



Álvaro Javier Ferriz Mejías

Secuenciación, ensamblaje *de novo* y
anotación del genoma mitocondrial
del ajo (*Allium sativum*)

Trabajo tutorizado por: Dr. Héctor Candela Antón y Ricardo Parreño

Montoro

Departamento de Biología Aplicada, Área de Genética

Grado en Biotecnología

Facultad de Ciencias Experimentales

Universidad Miguel Hernández de Elche

Curso 2016-2017

ÍNDICE DE MATERIAS

1. Resumen	4
2. Introducción.....	5
2.1. El ajo.....	5
2.2. La mitocondria	5
2.2.1. Estructura mitocondrial.....	6
2.2.2. ADN mitocondrial	7
2.2.3. Teoría endosimbiótica.....	8
2.2.4. Replicación mitocondrial	9
2.2.5. Transcripción.....	10
2.2.6. Traducción	11
2.3. Androesterilidad.....	12
2.4. Secuenciación y anotación de genomas mitocondriales	13
2.5. Herramientas moleculares para la mejora genética del ajo.....	13
3. Antecedentes y Objetivos.....	15
4. Materiales y Métodos	16
4.1. Cultivo en maceta de <i>Allium sativum</i>	16
4.2. Purificación de ADN genómico	16
4.3. Electroforesis de ácidos nucleicos en geles de agarosa	17
4.4. Análisis bioinformático	17
4.4.1. Ensamblaje del genoma.....	17
4.4.2. Anotación del genoma.....	18
4.5. Diseño de cebadores.....	18
4.5. Amplificaciones mediante PCR	18
4.6. Purificación de productos de PCR.....	19
4.7. Secuenciación de los productos de PCR.....	19
5. Resultados	20
5.1. Ensamblaje <i>de novo</i> del genoma del cloroplasto del ajo.....	20
5.2. Anotación del genoma.....	24
5.2.1. Genes necesarios para la expresión génica de la mitocondria	24
5.2.2. Genes implicados en la cadena de transporte de electrones.....	25
5.2.3. Otros genes.....	25
6. Discusión	27
7. Conclusiones y proyección futura.....	29
8. Bibliografía.....	30

ÍNDICE DE FIGURAS

Figura 1.- Grafos representados por Bandage.....	21
Figura 2.- Electroforesis en gel de agarosa de las amplificaciones por PCR realizadas con los diversos cebadores y contenido de las calles	22
Figura 3.- Electroforesis en gel de agarosa de la purificación de ADN de las amplificaciones por PCR realizadas con los diversos cebadores	22
Figura 4.- Cántigos conteniendo secuencias repetitivas marcadas en color rojo	23
Figura 5.- Representación de dos cántigos del genoma mitocondrial del ajo.....	25

ÍNDICE DE TABLAS

Tabla 1.- Nombre y secuencia de los cebadores empleados	19
Tabla 2.- Genes identificados en la secuencia del genoma mitocondrial del ajo.	24
Tabla 3.- Tamaños de genomas mitocondriales de angiospermas y número de genes.....	28



1. Resumen

En este trabajo hemos llevado a cabo la secuenciación, el ensamblaje *de novo* y la anotación de varios contigs, cuya longitud total es 300 kb, derivados del genoma mitocondrial del ajo (*Allium sativum*). La secuenciación fue realizada con un equipo HiSeq 2500 de Illumina, empleando lecturas emparejadas de 101 nucleótidos de longitud. El ensamblaje *de novo* de las secuencias obtenidas ha sido realizado con el programa Velvet, que utiliza una estrategia de ensamblaje basado en la utilización de grafos de De Bruijn. En los contigs ensamblados correspondientes al genoma mitocondrial hemos anotado 67 genes funcionales, incluyendo 40 genes que codifican proteínas, 24 genes de ARN de transferencia y 3 de ARN ribosómicos. De los genes que codifican proteínas, 8 parecen ser de origen cloroplástico, 19 codifican proteínas implicadas en complejos de la cadena transportadora de electrones y 6 codifican proteínas necesarias para la expresión génica mitocondrial. En la región analizada hemos detectado secuencias repetidas que parecen participar en la recombinación que da lugar a diferentes variantes estructurales del genoma mitocondrial.

Palabras clave: Ajo, *Allium sativum*, secuenciación masivamente paralela, genoma mitocondrial, recombinación.

In this work, we have carried out the sequencing, *de novo* assembly and annotation of several contigs derived from the mitochondrial genome of garlic (*Allium sativum*), with a combined length of 300kb. The genome was sequenced using the Illumina HiSeq 2500 next-generation sequencing platform, using 101-bp paired-end reads. The *de novo* assembly was carried out using Velvet software, which implements assembly algorithms based on the use of de Bruijn graphs. In the assembled contigs corresponding to the mitochondrial genome, we have annotated 67 functional genes (including 40 protein-coding genes, 24 transfer RNA genes and 3 ribosomal RNA genes). Eight of the protein-coding genes seem to derive from plastid sequences, 19 encode proteins involved in electron transport chain complexes and 6 encode proteins required for mitochondrial gene expression. In the analyzed region, we have detected repeated sequences that might participate in the recombination events that produce different structural variants of the mitochondrial genome.

Keywords: Garlic, *Allium sativum*, next-generation sequencing, mitochondrial genome, recombination.

2. Introducción

2.1. El ajo

El ajo *Allium sativum* es una planta monocotiledónea de la familia Amaryllidaceae, dentro del orden Asparagales. Las plantas de este género producen bulbos que son las partes comestibles de las mismas. El género *Allium* comprende más de 860 especies según la página web de los Royal Botanic Gardens, Kew (<http://apps.kew.org/wcsp/report-builder.do>). Este género comprende plantas con un gran interés económico y agronómico como la cebolla (*A. cepa*), el puerro (*A. ampeloprasum*) o el cebollino chino (*A. tuberosum*).

El ajo es una planta bianual que forma bulbos compuestos por un número variable de dientes, que suele oscilar entre 8 y 14, cada uno de ellos cubierto por una túnica de color blanquecino o violáceo. El bulbo se encuentra rodeado también por túnicas exteriores con pigmentación similar. En la inflorescencia también se forman propágulos, los denominados bulbilos, que permiten la propagación vegetativa de la planta. Las inflorescencias no suelen producir semillas, por lo que la propagación vegetativa se realiza a partir de los dientes y los bulbilos. La propagación vegetativa a menudo conlleva la presencia de patógenos que reducen la viabilidad de la futura planta y se evita normalmente con el cultivo in vitro de los meristemos.

Es una especie originaria de Asia Central que se cultiva mundialmente en zonas de clima templado (Mikaili *et al.*, 2013). Con una producción anual superior a 20 toneladas, China es actualmente el primer productor a nivel mundial (<http://www.fao.org>). En la Unión Europea, el primer productor es España con una producción de 180.000 toneladas en 2016, equivalente a la mitad de la producción europea. El mercado internacional del ajo se encuentra en crecimiento, como ilustra el incremento en las exportaciones ocurrido entre 2011 y 2015 (de 64.000 a 148.000; II Congreso Internacional del Ajo en Alberca de Záncara, Cuenca).

Su importancia gastronómica es notoria en la cocina mediterránea ya que se emplea como aderezo en muchas elaboraciones. Sus propiedades organolépticas se deben en gran medida a su contenido en compuestos azufrados. Uno de estos compuestos, la alicina, se produce a partir de aliina por medio de la enzima alinasa en respuesta a daños tisulares (Kourounakis y Rekká, 1991). Por su actividad antimicrobiana, la alicina se emplea como aditivo alimentario (Eja *et al.*, 2007; Singh *et al.*, 2009). Entre otras propiedades beneficiosas, los extractos de ajo parecen poseer propiedades antitumorales (Tsubura *et al.*, 2011), antiinflamatorias (Kim *et al.*, 2017), antioxidantes (Kim *et al.*, 2017), antifúngicas (Hayat *et al.*, 2016), e inmunomodulatorias (Clement *et al.*, 2010).

2.2. La mitocondria

Las mitocondrias son orgánulos subcelulares cuya función principal es realizar la respiración celular de forma aeróbica con la fin de sintetizar ATP. Por ello, se las considera las centrales energéticas de la célula (Siekevitz, 1957). Las mitocondrias se encuentran en el citoplasma de las células eucarióticas, excepto en el género *Monocercomonoides* (Roger *et al.*, 1998; Karnkowska *et al.*, 2016). El conjunto de mitocondrias de una célula se denomina condrioma (Logan, 2010). Cada célula presenta un número de mitocondrias que varía según la función de cada tipo celular. Las diferencias en el número de mitocondrias pueden deberse al reparto asimétrico del citoplasma entre las células hijas durante la

división celular. Estas diferencias, sumadas a la variabilidad en el tamaño y a los potenciales de membrana de las mitocondrias, pueden dar lugar a diferencias extrínsecas en los niveles de ATP (Neves *et al.*, 2010; Johnston *et al.*, 2012). Las mitocondrias forman una red asociada con el citoesqueleto dentro de la célula. Esta red se halla sometida a constantes procesos de transporte, fusión y fisión que pueden tener implicaciones en su función (Rappaport *et al.*, 1998). El campo que estudia el efecto de estos procesos en las mitocondrias recibe el nombre de dinámica mitocondrial (Hoitzing *et al.*, 2015).

2.2.1. Estructura mitocondrial

El tamaño de las mitocondrias suele oscilar entre 0,5 y 3 μm , por lo que es similar al de la bacteria *Escherichia coli* (Bereiter-Hahn, 1990; Rafelski y Marshall, 2008; Osiro *et al.*, 2012; Wiemerslage y Lee, 2016). Las mitocondrias poseen dos membranas constituidas por fosfolípidos y proteínas (Gohil y Greenberg, 2009), la externa y la interna, que delimitan el espacio intermembrana y la matriz mitocondrial. Estas membranas, y los compartimentos que definen, contienen proteínas específicas de acuerdo con las reacciones que tienen lugar en ellos.

La membrana externa delimita el perímetro exterior, separando la matriz citoplasmática o citosol de la propia mitocondria. La proporción de proteínas y fosfolípidos en su composición es semejante a la de la membrana plasmática, siendo los lípidos las moléculas más abundantes (Horvath y Daum, 2013). La membrana externa es crucial en el metabolismo de la célula ya que controla procesos como el paso de pequeñas moléculas y la conducción de señales hacia la mitocondria (Duncan *et al.*, 2013). Esta membrana contiene más de 120 proteínas diferentes, entre las que destacan las porinas, un tipo de proteína integral de membrana que forma canales por los que difunden moléculas de hasta 6 kDa (Benz, 1985; Weeber *et al.*, 2002). Las proteínas de tamaño superior pueden atravesar la membrana por medio de la unión de una secuencia señal en su extremo N-terminal a proteínas de membrana denominadas translocasas y a otros complejos de translocación (Herrmann y Neupert, 2000).

La membrana externa también posee otras proteínas implicadas en la transmisión de señales entre el citosol y la mitocondria, la síntesis de la membrana, el mantenimiento de su morfología y dinámica, y la catálisis de reacciones como la elongación de ácidos grasos o la oxidación y degradación de diversas moléculas (Giegé *et al.*, 2005; Tait y Green, 2012). La disrupción de la membrana externa puede llevar a la muerte celular por liberación al citosol de las proteínas que se encuentran en el espacio intermembrana (Chipuk *et al.*, 2006).

El espacio intermembrana se localiza entre la membrana externa y la interna. Su pH es bajo debido a la elevada concentración de protones, que se origina por el bombeo de los mismos en la cadena respiratoria (Porcelli *et al.*, 2005). Su composición es similar a la del citosol en cuanto a moléculas pequeñas, ya que la membrana externa es permeable a iones y azúcares, pero no a moléculas mayores, por lo que su contenido en proteínas es diferente. El espacio intermembrana contiene enzimas encargadas de la transferencia de ATP (adenilato y creatina quinasa) y del transporte de ácidos grasos hasta la matriz mitocondrial (McKelvey y Fioravanti, 1986; Fraser *et al.*, 1997).

La membrana interna de la mitocondria separa el espacio intermembrana de la matriz mitocondrial. Presenta numerosos pliegues que aumentan sustancialmente su superficie, de manera

que el área de la membrana interna puede llegar a ser hasta cinco veces superior al de la externa (Griparic y Van Der Bliet, 2001). Estos pliegues se denominan crestas mitocondriales, y su número está directamente relacionado con la demanda de ATP. Esta membrana presenta un tipo de lípido, la cardiolipina, que también se encuentra en las membranas bacterianas (Carranza *et al.*, 2017). La cardiolipina es necesaria para la fosforilación oxidativa y hace menos impermeable y más selectiva a esta membrana en comparación a la membrana externa (Ikon y Ryan, 2017).

La membrana interna posee un elevado contenido en proteínas (80%) ya que en ella residen complejos enzimáticos como el que realiza la fosforilación oxidativa, compuesto por ATP sintasas que acoplan el paso de protones de H⁺ con la síntesis de ATP, los cuatro complejos de la cadena de transporte electrónico, proteínas de fusión y fisión mitocondrial, y transportadores que contribuyen a la alta selectividad de esta membrana.

Por último, se encuentra la matriz mitocondrial que contiene dos tercios de las proteínas de la mitocondria. Presenta una densidad de proteínas de 0,8 µL de agua por cada mg de proteína, mayor que el espacio intermembrana (Pfaff *et al.*, 1968). El pH de la matriz mitocondrial es 7,8 (Porcelli *et al.*, 2005). En ella se dan las reacciones del ciclo de Krebs, el ciclo del ácido cítrico, la beta-oxidación de los ácidos grasos, la síntesis del grupo hemo y la oxidación de aminoácidos, entre otras. La matriz mitocondrial contiene un genoma de ADN y moléculas de ARN. Al igual que las células procarióticas de vida libre, las mitocondrias contienen la maquinaria necesaria para la expresión de su genoma, incluyendo moléculas de ARN de transferencia (ARNt) y ribosómico (ARNr) (Kühlbrandt, 2015), así como ribosomas con características similares a los de tipo 70S de los procariotas (Leaver y Harmey, 1976).

2.2.2. ADN mitocondrial

Las mitocondrias poseen su propio ADN, que se encuentra alojado en la matriz mitocondrial. El estudio del ADN mitocondrial (ADNmt) de las plantas es complejo y está considerado un reto en la biología molecular evolutiva (Galtier, 2011). Las mitocondrias de las plantas difieren de las mitocondrias de los animales en el tamaño de sus genomas. En las células animales, este tamaño oscila entre 15 y 17 kb (Ladoukais y Zouros, 2017). Los genomas mitocondriales más pequeños corresponden a animales del filo ctenóforos, con un tamaño de 10.326 bp para *Mnemiopsis leidyi* (número de acceso JF760210.1; Pett *et al.*, 2011) y de 11.016 bp para *Pleurobrachia bachei* (NC_016697.1; Kohn *et al.*, 2012). Por otra parte, los genomas mitocondriales animales de mayor tamaño corresponden a *Clathrina clathrus*, que posee un genoma mitocondrial con 6 cromosomas lineales y longitud total de 51 kb (Lavrov *et al.*, 2013), y *Petrobiona massiliana*, que posee más de 150 cromosomas mitocondriales y una longitud superior a 77 kb (Lavrov *et al.*, 2016).

Sin embargo, los genomas mitocondriales de plantas terrestres son mucho mayores, con tamaños que varían desde 100 kb (*Buxbaumia aphylla*; NC_024518; Liu *et al.*, 2014) y 2.740 kb (*Cucumis melo*; JF412792.1; Rodríguez-Moreno *et al.*, 2011), existiendo notables diferencias de tamaño incluso entre especies filogenéticamente próximas, como se aprecia en el buscador de genomas secuenciados (<https://www.ncbi.nlm.nih.gov/genome/browse/>).

