la molécula evolucionan a tasas diferentes, y por lo tanto una región apropiada se puede seleccionar para resolver la relación filogenética de los linajes maternos particulares bajo el estudio. Se pensaba que el orden de los genes mitocondriales eran altamente conservada dentro de los vertebrados en base al orden de los genes de los primeros genomas secuenciados (Anderson et al., 1981), pero el acomodo de los genes es variable entre, y en menor medida, dentro de los principales grupos de vertebrados (San Mauro et al., 2006).

Con el creciente número de genomas mitocondriales completos publicados hasta ahora, está claro que la estabilidad del acomodo de los genes es variable y específica de cada

grupo dentro de los vertebrados. Las nuevas herramientas y técnicas de secuenciación de ADN, revelan variaciones en la organización de los genomas mitocondriales entre los distintos niveles taxonómicos de vertebrados.

El acomodo del orden de los genes más conservado se encuentra en mamíferos placentarios, tortugas y peces. Por otro lado, algunas especies de reptiles, peces, anfibios y particularmente aves, tienen acomodos de genes que son menos conservados (Pereira, 2000).

De los distintos mecanismos que se han propuesto para poder explicar los reajustes de los genes, uno de ellos es la duplicación en tándem, seguida de pérdidas o deleciones aleatorias de los genes duplicados. Este modelo es generalmente considerado como el más importante en vertebrados (Brown et al., 1986; Arndt y Smith, 1998; Boore, 1999; San Mauro et al., 2006). Se considera que la duplicación de los genes mitocondriales puede tener cuatro destinos potenciales, tales como la no-funcionalización, sub-funcionalización, neo-funcionalización y el mantenimiento redundante (Force et al., 1999; Lynch et al., 2001; Zhang, 2003; Rastogi y Liberles, 2005; Schirtzinger et al., 2012).

La duplicación de genes es frecuente en el genoma mitocondrial y se han descrito varios arreglos de genes que involucran la duplicación de la RC mitocondrial o remanentes de ésta. La duplicación genómica ocurre concurrentemente alrededor de la RC, más que en cualquier otra parte del genoma mitocondrial (Fujita et al., 2007) e implica cambios en el orden de los genes adyacentes (Pereira, 2000; Verkuil et al., 2010).

De manera general, una molécula de ADNmt posee normalmente sólo una RC, pero en varios grupos de vertebrados se han observado genomas mitocondriales con reajustes que poseen dos RC's, e inclusive se han reportado genomas con tres RC's (Braband et al., 2010; Zheng et al., 2013). En la mayoría de los casos, las RC's duplicadas muestran alta similitud en las secuencias entre las copias y además las características típicas de una

RC funcional, lo cual puede ser un indicador de que estas evolucionan de manera concertada (Tatarenkov y Avise, 2007).

Los reacomodos en los genes mitocondriales son considerados como eventos evolutivos raros, y como tales, la existencia de un orden de genes derivados compartido entre taxones suele ser indicativo de un ancestro común (Boore, 1999). Las variaciones en el orden de genes han sido aplicados como marcadores tanto en los estudios genéticos y filogeográficos de poblaciones (Verkuil et al., 2010). Sin embargo, estas duplicaciones genéticas y los reacomodos complican el uso de la RC como un marcador porque esta variación puede generar resultados poco confiables si no se comparan las secuencias homólogas en estudios filogenéticos (Verkuil et al., 2010).

En los últimos 15 años, los análisis de secuenciaciones de genomas completos han revelado que los reordenamientos son comunes, particularmente en Aves, donde varios linajes diferentes han estado bajo reordenamientos repetidos (Eberhard y Wright, 2016). Se ha encontrado que existen al menos ocho ordenamientos en el genoma y muchos genomas mitocondriales poseen dos copias de algunos genes provocando duplicación de genes (Desjardins y Morais, 1990; Mindell et al., 1998; Eberhard et al., 2001; Abbott et al., 2005; Zhou et al., 2014).

Las secuencias obtenidas de la RC entre individuos, poblaciones o especies pueden no ser las secuencias homólogas, es decir, una RC puede ser amplificada y secuenciada en algunos individuos, y la segunda copia de la RC puede ser obtenida para otros individuos, lo que resulta en comparaciones entre secuencias no homólogas. En este caso, la variación observada en las secuencias de la RC puede ser tanto paróloga (comparaciones entre genes duplicados) como ortóloga (comparación entre individuos). Sin embargo, las RC duplicadas tienen potencial como marcadores filogenéticos a niveles taxonómicos, dentro de géneros y familias, o entre las familias. Por lo tanto, la presencia de RC

duplicadas en los genomas de ADN mitocondrial de Metazoarios es un fenómeno genético interesante y potencialmente útil (Ogoh y Ohmiya, 2007).

El primer genoma mitocondrial aviar completamente secuenciado fue el de *Gallus gallus* (Desjardins y Morais, 1990). El acomodo de los genes de esta especie es el observado con mayor frecuencia en las aves; este acomodo de genes es el denominado como el orden de genes "típico" aviar (Eberhard y Wright, 2016). El genoma mitocondrial típico aviar posee el mismo juego de 37 genes y la RC no codificante, pero difiere de la mayoría de otros vertebrados debido a un reordenamiento cerca de la RC.

A pesar de la creciente lista de taxones de Aves en los cuales se han encontrado reordenamientos en el orden de los genes mitocondriales, las preguntas persisten sobre los mecanismos por los cuales los reordenamientos se producen, el grado en que las duplicaciones se conservan en el tiempo evolutivo y el efecto que estos reordenamientos tienen sobre la función, replicación y evolución del genoma mitocondrial (Eberhard y Wright, 2016).

Un estudio realizado por Schirtzinger y colaboradores demostró que la duplicación de la región control ha ocurrido varias veces en el Orden Psittaciformes; con la reconstrucción ancestral sugieren que hubo seis orígenes independientes en la duplicación de la región control dentro de este Orden (Schirtzinger et al., 2012), este tipo de análisis permitió elucidar el posible surgimiento, antigüedad de las duplicaciones y ubicar este evento en un árbol filogenético de una Familia o un Orden de aves.

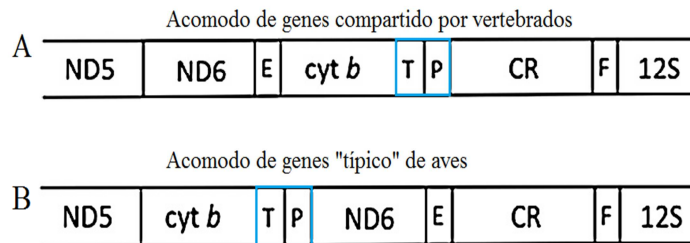
El Orden Charadriiformes (clase: Aves) es un clado de aves diverso que alberga a especies con una amplia variabilidad morfológica, conductual, crianza y otros caracteres ecológicos. Actualmente, en la Familia Scolopacidae (orden: Charadriiformes) existen dos acomodos de genes distintos en donde se describe un acomodo típico y además otro

con RC duplicada para *Arenaria interpres* y *Philomachus pugnax*, respectivamente (Paton et al., 2002; Verkuil et al., 2010). Ambas especies son basales en el árbol filogenético de la Familia. Este factor motiva el estudio debido a que sería interesante conocer si existen más acomodados dentro de Scolopacidae por ser un grupo tan diverso.

En este trabajo, se analizan algunas especies de escolopácidos, particularmente del género *Calidris*; especies migratorias en las que se ha estudiado ampliamente la conectividad de sus poblaciones mediante los haplotipos mitocondriales de la región control. Por lo mismo, es relevante evaluar la presencia de duplicaciones en esta región del genoma y las implicaciones que esto podría tener en los estudios poblacionales y evolutivos. Esta información podría contribuir para ayudarnos a comprender bajo qué mecanismo se generaron los reacomodos de los genes mitocondriales, si la duplicación de la RC en Charadriiformes tiene un origen singular o múltiple, y determinar si el acomodo de los genes podría servir como un carácter filogenético útil.

## Antecedentes

Lo que se llegó a considerar como el ordenamiento de genes "universal" para los vertebrados se descartó cuando el primer genoma aviar fue secuenciado por completo (Desjardins y Morais, 1990). En el genoma mitocondrial aviar típico existe un reacomodo que involucra el desplazamiento de una porción que precede a la región control (RC) y que involucra tanto al gen del citocromo b (*cyt b*), como a los genes correspondientes a los ARN de transferencia de la Treonina (T) y la Prolina (P), (Figura 1).



*Figura 1. Organización de los genes mitocondriales desde el gen ND5 hasta el 12S ribosomal en Vertebrados y aves. A) Muestra el acomodo de los genes que comparten los vertebrados. B) Acomodo de los genes de tipo ancestral en aves, se observa una diferencia en el orden del gen que codifica para la NADH deshidrogenasa 6 (ND6), ácido glutámico (E), citocromo b (Cytb), treonina (T) y prolina (P); los cuales, en aves, están invertidos ubicados adyacentemente al gen NADH deshidrogenasa 5 (ND5). La región control (CR), el gen que codifica para Fenilalanina (F) y el ARNr 12S no presentan ningún cambio en el orden.*

Aunque Desjardins y Morais (1990) sugirieron que esta organización confiere al grupo una ventaja selectiva asociada con la elevada demanda energética requerida para el vuelo, esta correlación aún no ha sido demostrada (Pereira, 2000). No obstante, se han propuesto distintos mecanismos que tratan de explicar cómo es que los genes mitocondriales se reorganizan para poder entender así la posible función y el contexto evolutivo de estas modificaciones (Eberhard y Wright, 2016). La mayoría de estas modificaciones en la organización, involucran desplazamientos o duplicaciones de zonas

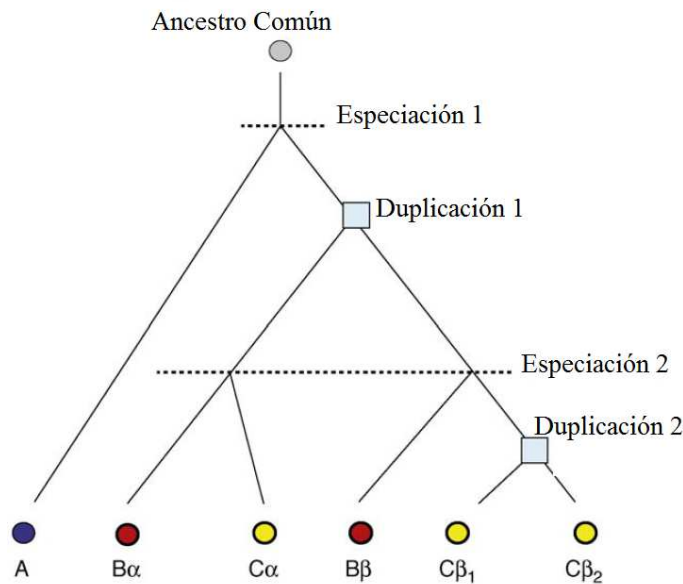
adyacentes a la región control e incluso a ella misma. Por lo anterior, los mecanismos involucrados se han clasificado en torno al tamaño del fragmento duplicado y si implica o no un intermediario de ARN.

Uno de estos mecanismos se denomina “retrotransposición” y describe la integración en el genoma de un ARN mensajero maduro que sufrió transcripción inversa espontánea (pseudogén procesado). El gen duplicado resultante (retrogén) carece de intrones y presenta una cola de poli-Adeninas; separado de sus elementos regulatorios, estas secuencias recién integradas rara vez generan codones que expresen en la completa longitud de esta secuencia. Sin embargo, algunos retrogenes funcionales ya se han identificado en la mayoría de los genomas (Hurles, 2004).

Otro mecanismo conocido es la “duplicación en tándem” de un segmento del genoma y ésta es debido a un posible resultado de una recombinación desigual (unequal crossing over), la cual resulta a partir de la recombinación homóloga entre secuencias parálogas (Hurles, 2004) . Este evento de recombinación puede también generar una delección o una inversión entre las secuencias involucradas (Bailey et al., 2003; Hurles, 2004).

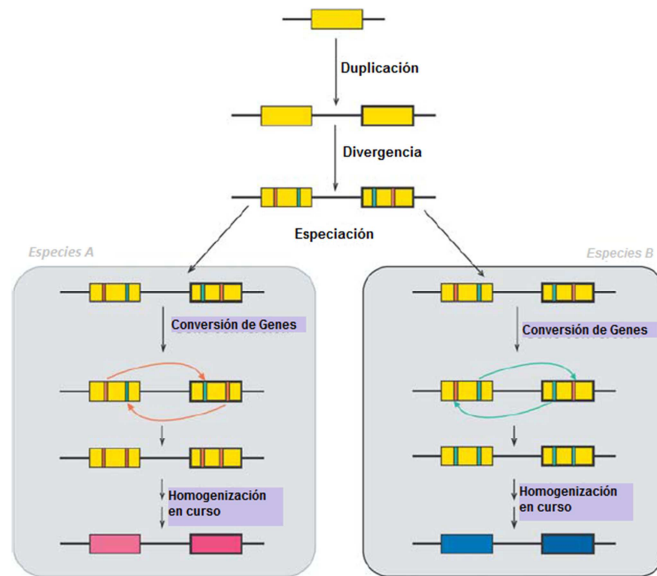
En función de su extensión, la duplicación puede generar un par de genes homólogos funcionales, o bien copias incompletas no funcionales, que sean ortólogos o parálogos (Koonin, 2005; Studer y Robinson-Rechavi, 2009). Los genes homólogos ortólogos han divergido desde un evento de especiación, mientras que los parálogos son genes homólogos que se han diferenciado desde que ocurrió su duplicación (Figura 2)

Una recombinación homóloga entre secuencias parálogas puede resultar en reordenamientos de la secuencia, incluyendo duplicaciones en tándem, así como en transferencia no recíproca de una secuencia paróloga hacia la otra, proceso conocido como “conversión de genes” (Tayebi et al., 2003; Hurles, 2004). Dicha conversión homogeniza las secuencias parálogas, retardando su divergencia evolutiva y complicando la determinación de la antigüedad de las copias. Esto conduce a la observación de un proceso denominado como “evolución concertada” en donde las copias pueden ser altamente similares pero divergir evolutivamente (Figura 3). Finalmente, una vez que los genes han divergido lo suficiente como para alterar su funcionalidad (o disfuncionalidad), los eventos de conversión de genes pueden resultar deletéreos.



*Figura 2: Formación de ortólogos y parálogos. El árbol muestra seis genes homólogos de tres especies designadas A, B y C. Los genes están representados por círculos y cada color representa una especie. Los genes parálogos se rodean con una línea más gruesa (Sólo el gen del linaje A no tiene un parálogos). Las cajas en los nodos representan eventos de duplicación. Duplicación 1 produce los parálogos  $\alpha$  y  $\beta$  en el ancestro de B y C; mientras que la duplicación 2 produce los parálogos  $\beta_1$  y  $\beta_2$  en el linaje C. Todos los genes de B y C son co-ortólogos del gen A. Los genes A y B son “in-paralogs” relativos a la especiación 1, pero son “out-paralogs” relativos a la especiación 2. Genes  $\beta_1$  y  $\beta_2$  están “in-paralogs” en relación con ambas especiaciones en el árbol. Genes  $B_\alpha$  y  $C_\alpha$  son “one-to-one orthologs” (Studer y Robinson-Rechavi, 2009; Gibson y Baker (2012)).*

En Aves se ha encontrado que existen al menos ocho ordenamientos de genes mitocondriales y varios de estos genomas mitocondriales poseen genes duplicados (Desjardins y Morais, 1990; Mindell et al., 1998; Eberhard et al., 2001; Abbott et al., 2005; Zhou et al., 2014).



*Figura 3: Diferentes eventos de Conversión de genes homogenizan al mínimo genes que se duplican y divergen en cada especie hija (A y B), de modo que, mientras que los genes parálogos son muy similares, los ortólogos divergen en el tiempo. (tomado de (Hurles, 2004))*

Mindell y colaboradores (1998) demostraron que en Aves hay dos ordenamientos de genes. Mediante la asignación de estos órdenes de dos genes en el árbol filogenético aviar, han demostrado que uno de estos ordenamientos de genes es ancestral (Figura 4A), el cual fue encontrado en gallinas (*Gallus gallus*) quienes son Aves que se sitúan en las ramas más basales del árbol filogenético. Este acomodo es el observado con mayor frecuencia en las aves; y es denominado como el orden de genes "típico" aviar.

El acomodo de genes (Figura 4A) que poseen las gallinas es compartido con reptiles, y se asume que el acomodo genético derivado de éste (Figura 4H), evolucionó por una duplicación en tándem de la región *Thr/Phe/ND6/Glu/CR*, seguida de una delección de las

regiones *Pro*<sup>1</sup>/*ND6*<sup>1</sup>/*Glu*<sup>2</sup> y *Thr*<sup>2</sup> (Gibb et al., 2007) esto generó además una duplicación de la región de control.

Si bien, el orden de los genes de las gallinas (*Gallus gallus*) se encuentra en muchos otros taxones de aves, tales como *Struthioniformes*, *Ciconiiformes*, *Piciformes*, *Passeriformes* y *Charadriiformes*; posteriormente varios investigadores (Mindell et al. (1998); Eberhard et al. (2001); Abbott et al. (2005); Verkuil et al. (2010); Zhou et al. (2014)) presentaron órdenes alternativos de genes mitocondriales de aves y discutieron su origen potencial (Figura 4).