Los genomas mitocondriales de las células animales poseen una elevada densidad de genes, debido a su tamaño reducido y la casi total ausencia de intrones (Boore, 1999; Lavrov y Pett, 2016).

Sin embargo, los genomas mitocondriales de las plantas son más grandes debido a factores como la abundancia de secuencias repetitivas, la presencia de intrones largos que contienen secuencias no codificantes susceptibles de expandirse, y la adquisición de secuencias provenientes de otros genomas que se han integrado en el genoma mitocondrial a lo largo de la evolución (como los genomas del núcleo y del cloroplasto, o genomas de virus y bacterias; Unseld *et al.*, 1997; Aono *et al.*, 2002; Alverson *et al.*, 2011). La abundancia en secuencias repetitivas es una característica de los genomas mitocondriales de las angiospermas, y ocasiona un gran dinamismo estructural (Gualberto and Newton, 2017). La elevada frecuencia de recombinación es la principal responsable de la rápida evolución estructural de los genomas mitocondriales de las plantas y de la generación de subgenomas lineales, circulares o ramificados (Bendich, 1993; Oldenburg y Bendich, 2001).

En las células animales, el genoma mitocondrial está presente en un número de copias superior al de las células vegetales (Woloszynska *et al.*, 2006; Preuten *et al.*, 2010). En las plantas, el estado de madurez puede afectar al número de copias, como sucede en el maíz, donde la senescencia se ha asociado a un menor número de moléculas de ADNmt por célula (Oldenburg *et al.*, 2013). También se ha observado que el número de copias del genoma mitocondrial varía según el tejido, como sucede en las raíces, que suelen contener más copias que las hojas (Kumar *et al.*, 2014). Las mitocondrias también pueden presentar transferencia de material genético entre ellas debido a su capacidad de fusionarse (Arimura *et al.*, 2004). Este hecho produce mitocondrias con diferente número de copias del genoma, e incluso dar lugar a mitocondrias con solo una parte del genoma o sin éste (Arimura *et al.*, 2004), como ilustra un estudio realizado en *Arabidopsis thaliana*, *Nicotiana tabacum* y *Hordeum vulgare* en el que, a pesar de que el número medio de proplastos de hojas maduras fue aproximadamente 450, el número de copias del gen *atp1*, que reside en el genoma mitocondrial y codifica la subunidad alfa de la ATP sintasa, fue de solo 280 (Preuten *et al.*, 2010). A pesar del gran tamaño del genoma mitocondrial de las plantas respecto a los animales, el número de genes es parecido, oscilando entre 50 y 60 (Gualberto *et al.*, 2014). Esto contrasta con las diferencias de tamaño existentes entre los genomas mitocondriales de las plantas.

La estructura de los genomas mitocondriales de plantas suele variar. Diversos experimentos, en los que se han empleado técnicas como la microscopía electrónica o los geles de campo pulsado, no han podido verificar una predominancia estructural del genoma (Oldenburg y Bendich, 2001; Backert y Börner, 2000; Cheng *et al.*, 2017). Por lo general, abundan las moléculas lineales de distinto tamaño frente a las moléculas circulares o ramificadas (Backert *et al.*, 1997). También se han descrito moléculas ramificadas cuya replicación se realiza mediante el mecanismo del círculo rodante, como posee la planta *Chenopodium album* (Backert *et al.*, 1996). Sin embargo, la estructura física de los genomas mitocondriales de plantas suele representarse como una molécula circular de doble cadena (Smith y Keeling, 2015).

2.2.3. Teoría endosimbiótica

La teoría endosimbiótica propone que los cloroplastos y mitocondrias descienden de células procariotas de vida libre que sobrevivieron a la endocitosis por parte de una célula eucariota (Taylor, 1974). La gran similitud fisiológica y bioquímica que presentan estos orgánulos respecto a las células

procariotas de vida libre, junto con la posesión de un genoma propio, son los dos grandes pilares en los que se sustenta esta teoría (John y Whatley, 1975).

Las mitocondrias contienen aproximadamente 2.000 proteínas (Meisinger *et al.*, 2008) y, sin embargo, los genomas mitocondriales únicamente codifican para 60 proteínas aproximadamente (Burger *et al.*, 2013). La teoría endosimbiótica propone que la diferencia entre el número de proteínas y de genes presentes en la mitocondria se debe a la transferencia de genes desde el genoma mitocondrial hacia los cromosomas del núcleo (Martin y Herrmann, 1998). De este modo, las proteínas mitocondriales codificadas en el núcleo serían transportadas a las mitocondrias. Las secuencias retenidas en el genoma mitocondrial son necesarias para la cadena transportadora de electrones, para la síntesis de los ribosomas y para la regulación de la expresión de estos genes (Allen, 1993).

Otras teorías propuestas planteaban un origen endógeno para los cloroplastos y mitocondrias que se encuentran en las células eucariotas a partir de invaginaciones de la membrana plasmática (Bogorad, 1975). Esta asunción asumía que el ADN de estos orgánulos debía ser más parecido al ADN nuclear que a cualquier otro puesto que el orgánulo al completo incluido su ADN derivaba de la célula que lo contenía. Más tarde, mediante comparaciones de secuencia se observó que el ADN de los cloroplastos presentaba similitud con secuencias nucleares de cianobacterias; y por medio de análisis de ARNr 16s, se descartó la teoría del origen endógeno para estos orgánulos, aceptándose las proteobacterias como organismo de origen para las mitocondrias (Yang *et al.*, 1985; Stackebrandt *et al.*, 1988; Giovannoni *et al.*, 1988).

La teoría endosimbiótica plantea que la supervivencia de una célula procariota a la endocitosis por parte de una eucariota hace 2.000 millones de años, pudo suponer la realización de la respiración para la obtención de energía para la célula eucariota (Zimorski *et al.*, 2014). Este hecho se debe a que las mitocondrias se conciben en su origen como células procariotas de vida libre capaces de realizar funciones oxidativas carentes en las células eucariotas (Sa y Bl, 2017). Esta nueva función pudo suponer una gran ventaja evolutiva a las células eucariotas que dependían de la glicólisis o de la fermentación para la obtención de energía.

Esta teoría es la más aceptada actualmente, y se ve apoyada por las numerosas características comunes que comparten las mitocondrias y las proteobacterias. Una de estas características es la gran similitud que existe entre genomas mitocondriales y la proteobacteria del género *Rickettsia* observada mediante comparaciones genómicas (Andersson *et al.*, 1998). La división de las mitocondrias también presenta similitud con la de células procariotas debido a que ocurre por fisión (Margolin, 2005). Las porinas, proteínas transportadoras que se encuentran en la membrana bacteriana, también se encuentran en la membrana externa mitocondrial (Zeth y Thein, 2010; Fairman *et al.*, 2011). Las proteínas que se generan en las mitocondrias contienen N-formilmetionina como primer aminoácido, siendo ésta, una característica de células procariotas (Schwartz *et al.*, 1967; Smith y Marcker, 1968). El lípido cardiolipina, que se encuentra en la membrana interna mitocondrial solo se encuentra también, en las membranas bacterianas (Mileykovskaya y Dowhan, 2009).

2.2.4. Replicación mitocondrial

Como se ha mencionado en apartados anteriores, los ADN mitocondriales de plantas son mucho más extensos que los genomas mitocondriales de animales, llegando a alcanzar las 11,3 Mb

que posee la planta *Silene conica* (Knoop, 2004; Gualberto *et al.*, 2014; Sloan *et al.*, 2012). El genoma mitocondrial de células animales consiste en una molécula circular, y se replica mediante un bucle de manera unidireccional, utilizando un transcrito procesado como cebador, aunque también puede ser bidireccional por otros mecanismos (Bogenhagen y Clayton, 2003). La polimerasa que realiza la replicación está codificada en el ADN nuclear. Sin embargo, los genomas mitocondriales de plantas poseen estructuras más complejas, a pesar de que su representación se ha realizado comúnmente mediante una secuencia circular de doble cadena (Oldenburg y Bendich, 2001; Backert y Börner, 2000; Cheng *et al.*, 2017).

Predominantemente, los genomas mitocondriales de las plantas son lineales y presentan diversas estructuras menores que dificultan un mecanismo de replicación similar al de los genomas mitocondriales animales. La replicación de los genomas mitocondriales en plantas es un proceso que todavía no se conoce completamente. Mediante técnicas proteómicas y métodos de búsqueda de homología en secuencias, se ha conseguido identificar algunos factores implicados en la replicación. Diversos estudios han propuesto el mecanismo del círculo rodante del bacteriófago T4 como aproximación al tipo de replicación que podría darse en el genoma mitocondrial de las plantas, junto con un mecanismo dependiente de recombinación (Backert *et al.*, 1996, 1997; Oldenburg y Bendich, 2001; Bogenhagen y Clayton, 2003; Manchekar *et al.*, 2006, 2009). Este hecho supone un tipo de replicación más compleja que el mecanismo de bucle de los genomas mitocondriales animales (Bogenhagen y Clayton, 2003). A pesar de ello, no se sabe si los genomas mitocondriales de las plantas contienen orígenes de replicación. Tampoco se conocen los mecanismos que controlan el número de copias del genoma mitocondrial (Cupp y Nielsen, 2014). Sin embargo, el número de copias varía sustancialmente dependiendo del tejido, de las necesidades energéticas y del estado de madurez de la planta (Rauwolf *et al.*, 2010; Oldenburg *et al.*, 2014). Por ello, se asume que debe existir un mecanismo regulador del número de copias del genoma distinto del de las células animales.

2.2.5. Transcripción

Durante la coevolución de las mitocondrias y sus hospedadores, muchos genes de origen mitocondrial han sido transferidos al núcleo celular (Brennicke *et al.*, 1993; Leister, 2005; Stegemann y Bock, 2006). Sin embargo, también se encuentran secuencias de origen nuclear en las mitocondrias, lo que pone de manifiesto que la transferencia de material genético del núcleo a la mitocondria también es posible (Sato, 2001; Hengeveld y Fedonkin, 2004). A pesar de sus similitudes con las α -proteobacterias, de las que derivan, la maquinaria transcripcional de las mitocondrias es mucho más compleja y parecida a la de las células eucariotas que a la de las células procariotas (Liere *et al.*, 2011).

La transcripción mitocondrial en plantas produce ARN mensajeros policistrónicos, como los de las células procariotas (Rossmann y Karwan, 1998). Sin embargo, no se han descrito genes cotranscritos en abundancia en plantas (Unselde *et al.*, 1997; Giegé *et al.*, 2000), lo que podría deberse a la elevada tasa de reorganización observada en los genomas mitocondriales de las plantas (Woloszynska *et al.*, 2006; Arrieta-Montiel *et al.*, 2009; Shedge *et al.*, 2010).

Como se ha comentado anteriormente, los genomas mitocondriales de los animales apenas poseen intrones. Sin embargo, los genomas mitocondriales de las plantas presentan intrones del grupo I e intrones del grupo II (Bonen, 2008). Algunos intrones de estos dos tipos codifican una enzima

llamada maturasa que promueve el procesamiento (*splicing*) y la integración de los intrones en nuevos sitios del genoma. La maturasa posibilita que los intrones de tipo I y II puedan ser móviles y capaces de auto-escindirse (Hammani y Giegé, 2014).

Además de la maturasa, en el núcleo celular también se encuentran muchos factores que contribuyen al *splicing* en los genomas mitocondriales de plantas (Hammani y Giegé, 2014). En el genoma nuclear de plantas superiores se encuentran cuatro homólogos de *matR* llamados *nMat1-4*, que codifican productos cuya diana se encuentra en la mitocondria (Mohr y Lambowitz, 2003; Keren *et al.*, 2009, 2). En dos de ellos, *nMAT1* y *nMAT2*, se ha confirmado su papel en el *splicing* mitocondrial en *Arabidopsis* (Keren *et al.*, 2009, 2, 2012). Aparte de estos elementos, otras proteínas con motivos "DEAD-box" (Asp-Glu-Ala-Asp) también pueden estar involucradas en el *splicing* (Köhler *et al.*, 2010).

2.2.6. Traducción

Debido al origen procariota de las mitocondrias y al alto grado de conservación encontrado en las regiones de los ARN ribosómicos con células procariotas, cabría esperar que la traducción de los genes mitocondriales en las plantas fuese muy similar a los mecanismos que ocurren en células procariotas (Gray, 1992). Sin embargo, ciertas características relacionadas con la traducción del ADN difieren entre las células procariotas y las mitocondrias de las células vegetales.

El reconocimiento del codón de inicio del ARN mensajero en las células procariotas depende de dos factores: el primero de ellos es el apareamiento de las bases entre el extremo 3' rico en pirimidinas de la subunidad pequeña del ARN ribosómico 16S; el segundo es la secuencia de Shine-Dalgarno, rica en purinas, que está presente en el ARN mensajero. Esta secuencia se localiza de 5 a 9 nucleótidos aguas arriba del codón de inicio de la traducción y suele constar de 4 a 5 nucleótidos. (Kozak, 1999; Marintchev y Wagner, 2004; Nakamoto, 2006). Sin embargo, un análisis realizado con 164 secuencias situadas aguas arriba de genes mitocondriales que codificaban proteínas reveló que la secuencia de Shine-Dalgarno AGGAGG no está presente en muchos genes (Hazle y Bonen, 2007)

Por otra parte, las mitocondrias de células vegetales poseen ribosomas similares a los de las bacterias que, sin embargo, han incorporado partes originarias de las células eucariotas hospedadoras (Bonen y Calixte, 2006). Otra peculiaridad es que los ARN mensajeros mitocondriales que carecen codon de terminación se traducen, como sucede con los ARNm de los genes *nad6* y *ccmC* de *Arabidopsis* y en la coliflor (Raczynska *et al.*, 2006).

Estas observaciones ponen de manifiesto que existen diferencias entre la maquinaria traduccional de las células procariotas y la de las mitocondrias de las plantas. Así mismo, estas características apoyan la idea de que la traducción de los genes mitocondriales en los ribosomas puede estar asistida por factores tipo trans (Hammani y Giegé, 2014). Esta hipótesis está apoyada por la existencia de proteínas PPR (proteínas con repeticiones de pentatricopéptidos), tales como MPPR6, que se une a la región 5' no traducida de la proteína ribosómica *rps3* en el maíz, o la proteína PPR336 que se asocia con los polisomas mitocondriales, lo que sugiere su participación en la traducción mitocondrial (Uyttewaal *et al.*, 2008; Manavski *et al.*, 2012). La implicación de las proteínas PPR en la traducción mitocondrial de plantas también está apoyada por su habilidad para restaurar la fertilidad en plantas con flores.

2.3. Androesterilidad

La androesterilidad hace referencia a la incapacidad de producir polen viable. Esta propiedad es especialmente atractiva para producir individuos híbridos. Para obtener híbridos, se cruzan plantas que difieren en rasgos de interés. Este procedimiento se realiza mediante la polinización cruzada de plantas de una misma especie, polinizando el ovario de una planta con el polen de otra. El interés por la obtención de híbridos se debe al mayor rendimiento que presentan en comparación con sus parentales. La mayor productividad puede atribuirse al vigor híbrido, también denominado heterosis. Los híbridos pueden llegar a ser entre un 15 y 50% más productivos que sus parentales, como sucede a ciertos híbridos del arroz cuyo rendimiento se incrementa en un 20% (Tester y Langridge, 2010). Este hecho es especialmente interesante para la sociedad actual, ya que se prevé un incremento del 70% en la demanda de alimentos en las próximas décadas, según datos de la Comisión Europea de 2011. Las variedades híbridas de cultivos como el maíz, arroz, girasol o sorgo suponen más del 50% de la producción agrícola global (Li *et al.*, 2007).

Para la obtención de plantas híbridas, es deseable impedir la autopolinización. Antiguamente, esto se lograba extirpando los estambres de las flores, para que la planta no pudiera producir polen. La emasculación puede realizarse manualmente o con ayuda de máquinas o agentes químicos, con el consiguiente incremento en el coste de producción. Las plantas androestériles no presentan estos inconvenientes, por lo que suponen una gran herramienta para la producción de híbridos.

La androesterilidad fue descubierta por J.G. Kölreuter en 1763 (Mayr, 1986) y se ha descrito en más de 600 especies (Kaul, 1988). Puede estar causada por interacción entre genes mitocondriales con genes nucleares, en cuyo caso se conoce como esterilidad masculina citoplasmática CMS; o por genes nucleares únicamente, también denominada esterilidad génica masculina GMS (Vedel *et al.*, 1994). La esterilidad citoplasmática masculina impide la producción de polen funcional, y suele deberse a reorganizaciones del genoma mitocondrial que se transmiten por vía materna (Eckardt, 2006; Chase, 2007; Chen y Liu, 2014). Estas reorganizaciones se deben frecuentemente a recombinación no homóloga, aunque también pueden ocurrir entre secuencias repetidas (Sugiyama *et al.*, 2005; Horn *et al.*, 2014). Las reorganizaciones suelen causar la expresión de pautas de lectura abierta (ORF; *open reading frames*) quiméricas que causan disfunción mitocondrial (Chase, 2007) e incompatibilidad con el genoma nuclear durante desarrollo del gametofito, provocando el aborto del polen (Hanson y Bentolila, 2004).

Esta androesterilidad es susceptible de ser revertida mediante genes nucleares denominados restauradores de la fertilidad (Rt) (Gabay-Laughnan y Laughnan, 1994). Estos genes suelen codificar proteínas PPR que reprimen la expresión de ORF tóxicas, como es el caso de la proteína PPR-B de *Raphanus sativus* (Uyttewaal *et al.*, 2008; Fujii *et al.*, 2011). Los genes restauradores de la fertilidad suelen manifestarse de manera dominante, aunque se han descrito casos de recesividad y superdominancia (Wen *et al.*, 2003; Fujii y Toriyama, 2009; Wang *et al.*, 2013). Los sistemas de restauración de fertilidad pueden actuar sobre el gametofito o sobre el esporofito de tal manera que una planta diploide CMS Rf/rf producirá dos tipos de gametos Rf y rf. Ambos tipos son funcionales si ocurre restauración esporofítica, mientras que sólo es funcional el tipo Rf si la restauración es gametofítica (Wise *et al.*, 1999). El estudio de la androesterilidad también despierta gran interés por su relación con

el desarrollo del polen y las interacciones entre el núcleo y el citoplasma (Schnable y Wise, 1998; Hanson y Bentolila, 2004; Chase, 2007; Guo y Liu, 2012).

2.4. Secuenciación y anotación de genomas mitocondriales

La secuenciación del genoma de un organismo consiste en la determinación de su secuencia completa de nucleótidos. La anotación de un genoma consiste en la identificación y descripción de los genes que contiene dicha secuencia. El número de genomas mitocondriales depositados en la base de datos del NCBI ha experimentado un incremento exponencial en los últimos años (Smith, 2016). Entre 2010 y 2012 se duplicó el número de secuencias mitocondriales depositadas en GenBank. Volvió a duplicarse en 2014, año en el que se describieron más de 1.000 nuevos genomas mitocondriales (Smith, 2016; NCBI). Actualmente, hay más de 8.080 secuencias pertenecientes a genomas mitocondriales en GenBank. Este gran número ha propiciado que sea el tipo de cromosoma eucariótico más secuenciado, lo que ha contribuido al avance de la ciencia forense, la medicina y la biología molecular evolutiva (Smith y Keeling, 2015; Smith, 2016).

El número de artículos científicos publicados sobre genomas mitocondriales también ha aumentado exponencialmente (Smith, 2016). Más de 4.900 artículos publicados en 2017 contienen las palabras "genoma mitocondrial completo", lo que supone un número mayor al de los años 2000 a 2014 combinados (Smith, 2016). Muchos de estos artículos estudian relaciones filogenéticas, enfermedades mitocondriales o describen la secuencia de genomas mitocondriales, lo que puede emplearse para caracterizar aspectos como su estructura o los mecanismos de reparación, replicación, transcripción y traducción (Dimauro y Davidzon, 2005; Witt *et al.*, 2015).

Este incremento en el número genomas mitocondriales secuenciados ha sido posible gracias a avances en las tecnologías de secuenciación y al desarrollo de programas bioinformáticos, que permiten su secuenciación y anotación de forma rápida, barata y eficaz. Muchas veces no es necesario purificar las mitocondrias para secuenciar y ensamblar el genoma mitocondrial. La secuenciación de una muestra de ADN total con las nuevas tecnologías de secuenciación suele ser suficiente para realizar el ensamblaje de un genoma mitocondrial (Smith, 2012).

2.5. Herramientas moleculares para la mejora genética del ajo

La reproducción vegetativa de la mayoría de las variedades comerciales de ajo es una de las grandes limitaciones que poseen los programas de mejora genética de esta especie. En la actualidad, solo algunas variedades de ajo poseen la capacidad de florecer y producir semillas. La ausencia de reproducción sexual y la consecuente falta de recombinación meiótica limitan las herramientas para la mejora genética de esta especie a la acumulación de mutaciones somáticas y su posterior selección clonal (Bhagyalakshmi *et al.*, 2009).

La secuenciación del genoma nuclear del ajo es una tarea compleja debido a su gran tamaño, que alcanza 15,9 Gb (Kamenetsky *et al.*, 2015). A pesar de ello, se han realizado diversos estudios transcriptómicos que permiten la realización de estudios sobre la función de los genes (Kim *et al.*, 2009; Shemesh-Mayer *et al.*, 2015). A estos trabajos se suman otros en los que se describen marcadores moleculares para el ajo, como microsatélites y AFLPs (Buso *et al.*, 2008; Cunha *et al.*, 2012; Jo *et al.*, 2012). Estos trabajos han permitido la realización de estudios sobre las relaciones filogenéticas entre

especies del género *Allium* y la creación de mapas de ligamiento genético del ajo, entre otros (Ipek et al., 2004; Zewdie et al., 2005).



3. Antecedentes y Objetivos

La línea de trabajo en la que se enmarca este proyecto tiene como objetivo desarrollar herramientas genéticas y moleculares para el estudio del ajo cultivado, *Allium sativum*. En el marco de esta línea de investigación, iniciada en 2015 en el laboratorio del tutor (Dr. Héctor Candela), se han realizado diversos experimentos para ampliar el conocimiento disponible sobre el genoma del ajo. Estos experimentos incluyen caracterización de los transcriptomas de diversos tejidos del ajo, que han sido secuenciados, ensamblados *de novo* y anotados. Por el interés básico del conocimiento del genoma del ajo y por sus posibles aplicaciones, iniciamos la caracterización del genoma de los orgánulos (cloroplasto y mitocondria) de esta especie.

El estudio a nivel molecular de *Allium sativum* facilitará el desarrollo de herramientas para comprender las bases moleculares de los problemas que limitan la mejora genética del ajo cultivado, como la infecundidad de las flores de casi todas las variedades o la infección de los recursos genéticos disponibles con agentes patógenos como nematodos o virus.

Algunos grupos de investigación han investigado la base molecular de la infertilidad en esta especie y han logrado la producción de flores fértiles en unas pocas variedades del ajo (Shemesh Mayer *et al.*, 2013; Mayer *et al.*, 2015; Shemesh-Mayer *et al.*, 2015). Estos estudios de restauración de la fertilidad posibilitarán en el futuro el desarrollo de variedades híbridas del ajo cultivado. El estudio del genoma mitocondrial podría facilitar la identificación de marcadores moleculares asociados a la esterilidad en el ajo (Shemesh-Mayer *et al.*, 2015). Esta aproximación es la que se ha seguido en una especie muy próxima, la cebolla, *Allium cepa*, y podría permitir el desarrollo de sistemas basados en androesterilidad para la obtención de variedades híbridas de ajo (Kim *et al.*, 2009).

En un trabajo anterior se completó la secuenciación, ensamblaje *de novo* y anotación del genoma del cloroplasto del ajo, *Allium sativum*. El objetivo principal de este Trabajo de Fin de Grado, complementario del trabajo anterior, ha sido realizar la secuenciación, ensamblaje *de novo* y anotación del genoma mitocondrial de *Allium sativum*. Este objetivo ha sido concretado en los siguientes objetivos particulares, que se corresponden con las etapas seguidas en la realización del trabajo:

1. Realizar la secuenciación y el ensamblaje *de novo* del genoma mitocondrial de *Allium sativum* mediante el uso de herramientas bioinformáticas adecuadas.
2. Realizar la anotación de los genes contenidos en la secuencia del genoma mitocondrial, utilizando para ello una combinación de métodos *ab initio* y métodos basados en homología.
3. Contribuir al conocimiento de las bases moleculares responsables de la diversidad estructural del genoma mitocondrial. Este tercer objetivo se planteó tras detectar hechos de recombinación entre las secuencias de varios cóntigos.

4. Materiales y Métodos

4.1. Cultivo en maceta de *Allium sativum*

Las plantas de ajo se cultivaron en macetas con dimensiones de 11×11,5 cm, todas ellas situadas en bandejas de plástico con dimensiones de 28×50 cm. El contenido de las macetas constaba de tres componentes principales: perlita, vermiculita y turba, con la proporción de 2:2:1 (v/v/v) respectivamente. En cada maceta se plantaron dos dientes de ajo. Las macetas fueron irrigadas con una disolución nutritiva comercial (abono Universal Carrefour), que contiene diferentes elementos en las concentraciones que se indican: nitrógeno (N): 6% (incluye N amoniacal: 2,8%; N ureico: 2,2%; y N nítrico: 1,0%); pentóxido de fósforo (P₂O₅): 6%; óxido de potasio (K₂O): 6%; micronutrientes: cobre (Cu): 0,002%; hierro (Fe): 0,025%; manganeso (Mn): 0,014%; molibdeno (Mo): 0,001%; y zinc (Zn) quelado con EDTA: 0,004%.

Las plantas se irrigaron con agua una vez a la semana y cada 15 días con la disolución nutritiva previamente mencionada en forma de disolución al 0,5% (v/v). En todo momento, el nivel de líquido presente en las bandejas se mantuvo de 2 a 3 cm. Las plantas se mantuvieron en una cámara climática con una temperatura que fue de 23±1°C. Esta cámara utiliza tubos fluorescentes F28 T5/D y Leuci T5-28W que proporcionan iluminación continua de 5.000 lx.

4.2. Purificación de ADN genómico

Para la purificación del ADN genómico de las plantas de ajo se empleó el kit *GeneJET Plant Genomic DNA Purification Mini Kit*, indicado para plantas. El protocolo indicado por el fabricante (Thermo Scientific) fue el procedimiento que se siguió y se resume a continuación. Se emplearon 200 mg de tejido foliar de plantas de ajo jóvenes pertenecientes a la variedad "Spring blanco" y que fueron proporcionadas por la empresa Coopaman S. C. L. Este tejido fue triturado empleando morteros estériles de porcelana con la ayuda de nitrógeno líquido. Después, el material vegetal fue trasladado a tubos *ependorf* y se adicionaron 350 µL de tampón de lisis A. Se empleó un vórtex para agitar dichos tubos durante 10-15 s antes y se agregaron 50 µL de tampón de lisis B junto con 20 µL de RNasa A. Seguidamente, se procedió a la incubación de los tubos en un baño con temperatura de 65°C durante 10 min con agitación periódica.

Tras este paso, fueron agregados a cada tubo 130 µL de disolución de precipitación, se agitaron los tubos por inversión y se incubaron 5 min en hielo. A continuación, los tubos fueron centrifugados durante 5 min a 13.300 rpm (16.300g). Los sobrenadantes de cada tubo se trasladaron a tubos limpios y se adicionaron 400 µL de disolución de unión (*binding solution*) y 400 µL de etanol al 96%. El volumen resultante fue transferido por fracciones de 600-700 µL a las columnas proporcionadas por el kit. Estas columnas fueron centrifugadas a 8.000 rpm (5.900g) durante 1 min, desechando el sobrenadante tras su paso por la misma. Seguidamente, las columnas fueron lavadas dos veces empleando 500 µL de tampón de lavado 1, que fue eliminado tras una centrifugación durante 1 min a 10.000 rpm (9.200g).

Posteriormente, se adicionaron 100 µL de tampón de elución al ADN genómico (ADNg) inmovilizado en la columna. Esta adición se realizó en el centro de la columna para obtener un mejor rendimiento en la elución del ADNg, tal y como indicaba el fabricante. Las columnas fueron incubadas a temperatura

ambiente durante 5 min. Finalmente, se centrifugaron a 10.000 rpm durante 5 min para eluir el contenido en tubos *ependorf* estériles que fueron conservados a 4°C hasta el momento de su secuenciación.

4.3. Electroforesis de ácidos nucleicos en geles de agarosa

Para evaluar la calidad del ADN purificado, se realizó una electroforesis en gel de agarosa al 1% (m/v) en tampón de electroforesis TAE 1x (1 g de agarosa en 100 mL de TAE 1x). La disolución de TAE 1x se preparó a partir una disolución madre a concentración 50x (Tris-HCl 2 M; ácido acético 5,71%; EDTA 50 mM; pH 8,0). La tinción del gel se realizó con 2,5 µL de *SimplySafe* (EURx, Molecular Biology Products), que se añadieron a la disolución de agarosa antes de su solidificación. En la primera calle de cada gel se cargaron 5 µL de marcador de peso molecular (*GeneRuler 1 Kb Plus DNA Ladder* de Thermo Scientific) a una concentración de 0,1 µg/ µL). Las muestras se prepararon con 2 µL de ADN, 8 µL de agua desionizada y 2 µL de tampón de carga 10x (*DNA Loading Buffer 10x* de 5prime). Por último, la electroforesis se llevó a cabo a 100 V durante 35 min con una fuente de alimentación (LabNet).

4.4. Análisis bioinformático

4.4.1. Ensamblaje del genoma

Las muestras de ADN genómico extraídas del tejido foliar fueron secuenciadas por la empresa STAB VIDA (Caparica, Portugal) para su secuenciación con un secuenciador masivamente paralelo Illumina HiSeq 2500 y lecturas emparejadas (*paired-end reads*) de 101 nucleótidos. Posteriormente, se empleó el programa FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) que evalúa la calidad de lecturas producidas y detecta la presencia en las mismas de secuencias provenientes de los adaptadores adicionados durante la elaboración de las muestras.

Para excluir las secuencias provenientes del genoma del cloroplasto, alineamos las lecturas a dicho genoma mediante el programa Bowtie2 (versión 2.1.0), que emplea un algoritmo de alineamiento muy eficiente basado en la transformación de Burrows-Wheeler (Langmead *et al.*, 2009; Langmead y Salzberg, 2012). El índice del genoma de referencia se preparó con el programa *bowtie2-build*, especificando la secuencia, determinada previamente en el laboratorio, con la opción *-f*. El alineamiento de las lecturas emparejadas se llevó a cabo mediante el programa *bowtie2*, con las siguientes opciones: *-x* (especifica el índice del genoma al que deben alinearse las lecturas, preparado anteriormente con *bowtie2-build*), *--qc-filter* (descarta las lecturas de baja calidad), *-p* (especifica el número de procesadores disponibles; en nuestro caso es igual a 32), y *--no-discordant* (descarta las parejas de lecturas cuyos alineamientos no son concordantes). Las parejas de lecturas no alineadas se utilizaron, a continuación, para realizar un ensamblaje *de novo*.