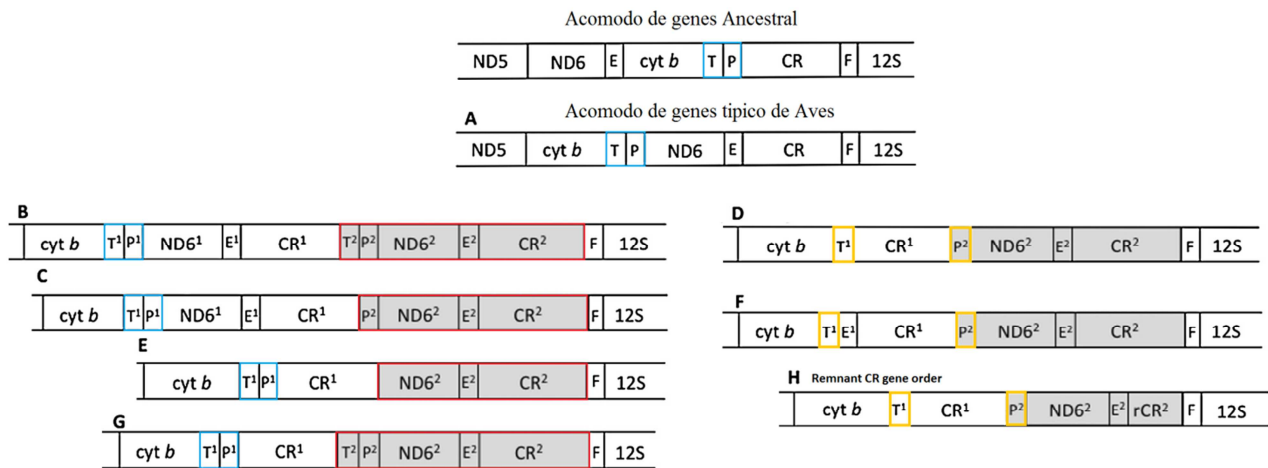


Figura 4: Distintos Acomodos en Aves alrededor de la Región Control (RC). En la parte superior se describe el acomodo de los genes que se comparte entre los Vertebrados. (A) Acomodo común reportado en Aves. (B) Acomodo encontrado en algunas Garzas (Ardeidae) (Boore (2000); Zhou et al., 2014) y en Albatros (*Phoebastria albatrus*) (Lounsbury et al., 2015). (C) Acomodo descrito para *Philomachus pugnax* (Verkuil et al., 2010). (D) Acomodo encontrado en Loros del Amazonas (Eberhard et al., 2001), (E) Acomodo encontrado en el lorito de Luzón (*Prioniturus luconensis*) (Eberhard y Wright, 2016). (F) Acomodo encontrado en *Ardea cinerea* y *Ardea purpurea* (Zhou et al., 2014). (G) Acomodo encontrado en el avetorillo canelo (*Ixobrychus cinnamomeus*) (Zhou et al., 2014). (H) Primer orden alternativo encontrado en Aves (Mindell et al., 1998; Bensch y Anna, 2000). (tomado de Eberhard y Wright (2016)).

Uno de los modelos que puede explicar de una manera más adecuada como es que las duplicaciones se originaron en Aves es el modelo de la “duplicación y pérdida al azar en tándem (TDRL)” (Boore, 2000; Kurabayashi et al., 2008). En este modelo, inicialmente una porción del genoma se duplica. Luego, uno de los genes copiados se puede convertir en un pseudogén o se escinde del genoma por procesos mutacionales subsecuentes. La copia que se pierde se determina al azar (se puede generar una mutación modificando la función del gen copiado o se puede perder). Por lo tanto, un cierto patrón de delección puede recuperar el acomodo de genes original, pero también puede conducir a la reorganización de genes (Kurabayashi et al., 2008) (Figura 5).



Figura 5: Duplicación en tándem mediada por errores de replicación y pérdida al azar (modelo TDRL; Boore, 2000).

Kurabayashi y colaboradores (2008) sugieren que la duplicación de los genomas mitocondriales es principalmente causado por errores durante la replicación del ADNmt, como el deslizamiento y mal complementación de las cadenas (slipped-strand mispairing), o por una sincronía en los puntos de iniciación y terminación de la replicación.

La región de control mitocondrial es una secuencia codificante de longitud variable que contiene los promotores para la transcripción (LSP, HSP), origen de replicación de la cadena pesada  $O_H$  y el bucle de desplazamiento (*D-loop*) en los vertebrados (Clayton, 1984; Chang y Clayton, 1986). En Aves, la RC se extiende por la región entre los genes

para  $tRNA^{Glu}$  y  $tRNA^{Phe}$  en la mayoría de las especies de aves estudiadas hasta ahora (Desjardins y Morais, 1990; Mindell et al., 1998). La RC también se ha encontrado flanqueada entre  $tRNA^{Thr}$  y  $tRNA^{Pro}$  en algunas especies de *Picidae*, *suboscines*, *Passeriformes*, *Falconiformes* y *Cuculidae* (Mindell et al., 1998; Bensch y Anna, 2000).

Basado en la distribución de las posiciones de nucleótidos variables y frecuencias diferenciales de los nucleótidos, la región de control mitocondrial se divide en tres dominios (Brown et al., 1986; Saccone et al., 1991), (Figura 6).

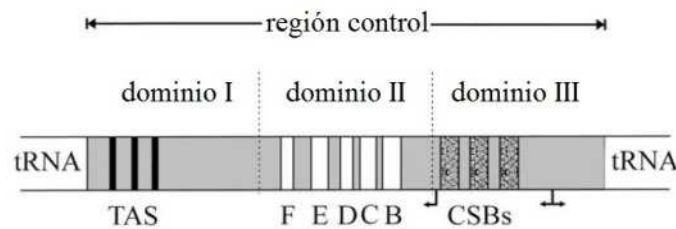


Figura 6: Estructura general de la región control mitocondrial de vertebrados. Las flechas indican la ubicación del origen de replicación para la cadena pesada (H-strand) y el promotor bidireccional de transcripción para la cadena ligera y pesada (L y H-strand). TAS, secuencias de terminación asociado; F-B, cajas de secuencias conservadas en el dominio central; CSBs, bloques de secuencia conservados. (Tomado de Ruokonen y Kvist (2002))

El dominio conservado central de la cadena ligera II es rica en guanina y pobre en adenina en comparación con los dominios flanqueantes I y III. La mayor parte de la variabilidad dentro de la región de control se concentra en los dominios I y III, que muestran tanto un elevado número de sustituciones de nucleótidos y una variación en la longitud en comparación con el dominio II (Saccone et al., 1991; Baker y Marshall, 1997; Ruokonen y Kvist, 2002).

Las aves playeras (Aves: Charadriiformes) son un grupo diverso y se considera que existen más de 360 especies en tres subórdenes, 19 familias y 96 géneros (Christian et al., 1992; Baker et al., 2007). Se asume que las especies han divergido por radiación a partir

de diferentes clados que son morfológicamente y conductualmente distintos (Van Tuinen et al., 2004). Se estima que se originaron hace más de 90 millones de años (Baker et al., 2007). Este grupo exhibe una notable variedad de rasgos de comportamiento, morfología y ciclo biológico. El soporte morfológico para monofilia de *Charadriiformes* es débil, se han propuesto hipótesis contradictorias de las relaciones taxonómicas basadas en análisis de características morfológicas, osteológicas y genéticas. Muchos estudios han investigado las relaciones entre estas especies, incluyendo a varios taxones excéntricos que son difíciles de colocar filogenéticamente, con el objetivo de comprender los procesos evolutivos que dieron lugar a la diversidad que se observa en la actualidad. Históricamente, las relaciones filogenéticas se infieren utilizando características morfológicas, tales como la forma del pico, tipo de pie, características del esqueleto, y los patrones de color (Strauch, 1978; Sibley et al., 1988; Björklund, 1994).

Por otro lado, los marcadores moleculares son una herramienta utilizada para obtener información sobre migración, comportamiento reproductivo, demografía, estructura poblacional y niveles de diversidad genética (Hoelzel, 1992). Uno de los marcadores moleculares por excelencia es el ADN mitocondrial (Avise et al., 1987; Moritz et al., 1987). Para mejorar los estudios de divergencias de antigüedad media en Aves (de 20-60 mda) es deseable encontrar marcadores moleculares con tasas evolutivas ligeramente más rápidas que los genes nucleares que codifican proteínas utilizadas hasta ahora, e intrones en los genes nucleares puede proporcionar tales marcadores.

Hasta el momento en el Orden *Charadriiformes* se han secuenciado los genomas mitocondriales completos de; *Jacana jacana* (KJ631049.1), *Jacana spinosa* (KJ631048.1), *Vanellus vanellus* (NC\_025637.1), *Vanellus cinereus* (NC\_025514.1), *Scolopax rusticola* (NC\_025521.1), *Recurvirostra avosetta* (KP757766.1), *Haematopus ater* (NC\_003713.2), *Larus brunnicephalus* (JX155863.1), *Larus crassirostris* (NC\_025556.1), *Larus dominicanus* (NC\_007006.1), *Synthliboramphus antiquus* (NC\_007978.1), *Saundersilarus saundersi* (JQ071443.1), *Stercorarius maccormicki* (NC\_026125.1), *Uria aalge* (AJ242686), *Pinguinus impennis* (AJ242685), *Alle alle* (AJ242684), *Uria lomvia* (AJ242687), *Alca torda* (AJ242683), *Arenaria interpres*

(NC\_003712.2) y para *Philomachus pugnax* (GQ255993) se describió secuencia parcial desde la región control hasta el *12S* ribosomal.

La mayoría de la especies que se han secuenciado hasta el momento, presentan un ordenamiento típico de los genes mitocondriales (Figura 4A). Sin embargo, en la Familia Scolopacidae se encontraron dos acomodos de genes distintos en donde se describe un acomodo típico y además otro con RC duplicada para *Arenaria interpres* y *Philomachus pugnax*, respectivamente (Paton et al., 2002; Verkuil et al., 2010). Las dos especies son basales en el árbol filogenético de la Familia.

Actualmente no existe evidencia sobre cuál es el estado del acomodo de los genes que flanquean la región control en las especies del género *Calidris*. Verkuil y colaboradores (2010) encontraron un quinto ordenamiento alternativo para *Philomachus pugnax* (Figura 4C); éste ordenamiento indica duplicación del segmento *tRNA<sup>Thr</sup>-RC* acompañado de una delección del gen duplicado de *tRNA<sup>Thr</sup>* ( $\psi$ -*Thr*). Estimamos que las especies del género *Calidris* podrían poseer un ordenamiento similar a *Philomachus pugnax* del cual ya se tiene estructurado el orden de los genes que comprenden desde el gen del *Citocromo b* hasta el *12s* ribosomal (Verkuil et al., 2010). Sin embargo, *P. pugnax* es la única especie del género *Philomachus* y no sería prudente generalizar que todas las especies de la familia *Scolopacidae* poseen éste ordenamiento. Además que, a pesar de la distancia relativa en las ramas del árbol filogenético, parece indicar que *P. pugnax* es una especie ancestral del género *Calidris* (Figura 7).

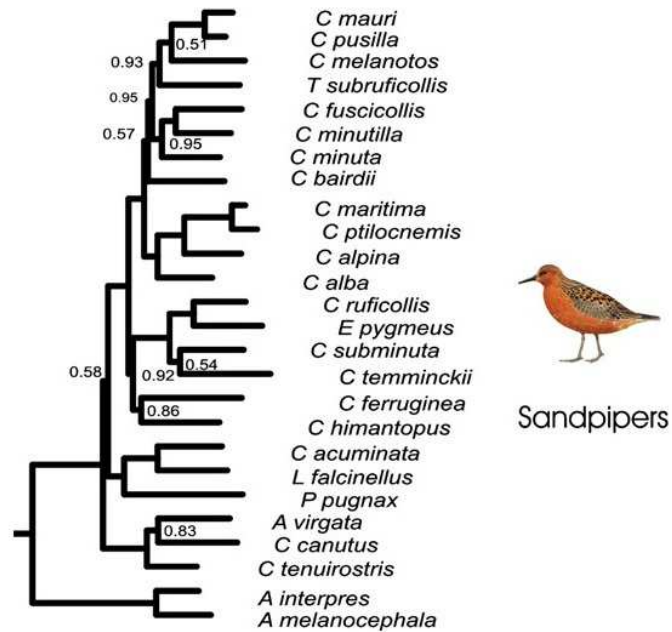


Figura 7: Filogenia basada en secuencias de cinco genes (*RAG1*, *CYT B*, *12S*, *ND2* y *COI*) de algunas especies de la Familia *Scolopacidae*. Tomado de Gibson y Baker (2012).

Es importante determinar cuál es el orden específico de los genes en esa zona del ADNmt en *Calidris*, ya que se ha utilizado a la RC como un marcador mitocondrial en estudios de filogenia, conectividad migratoria y tiempos de divergencia. Elucidar el contenido genómico nos brindará información valiosa para saber si previo a esta investigación se utilizó de manera objetiva la misma RC interindividualmente. Además nos permitirá diseñar marcadores específicos de las RC's y poder discriminar entre una u otra si hubiese duplicación.

Algunos estudios han revelado la funcionalidad de las duplicaciones de la RC que han persistido a través de múltiples eventos de especiación (Arndt y Smith, 1998; Kumazawa et al., 1998; Campbell y Barker, 1999; Eberhard et al., 2001; Abbott et al., 2005; Shao et al., 2005). Estos estudios demostraron que las RC duplicadas no mostraban signos de

degeneración, sino más bien parecían evolucionar de manera concertada. En el modelo TDRL, por lo general una copia de los genes o RC duplicadas pierde la función, se degrada o se elimina con el tiempo (Boore, 2000; Zhou et al., 2014). Sin embargo, en las aves es común que las duplicaciones persistan durante un largo tiempo evolutivo, en donde ambas copias son funcionales y tienen una alta similitud de secuencia. Tal situación en general se ha pensado que se debe a la evolución concertada y se han postulado distintos mecanismos que puedan explicar la evolución concertada (Zhou et al., 2014).

Para comprobar que las copias se encuentran bajo el mecanismo de la evolución concertada a nivel microevolutivo, es necesario en primer lugar, identificar las secuencias polimórficas duplicadas en una muestra de individuos conespecíficos, preferiblemente dentro de una población local. En segundo lugar, identificar que las diferencias genéticas en un locus parálogo dentro de un mismo individuo son mucho menores que las diferencias entre locus ortólogos entre los diferentes individuos (Tatarenkov y Avise, 2007). Si al comparar las RC's en un individuo estas son más similares que entre individuos es indicativo de que las RC's están bajo el proceso de evolución concertada. Si el resultado es contrario, significa que las RC's están evolucionando de manera independiente.

La mayoría de las aves presentan un acomodo de genes típico, mientras que otras presentan reacomodos. Es decir, una duplicación en tándem de una sección que incluye parte del *Cytb*, *Thr*, *Pro*, *ND6*, *Glu* y gran parte de la RC (*tRNA<sup>Thr</sup>-RC*). Algunas especies tienen sólo la duplicación de la RC y otras con duplicaciones y degeneración de la RC original (Wang et al., 2015).

Las variaciones de secuencias en la RC hipervariable ha demostrado ser muy informativa en la determinación de estructura genética de poblaciones en aves playeras (Wenink et al., 1994; Buehler y Baker, 2005; Buehler et al., 2006). Sin embargo, el poseer copias de

genes y RC duplicada complican el uso de la RC como un marcador fiable debido a que la co-amplificación de los genes duplicados puede resultar en signos de “heteroplasmia aparente” (más de un tipo de ADNmt dentro de un individuo) (Verkuil et al., 2010).

Si se realizan análisis utilizando una RC obtenida de un individuo que posea duplicación contra otro que no presente copia, las comparaciones podrían realizarse entre secuencias no homólogas (ortóloga vs paróloga). Por consecuencia, se generaría una confusión debido a que no existe heteroplasmia, sino más bien que, realmente se observan sustituciones de nucleótidos entre secuencias parálogas que están en un proceso de divergencia. Determinar la presencia de duplicaciones debería ser primordial antes de implementar análisis de filogenia en aves.

## Objetivos

### Objetivo general

Caracterizar la organización y la presencia de duplicaciones de los genes mitocondriales en especies representantes de la Familia Scolopacidae (Orden Charadriiformes).

### Objetivos particulares

- Caracterizar la región comprendida entre el *citocromo b* y el gen del ARN ribosomal *12S* en especies del genero *Calidris* y otras especies representantes de la Familia Scolopacidae y el Orden Charadriiformes.
- Determinar si existen diferentes arreglos de genes y duplicaciones asociadas a la región control mitocondrial en estas especies
- Evaluar el orden de los genes del ADN mitocondrial que comprenden la región desde el Citocromo b hasta el ARNr *12S* en *Charadrius nivosus*, *Charadrius semipalmatus*, *Charadrius wilsonia*, *Charadrius vociferus*, *haematopus palliatus*, *Himantopus mexicanus*, *Limosa fedoa*, *Limnodromus griseus*, *Limnodromus scolopaceus*, *Tringa flavipes*, *Tringa semipalmata*, *Calidris alpina*, *Calidris himantopus*, *Calidris minutilla*, *Calidris canutus*, *Thalasseus maximus* y *Leucophaeus atricilla*. \*.
- Diseñar cebadores específicos para discriminar entre las copias de la Región Control para análisis posteriores.

## Justificación

En primera instancia el objetivo de esta investigación es determinar el orden específico de los genes mitocondriales que comprenden desde el *citocromo b* hasta la región *12S* ribosomal, esto nos permitirá discernir si existe en todas las especies del estudio duplicaciones de alguno de sus genes. El contar con este tipo de información es fundamental debido a que, si la presentaran, se podría utilizar de manera objetiva la misma RC interindividualmente para estudios posteriores de filogenia. Además que nos permitirá elegir o diseñar marcadores específicos de las RC's y poder discriminar entre una u otra si hubiese duplicación para realizar análisis posteriores.

Como se mencionó anteriormente, al ocurrir eventos de duplicación de genes una de las copias generadas comúnmente van acompañada de deleciones o degeneraciones, por lo tanto, una vez generada una duplicación, la tasa mutacional de las copias es completamente distinta. Lo que afecta de manera desmesurada los resultados que se obtengan de los análisis filogenéticos si no se elige de manera objetiva la misma secuencia para todas las especies que se sometan en dicho estudio.

Al obtener cual es el orden específico de los genes mitocondriales la información generada podría servir para hacer inferencias sobre las posibles causas de las divergencias evolutivas, y al obtener las secuencias ortólogas, se podría generar un árbol filogenético más preciso de la Familia *Scolopacidae*.

## **Materiales y Métodos**

### **Muestras y extracción del ADN**

Las muestras empleadas en el presente estudio forman parte de la colección de del Laboratorio de Ecología de Aves Playeras del Instituto de Ciencias del Mar y Limnología, Universidad Autónoma de México – Unidad Mazatlán (ICMyL-Mazatlán), recolectadas como parte del proyecto “Monitoreo de Aves Playeras y Acuáticas del Noroeste de México” que coordina el Dr. Guillermo Fernández Aceves.

Las especies que se incluyeron para el análisis fueron las siguientes: *Calidris mauri*, *Calidris himantopus*, *Calidris minutilla*, *Calidris canutus*, *Calidris alpina*, *Charadrius nivosus*, *Charadrius semipalmatus*, *Charadrius wilsonia*, *Charadrius vociferus*, *Himantopus mexicanus*, *Tringa flavipes*, *Tringa semipalmata*, *Limnodromus griseus*, *Limnodromus scolopaceus*, *Limosa fedoa*, *haematopus palliatus*, *Thalasseus maximus* y *Leucophaeus atricilla*.