El ensamblaje *de novo* de las lecturas se llevó a cabo con el programa Velvet (versión 1.2.10) (Zerbino y Birney, 2008), ejecutando consecutivamente los programas *velveth* y *velvetg*. El programa *velveth* se ejecutó con las opciones siguientes: longitud de los *k*-meros de 61 nucleótidos (parámetro denominado *hash length* por los autores del programa), lecturas emparejadas cortas (*--shortPaired*), en formato Fastq (*--fastq*), suministradas en archivos separados (*--separate*) y almacenamiento de los datos en formato binario (*--create_binary*). El programa *velvetg* se ejecutó con las opciones

siguientes: cobertura esperada 27 (--exp_cov 27), cobertura máxima 54 (--max_cov 54), umbral de cobertura (--cov_cutoff 5) y otras (--scaffolding no --read_trkg yes). El grafo del ensamblaje, producido por velvetg, se visualizó con el programa Bandage (Wick *et al.*, 2015), que representa los solapamientos entre los contigs. Se realizaron búsquedas BLAST con el programa Bandage con un valor de E de 1×10^{-10} .

Las lecturas se alinearon a los contigs resultantes con ayuda de Bowtie2, por el procedimiento descrito. Los archivos SAM resultantes se convirtieron a formato BAM mediante el comando view de samtools (Li *et al.*, 2009). Los archivos BAM se ordenaron según la posición de las lecturas en el genoma de referencia mediante el comando sort de samtools. El archivo resultante se indexó con el comando index de samtools y se visualizó con ayuda del programa Tablet (Milne *et al.*, 2013).

4.4.2. Anotación del genoma

Se emplearon diversos programas para la anotación del genoma. Para la identificación de genes cuyo producto eran moléculas de ARN no codificante (ARNt y ARNr) se emplearon los programas ARAGORN (Laslett y Canback, 2004) y tRNA-scan (Schattner *et al.*, 2005) en caso de los ARNt, y RNAMmer 1.2 Server (Lagesen *et al.*, 2007) para los ARNr, con sus opciones por defecto. La anotación de genes correspondientes a proteínas se realizó mediante una aproximación basada en homología, empleando alineamientos efectuados con el conjunto de programas BLAST (Altschul *et al.*, 1990, 1997).

4.5. Diseño de cebadores

En este trabajo se han realizado amplificaciones mediante PCR con diferentes objetivos, entre los que se encuentra la confirmación de los resultados de los ensamblajes bioinformáticos y la secuenciación de fragmentos de ADN mitocondrial. Los cebadores utilizados en este trabajo han sido sintetizados por un proveedor comercial (Sigma-Aldrich) y se muestran en la Tabla 1. El programa Primer3web en su versión 4.0.0 (<http://primer3.ut.ee/>; Untergasser *et al.*, 2012; Koressaar *et al.*, 2007) se utilizó para diseñar los cebadores, de modo que no formaran estructura secundaria y su temperatura de hibridación fuese adecuada. El programa Bandage se utilizó como ayuda visual para escoger las regiones a amplificar.

4.5. Amplificaciones mediante PCR

Las reacciones de PCR se prepararon en tubos *ependorf* de 0,2 mL de pared fina. Un conjunto de reacciones se llevó a cabo en un volumen de 20 μ L y otro en un volumen de 10 μ L. La composición de las reacciones con volumen de 20 μ L fue la siguiente: 6,3 μ L de agua desionizada, 2 μ L de DreamTaq Buffer 10x, 1,6 μ L de una mezcla de los 4 dNTP a una concentración de 10 mM (con cada uno de ellos a 2,5 mM), 4 μ L de cada uno de los cebadores *forward* y *reverse* correspondientes a 2,5 μ M cada uno, 0,1 μ L de la enzima DreamTaq a 5 U/ μ L y 2 μ L de ADN genómico como molde. En las reacciones de 10 μ L se utilizaron las mismas concentraciones finales. Los reactivos se mantuvieron en hielo durante la preparación de las mezclas de reacción.

Se empleó un termociclador *T100 Thermal Cycler* (Bio-Rad) para la incubación de las reacciones de PCR con el siguiente programa de amplificación. La primera etapa se programó a 95°C con una duración de 2 min. Tras esta, se realizaron 35 ciclos en las siguientes condiciones: (1) una

etapa de desnaturalización a 95°C 30 s; (2) una etapa de hibridación a una temperatura de 58°C, correspondiente a la temperatura de fusión (T_m) más baja de los cebadores presentes en la mezcla de reacción, durante 30 s; y (3) una etapa de síntesis a 72°C durante 2 min y 30 s, tiempo que se seleccionó considerando la tasa de síntesis de ADN de la enzima DreamTaq, que es de 1 kb cada 30 s. Por último, se añadió una etapa de extensión final a 72°C durante 5 min.

Tabla 1.- Nombre y secuencia de los cebadores empleados

Cebador	Secuencia nucleotídica (5'→ 3')	T _m (°C)	Letra asignada
Contig_1	GAATTGATTGAATGCGGGCA	67,9	A
Contig_1_R	GTGAAAGACGGTCAACAAGC	62,3	B
Contig_9_R	TGTCGGAGTGAAAGAGCGAA	66,1	C
Contig_23	TCAGGTAAATGCGCATTCT	64,4	D
Contig_25	GGCAAGTTCAGTTGTGAGG	64,8	E
Contig_28_R	GTCGAGAAGGGAGGTGTGAA	64,3	F
Contig_16_R	CTGGCAGCTATGAGTCTAGC	59,4	G
Contig_2_R	GTACCTCTCTAGCATCCCCT	58,8	H
Contig_1_E	ATAGCCGCTAAGAAACCACG	63,0	I
Contig_51	CACACGATATAGGAAGCGGA	62,6	J
Contig_31	AGAGGTTTCGTCGTACAGGTG	61,8	K
Contig_6_4_F	GCTCGTGTGGAACCTTGGAT	66,3	L
Contig_6_4_R	GTTCTGCTCCGTTTTGAG	63,6	M
Contig_4_1_F	GCGAAAAGGACTCTGACCTG	63,9	N
Contig_4_1_R	CGAGACGGCGTGACATACTA	63,9	O
Contig_29_23_F	TCAACGCCCTTACGAGGTAG	64,0	P
Contig_29_23_R	CCAGATGCTTAGCCAGAAGC	63,9	Q

4.6. Purificación de productos de PCR

Tras la correcta visualización en un gel de agarosa de los resultados esperados, se realizó la purificación de los productos de PCR. Para eliminar dímeros de cebadores y cebadores no incorporados que se encontraran en la mezcla de reacción, empleamos el kit *GeneMATRIX PCR / DNA Clean-Up Purification Kit* siguiendo el protocolo del fabricante (EURx). El ADN purificado se almacenó a -20 °C.

4.7. Secuenciación de los productos de PCR

Los productos de PCR purificados fueron secuenciados mediante el método de Sanger por la empresa STAB VIDA (Caparica, Portugal) en reacciones preparadas con los mismos cebadores utilizados en la amplificación.

5. Resultados

5.1. Ensamblaje *de novo* del genoma del cloroplasto del ajo

En este trabajo hemos llevado a cabo la secuenciación, ensamblaje *de novo* y anotación de ~300 kb del genoma mitocondrial del ajo, con el objetivo de contribuir al desarrollo de nuevas herramientas moleculares para la caracterización genómica de variedades de esta especie. La secuenciación del genoma mitocondrial se ha llevado a cabo a partir de una muestra de ADN total purificada a partir de tejido foliar de plantas jóvenes de ajo, que ha sido utilizada previamente con éxito para la secuenciación del genoma completo del cloroplasto (Gallego-Zaragoza, 2016). La secuenciación fue llevada a cabo por un secuenciador masivamente paralelo de Illumina HiSeq2500 mediante un protocolo de secuenciación por síntesis. Dicha secuenciación produjo 84.919.528 lecturas emparejadas (*paired-end reads*) de 101 nucleótidos de longitud (producto de secuenciar 42.459.814 fragmentos por ambos extremos).

Para realizar el ensamblaje del genoma mitocondrial, en primer lugar excluimos las lecturas derivadas del genoma del cloroplasto. Para ello, realizamos el alineamiento de las lecturas a este genoma mediante el programa Bowtie2, según se describe en la sección de Materiales y Métodos. Las parejas de lecturas no alineadas se ensamblaron a continuación mediante el programa Velvet (versión 1.2.10). Este programa ensambla las secuencias mediante una estrategia basada en la representación de las lecturas mediante un grafo de De Bruijn (Zerbino y Birney, 2008). El ensamblaje se realizó con las opciones descritas en la sección de Materiales y Métodos. La selección de umbrales de cobertura mínima (5) y máxima (54) tiene como objetivo excluir posibles secuencias remanentes de origen cloroplástico, que deberían encontrarse a una cobertura mucho mayor, y de origen nuclear, ya que la cantidad de secuencia obtenida equivale a una cobertura 0,5x del genoma nuclear del ajo. Nuestro ensamblaje produjo 4.169 cóntigos, con una longitud acumulada de 1.946.224 pb. La longitud del cóntigo más largo fue 79.866 pb, y el valor N50 del ensamblaje fue 1.373 pb. Estos cóntigos incorporaron 2.581.065 lecturas (3,04% del total), y deben estar enriquecidos en secuencias de origen mitocondrial.

Posteriormente, empleamos el programa Bandage para representar el grafo producido por Velvet, que permite visualizar los solapamientos existentes entre los cóntigos producidos en el ensamblaje *de novo*. Al representar los cóntigos mediante Bandage, apreciamos ramificaciones en el grafo, en las que dos o más cóntigos se unen a un tercero por el mismo lugar, posibilitando diversas reconstrucciones distintas de la secuencia (Figura 1). Para determinar la conectividad entre los cóntigos e identificar la secuencia correcta, descartando conexiones artefactuales de las secuencias producidas durante el ensamblaje, realizamos amplificaciones mediante PCR con cebadores diseñados a partir de las secuencias de los posibles recorridos a través del grafo (Tabla 1). Las secuencias utilizadas para el diseño de los cebadores se localizan en los extremos de cóntigos adyacentes, o en posiciones en las que un mismo cóntigo podría conectarse con varios diferentes, como se aprecia en la Figura 1.

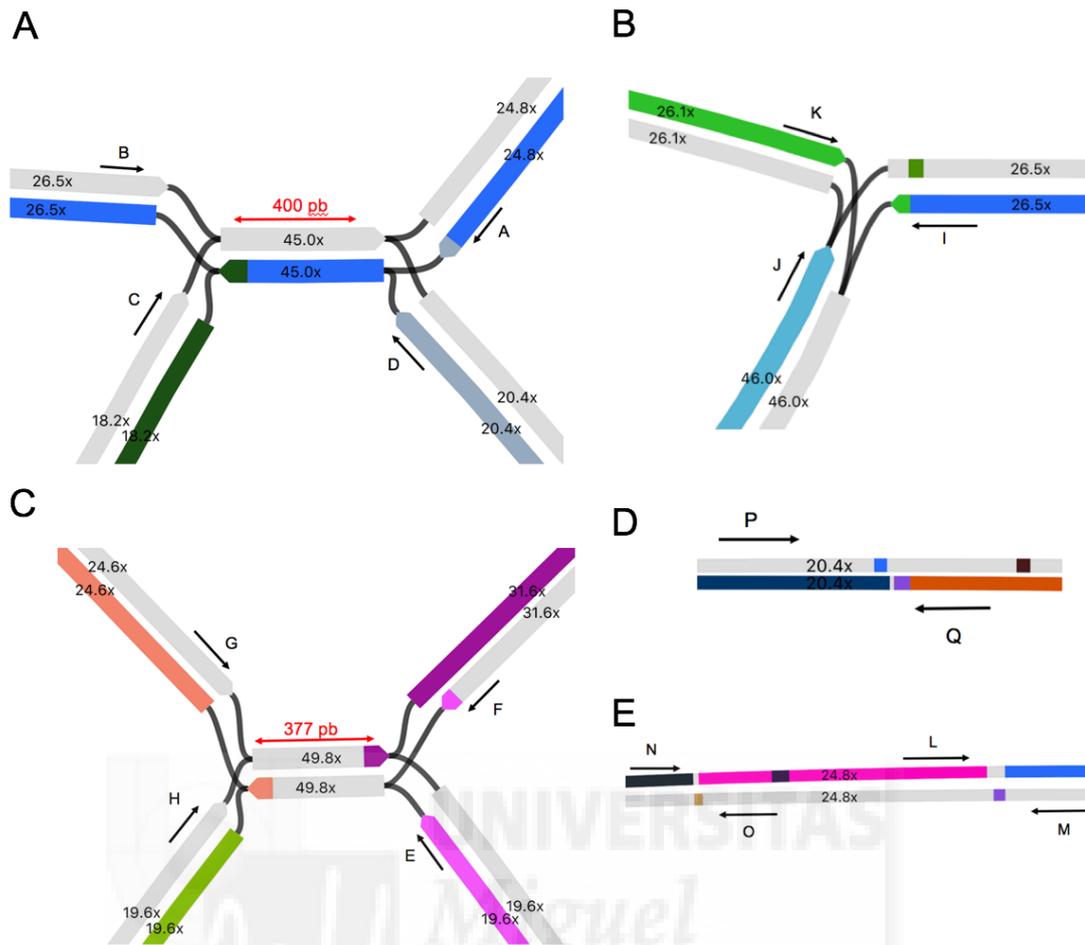


Figura 1.- Grafos representados por Bandage. Se representan en colores los cóntigos bicatenarios ensamblados que presentan similitud con el ensamblaje *de novo*. Se han destacado en las figuras A, B y C lugares de intersección entre cóntigos que dan lugar a diversos caminos posibles. Se muestran en las figuras D y E lugares de unión entre cóntigos separados por huecos susceptibles de ser rellenados. Las flechas negras representan la secuencia empleada para la construcción de los cebadores en dirección 5' a 3'. Las letras acompañando a cada flecha indican el nombre del cebador. La flecha bidireccional roja indica el tamaño de la secuencia común a los cóntigos. Los números que aparecen en cada cóntigo indican su cobertura.

Para el diseño de los cebadores se escogieron una cantidad de nucleótidos entre 500-700 pb aguas arriba o aguas abajo de los cóntigos involucrados en intersecciones, de tal manera que el producto de PCR a amplificar tuviera un tamaño de 1 kb aproximadamente. Una vez seleccionada la secuencia se empleó el programa Primer3web 4.0.0. para el óptimo diseño de los cebadores.

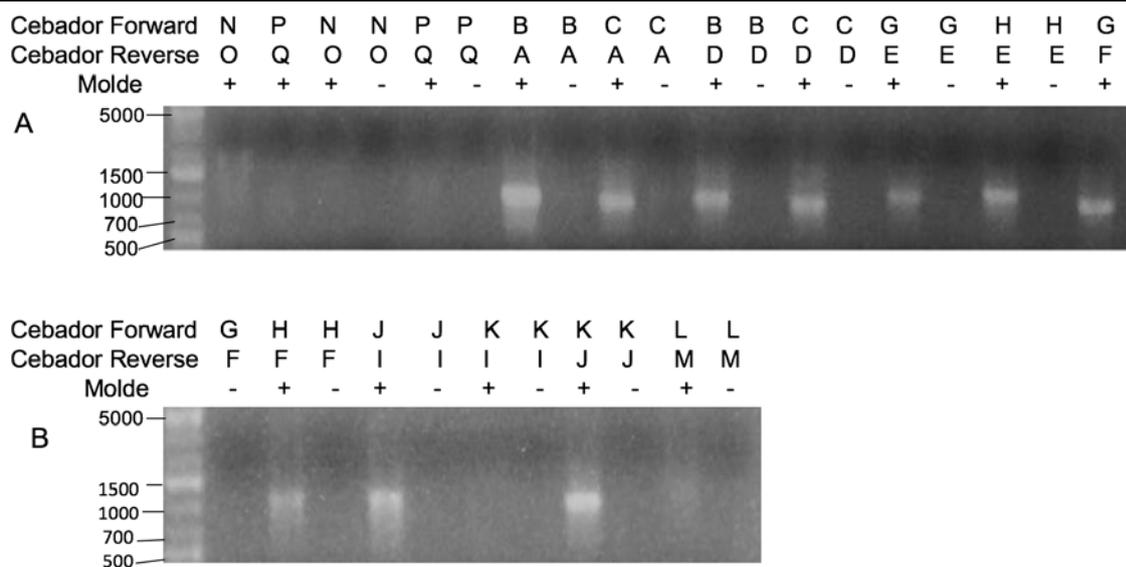


Figura 2.- Electroforesis en gel de agarosa de las amplificaciones por PCR realizadas con los diversos cebadores y contenido de las calles. La calle 1 de cada gel fue cargada con marcador de peso molecular y se indican los tamaños de las bandas en pb. El resto de las calles fueron cargadas con cada una de las combinaciones de cebadores que darían lugar a los caminos representados por Bandage en las intersecciones y con su respectivo control negativo carente de ADN genómico.