El ADN genómico se obtuvo de muestras de sangre preservadas en anticoagulante (EDTA tetrasódica 100mM, Tris 100 mM, NaCl 10 mM y SDS 1%), tomando 150  $\mu$ L de una mezcla de sangre-anticoagulante y siguiendo el protocolo universal de extracción rápida descrito por Aljanabi y Martinez (1997). La calidad y cantidad de ADN fueron evaluadas por electroforesis en geles de agarosa al 1.7%, teñidos con bromuro de etidio (0.5  $\mu$ g/mL). Para determinar la concentración relativa del material genético extraído se utilizó un marcador de peso molecular conocido (100 pb DNA Size Standard Invitrogen™) y posteriormente se llevaron a cabo diluciones de cada muestra para obtener una concentración final de ADN de entre 15 y 25 ng/uL.

## ***Amplificación por PCR y secuenciación***

Para explorar el orden de los genes alrededor de la región control y excluir la posibilidad de amplificar copias nucleares, se amplificaron por PCR (reacción en cadena de la polimerasa) fragmentos que se extendieran desde el gen del citocromo b (*Cytb*) y el gen del ARN ribosomal (*12S*). Para ello se utilizaron cebadores universales diseñados para aves, particularmente aquellos empleados por Verkuil et al. (2010) para el Ruff, *P. pugnax*, para el que se ha reportado duplicación de genes mitocondriales (Anexos 1 y 2).

Para minimizar la potencial amplificación de copias nucleares y aumentar la efectividad de los cebadores se amplificaron templates largos, utilizando pares de cebadores que tuvieran una distancia de entre 5000 y 1500 pb, polimerasa con mayor procesividad e incrementos paulatinos en el tiempo de reincorporación (annealing) en el perfil de PCR (Anexos 3 y 4).

Los amplicones así obtenidos, se emplearon como templates en una segunda ronda de amplificación utilizando cebadores internos.

Como primera aproximación, en apego a los resultados de Verkuil y colaboradores (2010), se empleó el par de cebadores que amplifican la región que comprende desde la región control 1 (*RC1*) hasta el gen duplicado de la prolina (*Pro*<sup>2</sup>) en la secuencia del Ruff. Este amplicón *RC*<sup>1</sup>-*Pro*<sup>2</sup> se produce cuando ocurre el anclaje del par de cebadores *L141* y *PropR* (véase Anexos 1 y 2). La ausencia del segundo gen de prolina evitaría la amplificación de este fragmento, lo que supondría que solo en aquellas especies en las que se produce el amplicón *RC*<sup>1</sup>-*Pro*<sup>2</sup>, existe duplicación de la *RC*.

Para aquellas especies en las que se obtuvo un producto de amplificación de la región *RC*<sup>1</sup>-*Pro*<sup>2</sup>, se prosiguió amplificando los productos de PCR de mayor longitud, empleando para ello los cebadores más externos al fragmento de interés (*Cytb* y *12S*).

Aquellos productos de PCR o amplicones de más de 3000 pares de bases, así como los productos obtenidos con los cebadores *CmaRCF* y *CmRCR* (véase Anexos 1 y 2) fueron

clonados para corroborar la presencia de duplicación de la región control mitocondrial. Para llevar a cabo las reacciones de clonación se empleó el kit TOPO TA Cloning Kit (Invitrogen) y el vector PCR TOPO 2.1, así como células electrocompetentes de *Escherichia coli*, cepa Top 10. La transformación de células electrocompetentes, se llevó a cabo en un electroporador para células procariontas Micropulser (BIO-RAD).

Después de incubar las células en 1ml de medio SOC a 37°C por 1 h, se inocularon dos cajas petrí con LB agar, ampicilina (50ug/ml) y X-gal (300µg/ml), 50 y 100 uL de células por cada amplicon de 3000-5000 pb obtenido. Una vez obtenidas las colonias, al cabo de 24 hr, se verificó la presencia de los fragmentos clonados mediante un PCR de colonia y empleando las mismas condiciones de amplificación del templete inicial.

Los productos clonados se emplearon como templetos para obtener productos de amplificación internos y corroborar el ordenamiento de los genes entre los genes *Cytb* y *12S* del genoma mitocondrial.

Todos los amplicones obtenidos se secuenciaron en ambas direcciones empleando nucleótidos terminadores de secuencia marcados con fluorocromos (BigDye® - Applied Biosystems) y un secuenciador automático ABI Prism® 3100 (Automated Capillary DNA Sequencer) en SeqXcel Inc. (San Diego, CA. <http://www.seqxcel.com>).

### **Análisis de Secuencias**

Las secuencias obtenidas para el amplicón *RC<sup>1</sup>-Pro<sup>2</sup>*, así como para los genes comprendidos entre el *Cytb* y el *12S de cada especie*, se editaron en el programa ChromasPro v1.7.5 y se alinearon manualmente utilizando el programa UltraEdit v20.0. Para ello se emplearon como base del alineamiento las regiones homólogas *Cytb-12S de Philomachus pugnax* (GenBank GQ255993.1), *Arenaria interpres* (NC\_003712.2) y de *Egretta eulophotes* (GenBank KJ190949). Posteriormente, las secuencias se analizaron y se alinearon con mayor detalle en el programa Genious v4.8.4 para obtener la secuencia consenso correspondiente a cada especie.

Con las secuencias del amplicón *RC<sup>1</sup>-Pro<sup>2</sup>* y, muy particularmente, con las secuencias clonadas de ambas regiones control, se realizó un análisis filogenético. Para ello se empleó el programa MEGA 4 (Tamura et al., 2007) y el modelo de sustitución de Kimura de dos parámetros (Kimura, 1980).

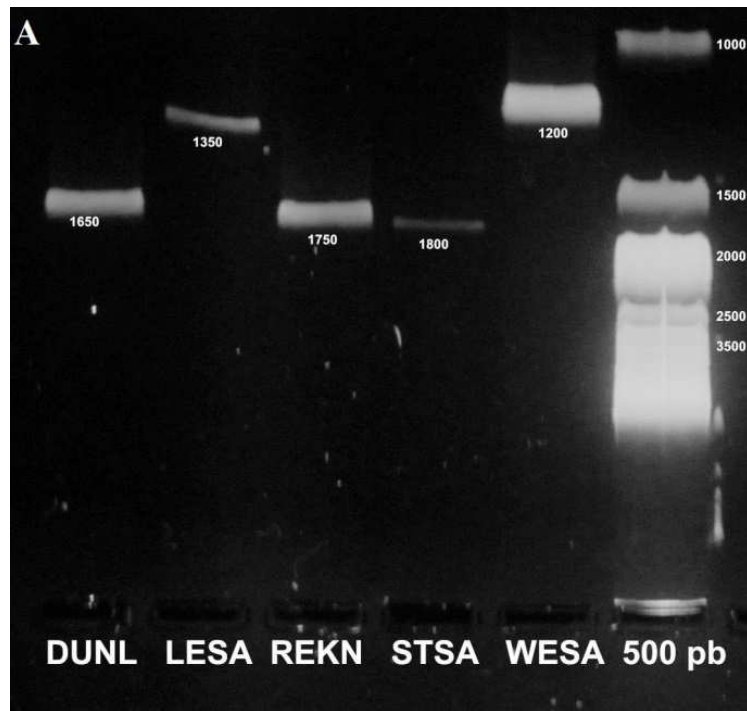
## Resultados

### Amplicón RC<sup>1</sup>-Pro<sup>2</sup>:

Como primera aproximación para evaluar si existían diferencias en la organización y el tamaño de la región control mitocondrial en la Familia Scolopacidae se amplificó y evaluó la región RC<sup>1</sup>-Pro<sup>2</sup> en 5 especies del género *Calidris*: *Calidris mauri* (Western Sandpiper / WESA), *Calidris himantopus* (Stilt Sandpiper / STSA), *Calidris minutilla* (Least Sandpiper / LESA), *Calidris canutus* (Red Knot / REKN) y *Calidris alpina* (Dunlin / DUNL). Los resultados de la electroforesis evidenciaron amplificación exitosa de la región RC<sup>1</sup>-Pro<sup>2</sup> en las cinco especies (*Figura 8*) y por ende la ocurrencia de duplicación de al menos algunos de los genes adyacentes a la región control mitocondrial.

Se determinó que *C. mauri* posee un amplicón RC<sup>1</sup>-Pro<sup>2</sup> de una longitud de 1140 pb, mientras que en *C. alpina* fue de 1650 pb, en *C. canutus* de 1750 pb, en *C. minutilla* de 1350 pb y en *C. himantopus* de 1800 pb. Tomando en cuenta que en *Philomachus pugnax* el peso molecular reportado de esta región es muy similar (1186 pb) al obtenido para estas especies, se consideró altamente probable que compartieran el mismo ordenamiento y patrón de duplicación.

Las secuencias de las cadenas ligeras y pesadas se ensamblaron y editaron para analizarlas a través de la herramienta BLAST (Basic Local Alignment Search Tool) que ofrece la base de datos del NCBI (National Center for Biotechnology Information). El porcentaje de identidad de las secuencias con respecto a la base de datos, fue de alrededor del 87% al 97% de similitud con las secuencias homólogas reportadas en las bases de datos para otras especies del género y para *Philomachus*. Del resto de las especies analizadas, los cebadores para el amplicón RC<sup>1</sup>-Pro<sup>2</sup> amplificaron positivamente solo para *Tringa semipalmata* (WILL) y *Tringa flavipes* (LEYE).



**B**



Figura 8. Amplicón  $RC^1$ - $Pro^2$  (L141-PropR). (A) Fotografía del gel de agarosa al 1.5%; las bandas indican la efectividad de amplificación de los cebadores, se indica en cada una de las bandas el peso aproximado en pares de bases (pb) con respecto a la escalera (500pb) el sentido de la corrida va de + a - en sentido vertical (de arriba abajo), las bandas más ligeras se sitúan en la parte superior. En la parte inferior se muestra el nombre de las aves en el carril sembrado; *C. alpina* (DUNL), *C. minutilla* (LESA), *C. canutus* (REKN), *C. himantopus* (STSA) y *C. mauri* (WESA). (B) Se muestra el esquema del orden de los genes de *P. pugnax* y el sitio de anclaje de los cebadores; directo (L141) e inverso (PropR).

## **Cobertura de la secuenciación de la región *Cytb-12S***

Empleando el conjunto de cebadores diseñados para amplificar la región *Cytb-12S* en *P. pugnax* y algunos cebadores diseñados con base en la secuencia del ADN mitocondrial completo de *Arenaria interpres* (GenBank NC\_003712), se obtuvieron amplicones de longitud y calidad de lectura de secuencia muy variables. La **Tabla 1** resume los amplicones y condiciones que permitieron el ensamblaje completo de la secuencia para *C. mauri*, mismas con las que se intentó amplificar el resto de las especies evaluadas.

**Tabla 1.** Resumen del control de calidad y evaluación de la amplificación y la secuenciación

Región	Cebadores		Programa	Temp. °C	Longitud (pb)	Control de calidad
	Cebador directo	Reverse				
<i>Cytb</i>	Cytb-end	ND6HR	2	55.1	1200	OK
<i>CR1-ND62</i>	LH402	ND6P2R	2	52.0	1000	OK
<i>CR1-ND62</i>	L716	ND6P2R	2	54.0	700	OK
<i>CR2-12S</i>	L438	H1537	2	54.0	500	OK
<i>Cytb</i>	CmCytb F	ND6HR	2	55.0	1200	OK
<i>RC's</i>	cmRC F	cmRC R	2	52.0	800	OK
<i>Cytb</i>	CmCytbF	CmCytbR	1	52.0	800	OK
<i>ND6-CR</i>	ND6L3	H451	2	56.0	1000	OK
<i>CR1-ND62</i>	LH402	ND6P6R	2	55.0	1000	OK
<i>Cytb-Pro2</i>	b71	PropR	2	60.0	450	OK
<i>Cytb-Pro2</i>	b3	PropR	2	60.0	947	OK
<i>Cytb-Pro2</i>	b5	PropR	2	60.0	700	OK
<i>Cytb-Pro2</i>	b86	PropR	2	60.0	267	Secuencia corta
<i>CR-12S</i>	L716	H1390	2	55.0	571	OK
<i>CR-12S</i>	L716	H1827	2	55.0	800	OK
<i>CR-12S</i>	LH402	H1827	2	55.0	1100	OK
<i>CR-12S</i>	LH402	H1390	2	55.0	1100	Secuencia no útil
<i>RC1-Pro2</i>	L141	PropR	1	52.0	1200	OK
<i>CR-12S</i>	LH402	H1537	3	66.0	1000	Picos múltiples región CAA
<i>cyt-CR1</i>	JCytbF	JRCRR	4	57.0	1014	Reverse no secuenciable
<i>CR-16S</i>	CmCRbF	16ScR	4	62.0		
<i>CR-12S</i>	CmCRbF	H1827	4	62.0	Contig	OK
		H1390	4	62.0		
<i>CAA-12S</i>	mtScoW3F	mtScoW2R	4	62.0	280	Banda ligera, debajo de las 500pb
<i>CAA-12S</i>	mtScoW4F	mtScoW4R	4	62.0	460	OK
<i>CAA-12S</i>	mtScoW5F	mtScoW5R	4	62.0	400	OK
<i>CR2-12s</i>	mtScoW1F	mtScoW1R	4	50.0	130	Regular
<i>CR2-12s</i>	mtScoW1F	mtScoW6R	4	50.0	800	Regular
<i>CR1-propR</i>	PreNF	PropR	4	64.0	457	no secuenciable
<i>CR-12S</i>	L141	H1390	4	67.0	250	OK
<i>CR-12S</i>	L141	H1537	3	60.0	450	OK
<i>CR-12S</i>	L141	mtScoW6R	4	56.0	460	OK

<i>CR1-CR2</i>	LH402	H451	5	58.0	Contig	Regular
<i>cytb-12s</i>	JCytbF	H1827	6	65.7	5000	Regular

]

La Figura 9 resume gráficamente la cobertura de secuenciación obtenida para las especies de la Familia Scolopacidae evaluadas.

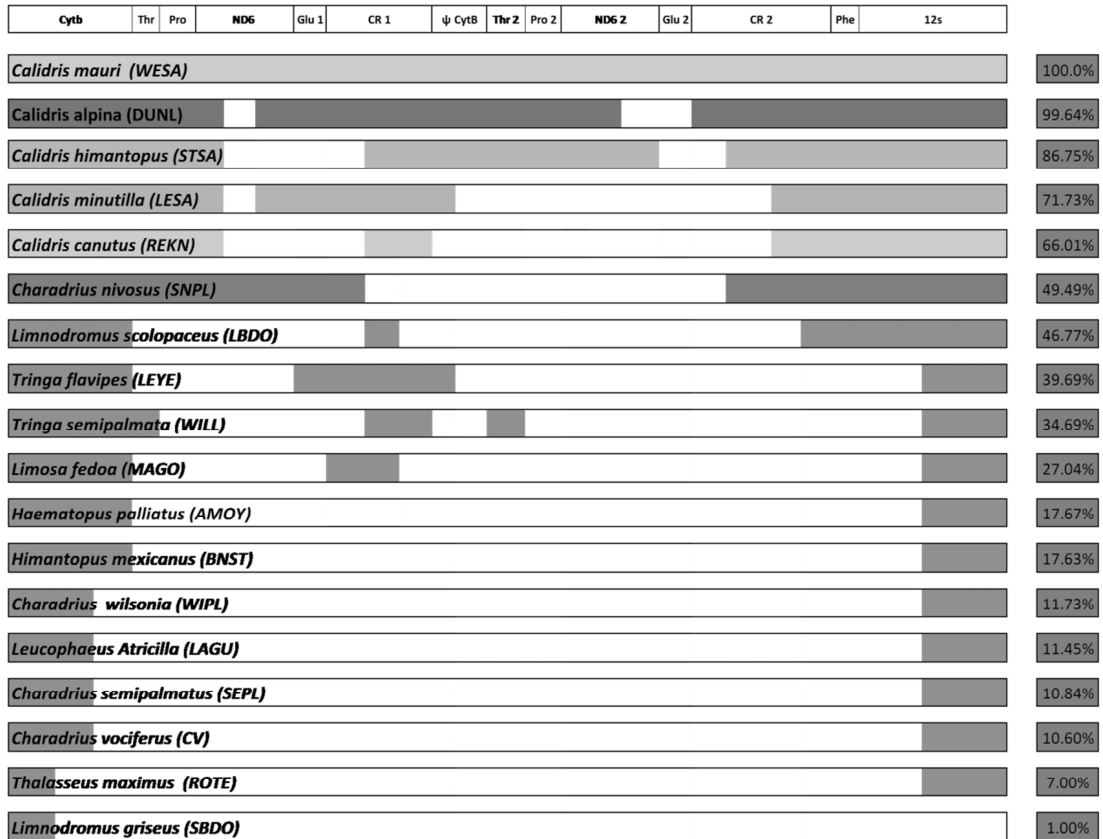


Figura 9. Muestra la cobertura de amplificaciones. En la parte superior se muestra el esquema del acomodo de los genes observado en *Calidris mauri*. Debajo del esquema se muestran los nombres de las aves playeras disponibles y debajo de estos los cuadros grises representan la cobertura de los amplicones obtenidos para cada una de ellas.

Pese a que se empleó una ADN polimerasa de amplio rango y de alta procesividad, la eficiencia y la especificidad de la amplificación fue pobre. Para la mayoría de los fragmentos largos (3500-5000 pb) se obtuvieron invariablemente bandas múltiples y secuencias con un elevado número de sitios con ambigüedades como producto de la inespecificidad de los cebadores.

De igual forma, muchos de los fragmentos de tamaño intermedio (1000-2500 pb) resultaron en secuencias de mala calidad o truncadas por la presencia de homopolímeros.

Es por ello que, como medida de control de calidad y como estrategia para ampliar la cobertura de la secuenciación, se corroboró el ordenamiento de los genes mitocondriales utilizando amplicones de una longitud de alrededor de 5000 pb que se obtenían usando los cebadores “*JCytbF*” y “*H1827*”. Usando estos amplicones como ADN template en una segunda reacción de PCR, y utilizando distintos pares de cebadores que estratégicamente se anclan internamente en este template, nos permitió generar amplicones de menor longitud que elucidaban las bases faltantes o ambiguas de estas regiones problemáticas. El resultado obtenido confirmó la posición, el tamaño y la homología de las secuencias, e incluso nos permitió completar algunas secciones faltantes y ampliar la cobertura de la secuencia total.

La única secuencia *Cytb-12S* que se logró completar fue la de *C. mauri* que contiene un total de 5920 pb (Figura 10), y en un alto porcentaje la de *C. alpina* (5800 pb). Para el resto de las especies del género *Calidris* el porcentaje de cobertura varió entre el 66% y el 99% y disminuyó drásticamente para los representantes de otros géneros.

GACCTCCCG	CCCCCTCAA	TATTTCCGCT	TGATGAACT	TTGGATCACT	CCTAGGTATC	TGCTAGGAA	CACAAATCCT	[ 80]
AACCGGTCTC	TTACTCGCTA	TACACTACAC	CGCAGACACA	ACCCTAGCCT	TCTCATCCGT	GGCCCATACA	TGTCGAAACG	[ 160]
TACAAATACGG	CTGACTAATT	CGCAACCTAC	ATGCAAAACGG	AGCCTCATTC	TTCTTCAATT	GCATCTACCT	CCACATCGGA	[ 240]
CGAGGTTTTT	ACTACGGCTC	ATACCTATAC	AAAGAGACAT	GAACACAGG	AGTCATCCTA	CTCTTAACCC	TAATAGCAAC	[ 320]
CGCTTTTGSTA	GGATATGTAC	TTCCATGAGG	ACAAATATCT	TTCTGAGGAG	CTACAGTCAT	CACCAATTTA	TTCTCAGCAA	[ 400]
TCCCATACAT	CGGCCAAACC	TTGGTAGAAT	GGGCATGAGG	GGGATTTTCC	GTAGACAACC	CAACACTAAC	CCGATTCCTT	[ 480]
GGCCTCCACT	TCCACTTCCC	ATTCATAAATC	GCTGGCCTTA	CCTTAATCCA	CCTTACCTTC	CTACACGAAA	CAGGCTCAA	[ 560]
CAATCCACTT	GGCATTGTAT	CGGATTGCGA	CAAAATCCCA	TTCCACCCTT	ACTTTTCACT	CAAAGATATC	CTAGGATTCA	[ 640]
TCATCATATT	CCTATTCCCTA	CTAATCTCTG	CCCTATTCTC	ACCTAATCTA	CTTGGAGACC	CAGAAAACCT	TACCCAGCA	[ 720]
AACCCCTCTAG	TTACACCTCC	CCATATCAAG	CCCGAATGAT	ATTTCCTGTT	CGCATATGCC	ATTCTACGTT	CAATCCCAA	[ 800]
CAAACATAGGA	GGTGTACTAG	CTTTAGCTGC	CTCCGTACTA	GTCTTATTC	TAATCCACT	TCTCCACAAA	TCCAAACAA	[ 880]
GGTCAATAAC	TTTCCGTC	CTATCCCAAC	TCTTATCTG	AACCCCTAGT	GCCAACCTCT	TCATCTTAA	TTGAGTAGGT	[ 960]
ACCAACCCCG	CAGAACACCC	ATTTCATCATC	ATTGGCCAAC	TAGCATCCCT	CACCTACTTT	ACAATCCCTC	TAATCCCTATT	[ 1040]
CCCCATTACG	GGAGCCCTAG	AAAATAAAAT	ACTCAACTAC	TAAACACTCT	AATAGTTTAT	AAAAAACATT	GGCTTTGTAA	[ 1120]
ACCAAAGACT	GAAGACTATC	CCTCTTCTTA	GAGTTTACCC	TAAACATAAA	TCAGAAAAG	AGGACTTAAA	CCTTATCTC	[ 1200]
CAACTCCCAA	AGCTGGTATT	TTATACTAAA	CTATTTTCTG	ACCTTCCCAA	CCCTAAACCG	CCGAAATAGC	CCCACGAGC	[ 1280]
AATCCCGCA	CCAACCTTAA	CACCACAAAC	AAAGTCAACA	ACAACCCCTA	CCCCGCTACT	AAAAACATCC	CACCCCCGTA	[ 1360]
CGAATAAAAC	ATAGCCACCC	CACTAAAATC	TAATCGCAC	GAAAACATTC	CACCCACTATC	AACAGTAACC	ACTCCACCT	[ 1440]
TTCAACATTC	AACAAGTCCA	CCAATAATCA	TACCAGCAAT	TACTACTAAA	ACAAGCTTG	CACCATACCC	AACACACGC	[ 1520]
CAACTCTCTC	AGGCTTCAGG	AAAAGGATCC	CGCCCAAAG	CACAGAAATA	GACAAAACC	ACTAACATTC	CCCCAAATA	[ 1600]
TACTATAAAC	AACACCAATG	ACACAAACGA	CACACCCAAA	CTTAACAACC	ACCCACACCC	TGCAACAGAA	GCCAACACCA	[ 1680]
AACCAACTAC	CCATAAATAT	GGAGAAGGAT	TAGATGCAAC	AGCCAAACCC	CCTAATACAA	AACACAAACC	CAATAAAAG	[ 1760]
ACAAAATAAG	TCCAACCCGG	TCTCGCTTGG	CTTTTTCCTA	AGATCTATGG	CCTGAAAAC	CATCGTTTGT	AAATTTCAA	[ 1840]
CACRCAGGAA	CATACACCTG	CCCGCCCAT	AAAATACTTG	CGGCCCCCC	CTACCCCCCC	CCCCCGCC	GCATGTAAT	[ 1920]
CGGGCATT	TTATGTACAA	TCATGTGTTG	TATTGCATAT	ACTTGTCTCT	CCCCATAATA	CATACCATCC	ATGTTCCAAA	[ 2000]
TCCATTAATA	TCCAACCCGG	CTATACACCT	CCACCCCTC	CACCTGCCTC	CAGAGACAAT	CGAGACAATG	AACCTAGGAA	[ 2080]
TATTCACATA	CTACGTAATA	AACCCATCAA	CTTGCTAGAA	TTATACATAA	CAAACCCCTG	AAAGTGTACG	GCAGTGCTT	[ 2160]
AACACACGCT	ATGAATGGTC	TAAGTGCGA	CAGTTCGAAA	ACTCTTGAAG	TGCACCCAG	TCGTACCAGG	TTATTTATA	[ 2240]
ATCGAGCTCC	TCACGTGAAA	TCAGCAACCC	GGCGTAAGTA	ATGTCCTGCG	TTACTAGCTT	CAGGCCCAT	CTTTCCCTC	[ 2320]
AAACCCTAGC	ACAATTTGCT	CTTTTGC	TCTGTTTCT	ATGTCAGGGC	CATAAATAGG	TTAATACTCA	TAATCTGCTC	[ 2400]
TTTACGAATA	CACCTGGTTG	GCTATATATC	ACCATTTCG	TCCGTTGATCG	CGGCATFCCA	AAATTCCTAT	ACTTTTGGTT	[ 2480]
CCTTTTTTTT	TTGGGGGGTC	TTCCACAGTT	GCCCTCCAG	CGCAACGGGT	GAATACAATC	TAAGACCTGT	GCATCTCATG	[ 2560]
CGTTGCGGCC	TTATTTTGGC	CCTCAGGCT	TACTGAAATGA	GACGGTTTCA	AGTGTTTGGG	GAATCATATC	AACACTGATG	[ 2640]
CACTTTGT	TACACTTGGT	TATGGCTCC	CGCAAGCAG	CGTTCATGCC	ACTGTTTAA	GAATGCTTGT	GGACATGAT	[ 2720]
TTTTCTATTT	TTCACTTCT	CTAATTTCT	TAACAACACT	AGTAGGTTT	AGCTAATTT	AAACTACAT	TCCATACAG	[ 2800]
ATTTTATTTT	CACATTTTTC	CATACGTTAC	CGCCTGAA	ATTACATTTAG	TAAACCCATC	CATCAATTC	GTCATCGATC	[ 2880]
ATCCACATTT	CACCTAACAA	AACAAGCTCA	AACAATCCAT	TTAACATTTG	ATCAGATTAC	AACAAAATCC	CACCTCTCAC	[ 2960]
CTACTTTACA	ATCCTCTCAA	TCCTATTC	CATTACGGGA	GCCCTAGAAA	ATAAAATACT	CAACTACTAA	ACACTCTAAT	[ 3040]
AGTTTATAAA	AAACATTTGGT	CTTGTAAC	AAAGACTGAA	GACTATCCCT	CTTCTAGAG	TTTACCCTAT	AAATAAATCA	[ 3120]
GAAAAGAGG	ACTTAAACCT	CTATCTCAA	CTCCAAAGC	TGGTATTTTA	TACTAAACTA	TTTTCTGACC	TCCCAACCC	[ 3200]
TAAACCGCCC	GAATAGCCCC	ACGAGACAAT	CCCCGCACA	ACTCTAACAC	CACAAAACAA	GTCAACAACA	ACCCTCACCC	[ 3280]
CGCTACTAAA	AACATCCCAA	CCCCGTACGA	ATAAAACATA	GCCACCCAC	TAAAATCTAA	TCGACCCGAA	AACATTCAC	[ 3360]
CACATCTAAC	AGTAACCACT	CCAACCTTTC	AACATTCAC	AAGTCCACCA	ATAATCATAC	CAGCAATTA	TACTAAAACA	[ 3440]
AGGCTTGAC	CATACCCAAC	AACACGCCAA	TCTCCTCAGG	CTTCAAGAAA	AGGATCCGCC	GCCAAAGACA	CAGAATAGAC	[ 3520]
AAAAACCACT	AACATTTCCC	CCAAATATAC	TATAAACAC	ACCAATGACA	CAAAACGAC	ACCCAAACTT	ACAACCCACC	[ 3600]
CACACCCCTG	ACAAGAAGCC	AACACCAAAC	CAACTACCC	ATAATATGGA	GAAGGATTAG	ATGCAACAGC	CAACTCCCT	[ 3680]
AATACAAAC	AACAACCCAA	TAAAAGAAACA	AAATAAGTCA	TAGCAGTCC	TGCTTGGCTT	TTTTCCAAGA	CTTATGGCTT	[ 3760]
GAAAACCAT	CGTTGTTAAA	TTTCAACCAC	RCAGGAACAT	ACACCTGCC	GCCCCATAA	ATACTTGGCG	CCCCCCCCTA	[ 3840]
CCCCCCCC	CCGCGCCGCA	TGTAATTCGG	GCATTTTTTA	TGTACAATCA	TGTGTTGTAT	TGCATATACT	TGTCCTTCCC	[ 3920]
CATAAATACG	ACCATCCATG	TTCCAAATCC	ATTAATATCC	AACCCGGCTA	TACACCTCCA	CACCCCTCAC	CTGCCCTCAG	[ 4000]
AGAACAACTG	AGCAAAATGAA	CCTAGGAATA	TTCCATACAT	ACGTAATAA	CCCATCAACT	TGCTAGAAAT	ATACATAACA	[ 4080]
AACCTCCGAA	AGTGTACGGC	AGTGTCTTAA	CACACGCTAT	GAATGGTCTA	AGTCAGACA	GTTCGAAAAC	TCTTGAAGTG	[ 4160]
CACACCAGTC	GTACCAGGTT	ATTTATTAAT	CGAGCTCCTC	ACGTGAAATC	AGCAACCCGG	CGTAAGTAAT	GTCCTGCGTT	[ 4240]
ACTAGCTTCA	GGCCCATCT	TTCCCCCTAA	ACCCTAGCAC	AACTTGCTCT	TTTGGCCCTC	TGGTTCCCTAT	GTCAGGGCCA	[ 4320]
TAAATAGGTT	AAATCTCATA	ACTTGTCTCT	TACGAATACA	TCTGGTTGGC	TATATATCAC	CATTTTCTGTC	CGTGATCGCC	[ 4400]
GCATTCCAAA	ATTCTTATAC	TTTTGGTCC	TTTTTTTTTT	GGGGGGTCTT	CACAGGTTGC	CCTCCAGCGC	CAACGGGTGA	[ 4480]
ATACAACTA	AGACCTGTGC	ATCTCATGCG	TTGGCGCCTT	ATTTTGGCCC	TCAGGCGTTA	CTGAATGAGA	CGGTTTCAAG	[ 4560]
TGTTTGGGGA	ATCATATCAA	CACCTGATGCA	CTTTGTTTTA	CACTTGGTTA	TGGCCCTCCC	GCAAGCAGCG	TTCATGCCAC	[ 4640]
TGTTTAAATGA	ATGCTTGTGG	GACATGATTT	TTCAATTTTT	CACCTTCTCT	AACTTTCTTA	ACAACACTAG	TAGGTTTACG	[ 4720]
CTAAATTTAA	ACTACATTTT	CATTACACAT	TTTATTTTCA	CATTTTTC	TACGTTACCG	GCACTGAAAT	TACATTAACA	[ 4800]
CAACAACCAT	AAACATTCAC	GTCCATTTAC	TTAAACATAA	CCCCCTGCCA	ATAAAAGAAA	CTCTCCTGCA	AAAACAAACA	[ 4880]
AACAACAAA	CAACAACA	AACAACAACA	CAACAACA	AACAACAACA	CAACAACA	AACAACAACA	CAACAACA	[ 4960]
TTCAATCCGT	CCCTGTAGCT	TAAACATCAA	AGCATGGCAC	TGAAGATGCC	AAGATGGCTG	TCATATAACA	CCCAGAGACA	[ 5040]
AAAGACTTAG	TCCTAACCTT	ACCGTAAAT	TTTGCTAAAT	ATATACATGC	AAGTATCTGC	GCTCCAGTGT	AAATGCCCTT	[ 5120]
AATCTCTTAC	TAAGAAGAAA	GGAGCGGGTA	TCAGGCACAC	CTAAACTGTA	GCCCAAGAGC	CCTTGCTCAG	CCACACCCCT	[ 5200]
ACGGGTACTC	AGCAGTAATT	AACATTAAGC	AATAAGTGAA	AACCTTGACTT	AGTTATAGCA	ACACCAGGGT	TGGTAAATCT	[ 5280]
TGTGCCAGCC	ACCCGGTCA	CACAAGAGAC	CCAAATTAAC	TGTATACGGC	GTAAGAGAGT	GTATCATGTT	ATCTTAAACA	[ 5360]
CTAAGACCAA	AGTCAACTA	AGCTGTCTATA	AGCCCAAGAT	GTACCTAAGA	CCCCTTAA	AACGATCTTA	ACGCTTTGA	[ 5440]
TTAATTTAAT	TCCACGAAG	CTAAGACACA	AACCTGGATT	AGATACCCCA	CTATGCTTAG	CCCTAAATCT	TGATGCTACC	[ 5520]
CCGTACTAAA	GCATCCGCC	GAGAACTACG	AGCACAAACG	CTTAAACTC	TAAGGACTTG	GCGGTGCCCT	AAACCCACT	[ 5600]
AGAGGAGCCT	GTTCTATAAT	CGATAACCCA	CGTTACACCC	GACCCCTC	TGCCAAAGCA	GCCTACATAC	CGCCGTCCGC	[ 5680]
AGCTCACCTC	TTCTGAGAGT	TCAACAGTGA	GCACAAATAGC	CCAAACCCGC	TAATAAGACA	GGTCAGGTA	TAGCCCATGG	[ 5760]
GGCGAAGAA	ATGGGCTACA	TTTTCTAAGA	TAGAAAATC	ACGGAAGGGG	GCATGAAACT	GGCCCCGCA	AGGCGGATTT	[ 5840]
AGTAGTAAAG	TGGGCAATA	TAGCCCTC	TTAAGTCGGC	CCTGGGGCAC	GTACATACCG	CCGTCACCC	TCCCTCATAA	[ 5920]

**Figura 10** Secuencia del ADN mitocondrial de *Calidris mauri* que comprende desde el Citocromo b hasta el ARNr 12S.

**Tabla 2.** Sitio de los genes desde el *Cytb* hasta el ARNr 12S. Los genes se encontraron utilizando el programa DOGMA (Dual Organellar Genome Annotator) para identificar los límites de codificación de proteínas, ARNr, ARNt y genes utilizando la configuración por defecto de BLAST (Wyman et al., 2004) basados en datos de ADNmt para vertebrados .

<b>Inicio</b>	<b>Fin</b>	<b>Gen</b>	<b>Cadena</b>
1	1080	<i>Cytb</i>	+
1086	1155	<i>tRNA<sup>Thr</sup></i>	+
1171	1240	<i>tRNA<sup>Pro</sup></i>	-
1256	1774	<i>ND6</i>	-
1778	1851	<i>tRNA<sup>Glu</sup></i>	-
2914	3027	<i>ψCytb</i>	+
3033	3102	<i>tRNA<sup>Thr2</sup></i>	+
3118	3187	<i>tRNA<sup>Pro</sup></i>	-
3203	3721	<i>ND6<sup>2</sup></i>	-
3725	3798	<i>tRNA<sup>Glu2</sup></i>	-
4969	5040	<i>tRNA<sup>Phe</sup></i>	+
5041	5912	<i>12S</i>	+

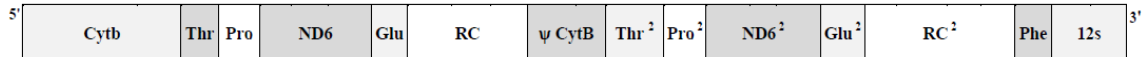


**Figura II.** Ubicación de los genes desde el Cytb hasta el ARNr 12S. La línea delgada muestra con números el sitio de las bases en esta sección del ADNmt. Las flechas de color verde indican el nombre de los genes (Cytb, ND6). Las flechas rosas indican el nombre de los ARNt (Thr, Pro, Glu y Phe). Las flechas de color negro indican las regiones no codificantes (RC<sup>1</sup> y RC<sup>2</sup>). La flecha en rojo indica el ARNr (12S). Las anotaciones se realizaron manualmente en el programa Geneious V4.7 (Drummond et al., 2009).

## Ordenamiento de genes de la región *CytB-12S* en *Calidris mauri*

El análisis detallado de los alineamientos de secuencias de los diferentes amplicones obtenidos a partir del ADN de *C. mauri* permitió detectar un segmento duplicado en tándem, mismo que comprende la porción entre el *tRNA<sup>Thr</sup>* y la *RC*. Esta duplicación o copia es contigua a la primera *RC* y se encuentran separadas tan solo por una región de aproximadamente 145 pb (Figura 12). A través del análisis con la herramienta BLAST (NCBI) y usando como base el genoma mitocondrial de *A. interpres* y otras especies de aves playeras, fue posible identificar las correspondencias con los genes mitocondriales típicos del genoma aviar (*Cytb*, *Thr*, *Pro*, *ND6*, *Glu*, *RC*, *Phe* y *12S*).