Los resultados de las amplificaciones por PCR se muestran en la Figura 2. Como se puede apreciar en los geles de electroforesis, se amplificaron todos los caminos posibles que unían a los cóntigos implicados en intersecciones mostradas en la Figura 1 excepto el camino unido por los cebadores I y K. De los lugares con secuencia entre cóntigos no resuelta por Bandage únicamente no se amplificó el camino que une los cebadores P y Q, indicando que no se encuentra en el genoma mitocondrial. El hecho de que se produzca amplificación en el resto de caminos verifica la existencia de los mismos en el genoma mitocondrial del ajo. Posteriormente, se purificaron los productos de ADN de la PCR y se corrieron en un gel para comprobar su calidad antes de mandarlos a secuenciar (Figura 3).

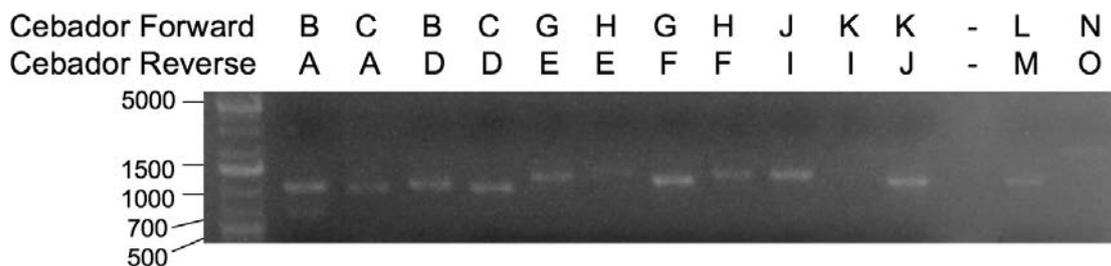


Figura 3.- Electroforesis en gel de agarosa de la purificación de ADN de las amplificaciones por PCR realizadas con los diversos cebadores. La calle 1 fue cargada con marcador de peso molecular y se indican los tamaños de las bandas en pb.

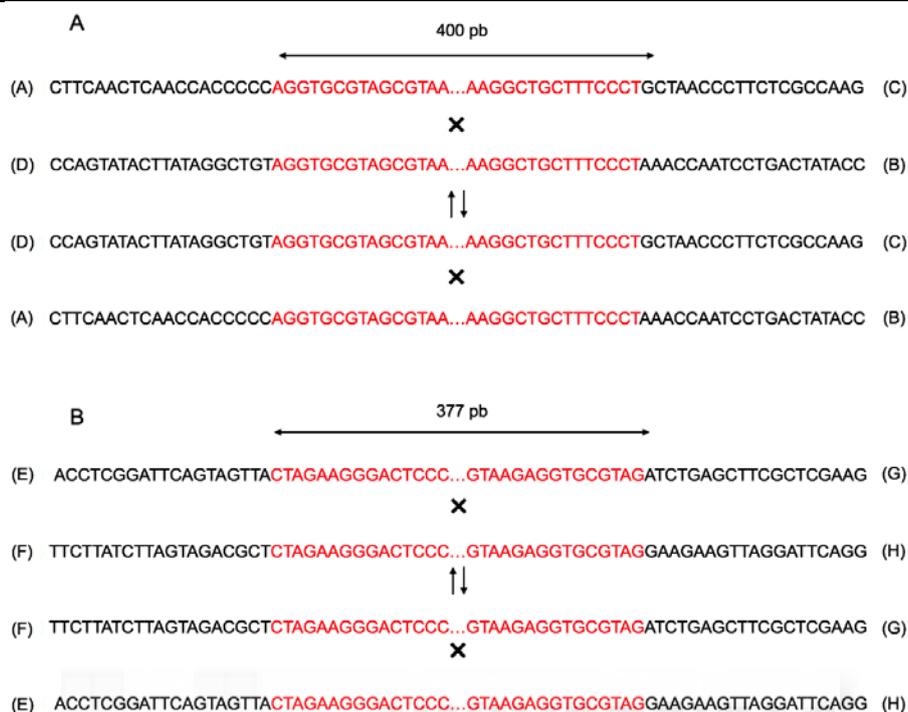


Figura 4.- Cántigos conteniendo secuencias repetitivas marcadas en color rojo. Se representan las posibles reorganizaciones encontradas en el genoma debidas a las recombinaciones causadas por las secuencias repetitivas. En la imagen A se ilustran las secuencias pertenecientes a los cántigos involucrados en la intersección de la Figura 1 A. En la imagen B se ilustran las secuencias pertenecientes a los cántigos involucrados en la intersección de la Figura 1 C. Se indica entre paréntesis el cebador que amplifica el cántigo correspondiente.

Tras la secuenciación por Sanger, se analizaron las secuencias obtenidas para delimitar con precisión las secuencias repetidas asociadas a bifurcaciones en el grafo de los cántigos (Figura 4). Hemos planteado que las secuencias repetitivas generan hechos de recombinación que dan lugar a diferentes reorganizaciones estructurales y subgenomas que pueden explicar la variedad de posibles caminos encontrados en los ensamblajes mitocondriales de la mitocondria del ajo.

En este trabajo se ha realizado una aproximación a una secuenciación, no hemos conseguido ensamblar el genoma mitocondrial completo por diversos motivos, a diferencia del genoma cloroplástico del ajo, ya realizado en el mismo laboratorio. Una de las razones es la menor cobertura del genoma mitocondrial frente al genoma del cloroplasto que aparece con una cobertura media superior a 1 orden de magnitud. Otro de los hechos que ha podido complicar el ensamblaje es que el ADN empleado ha sido ADN genómico, que presenta secuencias de núcleo, cloroplasto y mitocondria, en vez de ADN mitocondrial purificado. Otra característica que complica el ensamblaje es la compleja estructura que poseen las mitocondrias, debido a que poseen numerosas secuencias repetitivas que dan lugar a reorganizaciones genómicas, como se ha comentado en este trabajo.

5.2. Anotación del genoma

Hemos identificado los genes contenidos en el genoma mitocondrial del ajo mediante aproximaciones *ab initio* y basadas en homología. Hemos empleado la herramienta BLAST (Altschul *et al.*, 1990; 1997) para identificar secuencias codificantes parecidas a las presentes en otros genomas mitocondriales de angiospermas (Tabla 3). Para la identificación de genes correspondientes a moléculas de ARN no codificantes utilizamos el programa ARAGORN (Laslett y Canback, 2004) en caso de los ARNr y los programas RNAamner 1.2 Server (<http://www.cbs.dtu.dk/services/RNAamner/>; Lagesen *et al.*, 2007) y tRNA-scan (Schattner *et al.*, 2005) para la búsqueda de ARNt.

Si bien no hemos conseguido ensamblar el genoma mitocondrial completo, hemos identificado un total de 67 genes en los cóntigos que hemos caracterizado (Tabla 2). Este conjunto de genes incluye 40 genes que codifican proteínas implicadas en diversos procesos, 24 que corresponden a moléculas de ARNt y 3 a moléculas de ARNr, como se describe con mayor detalle en las secciones siguientes.

Tabla 2.- Genes identificados en la secuencia del genoma mitocondrial del ajo.

Complejos de la cadena	Complejo I NADH deshidrogenasa	<i>nad1, nad2, nad3, nad4, nad4L, nad5, nad6, nad7 y nad9</i>
	Complejo II succinato deshidrogenasa	<i>sdh4</i>
	Complejo III citocromo bc1	<i>cob</i>
	Complejo IV citocromo c oxidasa	<i>cox1, cox2 y cox3</i>
	Complejo V atp sintasa	<i>atp1, atp4, atp6, atp8 y atp9</i>
Control de la expresión génica	Genes de origen nuclear	<i>rpo</i>
	Subunidad grande ribosomal	<i>rpl2</i>
	Subunidad pequeña ribosomal	<i>rps2, rps12, rps14a y rps14b</i>
	Genes ARNr	<i>5sRNA, 16s RNA y 23 sRNA</i>
	Genes ARNt	<i>trnA-UGC, trnR-ACG, trnR-UCU, trnN-GUU, trnD-GUC (x2), trnC-GCA (x2), trnE-UUC (x4), trnQ-UUG, trnG-GCC, trnK-UUU, trnL-UAA, trnL-UAG, trnL-UAG, trnL-CAA, trnM-CAU (x4), trnF-GAA, trnP-UGG, trnS-CGA, trnS-UGA, trnS-GCU, trnT-GGU (x3), trnW-CCA, trnY-GUA (x5) y trnV-GAC</i>
Otros genes	Genes de origen cloroplástico	<i>atpA, lhbA, petG, psaJ, psbA, petL, rpl33 y ycf4</i>
	Biogenesis del citocromo c	<i>ccmB, ccmC, ccmFc, ccmFn1 y ccmFn2</i>
	Intrón maturasa	<i>matR</i>
	Transportador independiente Sec-Y	<i>mttB</i>

5.2.1. Genes necesarios para la expresión génica de la mitocondria

Un gran conjunto de los genes identificados posee funciones implicadas en la expresión de los genes del genoma mitocondrial. Sólo hemos encontrado 5 genes (de los 24 descritos en otros genomas mitocondriales de plantas) que codifican proteínas de las subunidades grande y pequeña del ribosoma. Este número es superior al hallado en el genoma mitocondrial de la cebolla, donde sólo se ha

encontrado *rps12*. En cuanto a los ARN de transferencia, hemos anotado 24, un número superior al encontrado en plantas como la cebolla, la colza o el cocotero y otras (Tabla 3). También hemos encontrado un fragmento con similitud a la ARN polimerasa (*rpo*) descrita en el genoma mitocondrial del cocotero (Aljohi *et al.*, 2016).

5.2.2. Genes implicados en la cadena de transporte de electrones

Otro gran conjunto de genes anotado en el genoma mitocondrial del ajo está implicado en los complejos de la cadena transportadora de electrones (Tabla 2). A pesar de no haber realizado un ensamblaje completo del genoma mitocondrial, la secuencia obtenida contiene todos los genes encontrados previamente en otros genomas mitocondriales de plantas para los complejos I, III, IV y V.

Destaca la presencia en el genoma mitocondrial del gen *sdh4*, que codifica una proteína del complejo II succinato deshidrogenasa, ya que está ausente en otros genomas mitocondriales de monocotiledóneas, como en *Oryza sativa*, donde se encuentra codificado en el núcleo.

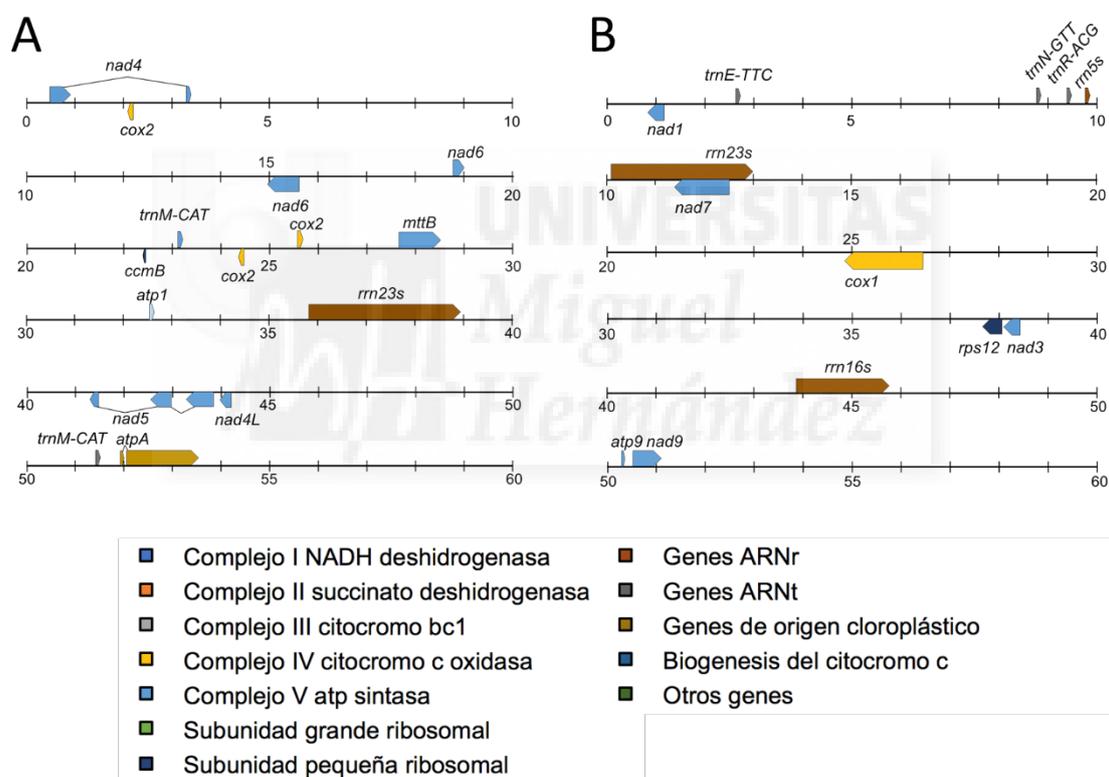


Figura 5.- Representación de dos cóntigos del genoma mitocondrial del ajo. Los genes representados por encima de cada línea se transcriben en sentido 5' a 3', y los genes representados por debajo se transcriben en sentido 5' a 3' de la hebra complementaria. Los arcos representan unen exones del mismo gen.

5.2.3. Otros genes

En el ensamblaje realizado del genoma mitocondrial del ajo hemos encontrado otros genes entre los que destacan los pertenecientes a la biogénesis del citocromo C y los de origen cloroplástico. Resulta de interés uno de los genes encargados de la biogénesis del citocromo C, *ccmFn*. Este gen se encuentra como tal en los genomas mitocondriales de todas las plantas monocotiledones secuenciadas

hasta la fecha excepto en cinco: *Philodendron oxycardium*, *Agave attenuata*, *Eichhornia crassipes*, *Cocos nucifera* y *Allium cepa*, en las que se encuentra dividido en dos genes, *ccmFn1* y *ccmFn2*. En las secuencias estudiadas, hemos encontrado los genes *ccmFn1* y *ccmFn2* separados en diferentes cóntigos, indicando que posiblemente este gen también se encuentre separado en esta especie.

En el ensamblaje realizado se han encontrado secuencias de origen cloroplástico. Los genomas mitocondriales y cloroplásticos suelen presentar cierta similitud debido a la transferencia de secuencias entre los mismos (Stern and Lonsdale, 1982; Stern and Palmer, 1984). Esta transferencia se estima que ha ocurrido durante 300 millones de años (Wang *et al.*, 2007), hecho que dificulta el ensamblaje del genoma mitocondrial a partir de una muestra de ADN total, en la que muchas lecturas podrían derivar igualmente de los genomas mitocondrial y cloroplástico. En nuestro ensamblaje, hemos determinado la presencia de ocho genes de origen cloroplástico (Tabla 2).