El fragmento de 145 pb que separa ambas copias exhibió un elevado porcentaje de identidad con la secuencia del *Cytb*, pero al estar truncada esta región no codificante se consideró un pseudogén de *Cytb* (*ψ Cytb*) (Figura 12).



**Figura 12:** Orden de los genes en *Calidris mauri*; se muestra la región que comprende desde el citocromo b (*Cytb* B) hasta el ARN ribosomal 12S (*12S*).

La comprobación de los pesos de las bandas fue indispensable ya que permitió depurar errores en las condiciones del programa de PCR en el termociclador y además sirvió de referencia para ubicar las secuencias obtenidas en el alineamiento global en construcción. Si en el genoma mitocondrial existe duplicación de genes similar al que presenta *C. mauri*, los genes duplicados intermedios entre la *RC*<sup>1</sup> y *RC*<sup>2</sup> impedirían amplificar de manera selectiva a la *RC*<sup>2</sup>, ya que, es posible ubicar de manera objetiva un cebador inverso en el gen de Fenilalanina (*Phe*) o en el *12S*, pero no se puede ubicar al cebador directo en algún sitio específico en los genes anteriores a la *RC*<sup>2</sup>, ya que el anclaje puede darse en cualquiera de las copias de los genes duplicados (el más cercano al extremo 5' o al que se encuentra ubicado hacia el extremo 3'), generando productos múltiples, y por lo tanto, las secuencias obtenidas serán de baja calidad, ilegibles o inespecíficas. Véase Figura 16.

La elección del amplicón RC<sup>1</sup>-Pro<sup>2</sup> utilizando el cebador directo “L141” e inverso “PropR” nos permiten discriminar entre las Regiones Control duplicadas. El anclaje de los cebadores permiten amplificar la RC<sup>1</sup> debido a que el gen de Pro<sup>2</sup> flanquea únicamente a la RC<sup>1</sup> mientras que la RC<sup>2</sup> se encuentra flanqueada por el gen de *Phe* seguido de *12S* al final de su extremo 3’ (véase Figura 8).

### **Comparación de las copias de la región control en *Calidris mauri***

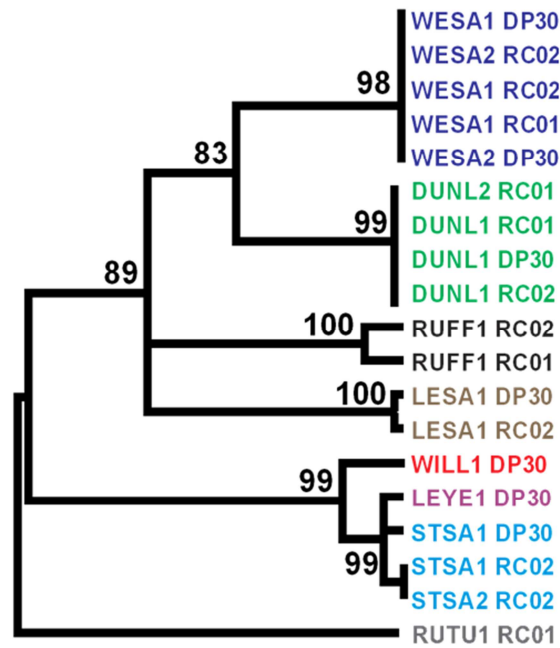
*Calidris mauri* presentó una duplicación en la RC y ambas poseen las regiones conservadas que caracterizan a la RC. La horquilla de replicación (poli C Stretch) la cual consiste en una cadena mono nucleotídica de repeticiones de Citosina la cual es característica de la región inicial (5’ → 3’) de la RC adyacente al *tRNA<sup>Glu</sup>* se encontró en ambas copias. El análisis de RC’s duplicadas indicaron que poseen las secuencias asociadas a los sitios de terminación (TAS), las cuales son características del dominio I. También se ubicaron las secuencias características del dominio II intermedio tales como la F box, E box, D box, C box D box BSB (bird similarity box) y similitud en las secuencias del dominio III tales como el CSB1 (Conserved Sequence Block 1) (Figura 13).

El alineamiento mostró que existe una diferencia en longitud entre las RC’s. La primer RC (la más cercana al extremo 5’) tiene una longitud de ~1041 pb, mientras que la copia 2 (la más cercana al extremo 3’) tiene una longitud de ~1176 pb. El alineamiento demostró que en un mismo individuo, ambas copias de la RC poseen un porcentaje de similitud de hasta el 99.80% desde el nucleótido 1 hasta el 1005, pero a diferencia de la RC<sup>1</sup>, la RC<sup>2</sup> posee a partir del nucleótido 1080 repeticiones de tetranucleótidos CAAA aumentando de longitud por ~135 pb, que en algunas aves se ha presentado en la CR<sup>2</sup> en el extremo 3’ del dominio III.

				C-Stretch			TAS		TAS		
#mauriRC1	AGGAACATAC	ACCTGCCCGC	CCCATAAAAT	ACTTGCGGGC	CCCCCTACC	CCCCCCCCCC	GCGCCGCATG	TAATTCGGGC	ATTTTTTATG	TACAATCATG	[ 100]
#mauriRC2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	[ 100]
				TAS							
#mauriRC1	TGTTGTATTG	CATATACTTG	TCCTCCCCCA	TAATACATAC	CATCCATGTT	CCAAATCCAT	TAATATCCAA	CGGGGTATA	CACCTCCACA	CCCCTCACCT	[ 200]
#mauriRC2	.....	.....	.....T.....	.....	.....	.....	.....	.....	.....	.....	[ 200]
									TAS		
#mauriRC1	GCCTCCAGAG	A-CAATCGAG	ACAATGAACC	TAGGAATATT	CACATACTAC	GTAATAAACC	CATCAACTTG	CTAGAATTAT	ACATAACAAA	CCCTCGAAAG	[ 300]
#mauriRC2	.....	.A.....	.....	.....	.....	.....	.....	.....	.....	.....	[ 300]
#mauriRC1	TGTACGGCAG	TGCTTTAACA	CACGCTATGA	ATGGTCTAAG	TGCAGACAGT	TCGAAAACCTC	TTGAAGTGCA	CACCAGTCGT	ACCAGGTTAT	TTATTAATCG	[ 400]
#mauriRC2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	[ 400]
		F-BOX					E-BOX				
#mauriRC1	AGCTCCTCAC	GTGAAATCAG	CAACCCGGCG	TAAGTAATGT	CCTGCGTTAC	TAGCTTCAGG	CCCATCTCTT	CCCCCTAAAC	CCTAGCACAA	CTTGCTCTTT	[ 500]
#mauriRC2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	[ 500]
		D-BOX					C-BOX				
#mauriRC1	TGCGCCTCTG	GTTCCATATG	CAGGGCCATA	AATAGGTAA	TACTCATAAC	TTGCTCTTTA	CGAATACATC	TGGTTGGCTA	TATATCACCA	TTTTCGTCCG	[ 600]
#mauriRC2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	[ 600]
#mauriRC1	TGATCGCGGC	ATTCCAAAAT	TCTTATACTT	TTGGTTCCTT	TTTTTTTTGG	GGGTCTTCA	CAGGTTGCC	TCCCAGCGCA	ACGGGTGAAT	ACAATCTAAG	[ 700]
#mauriRC2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	[ 700]
										BSB	
#mauriRC1	ACCTGTGCAT	CTCATGCGTT	GCGGCCTTAT	TTTGGCCCTC	AGGCGTTACT	GAATGAGACG	GTTTCAAGTG	TTTGGGAAT	CATATCAACA	CTGATGCACT	[ 800]
#mauriRC2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	[ 800]
		D-BOX					CSB 1				
#mauriRC1	TTGTTTACA	CTTGGTTATG	GCCTCCCCGC	AAGCAGCGTT	CATGCCACTG	TTTAATGAAT	GCTTGTGGGA	CATGATTTTT	CATTTTTTCA	CTTCTCTAA	[ 900]
#mauriRC2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	[ 900]
										TAS	
#mauriRC1	CTTCTTAAC	AACACTAGTA	GGTTTCAGCT	AAATTTAAAC	TACATTTCCA	TTACACATTT	TATTTTCACA	TTTTTCCATA	CGTTACGGC	ACTGAAATTA	[1000]
#mauriRC2	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	[1000]
		TAS									
#mauriRC1	CATTAGTAAA	CCCATCCATC	AAATTCGTCA	TCGATCATCC	C-----	-----	-----	-----	-----	-----	[1100]
#mauriRC2	....AC.C.	A.A.-....A	..CA.TCA.G	..C..TTA.T	TAAACATAAC	CCCCTGCCAA	TAAAAGAAAC	TCTCCTGCAA	AAACAAACAA	ACAAACAAAC	[1100]
#mauriRC1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	[1177]
#mauriRC2	AAACAAACAA	ACAAACAAAC	AAACAAACAA	ACAAACAAAC	AAACAAACAA	ACAAACAAAC	AAACAAACAT	TCAATCC	-----	-----	[1177]

**Figura 13:** Alineamiento de las regiones control de *Calidris mauri*, los puntos indican nucleótidos idénticos a la secuencia de referencia, y los guiones indican huecos (gaps). Las secuencias que contienen los sitios característicos de la RC como la poli C Stretch, TAS motif, F box, E box, D box, C box D box BSB y el CSB1 están resaltados en color azul.

Para corroborar bajo qué proceso evolutivo la RC duplicada se mantiene en el tiempo, se realizó un análisis filogenético entre las RC duplicadas. Se midieron las distancias genéticas calculadas por el software MEGA 4.0 (Tamura et al., 2007) utilizando las secuencias de los dominios I y II de las RC's de *C. mauri*, *C. alpina*, *C. minutilla*, *C. himantopus*, *Tringa semipalmata*, *Tringa flavipes*, *A. interpres* y *P. pugnax* (Zheng et al., 2013) (Figura 14).



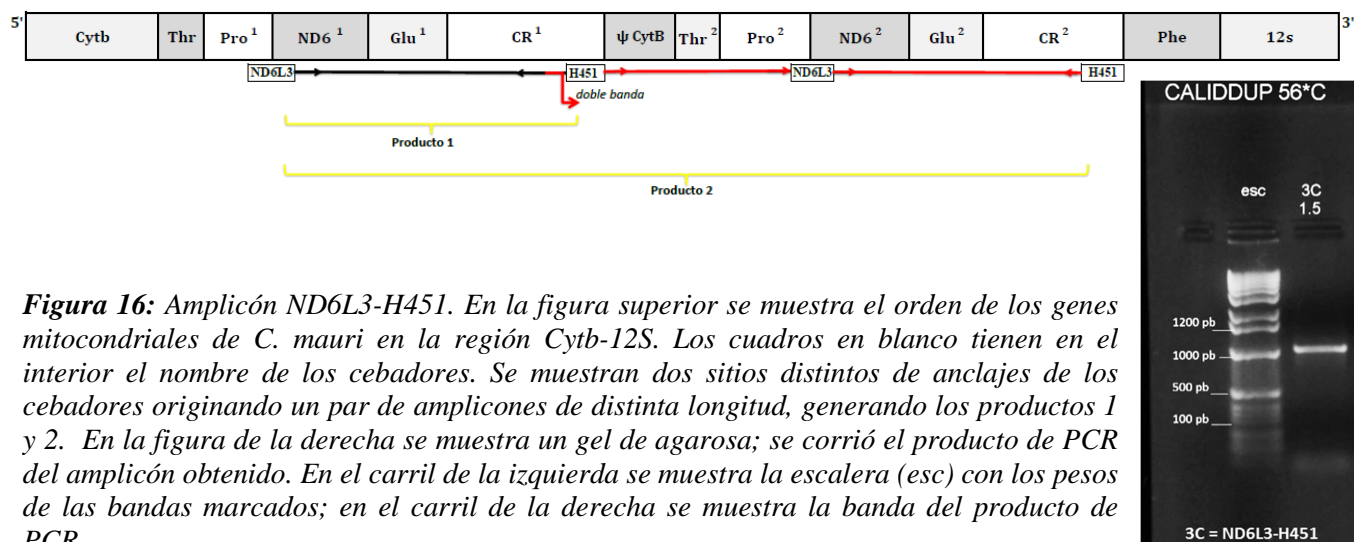
**Figura 14:** Árbol de distancias filogenéticas entre las RC's. El árbol muestra las secuencias de la RC<sup>1</sup> y la RC<sup>2</sup> de *C. mauri* (WESA), *C. alpina* (DUNL), *C. minutilla* (LESA), *C. himantopus* (STSA), *Tringa semipalmata* (WILL), *Tringa flavipes* (LEYE), *A. interpres* (RUTU) y *P. pugnax* (RUFF). Los números en los nodos indican el porcentaje de identidad entre las secuencias agrupadas. RUTU, al poseer sólo una RC, funcionó de grupo externo.

El análisis de divergencia genética relativa entre las RC's duplicadas indicó que las copias parálogas (RC<sup>1</sup> vs RC<sup>2</sup> dentro de las especies = 0,6%) eran más similares entre sí que a sus



El previo conocimiento de que el segmento RC<sup>1</sup>-Pro<sup>2</sup> amplificaba para *C. mauri*, nos indicaba presencia de una RC duplicada. Este supuesto permitió utilizar el ordenamiento descrito para *P. pugnax* como referencia para amplificar de manera discriminatoria la región que comprende desde el *Cytb* hasta la zona media de la RC<sup>1</sup>. También permitió utilizar la zona media de la RC<sup>1</sup> hacia el extremo final del *12S*. La electroforesis indicó el peso aproximado de los fragmentos obtenidos, si por ejemplo, el peso de las bandas de los amplicones excedía los 2500 pb mostraba una amplificación desde la parte media hacia el extremo inicial o final.

Por ejemplo, al amplificar un producto de PCR utilizando el cebador directo *ND6L3* y el inverso *H451* podía generar un par de productos con diferente longitud. En la electroforesis se pudo observar un par de bandas de pesos distintos, la primera banda tendría un peso aproximado de 1100 pb, mientras que la segunda banda podría exceder los 2500 pb. El peso de las bandas permitiría saber la posición en la que los cebadores se anclaban. El alineamiento global de los amplicones obtenidos permitiría determinar la longitud esperada de las combinaciones de los cebadores elegidos. En la Figura 16 se muestra la foto de la electroforesis del producto de PCR del amplicón *ND6L3- H451*, el peso de la banda es de aproximadamente 1100 pb, lo que indica que el amplicón obtenido es el producto 1, como se muestra en el esquema.



**Figura 16:** Amplicón *ND6L3-H451*. En la figura superior se muestra el orden de los genes mitocondriales de *C. mauri* en la región *Cytb-12S*. Los cuadros en blanco tienen en el interior el nombre de los cebadores. Se muestran dos sitios distintos de anclajes de los cebadores originando un par de amplicones de distinta longitud, generando los productos 1 y 2. En la figura de la derecha se muestra un gel de agarosa; se corrió el producto de PCR del amplicón obtenido. En el carril de la izquierda se muestra la escalera (*esc*) con los pesos de las bandas marcados; en el carril de la derecha se muestra la banda del producto de PCR.

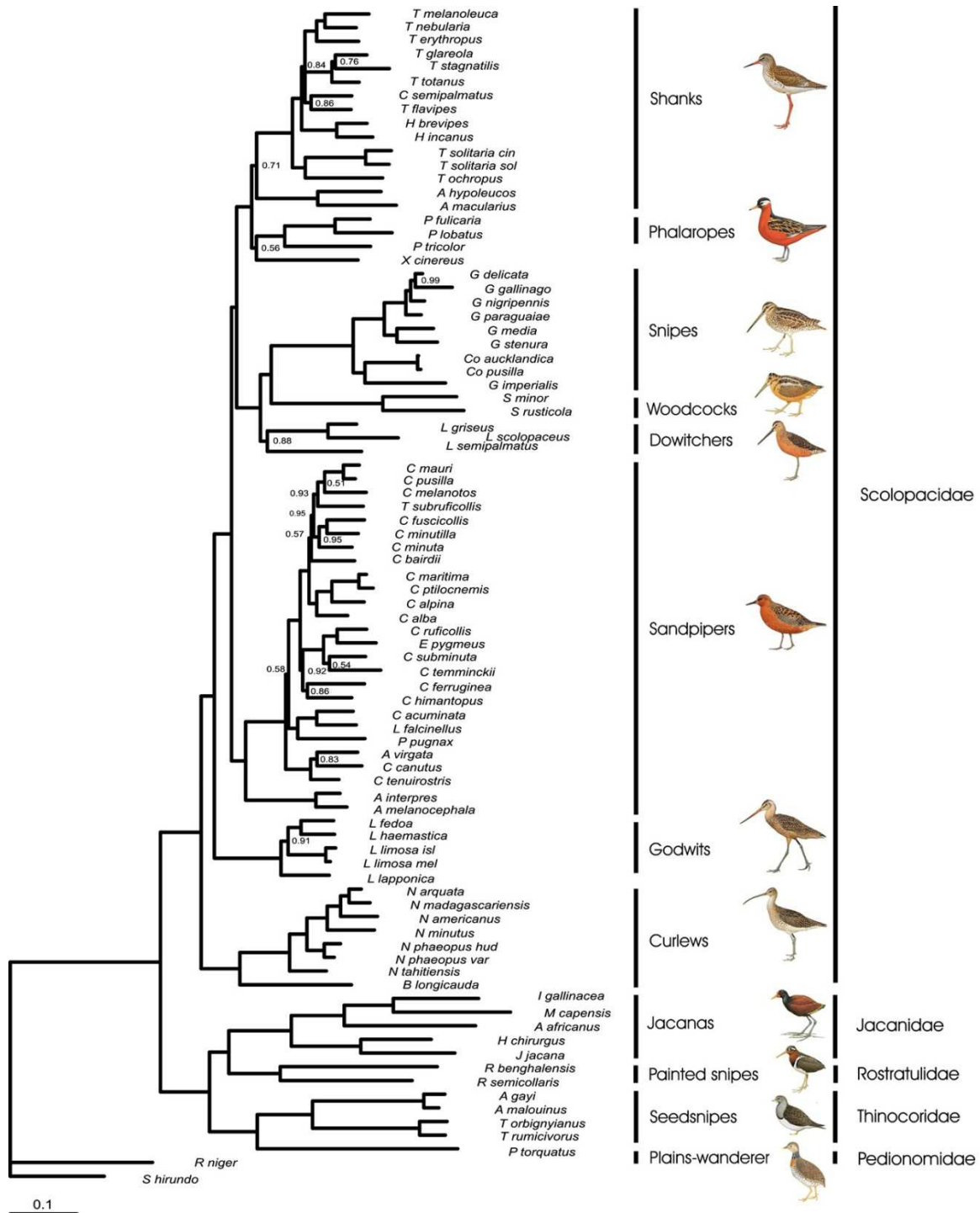
## Discusión

### Aspectos metodológicos

El porcentaje de éxito en la obtención de amplicones disminuyó de acuerdo a la distancia filogenética entre individuos debido a una alta especificidad de los cebadores. Por ejemplo, se obtuvieron las secuencias del Amplicón  $RC^1-Pro^2$  (LI41-PropR) para *C. minutilla*, *C. alpina*, *C. canutus*, *C. himantopus*, *Tringa semipalmata* y *Tringa flavipes*. Hubo una baja calidad de las secuencias en *Limnodromus griseus*, *Limnodromus scolopaceus* y *Limosa fedoa*.