6. Discusión

En este Trabajo de Fin de Grado hemos efectuado la caracterización de una parte del genoma mitocondrial de una planta cultivada, el ajo (*Allium sativum*). Los cóntigos que hemos ensamblado suman un tamaño aproximado al del genoma mitocondrial de la cebolla (*Allium cepa*) 316.363 pb, que ha sido publicado recientemente (Kim *et al.*, 2016). La línea de investigación desarrollada en el laboratorio del tutor en la que se engloba este trabajo, tiene como propósito contribuir al conocimiento del genoma del ajo. Con el objetivo de desarrollar herramientas para el análisis genético y genómico de esta planta, se han realizado experimentos en el laboratorio del tutor, que incluyen la secuenciación, ensamblaje *de novo* y anotación del transcriptoma de varios tejidos del ajo y del cloroplasto (Gallego-Zaragoza, 2016). Este tipo de estudios sobre los genomas de la mitocondria y el cloroplasto suponen un punto de partida para futuros estudios sobre el genoma nuclear del ajo.

En este trabajo se ha empleado ADN genómico extraído de tejido foliar de plantas jóvenes de *Allium sativum* para realizar el ensamblaje de su genoma mitocondrial. El ADN genómico contiene genomas extranucleares como los del cloroplasto y la mitocondria. Es por ello, que se emplearon las lecturas obtenidas de la secuenciación de ADN genómico para ensamblar tanto el genoma cloroplástico como el genoma mitocondrial de *Allium sativum*. Las lecturas pertenecientes al genoma mitocondrial pudieron ser seleccionadas empleando filtros de cobertura, debido a que las lecturas pertenecientes a cloroplasto aparecían con una cobertura de 10 a 15 veces superior a las pertenecientes a la mitocondria. Diversos genomas mitocondriales de plantas han sido ensamblados empleando ADN genómico (Wang *et al.*, 2012; Aljohi *et al.*, 2016). A pesar de este hecho, la transferencia de material genético entre cloroplastos y mitocondrias complica la determinación de las secuencias de origen cloroplástico que se encuentran insertadas en el genoma mitocondrial.

La existencia de secuencias repetitivas también complica el ensamblaje final, es por ello que en otros trabajos se han empleado clones de gran tamaño como fásmidos y clones BAC para ensamblar secuencias repetitivas (Hecht *et al.*, 2011; Park *et al.*, 2013). Aun así, hemos ensamblado diversos cóntigos con un tamaño total de 300 kb aproximadamente. En estos cóntigos se han detectado secuencias repetitivas que originan diversos caminos entre cóntigos. Los diversos caminos entre los cóntigos han sido verificados mediante PCR y se han determinado las secuencias repetitivas con exactitud mediante secuenciación por Sanger. Hemos planteado la existencia de diversos subgenomas mitocondriales debidos a hechos de recombinación que ocurren mediante dichas secuencias repetitivas.

Hemos anotado 67 genes que se dividen en 40 genes que codifican proteínas involucradas en diferentes procesos, 24 que corresponden a moléculas de ARNt y 3 a moléculas de ARNr. Además, se ha detectado la separación del gen *ccmFn* en los genes *ccmFn1* y *ccmFn2*, estos genes también aparecen separados en la cebolla. Este resultado pendiente de ser comprobado mediante aproximaciones experimentales, podría indicar una posible separación de dicho gen en un ancestro común de la cebolla y el ajo. También se ha detectado la presencia del gen *sdh4*, carente en otros genomas mitocondriales de monocotiledoneas.

La anotación de los cóntigos ensamblados ha resultado en un número mayor de genes al descrito en el genoma mitocondrial de una planta del mismo género, la cebolla (*Allium cepa*). En el genoma mitocondrial de la cebolla completo se encontraron 25 genes que codifican proteínas (Tabla

2). Como se puede apreciar, este número es muy inferior al encontrado en otras plantas angiospermas. También hemos encontrado más ARN de transferencia en la secuencia que hemos anotado que los descritos en el genoma mitocondrial completo de la cebolla. En la Figura 5 hemos representado dos de los cóntigos más largos obtenidos con los genes que contienen.

Tabla 3.- Tamaños de genomas mitocondriales de angiospermas y número de genes

Especie	Tamaño (pb)	% de AT	Número de genes			
			Totales	Proteínas	ARNt	ARNr
<i>Allium cepa</i> ¹	316,363	-	43	25	16	3
<i>Arabidopsis thaliana</i> ²	366,924	55.2	131	117	21	3
<i>Beta vulgaris</i> ³	368,799	56.1	171	140	26	5
<i>Brassica napus</i> ⁴	221,853	54.8	100	79	17	3
<i>Cocos nucifera</i> ⁵	678,653	54.5	98	72	23	3
<i>Cucumis sativus</i> ⁶	1,684,592	44.4	63	37	23	3
<i>Nicotiana tabacum</i> ⁷	430,597	55	183	156	23	4
<i>Oryza sativa</i> ⁸	490,520	56.2	81	53	22	3
<i>Phoenix dactylifera</i> ⁹	715,001	54.8	90	43	23	3
<i>Vitis vinifera</i> ¹⁰	773,279	55.9	161	74	31	3
<i>Zea mays</i> ¹¹	569,630	56.1	213	163	33	4

Los tamaños correspondientes a cada genoma y su número de genes han sido tomados de las referencias indicadas mediante superíndices: ¹ (Kim *et al.*, 2016); ² (Unsel *et al.*, 1997); ³ (Kubo *et al.*, 2000); ⁴ (Handa, 2003); ⁵ (Aljohi *et al.*, 2016); ⁶ (Alverson *et al.*, 2011); ⁷ (Sugiyama *et al.*, 2005); ⁸ (Notsu *et al.*, 2002); ⁹ (Fang *et al.*, 2012); ¹⁰ (Goremykin *et al.*, 2009); ¹¹ (Clifton *et al.*, 2004).

La secuenciación del genoma mitocondrial del ajo se suma a los esfuerzos por desarrollar herramientas moleculares para el estudio de la androesterilidad que se observa en esta especie. En muchos casos, esta falta de fertilidad se debe a pautas de lectura abierta quiméricas que interfieren en la función de genes relacionados con el balance de energía en la mitocondria, entre los que se encuentran los genes implicados en la cadena de transporte de electrones (Tabla 2). Este hecho ha sido descrito también en otras plantas como *Nicotiana sylvestris*, *Brassica napus*, la petunia y el maíz (Rasmussen and Hanson, 1989; Pla *et al.*, 1995; Leino *et al.*, 2005; Wen and Chase, 1999).

7. Conclusiones y proyección futura

Este trabajo de Fin de Grado está enmarcado en la línea de investigación que se está desarrollando sobre el ajo en el laboratorio del tutor.

Para el ensamblaje *de novo* de las secuencias obtenidas, hemos utilizado el programa Velvet, con el que hemos obtenido numerosos cóntigos de diversos tamaños para su ensamblaje.

La suma de la longitud de los cóntigos producidos resulta en 300 kb aproximadamente. A pesar de no obtener un ensamblaje completo, este tamaño es muy similar al del genoma mitocondrial de la cebolla (316.636 pb) y además hemos anotado un mayor número de genes.

Hemos realizado experimentos para verificar las uniones entre diversos cóntigos, que además han puesto de manifiesto la existencia de secuencias repetitivas que dan lugar hechos de recombinación en el genoma mitocondrial del ajo.

Hemos anotado el genoma del cloroplasto mediante una combinación de métodos *ab initio* y basados en homología. La anotación nos ha permitido determinar que los cóntigos obtenidos poseen 67 genes funcionales, de los cuales 40 codifican proteínas, 24 ARN transferentes y 3 ARN ribosómicos.

Como sucede con otros genomas mitocondriales de plantas, el genoma mitocondrial del ajo contiene genes relacionados con la expresión génica propia y con la función respiratoria.

Como continuación a este trabajo, pretendemos purificar ADN mitocondrial. Esto consiste en el aislamiento de las mitocondrias del tejido vegetal del ajo y en la obtención de ADN mitocondrial libre de secuencias cloroplásticas y nucleares. Esto permitirá realizar una secuenciación, ensamblaje *de novo* y anotación completa del genoma mitocondrial del ajo. Las ventajas de esta aproximación frente a emplear ADN genómico son que todas las lecturas generadas de la secuenciación pertenecerán a mitocondria y se encontrarán a alta cobertura, evitando interferencias producidas por secuencias cloroplásticas o nucleares. Esta aproximación es especialmente útil cuando el genoma mitocondrial posee secuencias de origen cloroplástico como ocurre en muchas otras especies (Stern and Palmer, 1984). Tras ello, se procederá a la finalización de la anotación para posteriormente depositarla en la base de datos GenBank para que quede a disposición de la comunidad científica. Adicionalmente, se realizarán estudios filogenéticos y transcriptómicos con el objetivo de profundizar en el desarrollo de herramientas moleculares para el estudio de esta especie. Planeamos redactar un manuscrito en el que se detallen los resultados que hemos obtenido.

8. Bibliografía

- Aljohi, H.A., Liu, W., Lin, Q., Zhao, Y., Zeng, J., Alamer, A., Alanazi, I.O., Alawad, A.O., Al-Sadi, A.M., Hu, S. y Yu, J. (2016). Complete sequence and analysis of coconut palm (*Cocos nucifera*) Mitochondrial Genome. *PLoS One* **11**: e0163990.
- Allen, J.F. (1993). Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. *J. Theor. Biol.* **165**: 609-631.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. y Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. y Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389-3402.
- Alverson, A.J., Rice, D.W., Dickinson, S., Barry, K. y Palmer, J.D. (2011). Origins and recombination of the bacterial-sized multichromosomal mitochondrial genome of cucumber. *Plant Cell* **23**: 2499-2513.
- Andersson, S.G.E., Zomorodipour, A., Andersson, J.O., Sicheritz-Pontén, T., Alsmark, U.C.M., Podowski, R.M., Näslund, A.K., Eriksson, A.-S., Winkler, H.H. y Kurland, C.G. (1998). The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**: 133-140.
- Aono, N., Shimizu, T., Inoue, T. y Shiraishi, H. (2002). Palindromic repetitive elements in the mitochondrial genome of *Volvox*. *FEBS Letters* **521**: 95-99.
- Arimura, S., Yamamoto, J., Aida, G.P., Nakazono, M. y Tsutsumi, N. (2004). Frequent fusion and fission of plant mitochondria with unequal nucleoid distribution. *Proc. Natl. Acad. Sci. USA* **101**: 7805-7808.
- Arrieta-Montiel, M.P., Shedge, V., Davila, J., Christensen, A.C. y Mackenzie, S.A. (2009). Diversity of the Arabidopsis mitochondrial genome occurs via nuclear-controlled recombination activity. *Genetics* **183**: 1261-1268.
- Backert, S. y Börner, T. (2000). Phage T4-like intermediates of DNA replication and recombination in the mitochondria of the higher plant *Chenopodium album* (L.). *Curr. Genet.* **37**: 304-314.
- Backert, S., Dörfel, P., Lurz, R. y Börner, T. (1996). Rolling-circle replication of mitochondrial DNA in the higher plant *Chenopodium album* (L.). *Mol. Cell Biol.* **16**: 6285-6294.
- Backert, S., Nielsen, B.L. y Börner, T. (1997). The mystery of the rings: structure and replication of mitochondrial genomes from higher plants. *Trends Plant Sci.* **2**: 477-483
- Bendich, A.J. (1993). Reaching for the ring: the study of mitochondrial genome structure. *Curr. Genet.* **24**: 279-290.
- Benz, R. (1985). Porin from bacterial and mitochondrial outer membranes. *Crit. Rev. Biochem.* **19**: 145-190.
- Bereiter-Hahn, J. (1990). Behavior of Mitochondria in the Living Cell. *Int Rev. Cytol.* **122**: 1-63.
- Bhagyalakshmi, N., Thimmaraju, R., Venkatachalam, L., Chidambara, Murthy, K.N., Sreedhar, R.V. (2005). Nutraceutical applications of garlic and the intervention of biotechnology. *Crit. Rev. Food Sci.* **45**: 609-621.

- Bogenhagen, D.F. y Clayton, D.A. (2003). The mitochondrial DNA replication bubble has not burst. *Trends Biochem. Sci.* **28**: 357-360.
- Bogorad, L. (1975). Evolution of organelles and eukaryotic genomes. Separation of genes for chloroplast ribosomes in two genomes suggests principles of organelle biology. *Science* **188**: 891-898.
- Bonen, L. (2008). Cis- and trans-splicing of group II introns in plant mitochondria. *Mitochondrion* **8**: 26-34.
- Bonen, L. y Calixte, S. (2006). Comparative analysis of bacterial-origin genes for plant mitochondrial ribosomal proteins. *Mol. Biol. Evol.* **23**: 701-712.
- Boore, J.L. (1999). Animal mitochondrial genomes. *Nucleic Acids Res.* **27**: 1767-1780.
- Brennicke, A., Grohmann, L., Hiesel, R., Knoop, V. y Schuster, W. (1993). The mitochondrial genome on its way to the nucleus: different stages of gene transfer in higher plants. *FEBS Letters* **325**: 140-145.
- Burger, G., Gray, M.W., Forget, L. y Lang, B.F. (2013). Strikingly bacteria-like and gene-rich mitochondrial genomes throughout *Jakobid* protists. *Genome Biol Evol.* **5**: 418-438.
- Buso, G.S., Paiva, M.R., Torres, A.C., Resende, F.V., Ferreira, M.A., Buso, J.A., y Dusi, A.N. (2008). Genetic diversity studies of Brazilian garlic cultivars and quality control of garlic- clover production. *Genet. Mol. Res.* **7**: 534-541
- Carranza, G., Angius, F., Ilioaia, O., Solgadi, A., Miroux, B. y Arechaga, I. (2017). Cardiolipin plays an essential role in the formation of intracellular membranes in *Escherichia coli*. *Biochim. Biophys. Acta* **1859**: 1124-1132.
- Chase, C.D. (2007). Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. *Trends Genet.* **23**: 81-90.
- Chen, L. y Liu, Y.-G. (2014). Male sterility and fertility restoration in crops. *Annu. Rev. Plant Biol.* **65**: 579-606.
- Cheng, N., Lo, Y.-S., Ansari, M.I., Ho, K.-C., Jeng, S.-T., Lin, N.-S. y Dai, H. (2017). Correlation between mtDNA complexity and mtDNA replication mode in developing cotyledon mitochondria during mung bean seed germination. *New Phytol.* **213**: 751-763.
- Chipuk, J.E., Bouchier-Hayes, L. y Green, D.R. (2006). Mitochondrial outer membrane permeabilization during apoptosis: the innocent bystander scenario. *Cell Death Differ.* **13**: 1396-1402.
- Clement, F., Pramod, S.N. y Venkatesh, Y.P. (2010). Identity of the immunomodulatory proteins from garlic (*Allium sativum*) with the major garlic lectins or agglutinins. *Int Immunopharmacol.* **10**: 316-324.
- Clifton, S.W., Minx, P., Fauron, C.M.-R., Gibson, M., Allen, J.O., Sun, H., Thompson, M., Barbazuk, W.B., Kanuganti, S., Tayloe, C., Meyer, L., Wilson, R.K. y Newton, K.J. (2004). Sequence and comparative analysis of the maize NB mitochondrial genome. *Plant Physiol.* **136**: 3486-3503.
- Cunha, C.P., Hoogerheide, E.S., Zucchim, M.I., Monteiro, M. y Pinheiro, J.B. (2012). New microsatellite markers for garlic, *Allium sativum* (Alliaceae). *Am. J. Bot.* **99**: e17-19.