El amplicón  $RC^1-Pro^2$  también se intentó amplificar en *Himantopus mexicanus*, *Tringa flavipes*, *Tringa semipalmata*, *Charadrius semipalmatus*, *Charadrius wilsonia*, *Charadrius vociferus*, *haematopus palliatus*, *Thalasseus maximus* y *Leucophaeus atricilla*, pero no hubo efectividad en la amplificación.

Suponemos que esto puede deberse a dos razones. La primera es que los cebadores al haber sido originalmente diseñados para *Philomachus pugnax*, las especies más divergentes a esta especie (Figura 17), tendrán una mayor diferencia nucleotídica y los sitios de anclaje de los cebadores serán cada vez más inespecíficos conforme haya presencia de cambios en las bases donde los cebadores hibridizan. La segunda razón es que la inefectividad en la obtención del amplicón se puede deber a la ausencia de duplicación y por lo tanto no se logra un anclaje de los cebadores.



**Figura 17** Filogenia basada en secuencias de cinco genes (*RAG1*, *CYT B*, *12S*, *ND2* y *COI*) estimados con el análisis bayesiano con particiones de 84 especies del suborden Scolopaci (Charadriiformes). Tomado de Gibson y Baker (2012).

Otra dificultad metodológica fue la presencia de repeticiones mononucleotídicas en el Dominio I de las RC's (Figura 10, alrededor del sitio 1881), así como en la región del ARNr 12S (Figura 10, sitios 4541-5020). Pese a que en la electroforesis en agarosa se observaron productos de PCR bien definidos y específicos, la presencia de estas repeticiones mononucleotídicas generó lecturas de baja calidad en estas regiones. Si bien la estrategia de diseñar nuevos cebadores en regiones más adelantadas en la secuencia o "primer walking" (Azhikina et al., 1993), funcionó para obtener la secuencia completa de *C. mauri* y *C. alpina*, no se obtuvo el mismo éxito para otras especies del género.

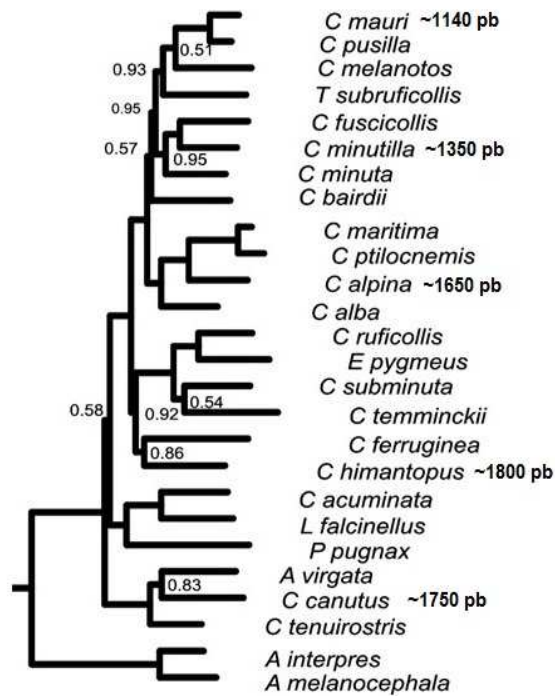
Adicionalmente, se presentó dificultad en el proceso de secuenciación cuando los productos de PCR incluían al Dominio III de la RC<sup>2</sup> debido a la presencia de repeticiones de Tetranucleótidos CAAA (en el sitio de las 4010 pb), de igual manera durante la secuenciación a partir de la zona CAAA la secuencia perdía calidad y generaba picos múltiples en el cromatograma, dando como resultado secuencias ambiguas que se volvieron a corroborar.

Para la comprobación en zonas conflictivas como la C-Stretch y repeticiones CAAA se amplificaron productos que incluyeran esas regiones y posteriormente se utilizó la técnica de "primer Walking". En esta técnica el primer cebador complementario a una porción conocida del ADN que se analizará se utiliza para la secuenciación de la región adyacente. La nueva información de la secuencia obtenida se utiliza luego para solicitar un nuevo cebador para su secuenciación posterior y secuenciar más allá de las regiones conocidas. La aplicación de este procedimiento varias veces, nos permitió conocer secuencias largas de ADN (ver Figura 15). Los mononucleótidos causan deslizamiento de la polimerasa generándose una secuencia ilegible, generando picos múltiples o secuencias truncadas. La obtención de nuevas secuencias por este método nos permitió ensamblar de manera más precisa las secuencias que se "ensuciaban" al alcanzar estas regiones conflictivas.

## **Ordenamiento de genes en el ADNmt en el género *Calidris*: Duplicación de la región control y los genes adyacentes**

La obtención de las secuencias del Amplicón RC<sup>1</sup>-Pro<sup>2</sup> en las cinco especies de *Calidris* analizadas indicaron diferencias en longitud y peso (Figura 8). El segmento de los genes que conforman la región *Cytb*-RC<sup>2</sup> (RC<sup>1</sup>, *Cytb*<sup>2</sup>, *Thr*<sup>2</sup> y *Pro*<sup>2</sup>) se originó por un evento de duplicación en tándem pero de manera parcial, una evidencia que soporta esta idea es la presencia de un pseudogén del *Cytb* en las cinco especies de *Calidris*. La extensión de la secuencia parcial del *Cytb* ( $\psi$ *Cytb*) duplicada difirió considerablemente entre estas especies y es variable de tal manera que existe una diferencia notable en la longitud de ese segmento duplicado entre estas especies.

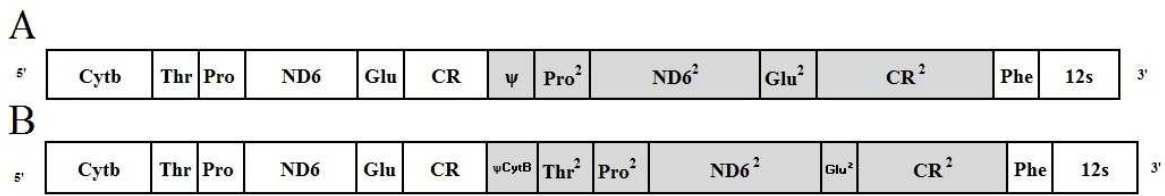
Al contrastar la posición filogenética de las especies del género *Calidris* con respecto a las diferencias en la longitud de la secuencia del  $\psi$ *Cytb*, aparentemente estas últimas podrían explicarse en términos de una degeneración del pseudogén en las ramas más recientes del árbol (Figura 18). A mayor profundidad de las ramas (tiempo de divergencia), mayor longitud del pseudogén *Cytb*.



**Figura 18** El árbol filogenético muestra la agrupación del género *Calidris* junto con las especies más cercanas a este género en la Familia Scolopacidae (tomado de Gibson y Baker (2012) ). Se indica la longitud aproximada del amplicón  $RC^1$ - $Pro^2$  para las cinco especies de Calidridos de las que se obtuvo esta secuencia

Verkuil y Colaboradores (2010) reportaron un nuevo acomodo en *Philomachus pugnax* (**Figura 19A**), hasta ese año ninguna otra Ave compartía este acomodo de genes. Sin embargo, nuestros resultados demuestran que al menos dentro de la Familia Scolopacidae este mismo acomodo es compartido. En la **Figura 19** se pueden observar ligeras diferencias en el acomodo de los genes entre *P. pugnax* y el acomodo encontrado en los Calidridos analizados. Con base a las secuencias obtenidas, y observando estos acomodos, se sugiere que en *P. pugnax* la secuencia no codificante que está flanqueada por la  $RC^1$  y  $Pro^2$  tiene un grado de degeneración mayor que los Calidridos. Por lo tanto *P. pugnax* ya no posee las secuencias correspondientes a  $\psi Cytb$  y  $Thr^2$ . Sin embargo, al observar el árbol filogenético de Scolopacidae se puede sugerir que en el nodo que separa a *Arenaria interpres* de los Calidridos y a *P. pugnax* debió existir un evento de duplicación el cual dio origen al acomodo de los genes que presentan los acomodos descritos en la **Figura 19**.

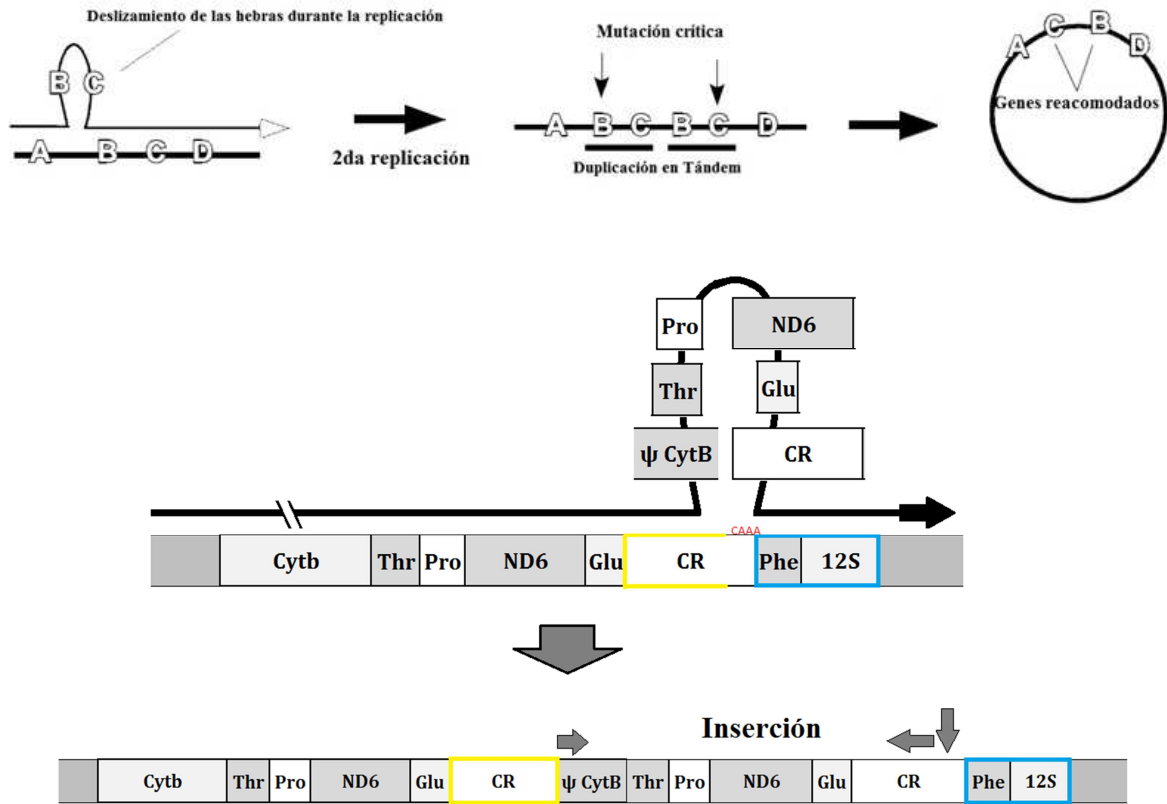
Con base a los resultados obtenidos se puede sugerir que las especies descendientes al nodo que alberga a *P. pugnax* y a los Calidridos comparten el mismo acomodo pero existe una variación en la longitud de  $\psi$ Cytb y Thr<sup>2</sup>. Para darle soporte a esta idea, hace falta analizar a cada una de las especies en esta rama del árbol ya que actualmente no existen secuencias completas de la sección Cytb-12S en el GenBank para las demás especies.



**Figura 19** Acomodo de los genes en la region Cytb-12S. La figura **A** muestra el acomodo de los genes en *P. pugnax*. La figura **B** muestra el acomodo encontrado en Calidridos.

## **Duplicación de la Región Control en *Calidris mauri* y otras especies del género**

La presencia de duplicación en tándem y el pseudogén de *Cytb* en las especies del género *Calidris* sugiere que los ordenamientos mitocondriales se originaron por un mecanismo correspondiente al modelo de duplicación y pérdida aleatoria en tándem (TDRL Zhou et al. (2014)). El ordenamiento *tRNA<sup>Thr</sup>-RC* presente en estas especies pudo haberse originado a partir del ordenamiento aviar ancestral por una duplicación que abarcó desde la parte final del *Cytb* hasta la *RC*, seguida de una delección de la porción 3' de la *RCI* y de la porción 5' del *Cytb* (Figura 20).



**Figura 20** Duplicación en tándem mediada por errores de replicación y pérdida aleatoria (modelo TDRL; (Boore, 2000)(Tomada de Kurabayashi et al. (2008)). El esquema superior muestra de manera simplificada el mecanismo por el cual se duplican varios genes mediante el modelo de TDRL, en este caso por el deslizamiento y mal complementación de las cadenas (slipped-strand mispairing). En el esquema inferior se muestra el acomodo final del segmento duplicado Cytb--CR en *Calidris mauri* conforme al modelo TDRL.

El análisis de las secuencias de ambas copias de la RC en *Calidris mauri* indica que ambas poseen los sitios característicos de una RC funcional (Figura 6) y sugiere que la replicación se pudo haber originado a partir de cualquiera de ellas. Se ha sugerido que el contar con dos copias funcionales podría representar una ventaja al funcionar como una protección contra el estrés oxidativo o la edad de la mitocondria (Schirtzinger et al., 2012), o como un mecanismo de replicación del ADNmt más rápida y efectiva ((Arndt y Smith (1998); Campbell y Barker (1999); Shao et al. (2005); Kumazawa et al. (2014)).

Las copias de la RC's resultaron casi idénticas en cada individuo, siendo la principal diferencia entre ambas una región de repeticiones CAAA en el extremo 3' de la RC<sup>2</sup> (Figura 6 y Figura 13). El análisis detallado de la secuencia de las RC's (Figura 14) reveló que, al menos en *C. mauri*, las variaciones entre individuos se encuentran concentradas en el extremo de los dominios I y III de la RC (Figura 6).

Las diferencias en el dominio II de las copias de la RC de cada individuo resultaron ser mínimas, por lo que es posible que la duplicación haya ocurrido recientemente en el tiempo evolutivo y que no se hayan acumulado suficientes sustituciones nucleotídicas para distinguir cada copia, o bien que las copias hayan evolucionado de manera concertada, lo que explicaría la acumulación de bloques CAAA en el extremo 3'. Este último escenario implicaría una conversión homóloga que podría introducir un sesgo en la estimación del número de sustituciones en la copia.

De acuerdo con las observaciones de Eberhard y colaboradores (2015) en su análisis del Orden Psittaciformes, si una de las RC duplicadas muestra signos de degeneración ésta es siempre la RC<sup>2</sup>. Sin embargo, reportaron que en Psittaciformes la RC<sup>2</sup> es la original pero esta se encuentra ubicada en la posición ancestral, situada entre *Glu* y *Phe* (Figura 4D); dato que les pareció sorprendente debido a que dicha region control (RC<sup>2</sup>) mostró signos de degeneración.

De la secuencia completa *Cytb-12S* de *Calidris mauri* y el ordenamiento obtenido en el presente estudio, contrario a lo planteado por Eberhard y Wright (2016), en el género *Calidris* la  $RC^I$  que se encuentra situada más al extremo 5' es la region control original. Tomando como base el modelo TDRL; la mala complementación y el deslizamiento de las cadenas provocaron la duplicación del extremo 5' del *Cytb* hasta la RC original; una vez copiado este segmento se insertó justo antes de las repeticiones CAAA al final del dominio III de la RC desplazando a la RC original y el  $tRNA^{Phe}$  hacia el extremo 3'.

Sumado a esto, los eventos de conversión de génica han mantenido las RC's en los dominios I y II casi idénticas. La RC situada al extremo 5' está flanqueada por el  $tRNA^{Glu}$  y la copia degenerada del Citocromo b ( $\psi Cytb$ ) indica evidencia que soporta este escenario (**Figura 20**).

Anteriormente se mencionó que se había reportado un ordenamiento de tipo ancestral en el ADNmt de *Arenaria interpres* (Paton et al., 2002), (Figura 4A); por lo consiguiente *A. interpres* tiene sólo una RC en su genoma mitocondrial. Sin embargo *A. interpres* posee repeticiones CAAA en el dominio III de su RC, similar a la  $RC^2$  de *C. mauri*, pero con un número de repeticiones distinta esta característica da soporte al mecanismo en que se generó la duplicación de genes en los Calidridos.

Por otra parte, al realizar un análisis de la RC de *Gallus gallus* se observó que en el Dominio III no presenta repeticiones CAAA; similar la  $RC^1$  de *C. mauri* la cual es la region control original. Si la ausencia de las repeticiones CAAA en el dominio III de la RC, corresponde al estado ancestral, tal como lo presenta *G. gallus*, da soporte a la idea de que la  $RC^1$  es la region control original, pero crea incertidumbre en el modo en que las repeticiones CAAA surgieron, ya que al menos en los *Charadriiformes* que se encuentran secuenciados mencionados anteriormente, todos poseen esta característica del dominio III.

## Mantenimiento y evolución de los genes duplicados

Las regiones de genes que se reorganizan en el ADNmt de aves también se codifican predominantemente por la cadena pesada y están cerca del origen de replicación de la misma. Es posible que cualquier combinación de estos mecanismos pudiera ser responsable del orden de los genes observados en las aves (Mueller y Boore, 2005).