- Cupp, J.D. y Nielsen, B.L. (2014). Minireview: DNA replication in plant mitochondria. *Mitochondrion* **19**: 231-237.
- Dimauro, S. y Davidzon, G. (2005). Mitochondrial DNA and disease. *Ann. Med.* **37**: 222-232.
- Duncan, O., van der Merwe, M.J., Daley, D.O. y Whelan, J. (2013). The outer mitochondrial membrane in higher plants. *Trends Plant Sci.* **18**: 207-217.
- Eckardt, N.A. (2006). Cytoplasmic Male Sterility and Fertility Restoration. *Plant Cell* **18**: 515-517.
- Eja, M.E., Asikong, B.E., Abriba, C., Arikpo, G.E., Anwan, E.E. y Enyi-Idoh, K.H. (2007). A comparative assessment of the antimicrobial effects of garlic (*Allium sativum*) and antibiotics on diarrheagenic organisms. *Southeast Asian J. Trop. Med. Public Health.* **38**: 343-348.
- Fairman, J.W., Noinaj, N. y Buchanan, S.K. (2011). The structural biology of β -barrel membrane proteins: a summary of recent reports. *Curr. Opin. Struct. Biol.* **21**: 523-531.
- Fang, Y., Wu, H., Zhang, T., Yang, M., Yin, Y., Pan, L., Yu, X., Zhang, X., Hu, S., Al-Mssallem, I.S. y Yu, J. (2012). A complete sequence and transcriptomic analyses of date palm (*Phoenix dactylifera* L.) mitochondrial genome. *PLoS ONE* **7**: e37164.
- Fraser, F., Corstorphine, C.G. y Zammit, V.A. (1997). Topology of carnitine palmitoyltransferase I in the mitochondrial outer membrane. *Biochem. J.* **323**: 711-718.
- Fujii, S., Bond, C.S. y Small, I.D. (2011). Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. *Proc. Natl. Acad. Sci. USA* **108**: 1723-1728.
- Fujii, S. y Toriyama, K. (2009). Suppressed expression of retrograde-regulated male sterility restores pollen fertility in cytoplasmic male sterile rice plants. *Proc. Natl. Acad. Sci. USA* **106**: 9513-9518.
- Gabay-Laughnan, S. y Laughnan, J.R. (1994). Male Sterility and Restorer Genes in Maize. *The Maize Handbook* Springer Lab Manuals. (eds M. Freeling), & V. Walbot), pp. 418-423. Springer New York.
- Gallego Zaragoza, A. (2016). Secuenciación, ensamblaje de novo y anotación del genoma del cloroplasto del ajo (*Allium sativum*). Trabajo de Fin de Grado. Facultad de Ciencias Experimentales. Universidad Miguel Hernández de Elche.
- Galtier, N. (2011). The intriguing evolutionary dynamics of plant mitochondrial DNA. *BMC Biol.* **9**: 61.
- Giegé, P., Hoffmann, M., Binder, S. y Brennicke, A. (2000). RNA degradation buffers asymmetries of transcription in Arabidopsis mitochondria. *EMBO Rep.* **1**: 164-170.
- Giegé, P., Sweetlove, L.J., Cognat, V. y Leaver, C.J. (2005). Coordination of nuclear and mitochondrial genome expression during mitochondrial biogenesis in Arabidopsis. *Plant Cell* **17**: 1497-1512.
- Giovannoni, S.J., Turner, S., Olsen, G.J., Barns, S., Lane, D.J. y Pace, N.R. (1988). Evolutionary relationships among cyanobacteria and green chloroplasts. *J. Bacteriol.* **170**: 3584-3592.
- Gohil, V.M. y Greenberg, M.L. (2009). Mitochondrial membrane biogenesis: phospholipids and proteins go hand in hand. *J. Cell Biol.* **184**: 469-472.
- Goremykin, V.V., Salamini, F., Velasco, R., Viola, R. (2009). Mitochondrial DNA of *Vitis vinifera* and the issue of rampant horizontal gene transfer. *Mol. Biol. Evol.* **26**: 99-110.

- Gray, M.W. (1992). The endosymbiont hypothesis revisited. *Int. Rev. Cytol.* **141**: 233-357.
- Griparic, L. y Van Der Bliek, A.M. (2001). The many shapes of mitochondrial membranes. *Traffic* **2**: 235-244.
- Gualberto, J.M., Mileschina, D., Wallet, C., Niazi, A.K., Weber-Lotfi, F. y Dietrich, A. (2014). The plant mitochondrial genome: dynamics and maintenance. *Biochimie* **100**: 107-120.
- Gualberto, J.M. y Newton, K.J. (2017). Plant mitochondrial genomes: dynamics and mechanisms of mutation. *Annu. Rev. Plant Biol.* **68**: 225-252.
- Guo, J.-X. y Liu, Y.-G. (2012). Molecular control of male reproductive development and pollen fertility in rice. *J. Integr. Plant. Biol.* **54**: 967-978.
- Hammani, K. y Giegé, P. (2014). RNA metabolism in plant mitochondria. *Trends Plant Sci.* **19**: 380-389.
- Handa, H. (2003). The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Res.* **31**: 5907-5916.
- Hanson, M.R. y Bentolila, S. (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell* **16**: S154-S169.
- Hayat, S., Cheng, Z., Ahmad, H., Ali, M., Chen, X. y Wang, M. (2016). Garlic, from remedy to stimulant: evaluation of antifungal potential reveals diversity in phytoalexin allicin content among garlic cultivars; allicin containing aqueous garlic extracts trigger antioxidants in cucumber. *Front. Plant Sci.* **7**: 1235.
- Hazle, T. y Bonen, L. (2007). Comparative analysis of sequences preceding protein-coding mitochondrial genes in flowering plants. *Mol. Biol. Evol.* **24**: 1101-1112.
- Hecht, J., Grewe, F., Knoop, V. (2011) Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of *Selaginella moellendorffii* mitochondria: the root of frequent plant mtDNA recombination in early tracheophytes. *Genome Biol. Evol.* **3**: 344-358.
- Hengeveld, R. y Fedonkin, M.A. (2004). Causes and consequences of eukaryotization through mutualistic endosymbiosis and compartmentalization. *Acta Biotheoretica* **52**: 105-154.
- Herrmann, J.M. y Neupert, W. (2000). Protein transport into mitochondria. *Curr. Opin. Microbiol.* **3**: 210-214.
- Hoitzing, H., Johnston, I.G. y Jones, N.S. (2015). What is the function of mitochondrial networks? A theoretical assessment of hypotheses and proposal for future research. *BioEssays* **37**: 687-700.
- Horn, R., Gupta, K.J. y Colombo, N. (2014). Mitochondrion role in molecular basis of cytoplasmic male sterility. *Mitochondrion* **19**: 198-205.
- Horvath, S.E. y Daum, G. (2013). Lipids of mitochondria. *Prog. Lipid Res.* **52**: 590-614.
- Ikon, N. y Ryan, R.O. (2017). Cardiolipin and mitochondrial cristae organization. *Biochim. Biophys. Acta.* **1859**: 1156-1163.
- Ipek, M., Ipek, A., Almquist, S.G., y Simon, P.W. (2004). Demonstration of linkage and development of the first low-density genetic map of garlic, based on AFLP markers. *Theoretical Applied Genet.* **110**: 228-236

- Jo, M.H., Ham, I.K., Moe, K., Kwon, S.-W., Lu, F.-H., Park, Y.-J., Kim, W.S., Kim, M.K., Kim, T., y Lee, E.M. (2012). Classification of genetic variation in garlic (*Allium sativum* L.) using SSR markers. *Australian J. Crop Sci.* **6**: 625-631.
- John, P. y Whatley, F.R. (1975). *Paracoccus denitrificans* and the evolutionary origin of the mitochondrion. *Nature* **254**: 495-498.
- Johnston, I.G., Gaal, B., das Neves, R.P., Enver, T., Iborra, F.J. y Jones, N.S. (2012). Mitochondrial Variability as a Source of Extrinsic Cellular Noise. *PLoS Comput. Biol.* **8**: e1002416.
- Kamenetsky, R., Faigenboim, A., Shemesh Mayer, E., Ben Michael, T., Gershberg, C., Kimhi, S., Esqira, I., Rohkin Shalom, S., Eshel, D., Rabinowitch, H.D., y Sherman, A. (2015). Integrated transcriptome catalogue and organ-specific profiling of gene expression in fertile garlic (*Allium sativum* L.). *BMC Genomics* **16**: 12.
- Karnkowska, A., Vacek, V., Zubáčová, Z., Treitli, S.C., Petrželková, R., Eme, L., Novák, L., Žárský, V., Barlow, L.D., Herman, E.K., Soukal, P., Hroudová, M., Doležal, P., Stairs, C.W., Roger, A.J., Eliáš, M., Dacks, J.B., Vlček, Č. y Hampl, V. (2016). A eukaryote without a mitochondrial organelle. *Curr. Biol.* **26**: 1274-1284.
- Kaul, M. L. H. (1988). Male sterility in higher plants. *Springer-Verlag*.
- Keren, I., Bezawork-Geleta, A., Kolton, M., Maayan, I., Belausov, E., Levy, M., Mett, A., Gidoni, D., Shaya, F. y Ostersetzer-Biran, O. (2009). AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in Arabidopsis mitochondria. *RNA* **15**: 2299-2311.
- Keren, I., Tal, L., Des, F.-S., Araújo, W.L., Shevtsov, S., Shaya, F., Fernie, A.R., Small, I. y Ostersetzer-Biran, O. (2012). NMAT1, a nuclear-encoded maturase involved in the trans-splicing of *nad1* intron 1, is essential for mitochondrial complex I assembly and function. *Plant J.* **71**: 413-426.
- Kim, D.-G., Kang, M.J., Hong, S.S., Choi, Y.-H. y Shin, J.H. (2017a). Antiinflammatory effects of functionally active compounds isolated from aged black garlic. *Phytother. Res.* **31**: 53-61.
- Kim, S., Kim, D.-B., Jin, W., Park, J., Yoon, W., Lee, Y., Kim, S., Lee, S., Kim, S., Lee, O.-H., Shin, D. y Yoo, M. (2017b). Comparative studies of bioactive organosulphur compounds and antioxidant activities in garlic (*Allium sativum* L.), elephant garlic (*Allium ampeloprasum* L.) and onion (*Allium cepa* L.). *Nat. Prod. Res.* **5**: 1-5.
- Kim, B., Kim, K., Yang, T.-J. y Kim, S. (2016). Completion of the mitochondrial genome sequence of onion (*Allium cepa* L.) containing the CMS-S male-sterile cytoplasm and identification of an independent event of the *ccmFn* gene split. *Curr. Genet.* **62**: 873-885.
- Kim, S., Lee, E.-T., Cho, D.Y., Han, T., Bang, H., Patil, B.S., Ahn, Y.K. y Yoon, M.-K. (2009). Identification of a novel chimeric gene, *orf725*, and its use in development of a molecular marker for distinguishing among three cytoplasm types in onion (*Allium cepa* L.). *Theor. Appl. Genet.* **118**: 433-441.
- Knoop, V. (2004). The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Curr. Genet.* **46**: 123-139.

- Köhler, D., Schmidt-Gattung, S. y Binder, S. (2010). The DEAD-box protein PMH2 is required for efficient group II intron splicing in mitochondria of *Arabidopsis thaliana*. *Plant Mol. Biol.* **72**: 459-467.
- Kohn, A.B., Citarella, M.R., Kocot, K.M., Bobkova, Y.V., Halanych, K.M. y Moroz, L.L. (2012). Rapid evolution of the compact and unusual mitochondrial genome in the ctenophore, *Pleurobrachia bachei*. *Mol. Phylogenet. Evol.* **63**: 203-207.
- Koressaar T, Remm M (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**: 1289-91.
- Kourounakis, P.N. y Rekka, E.A. (1991). Effect on active oxygen species of alliin and *Allium sativum* (garlic) powder. *Res. Commun. Chem. Pathol. Pharmacol.* **74**: 249-252.
- Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**: 187-208.
- Kubo, T., Nishizawa, S., Sugawara, A., Itchoda, N., Estiati, A. y Mikami, T. (2000). The complete nucleotide sequence of the mitochondrial genome of sugar beet (*Beta vulgaris* L.) reveals a novel gene for tRNACys(GCA). *Nucleic Acids Res.* **28**: 2571-2576.
- Kühlbrandt, W. (2015). Structure and function of mitochondrial membrane protein complexes. *BMC Biol.* **13**.
- Kumar, R.A., Oldenburg, D.J. y Bendich, A.J. (2014). Changes in DNA damage, molecular integrity, and copy number for plastid DNA and mitochondrial DNA during maize development. *J. Exp. Bot.* **65**: 6425-6439.
- Ladoukakis, E. D., & Zouros, E. (2017). Evolution and inheritance of animal mitochondrial DNA: rules and exceptions. *J. Biol. Res.* **24**: 2.
- Lagesen, K., Hallin, P., Rødland, E.A., Staerfeldt, H.-H., Rognes, T. y Ussery, D.W. (2007). RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**: 3100-3108.
- Langmead, B. y Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**: 357-359.
- Langmead, B., Trapnell, C., Pop, M. y Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**: R25.
- Laslett, D. y Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* **32**: 11-16.
- Lavrov, D.V., Adamski, M., Chevaldonné, P. y Adamska, M. (2016). Extensive mitochondrial mRNA editing and unusual mitochondrial genome organization in *Calcarean* sponges. *Curr. Bio: CB.* **26**: 86-92.
- Lavrov, D.V. y Pett, W. (2016). Animal mitochondrial DNA as we do not know it: mt-genome organization and evolution in nonbilaterian lineages. *Genome Biol. Evol.* **8**: 2896-2913.
- Lavrov, D.V., Pett, W., Voigt, O., Wörheide, G., Forget, L., Lang, B.F. y Kayal, E. (2013). Mitochondrial DNA of *Clathrina clathrus* (Calcarea, Calcinea): six linear chromosomes, fragmented rRNAs, tRNA editing, and a novel genetic code. *Mol. Biol. Evol.* **30**: 865-880.
- Leaver, C.J. y Harmey, M.A. (1976). Higher-plant mitochondrial ribosomes contain a 5S ribosomal ribonucleic acid component. *Biochem. J.* **157**: 275-277.

- Leino, M., Landgren, M. y Glimelius, K. (2005). Alloplasmic effects on mitochondrial transcriptional activity and RNA turnover result in accumulated transcripts of Arabidopsis orfs in cytoplasmic male-sterile *Brassica napus*. *Plant J.* **42**: 469-480.
- Leister, D. (2005). Origin, evolution and genetic effects of nuclear insertions of organelle DNA. *Trends Genet.* **21**: 655-663.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.
- Li, S., Yang, D. y Zhu, Y. (2007). Characterization and use of male sterility in hybrid rice breeding. *J. Integr. Plant. Biol.* **49**: 791-804.
- Liere, K., Weihe, A. y Börner, T. (2011). The transcription machineries of plant mitochondria and chloroplasts: Composition, function, and regulation. *J. Plant Physiol.* **168**: 1345-1360.
- Liu, Y., Medina, R. y Goffinet, B. (2014). 350 my of mitochondrial genome stasis in mosses, an early land plant lineage. *Mol. Biol. Evol.* **31**: 2586-2591.
- Logan, D.C. (2010). The dynamic plant chondriome. *Semin. Cell. Dev. Biol.* **21**: 550-557.
- Manavski, N., Guyon, V., Meurer, J., Wienand, U. y Brettschneider, R. (2012). An essential pentatricopeptide repeat protein facilitates 5' maturation and translation initiation of *rps3* mRNA in maize mitochondria. *Plant Cell.* **24**: 3087-3105.
- Manchekar, M., Scissum-Gunn, K.D., Hammett, L.A., Backert, S. y Nielsen, B.L. (2009). Mitochondrial DNA recombination in *Brassica campestris*. *Plant. Sci.* **177**: 629-635.
- Manchekar, M., Scissum-Gunn, K., Song, D., Khazi, F., McLean, S.L. y Nielsen, B.L. (2006). DNA recombination activity in soybean mitochondria. *J. Mol. Biol.* **356**: 288-299.
- Margolin, W. (2005). FtsZ and the division of prokaryotic cells and organelles. *Nat. Rev. Mol. Cell Biol.* **6**: 862-871.
- Marintchev, A. y Wagner, G. (2004). Translation initiation: structures, mechanisms and evolution. *Q Rev Biophys.* **37**: 197-284.
- Martin, W. y Herrmann, R.G. (1998). Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.* **118**: 9-17.
- Mayer, E.S., Ben-Michael, T., Kimhi, S., Forer, I., Rabinowitch, H.D. y Kamenetsky, R. (2015). Effects of different temperature regimes on flower development, microsporogenesis and fertility in bolting garlic (*Allium sativum*). *Funct. Plant Biol.* **42**: 514-526.
- Mayr E. (1986). Joseph Gottlieb Kölreuter's contributions to biology. *Osiris* **2**: 135-76
- McKelvey, J.R. y Fioravanti, C.F. (1986). Localization of cytochrome C oxidase and cytochrome C peroxidase in mitochondria of *Hymenolepis diminuta* (Cestoda). *Comp. Biochem. Physiol. B.* **85**: 333-335.
- Meisinger, C., Sickmann, A. y Pfanner, N. (2008). The mitochondrial proteome: from inventory to function. *Cell.* **134**: 22-24.
- Mikaili, P., Maadirad, S., Moloudizargari, M., Aghajanshakeri, S. y Sarahroodi, S. (2013). Therapeutic uses and pharmacological properties of garlic, shallot, and their biologically active compounds. *Iran. J. Basic Med. Sci.* **16**: 1031-1048.