Además, como se mencionó anteriormente el genoma mitocondrial está sujeto además a un segundo mecanismo: la “recombinación homóloga”, en donde las cadenas de ADNmt de dos porciones genómicas, con secuencia nucleotídica igual o similar entre cromosomas o entre moléculas de ADN son intercambiadas (Ujvari et al., 2007; Kurabayashi et al., 2008). Cuando la recombinación entre dos cadenas que contienen un set de genes o una secuencia nucleotídica similar este proceso no crea una duplicación, pero la recombinación puede homogenizar la secuencia entre las regiones intercambiadas conocida como conversión de genes (Figura 3). Por el contrario cuando alguna de las cadenas que se recombinan posee un set desigual de genes, una de las moléculas resultantes o de las porciones genómicas podría tener una copia genómica extra (unequal crossing over) (Kurabayashi et al., 2008; Sammler et al., 2011).

El análisis de divergencia genética relativa entre las RC's duplicadas (**Figura 14**) indicó que RC's parálogas de los diferentes individuos eran genéticamente más similares entre sí que las RC's ortólogas de los individuos dentro de las especies analizadas. Este patrón podría sugerir que estas copias tienen un origen reciente o que se mantienen idénticas gracias al proceso de evolución concertada y recombinación homóloga entre los genes duplicados en las especies de Calidridos analizados.

La suposición de que el ADN mitocondrial de los animales no sufre recombinación homóloga está basado en pruebas indirectas. Sin embargo, esta suposición ha tenido una influencia importante en nuestra comprensión de la reparación de ADN mitocondrial, la acumulación de mutaciones e inferencias biohistóricas realizadas a partir de datos poblacionales. Si la recombinación homóloga ocurre en el ADNmt animal, esto tendrá un efecto importante en nuestra comprensión de los mecanismos de mutación, reparación y tasas de acumulación de mutaciones en el ADNmt

(Ladoukakis y Zouros, 2001). Las implicaciones son igualmente graves para el uso de ADN mitocondrial en los estudios de filogenia, filogeografía y fechado en donde se asume que el genoma mitocondrial es transmitido vía clonal.

Estudios de simulación realizados por Schierup y Hein (2000) y Abbott et al. (2005) demostraron que cuando se asume que no existe recombinación cuando si la hay, conlleva a ramas basales más largas y ramas terminales cortas que en una situación real; y podría también aparentar una violación al reloj molecular. Para esto Ladoukakis y Zouros (2001) sugieren además otras dos complicaciones que podrían ser de especial importancia en la reconstrucción de filogenias de taxones estrechamente relacionados pero reproductivamente aislados. Una de ellas es la posibilidad de evolución concertada en el ADNmt; y la segunda complicación, puede resultar por la variación (ya sea entre individuos o entre las poblaciones) en la actividad de recombinación de la mitocondria, que debe ser controlada por genes nucleares. Todo esto tendrá un gran efecto de ruido alrededor de la señal filogenética.

Las RC's son ampliamente utilizadas como marcador molecular en análisis de genética de poblaciones, filogeografía y filogenética. Sin embargo, en los taxones en donde se presentan duplicación de la RC se debe examinar cuidadosamente para el uso de la misma en cualquier tipo de análisis, debido a la evolución concertada (Wang et al., 2015). Para poder encontrar algún patrón, bajo el cual el ADNmt de las Aves está evolucionando se necesita una mayor cobertura taxonómica para así hacer comparaciones en los acomodos de los genes dentro y entre taxones de Aves secuenciados.

Con las nuevas herramientas de secuenciación de nueva generación (NGS) se pueden obtener las secuencias completas de los genomas mitocondriales en Aves. Este recurso será de utilidad en estudios evolutivos comparativos. Los resultados de la NGS pueden proveer aceleradamente colecciones de genes y genomas mitocondriales de aves, ya que es una herramienta adecuada para la generación de datos debido a que estos genomas evolucionan rápidamente (Avise et al., 1987; Cooke et al., 2012).

Lounsberry y colaboradores (2015) reportaron el diseño del flujo de trabajo para la secuenciación del mitogenoma completo de Albatros (*Phoebastria* spp.) con el sistema secuenciación de nueva generación “miSeq” de Illumina. En este trabajo se emplea un procedimiento iterativo que permite el ensamblaje de novo. Además es posible detectar duplicaciones y determinar el acomodo de los genes. Este procedimiento podría ser aplicado para cualquier otro taxón de Aves. Sin embargo, el ensamblaje de mitogenomas de especies no modelo podrían presentar desafíos considerables. Los resultados de las secuencias obtenidas a partir de la NGS podrían presentar complicaciones debido a la presencia de regiones repetitivas tal como sucede en las RC's. Además la presencia de genes codificantes duplicados, estas secuencias al ser básicamente idénticas, generarían traslapes en los alineamientos entre las copias. Estos genes duplicados podrían pasar desapercibidos si no se encontraran adyacentes a algún otro gen no duplicado. Estas secuencias se podrían ensamblar en un sitio único de manera inapropiada enmascarando la presencia de uno o más genes con esta característica.

Nuevamente, la recomendación de utilizar el amplicón “RC<sup>1</sup>-Pro<sup>2</sup>” a manera de diagnóstico, permitiría tomar algunas medidas para evitar este error. Adicionalmente, la información obtenida de los genomas secuenciados nos podría sugerir el origen y la relación de los acomodos encontrados hasta ahora, establecer con mayor confianza las relaciones filogenéticas en Aves e incluso descubrir nuevos acomodos presentes.

## Conclusiones

En esta investigación se secuenció la sección completa del genoma mitocondrial de *Calidris mauri* que comprende desde el gen que codifica para *Cytb* hasta el *12S* ribosomal, encontramos que presentó dos copias de la sección Thr-RC (*Cytb/Thr/Pro/ND6/Glu/CR*), separadas por una copia degenerada *Cytb* que difieren en el extremo 3' de la RC<sub>2</sub> por varias repeticiones CAAA.

*Calidris mauri* presenta el ordenamiento genético que muestra la **Figura 12**. Este mismo acomodo de genes fue reportado en *Philomachus pugnax* por Verkuil et al. (2010) y en *Phoebastria albatrus* por Lounsberry et al. (2015).

Se obtuvieron secuencias parciales de *Calidris alpina*, *Calidris himantopus*, *Calidris minutilla*, *Calidris canutus*, *Tringa flavipes*, *Tringa semipalmata*, *Limnodromus griseus*, *Limnodromus scolopaceus* y *Limosa fedoa* que sugieren la presencia de genes duplicados.

Las secuencias parciales de *Calidris alpina*, *Calidris himantopus*, *Calidris minutilla* y *Calidris canutus* sugieren que estas especies comparten un ordenamiento de genes similar al de *Calidris mauri*.

No se obtuvieron secuencias de *Charadrius nivosus*, *Charadrius semipalmatus*, *Charadrius wilsonia*, *Charadrius vociferus*, *Himantopus mexicanus*, *haematopus palliatus*, *Thalasseus maximus* y *Leucophaeus atricilla*. Con lo cual no se pudo determinar si presentaban duplicación de alguno de sus genes. La distancia filogenética aumenta la variabilidad nucleotídica entre las especies, por lo tanto la efectividad de los cebadores utilizados disminuyó.

Se observó que el mecanismo por el cual se generó la duplicación de los genes en *Calidris mauri* es en tándem generada por el deslizamiento y mal complementación de las cadenas de acuerdo al modelo propuesto por Boore (2000), (TDRL model).

Las RC duplicadas se han mantenido prácticamente idénticas por dos posibles razones. Primero es que la antigüedad de las duplicaciones es relativamente reciente, por lo tanto no se han podido observar un número relevante de sustituciones. Por otro lado las regiones duplicadas están bajo el proceso de la conversión de genes.

Debido al análisis entre individuos de la Familia Scolopacidae, y al porcentaje de las distancias genéticas, observamos que las RC's dentro de cada individuo evolucionan de manera concertada.

De acuerdo al modelo propuesto por Boore (2000), la RC situada hacia el extremo 5', entre el *tRNA<sup>Glu</sup>* y *ψCytb* es la región control original.

## **RECOMENDACIONES**

Se evidenció que hay duplicación de región control, los cebadores empleados podrían amplificar ambas copias, por lo tanto se propone una revisión de los análisis filogenético del género *Calidris* que se han efectuado hasta la fecha.

Para ello se debería emplear el amplicón RC<sup>1</sup>-Pro<sup>2</sup> (L141-PropR), ya que este nos permite discriminar al menos en *Calidris* la RC's que de acuerdo a los resultados obtenidos en este trabajo, corresponde a la RC<sup>1</sup> ancestral.

## Referencias

- Abbott, C.L., Double, M.C., Trueman, J.W., Robinson, A. y Cockburn, A. (2005). An unusual source of apparent mitochondrial heteroplasmy: Duplicate mitochondrial control regions in thalassarche albatrosses. *Molecular Ecology* 14, 3605-3613.
- Aljanabi, S.M. y Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for pcr-based techniques. *Nucleic Acids Research* 25, 4692-4693.
- Anderson, S., Bankier, A.T., Barrell, B.G., De Bruijn, M., Coulson, A.R., Drouin, J., Eperon, I., Nierlich, D., Roe, B.A. y Sanger, F. (1981). Sequence and organization of the human mitochondrial genome. *Nature* 290, 457-465.
- Arndt, A. y Smith, M. (1998). Mitochondrial gene rearrangement in the sea cucumber genus cucumaria. *Molecular biology and evolution* 15, 1009-1016.
- Avise, J.C. (1989). Gene trees and organismal histories: A phylogenetic approach to population biology. *Evolution*, 1192-1208.
- Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. y Saunders, N.C. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual review of ecology and systematics*, 489-522.
- Azhikina, T., Veselovskaya, S., Myasnikov, V., Potapov, V., Ermolayeva, O. y Sverdlov, E. (1993). Strings of contiguous modified pentanucleotides with increased DNA-binding affinity can be used for DNA sequencing by primer walking. *Proceedings of the National Academy of Sciences* 90, 11460-11462.
- Bailey, J.A., Liu, G. y Eichler, E.E. (2003). An alu transposition model for the origin and expansion of human segmental duplications. *The American Journal of Human Genetics* 73, 823-834.
- Baker, A. y Marshall, H. (1997). Mitochondrial control region sequences as tools for understanding evolution. *Avian molecular evolution and systematics*, 51-82.
- Baker, A.J., Pereira, S.L. y Paton, T.A. (2007). Phylogenetic relationships and divergence times of charadriiformes genera: Multigene evidence for the cretaceous origin of at least 14 clades of shorebirds. *Biology Letters* 3, 205-210.
- Bensch, S. y Anna, H. (2000). Mitochondrial genomic rearrangements in songbirds. *Molecular biology and evolution* 17, 107-113.
- Björklund, M. (1994). Phylogenetic relationships among charadriiformes: Reanalysis of previous data. *The Auk*, 825-832.
- Boore, J.L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research* 27, 1767-1780.
- Boore, J.L. (2000) The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals. En: *Comparative genomics*, Pp. 133-147. Springer.
- Braband, A., Podsiadlowski, L., Cameron, S.L., Daniels, S. y Mayer, G. (2010). Extensive duplication events account for multiple control regions and pseudo-genes in the mitochondrial genome of the velvet worm *metaperipatus inae* (onychophora, peripatopsidae). *Molecular phylogenetics and evolution* 57, 293-300.
- Brown, G.G., Gadaleta, G., Pepe, G., Saccone, C. y Sbisà, E. (1986). Structural conservation and variation in the d-loop-containing region of vertebrate mitochondrial DNA. *Journal of molecular biology* 192, 503-511.
- Buehler, D.M. y Baker, A.J. (2005). Population divergence times and historical demography in red knots and dunlins. *The Condor* 107, 497-513.
- Buehler, D.M., Baker, A.J. y Piersma, T. (2006). Reconstructing palaeoflyways of the late pleistocene and early holocene red knot *calidris canutus*. *Ardea-Wageningen* 94, 485.

- Campbell, N. y Barker, S.C. (1999). The novel mitochondrial gene arrangement of the cattle tick, *Boophilus microplus*: Fivefold tandem repetition of a coding region. *Molecular biology and evolution* 16, 732-740.
- Clayton, D.A. (1984). Transcription of the mammalian mitochondrial genome. *Annual review of biochemistry* 53, 573-594.
- Cooke, G.M., King, A.G., Johnson, R.N., Boles, W.E. y Major, R.E. (2012). Rapid characterization of mitochondrial genome rearrangements in Australian songbirds using next-generation sequencing technology. *Journal of Heredity* 103, 882-886.
- Chang, D.D. y Clayton, D.A. (1986). Identification of primary transcriptional start sites of mouse mitochondrial DNA: Accurate in vitro initiation of both heavy- and light-strand transcripts. *Molecular and cellular biology* 6, 1446-1453.
- Christian, P., Christidis, L. y Schodde, R. (1992). Biochemical systematics of the Charadriiformes (shorebirds)-relationships between the Charadrii, Scolopaci and Lari. *Australian Journal of Zoology* 40, 291-302.
- Desjardins, P. y Morais, R. (1990). Sequence and gene organization of the chicken mitochondrial genome: A novel gene order in higher vertebrates. *Journal of molecular biology* 212, 599-634.
- Drummond, A., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas, S., Thierer, T. y Wilson, A. (2009) Geneious v4. 7.
- Eberhard, J.R. y Wright, T.F. (2016). Rearrangement and evolution of mitochondrial genomes in parrots. *Molecular phylogenetics and evolution* 94, 34-46.
- Eberhard, J.R., Wright, T.F. y Bermingham, E. (2001). Duplication and concerted evolution of the mitochondrial control region in the parrot genus Amazona. *Molecular biology and evolution* 18, 1330-1342.
- Ericson, P.G., Envall, I., Irestedt, M. y Norman, J.A. (2003). Inter-familial relationships of the shorebirds (Aves: Charadriiformes) based on nuclear DNA sequence data. *BMC evolutionary biology* 3, 16.
- Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.-I. y Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151, 1531-1545.
- Fujita, M.K., Boore, J.L. y Moritz, C. (2007). Multiple origins and rapid evolution of duplicated mitochondrial genes in parthenogenetic geckos (Heteronotia binoei; Squamata, Gekkonidae). *Molecular biology and evolution* 24, 2775-2786.
- Gibson, R. y Baker, A. (2012). Multiple gene sequences resolve phylogenetic relationships in the shorebird suborder Scolopaci (Aves: Charadriiformes). *Molecular phylogenetics and evolution* 64, 66-72.
- Gyllenstein, U., Wharton, D., Josefsson, A. y Wilson, A.C. (1991). Paternal inheritance of mitochondrial DNA in mice.
- Hoelzel, A.R. (1992). *Molecular genetic analysis of populations: A practical approach* Irl Press,
- Hurles, M. (2004). Gene duplication: The genomic trade in spare parts. *PLoS Biol* 2, e206.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution* 16, 111-120.
- Koonin, E.V. (2005). Orthologs, paralogs, and evolutionary genomics 1. *Annu. Rev. Genet.* 39, 309-338.
- Kumazawa, Y., Miura, S., Yamada, C. y Hashiguchi, Y. (2014). Gene rearrangements in gekkonid mitochondrial genomes with shuffling, loss, and reassignment of trna genes. *BMC genomics* 15, 930.
- Kumazawa, Y., Ota, H., Nishida, M. y Ozawa, T. (1998). The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. *Genetics* 150, 313-329.
- Kurabayashi, A., Sumida, M., Yonekawa, H., Glaw, F., Vences, M. y Hasegawa, M. (2008). Phylogeny, recombination, and mechanisms of stepwise mitochondrial genome

- reorganization in mantellid frogs from madagascar. *Molecular biology and evolution* 25, 874-891.
- Ladoukakis, E.D. y Zouros, E. (2001). Direct evidence for homologous recombination in mussel (*mytilus galloprovincialis*) mitochondrial DNA. *Molecular biology and evolution* 18, 1168-1175.
- Lounsberry, Z., Brown, S., Collins, P., Henry, R., Newsome, S. y Sacks, B. (2015). Next-generation sequencing workflow for assembly of nonmodel mitogenomes exemplified with north pacific albatrosses (*phoebastria* spp.). *Molecular ecology resources*.
- Lynch, M., O'Hely, M., Walsh, B. y Force, A. (2001). The probability of preservation of a newly arisen gene duplicate. *Genetics* 159, 1789-1804.
- Mindell, D.P., Sorenson, M.D. y Dimcheff, D.E. (1998). Multiple independent origins of mitochondrial gene order in birds. *Proceedings of the National Academy of Sciences* 95, 10693-10697.
- Moritz, C., Dowling, T. y Brown, W. (1987). Evolution of animal mitochondrial DNA: Relevance for population biology and systematics. *Annual review of ecology and systematics*, 269-292.
- Mueller, R.L. y Boore, J.L. (2005). Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders. *Molecular biology and evolution* 22, 2104-2112.
- Ogoh, K. y Ohmiya, Y. (2007). Concerted evolution of duplicated control regions within an ostracod mitochondrial genome. *Molecular biology and evolution* 24, 74-78.
- Pakendorf, B. y Stoneking, M. (2005). Mitochondrial DNA and human evolution. *Annu. Rev. Genomics Hum. Genet.* 6, 165-183.
- Paton, T., Haddrath, O. y Baker, A.J. (2002). Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. *Proceedings of the Royal Society of London B: Biological Sciences* 269, 839-846.
- Pereira, S.L. (2000). Mitochondrial genome organization and vertebrate phylogenetics. *Genetics and Molecular Biology* 23, 745-752.
- Rastogi, S. y Liberles, D.A. (2005). Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC evolutionary biology* 5, 28.
- Ruokonen, M. y Kvist, L. (2002). Structure and evolution of the avian mitochondrial control region. *Molecular phylogenetics and evolution* 23, 422-432.
- Saccone, C., Pesole, G. y Sbisá, E. (1991). The main regulatory region of mammalian mitochondrial DNA: Structure-function model and evolutionary pattern. *Journal of molecular evolution* 33, 83-91.
- Sammler, S., Bleidorn, C. y Tiedemann, R. (2011). Full mitochondrial genome sequences of two endemic philippine hornbill species (aves: Bucerotidae) provide evidence for pervasive mitochondrial DNA recombination. *BMC genomics* 12, 35.
- San Mauro, D., Gower, D.J., Zardoya, R. y Wilkinson, M. (2006). A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. *Molecular biology and evolution* 23, 227-234.
- Schierup, M.H. y Hein, J. (2000). Recombination and the molecular clock. *Molecular biology and evolution* 17, 1578-1579.
- Schirtzinger, E.E., Tavares, E.S., Gonzales, L.A., Eberhard, J.R., Miyaki, C.Y., Sanchez, J.J., Hernandez, A., Müeller, H., Graves, G.R. y Fleischer, R.C. (2012). Multiple independent origins of mitochondrial control region duplications in the order psittaciformes. *Molecular phylogenetics and evolution* 64, 342-356.
- Shao, R., Barker, S.C., Mitani, H., Aoki, Y. y Fukunaga, M. (2005). Evolution of duplicate control regions in the mitochondrial genomes of metazoa: A case study with australasian ixodes ticks. *Molecular biology and evolution* 22, 620-629.
- Sibley, C.G., Ahlquist, J.E. y Monroe Jr, B.L. (1988). A classification of the living birds of the world based on DNA-DNA hybridization studies. *The Auk*, 409-423.

- Strauch, J.G. (1978). The phylogeny of the charadriiformes (aves): A new estimate using the method of character compatibility analysis. *The Transactions of the Zoological Society of London* 34, 263-345.
- Studer, R.A. y Robinson-Rechavi, M. (2009). How confident can we be that orthologs are similar, but paralogs differ? *Trends in Genetics* 25, 210-216.
- Tamura, K., Dudley, J., Nei, M. y Kumar, S. (2007). Mega4: Molecular evolutionary genetics analysis (mega) software version 4.0. *Molecular biology and evolution* 24, 1596-1599.
- Tatarenkov, A. y Avise, J.C. (2007). Rapid concerted evolution in animal mitochondrial DNA. *Proceedings of the Royal Society of London B: Biological Sciences* 274, 1795-1798.
- Tayebi, N., Stubblefield, B.K., Park, J.K., Orvisky, E., Walker, J.M., LaMarca, M.E. y Sidransky, E. (2003). Reciprocal and nonreciprocal recombination at the glucocerebrosidase gene region: Implications for complexity in gaucher disease. *The American Journal of Human Genetics* 72, 519-534.
- Ujvari, B., Dowton, M. y Madsen, T. (2007). Mitochondrial DNA recombination in a free-ranging australian lizard. *Biology Letters* 3, 189-192.
- Van Tuinen, M., Waterhouse, D. y J Dyke, G. (2004). Avian molecular systematics on the rebound: A fresh look at modern shorebird phylogenetic relationships. *Journal of Avian Biology* 35, 191-194.
- Verkuil, Y.I., Piersma, T. y Baker, A.J. (2010). A novel mitochondrial gene order in shorebirds (scolopacidae, charadriiformes). *Molecular phylogenetics and evolution* 57, 411-416.
- Wang, X., Huang, Y., Liu, N., Yang, J. y Lei, F. (2015). Seven complete mitochondrial genome sequences of bushtits (passeriformes, aegithalidae, aegithalos): The evolution pattern in duplicated control regions. *Mitochondrial DNA*, 1-7.
- Wenink, P.W., Baker, A.J. y Tilanus, M. (1994). Mitochondrial control-region sequences in two shorebird species, the turnstone and the dunlin, and their utility in population genetic studies. *Molecular biology and evolution* 11, 22-31.
- Wyman, S.K., Jansen, R.K. y Boore, J.L. (2004). Automatic annotation of organellar genomes with dogma. *Bioinformatics* 20, 3252-3255.
- Zhang, J. (2003). Evolution by gene duplication: An update. *Trends in ecology & evolution* 18, 292-298.
- Zheng, C., Nie, L., Wang, J., Zhou, H., Hou, H., Wang, H. y Liu, J. (2013). Recombination and evolution of duplicate control regions in the mitochondrial genome of the asian big-headed turtle, platysternon megacephalum.
- Zhou, X., Lin, Q., Fang, W. y Chen, X. (2014). The complete mitochondrial genomes of sixteen ardeid birds revealing the evolutionary process of the gene rearrangements. *BMC genomics* 15, 573.

## Lista de Anexos

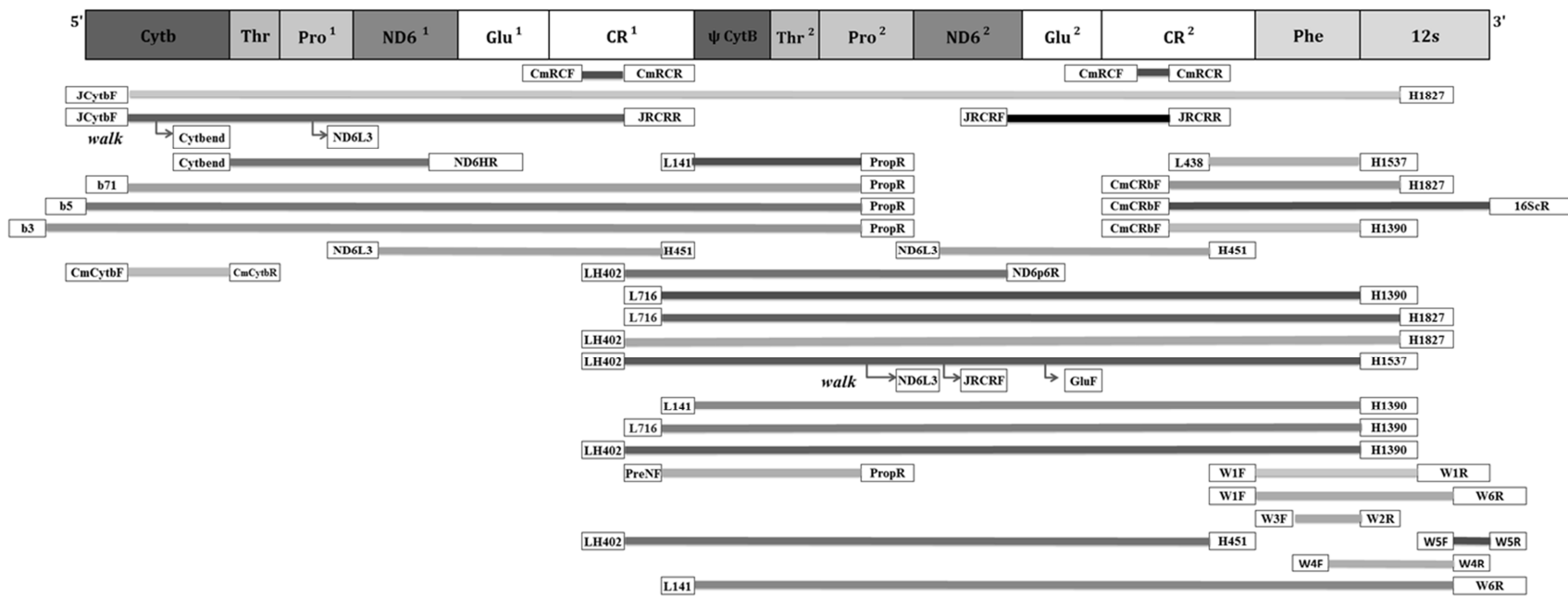
### Anexo 1

Listado de cebadores empleados en este estudio utilizados para la amplificación de ADN mitocondrial de especies del género *Calidris* y otros escolopácidos. Se indica el nombre del cebador, la secuencia específica y el gen correspondiente

Región	Cebador	Secuencia (5' → 3')	Fuente
Citocromo b	JCytbF	TGGCCCCAAACCTACGAAAA	Éste estudio
	JCytbR	CTGGCTGCCCACTCATGTTA	Éste estudio
	CmCytb F	ACAACCTAGCCTTCTCAT	Éste estudio
	CmCytbR	CCTAGTTTGTGGGATTG	Éste estudio
	b3	GGACGAGGCTTTACTACGGCTC	Tomado de Verkuil et al. (2010)
	b86	TGAATNGGNAGCCARCCNGTAGAACACCC	Tomado de Verkuil et al. (2010)
	b71	TGGGAGGAGTACTAGCTCTGGCAGCCTC	Tomado de Verkuil et al. (2010)
	b5	TTCCACCCCTACTTCTACTAAAGA	Tomado de Verkuil et al. (2010)
	cytbend	CACGACGTTGTAACGACAGCCAACCYAGAACACCC	Tomado de Verkuil et al. (2010)
	PropR	AATACCAGCTTTGGGAGTTGG	Tomado de Verkuil et al. (2010)
ND6p2R	ND6p2R	TGGTTACTGTTGACAGTGGGGGA	Tomado de Verkuil et al. (2010)
	ND6p6R	CTGTTGCAGGATGTGGTTGGCT	Tomado de Verkuil et al. (2010)
tRNAPro	ND6L3	ACTGCTCGAATCGCCCCACGAG	Tomado de Verkuil et al. (2010)
tRNA <sup>Glu</sup>	cmGluF	TAGCAGTTCTGCTTGGCTTT	Éste estudio
	cmGluR	GCGGGCAGGTGTATGTTCC	Éste estudio
ND6	ND6HR	TYCGNATRGATTTTAGYGGGGT	Tomado de Verkuil et al. (2010)
	H451	CCTGAAGCTAGTAACGCAGGAC	Tomado de Verkuil et al. (2010)
	L141	TCCATTAATCTACAACCGGGCT	Tomado de Verkuil et al. (2010)
	L/H402	TGAAATCAGCAACCCGGCGTAAG	Tomado de Verkuil et al. (2010)
Región Control	L438	TCACGTGAAATCAGCAACCC	Wenink et al., 1993
	L716	ACTTTGGCCCTCAGGCGTACTG	Tomado de Verkuil et al. (2010)
	JRCRR	AGGCCAAAATAAGGACGCA	Éste estudio
	preNF	CGCCCCATAAAATACTTGCG	Éste estudio
	preNR	AATTTGATGATGGGTTACTAATGT	Éste estudio
	cmaRCF	CCCCATACTACATACCATC	Éste estudio
	cmCRbF	CAACCGGGCTATACACCTCTCCAC	Éste estudio
	cmRCR	GTCCACAAGCATTTCATT	Éste estudio
tRNA <sup>Phe</sup>	mtScoW1F	TGCCAATAAAAGAACTCTC	Éste estudio
	mtScoW3F	TGGCACTGAAGATGCCAAGA	Éste estudio
12SrRNA	mtScoW4F	AGATGCCAAGATGGCTGTCA	Éste estudio
	mtScoW1R	CTAAGCATAGTGGGGTATCT	Éste estudio
	mtScoW2R	TAATTTGGGTCTCTTGTGTG	Éste estudio
	mtScoW4R	AGGGCTAAGCATAGTGGGGT	Éste estudio
	mtScoW5F	TACCCCACTATGCTTAGCCCT	Éste estudio
	mtScoW5R	CTTATGAGGAGGGTGACGGG	Éste estudio
	mtScoW6R	TACTAAATCCGCCTTCGGGG	Éste estudio
	H1390	TCAAGGGCATTACACTGGGGCGCAGATAC	Tomado de Verkuil et al. (2010)
	JRCRF	AAAGGATCTGCCCAAAGA	Éste estudio
	H1537	TGACCGCGTGGCTGGCACAAG	Tomado de Verkuil et al. (2010)
H1827	GCACCGCCAAGTCTTAGAG	Tomado de Verkuil et al. (2010)	
16ScR	TTCTCCAAGGTCGCCCAACC	Tomado de Verkuil et al. (2010)	

## Anexo 2

Orden de los genes mitocondriales que flanquean la región control descrito para *P. pugnax* (*Cytb* - *12S ribosomal*). Los rectángulos en blanco tienen en el interior los códigos de los cebadores (*JCytbF*, *Cytbend*, *b71*, etc). Las líneas gruesas indican la longitud de los fragmentos y el par de cebadores que las generan. Las flechas indican amplificaciones utilizando la técnica de “primer walking” utilizando los amplicones largos como template. Obsérvese la inespecificidad de algunos de los cebadores que flanquean las regiones control.



Anexo 3

Programas de PCR empleados para la amplificación de los fragmentos del ADNmt para establecer el orden de los genes en la región que comprende desde el *Cytb* hasta el *12S*.

Región	Cebadores		Programa	Temp. °C	Longitud (pb)	Control de calidad
	Forward	Reverse				
<i>Cytb</i>	Cytb-end	ND6HR	2	55.1	1200	OK
<i>CR1-ND62</i>	LH402	ND6P2R	2	52.0	1000	OK
<i>CR1-ND62</i>	L716	ND6P2R	2	54.0	700	OK
<i>CR2-12S</i>	L438	H1537	2	54.0	500	OK
<i>Cytb</i>	CmCytb F	ND6HR	2	55.0	1200	OK
<i>RC's</i>	CmaRC F	CmRC R	2	52.0	800	OK
<i>Cytb</i>	CmCytbF	CmCytbR	1	52.0	800	OK
<i>ND6-CR</i>	ND6L3	H451	2	56.0	1000	OK
<i>CR1-ND62</i>	LH402	ND6P6R	2	55.0	1000	OK
<i>Cytb-Pro2</i>	b71	PropR	2	60.0	450	OK
<i>Cytb-Pro2</i>	b3	PropR	2	60.0	947	OK
<i>Cytb-Pro2</i>	b5	PropR	2	60.0	700	OK
<i>Cytb-Pro2</i>	b86	PropR	2	60.0	267	Secuencia corta
<i>CR-12S</i>	L716	H1390	2	55.0	571	OK
<i>CR-12S</i>	L716	H1827	2	55.0	800	OK
<i>CR-12S</i>	LH402	H1827	2	55.0	1100	OK
<i>CR-12S</i>	LH402	H1390	2	55.0	1100	Secuencia no útil
<i>RC1-Pro2</i>	L141	PropR	1	52.0	1200	OK
<i>CR-12S</i>	LH402	H1537	3	66.0	1000	Picos múltiples región CAAA
<i>cyt-CR1</i>	JCytbF	JRCRR	4	57.0	1014	Reverse no

						secuenciab le
<i>CR-16S</i>	CmCRbF	16ScR	4	62.0	Contig	OK
<i>CR-12S</i>	CmCRbF	H1827	4	62.0		
		H1390	4	62.0		
<i>CAAA-12S</i>	mtScoW3F	mtScoW2R	4	62.0	280	Banda ligera, debajo de las 500pb
<i>CAAA-12S</i>	mtScoW4F	mtScoW4R	4	62.0	460	OK
<i>CAAA-12S</i>	mtScoW5F	mtScoW5R	4	62.0	400	OK
<i>CR2-12s</i>	mtScoW1F	mtScoW1R	4	50.0	130	Regular
<i>CR2-12s</i>	mtScoW1F	mtScoW6R	4	50.0	800	Regular
<i>CR1-propR</i>	PreNF	PropR	4	64.0	457	Forward no secuenciab le
<i>CR-12S</i>	L141	H1390	4	67.0	250	OK
<i>CR-12S</i>	L141	H1537	3	60.0	450	OK
<i>CR-12S</i>	L141	mtScoW6R	4	56.0	460	OK
<i>CR1-CR2</i>	LH402	H451	5	58.0	Contig	Regular
<i>cytb-12s</i>	JCytbF	H1827	6	65.7	5000	Regular

Anexo 4

Programas empleados para la amplificación, así como las modificaciones a los perfiles de amplificación específicas para cada par de cebadores empleado en el presente estudio.

Configuración del Programa #1 "CALISDUP"

Número de Ciclos	Tiempo	Temperatura	
1	4 minutos	94 °C	Desnaturalización inicial
35	30 Segundos	94 °C	Desnaturalización
	30 Segundos	52 °C	Reincorporación
	45 segundos	70 °C	Extensión (Síntesis)
1	10 minutos	70 °C	Extensión final
1	∞	15 °C	Finalización

Configuración del Programa #2 "CALIDDUP"

Número de Ciclos	Tiempo	Temperatura	
1	4 minutos	94 °C	Desnaturalización inicial
35	1 minuto	94 °C	Desnaturalización
	1 minuto	52 °C	Reincorporación
	90 segundos	70 °C	Extensión (Síntesis)
1	10 minutos	70 °C	Extensión final
1	∞	15 °C	Finalización

<b>Configuración del Programa #3 “CALIDLON”</b>			
<b>Número de Ciclos</b>	<b>Tiempo</b>	<b>Temperatura</b>	
<b>1</b>	4 minutos	94 °C	Desnaturalización inicial
<b>35</b>	30 Segundos	94 °C	Desnaturalización
	1.30 minutos	57 °C	Reincorporación
	3 minutos	68 °C	Extensión (Síntesis)
<b>1</b>	10 minutos	70 °C	Extensión final
<b>1</b>	∞	15 °C	Finalización
<b>Configuración del Programa #4 “CALISTAQ”</b>			
<b>Número de Ciclos</b>	<b>Tiempo</b>	<b>Temperatura</b>	
<b>1</b>	30 segundos	94 °C	Desnaturalización inicial
<b>30</b>	30 segundos	94 °C	Desnaturalización
	45 segundos	57 °C	Reincorporación
	3.30 minutos	65 °C	Extensión (Síntesis)
<b>1</b>	10 minutos	65 °C	Extensión final
<b>1</b>	∞	15 °C	Finalización
<b>Configuración del Programa #5 “ CALISTAQ2”</b>			
<b>Número de Ciclos</b>	<b>Tiempo</b>	<b>Temperatura</b>	
<b>1</b>	30 segundos	94 °C	Desnaturalización inicial
<b>30</b>	30 segundos	94 °C	Desnaturalización
	45 segundos	59 °C	Reincorporación
	4.20 minutos	65 °C	Extensión (Síntesis)
<b>1</b>	10 minutos	65 °C	Extensión final
<b>1</b>	∞	15 °C	Finalización

---

<b>Configuración del Programa #6 “ CALISTAQ18”</b>			
<b>Número de Ciclos</b>	<b>Tiempo</b>	<b>Temperatura</b>	
<b>1</b>	30 segundos	94 °C	Desnaturalización inicial
<b>30</b>	30 segundos	94 °C	Desnaturalización
	45 segundos	57 °C	Reincorporación
	10min	65 °C	Extensión (Síntesis)
<b>1</b>	5minutos	65 °C	Extensión final
<b>1</b>	∞	15 °C	Finalización

---

El volumen final de las reacciones de amplificación fue de 50µL, los cuales contenían 2.5U de polimerasa (LongAmp® Taq DNA Polymerase - New England Biolabs), amortiguador de amplificación específico para la enzima (LongAmp Taq Reaction – New England Biolabs), dNTP's 300mM y cebadores a una concentración final de 0.5 mM Todas las amplificaciones se llevaron a cabo a partir de 25 -30 ng/µL de ADN y en un termociclador Eppendorf Mastercycler™.