- Mileykovskaya, E. y Dowhan, W. (2009). Cardiolipin membrane domains in prokaryotes and eukaryotes. *Biochim. Biophys. Acta (BBA) - Biomembranes* **1788**: 2084-2091.
- Milne, I., Stephen, G., Bayer, M., Cock, P.J.A., Pritchard, L., Cardle, L., Shaw, P.D. y Marshall, D. (2013). Using Tablet for visual exploration of second-generation sequencing data. *Brief. Bioinform.* **14**: 193-202.
- Mohr, G. y Lambowitz, A.M. (2003). Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. *Nucleic Acids Res.* **31**: 647-652.
- Nakamoto, T. (2006). A unified view of the initiation of protein synthesis. *Biochem. Biophys. Res. Commun.* **341**: 675-678.
- Neves, R.P. das, Jones, N.S., Andreu, L., Gupta, R., Enver, T. y Iborra, F.J. (2010). Connecting variability in global transcription rate to mitochondrial variability. *PLOS Biol.* **8**: e1000560.
- Notsu, Y., Masood, S., Nishikawa, T., Kubo, N., Akiduki, G., Nakazono, M., Hirai, A. y Kadowaki, K. (2002). The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Mol. Genet. Genomics* **268**: 434-445.
- Oldenburg, D.J. y Bendich, A.J. (2001). Mitochondrial DNA from the liverwort *Marchantia polymorpha*: circularly permuted linear molecules, head-to-tail concatemers, and a 5' protein. *J. Mol. Biol.* **310**: 549-562.
- Oldenburg, D.J., Kumar, R.A. y Bendich, A.J. (2013). The amount and integrity of mtDNA in maize decline with development. *Planta* **237**: 603-617.
- Oldenburg, D.J., Rowan, B.A., Kumar, R.A. y Bendich, A.J. (2014). On the fate of plastid DNA molecules during leaf development: response to the Golczyk et al. Commentary. *Plant Cell.* **26**: 855-861.
- Osiro, D., Bernardes Filho, R., Assis, O.B.G., Jorge, L.A. de C. y Colnago, L.A. (2012). Measuring bacterial cells size with AFM. *Braz. J. Microbiol.* **43**: 341-347.
- Park, J.Y., Lee, Y., Lee, J., Choi, B., Kim, S., Yang, T. (2013) Complete mitochondrial genome sequence and identification of a candidate gene responsible for cytoplasmic male sterility in radish (*Raphanus sativus* L.) containing DCGMS cytoplasm. *Theor. Appl. Genet.* **126**: 1763-1774.
- Pett, W., Ryan, J.F., Pang, K., Mullikin, J.C., Martindale, M.Q., Baxevanis, A.D. y Lavrov, D.V. (2011). Extreme mitochondrial evolution in the ctenophore *Mnemiopsis leidyi*: insight from mtDNA and the nuclear genome. *Mitochondrial DNA.* **22**: 130-142.
- Pfaff, E., Klingenberg, M., Ritt, E. y Vogell, W. (1968). Korrelation des unspezifisch permeablen mitochondrialen Raumes mit dem „Intermembran-Raum". *Eur. J. Biochem.* **5**: 222-232.
- Pla, M., Mathieu, C., De Paepe, R., Chétrit, P. y Vedel, F. (1995). Deletion of the last two exons of the mitochondrial *nad7* gene results in lack of the NAD7 polypeptide in a *Nicotiana glauca* CMS mutant. *Mol. Gen. Genet.* **248**: 79-88.

- Porcelli, A.M., Ghelli, A., Zanna, C., Pinton, P., Rizzuto, R. y Rugolo, M. (2005). pH difference across the outer mitochondrial membrane measured with a green fluorescent protein mutant. *Biochem. Biophys. Res. Commun.* **326**: 799-804.
- Preuten, T., Cincu, E., Fuchs, J., Zoschke, R., Liere, K. y Boerner, T. (2010). Fewer genes than organelles: extremely low and variable gene copy numbers in mitochondria of somatic plant cells. *Plant J.* **64**: 948-959.
- Raczynska, K.D., Le Ret, M., Rurek, M., Bonnard, G., Augustyniak, H. y Gualberto, J.M. (2006). Plant mitochondrial genes can be expressed from mRNAs lacking stop codons. *FEBS Letters* **580**: 5641-5646.
- Rafelski, S.M. y Marshall, W.F. (2008). Building the cell: design principles of cellular architecture. *Nat. Rev. Mol. Cell Biol.* **9**: 593-602.
- Rappaport, L., Oliviero, P. y Samuel, J.L. (1998). Cytoskeleton and mitochondrial morphology and function. *Mol. Cell. Biochem.* **184**: 101-105.
- Rasmussen, J. y Hanson, M.R. (1989). A NADH dehydrogenase subunit gene is co-transcribed with the abnormal *Petunia* mitochondrial gene associated with cytoplasmic male sterility. *Mol. Gen. Genet.* **215**: 332-336.
- Rauwolf, U., Golczyk, H., Greiner, S. y Herrmann, R.G. (2010). Variable amounts of DNA related to the size of chloroplasts III. Biochemical determinations of DNA amounts per organelle. *Mol. Genet. Genomics.* **283**: 35-47.
- Rodríguez-Moreno, L., González, V.M., Benjak, A., Martí, M.C., Puigdomènech, P., Aranda, M.A. y Garcia-Mas, J. (2011). Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. *BMC Genomics.* **12**: 424.
- Roger, A.J., Svärd, S.G., Tovar, J., Clark, C.G., Smith, M.W., Gillin, F.D. y Sogin, M.L. (1998). A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: Evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. *Proc. Nat. Acad. Sci. USA* **95**: 229-234.
- Rossmannith, W. y Karwan, R.M. (1998). Characterization of human mitochondrial RNase P: novel aspects in tRNA processing. *Biochem. Biophys. Res. Commun.* **247**: 234-241.
- Sa, M. y Bl, N. (2017). Plant mitochondrial DNA. *Front. Biosci.* **22**: 1023-1032.
- Sato, N. (2001). Was the evolution of plastid genetic machinery discontinuous? *Trends Plant Sci.* **6**: 151-155.
- Schattner, P., Brooks, A.N. y Lowe, T.M. (2005). The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* **33**: W686-689.
- Schnable, P.S. y Wise, R.P. (1998). The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends Plant Sci.* **3**: 175-180.
- Schwartz, J.H., Meyer, R., Eisenstadt, J.M. y Brawerman, G. (1967). Involvement of N-formylmethionine in initiation of protein synthesis in cell-free extracts of *Euglena gracilis*. *J. Mol. Biol.* **25**: 571-574.

- Shedge, V., Davila, J., Arrieta-Montiel, M.P., Mohammed, S. y Mackenzie, S.A. (2010). Extensive rearrangement of the Arabidopsis mitochondrial genome elicits cellular conditions for thermotolerance. *Plant Physiol.* **152**: 1960-1970.
- Shemesh Mayer, E., Winiarczyk, K., Błaszczyk, L., Kosmala, A., Rabinowitch, H.D. y Kamenetsky, R. (2013). Male gametogenesis and sterility in garlic (*Allium sativum* L.): barriers on the way to fertilization and seed production. *Planta* **237**: 103-120.
- Shemesh-Mayer, E., Ben-Michael, T., Rotem, N., Rabinowitch, H.D., Doron-Faigenboim, A., Kosmala, A., Perlikowski, D., Sherman, A. y Kamenetsky, R. (2015). Garlic (*Allium sativum* L.) fertility: transcriptome and proteome analyses provide insight into flower and pollen development. *Front. Plant Sci.* **6**.
- Siekevitz, P. (1957). Powerhouse of the Cell. *Sci. Am.* **197**: 131-144.
- Singh, T.U., Kumar, D., Tandan, S.K. y Mishra, S.K. (2009). Inhibitory effect of essential oils of *Allium sativum* and *Piper longum* on spontaneous muscular activity of liver fluke, *Fasciola gigantica*. *Exp. Parasitol.* **123**: 302-308.
- Sloan, D.B., Alverson, A.J., Chuckalovcak, J.P., Wu, M., McCauley, D.E., Palmer, J.D. y Taylor, D.R. (2012). Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLOS Biol.* **10**: e1001241.
- Smith, D.R. (2012). Not seeing the genomes for the DNA. *Brief. Funct. Genomics* **11**: 289-290.
- Smith, D.R. (2016). The past, present and future of mitochondrial genomics: have we sequenced enough mtDNAs? *Brief. Funct. Genomics* **15**: 47-54.
- Smith, D.R. y Keeling, P.J. (2015). Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. *Proc. Natl. Acad. Sci. USA* **112**: 10177-10184.
- Smith, A.E. y Marcker, K.A. (1968). N-formylmethionyl transfer RNA in mitochondria from yeast and rat liver. *J. Mol. Biol.* **38**: 241-243.
- Stackebrandt, E., Murray, R.G.E. y Trüper, H.G. (1988). *Proteobacteria* classis nov., a Name for the Phylogenetic Taxon That Includes the "Purple Bacteria and Their Relatives." *Int. J. Syst. Evol. Microbiol.* **38**: 321-325.
- Stegemann, S. y Bock, R. (2006). Experimental Reconstruction of Functional Gene Transfer from the Tobacco Plastid Genome to the Nucleus. *Plant Cell.* **18**: 2869-2878.
- Stern, D.B. y Lonsdale, D.M. (1982). Mitochondrial and chloroplast genomes of maize have a 12-kilobase DNA sequence in common. *Nature* **299**: 698-702.
- Stern, D.B. y Palmer, J.D. (1984). Extensive and widespread homologies between mitochondrial DNA and chloroplast DNA in plants. *Proc. Natl. Acad. Sci. USA* **81**: 1946-1950.
- Sugiyama, Y., Watase, Y., Nagase, M., Makita, N., Yagura, S., Hirai, A. y Sugiura, M. (2005). The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: comparative analysis of mitochondrial genomes in higher plants. *Mol. Genet. Genomics* **272**: 603-615.
- Tait, S.W.G. y Green, D.R. (2012). Mitochondria and cell signalling. *J. Cell Sci.* **125**: 807-815.

- Taylor, F.J.R. (1974). Implications and Extensions of the Serial Endosymbiosis Theory of the Origin of Eukaryotes. *Taxon* **23**: 229-258.
- Tester, M. y Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science (New York, N.Y.)*. **327**: 818-822.
- Tsubura, A., Lai, Y.-C., Kuwata, M., Uehara, N. y Yoshizawa, K. (2011). Anticancer effects of garlic and garlic-derived compounds for breast cancer control. *Anticancer Agents Med. Chem.* **11**: 249-253.
- Unsold, M., Marienfeld, J.R., Brandt, P. y Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.* **15**: 57-61.
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012). Primer3 - new capabilities and interfaces. *Nucleic Acids Res.* **40**: e115.
- Uyttewaal, M., Arnal, N., Quadrado, M., Martin-Canadell, A., Vrielynck, N., Hiard, S., Gherbi, H., Bendahmane, A., Budar, F. y Mireau, H. (2008a). Characterization of *Raphanus sativus* pentatricopeptide repeat proteins encoded by the fertility restorer locus for Ogura cytoplasmic male sterility. *Plant Cell*. **20**: 3331-3345.
- Uyttewaal, M., Mireau, H., Rurek, M., Hammani, K., Arnal, N., Quadrado, M. y Giegé, P. (2008b). PPR336 is associated with polysomes in plant mitochondria. *J. Mol. Biol.* **375**: 626-636.
- Vedel, F. (Universite de P. 11, Pla, M., Vitart, V., Gutierrez, S., Chetrit, P. y De Paepe, R. (1994). Molecular basis of nuclear and cytoplasmic male sterility in higher plants. *Plant Physiol Biochem.* **32**: 601-618.
- Wang, Z.W., De Wang, C., Wang, C., Gao, L., Mei, S.Y., Zhou, Y., Xiang, C.P. y Wang, T. (2013). Heterozygous alleles restore male fertility to cytoplasmic male-sterile radish (*Raphanus sativus* L.): a case of overdominance. *J. Exp. Bot.* **64**: 2041-2048.
- Wang, D., Wu, Y.-W., Shih, A.C.-C., Wu, C.-S., Wang, Y.-N. y Chaw, S.-M. (2007). Transfer of chloroplast genomic DNA to mitochondrial genome occurred at least 300 MYA. *Mol. Biol. Evol.* **24**: 2040-2048.
- Weeber, E.J., Levy, M., Sampson, M.J., Anflous, K., Armstrong, D.L., Brown, S.E., Sweatt, J.D. y Craigen, W.J. (2002). The Role of Mitochondrial Porins and the Permeability Transition Pore in Learning and Synaptic Plasticity. *J. Biol. Chem.* **277**: 18891-18897.
- Wen, L.-Y. y Chase, C.D. (1999). Mitochondrial gene expression in developing male gametophytes of male-fertile and S male-sterile maize. *Sex Plant Reprod.* **11**: 323-330.
- Wen, L., Ruesch, K.L., Ortega, V.M., Kamps, T.L., Gabay-Laughnan, S. y Chase, C.D. (2003). A nuclear restorer-of-fertility mutation disrupts accumulation of mitochondrial ATP synthase subunit alpha in developing pollen of S male-sterile maize. *Genetics* **165**: 771-779.
- Wick R.R., Schultz M.B., Zobel J. & Holt K.E. (2015). Bandage: interactive visualisation of *de novo* genome assemblies. *Bioinformatics* **31**: 3350-3352.
- Wiemerslage, L. y Lee, D. (2016). Quantification of mitochondrial morphology in neurites of dopaminergic neurons using multiple parameters. *J. Neurosci. Methods* **262**: 56-65.
- Wise, R.P., Gobelman-Werner, K., Pei, D., Dill, C.L. y Schnable, P.S. (1999). Mitochondrial transcript processing and restoration of male fertility in T-cytoplasm maize. *J. Hered.* **90**: 380-385.

- Witt, K.E., Judd, K., Kitchen, A., Grier, C., Kohler, T.A., Ortman, S.G., Kemp, B.M. y Malhi, R.S. (2015). DNA analysis of ancient dogs of the Americas: identifying possible founding haplotypes and reconstructing population histories. *J. Hum. Evol.* **79**: 105-118.
- Woloszynska, M., Kmiec, B., Mackiewicz, P. y Janska, H. (2006). Copy Number of Bean Mitochondrial Genes Estimated by Real-time PCR does not Correlate with the Number of Gene Loci and Transcript Levels. *Plant Mol. Biol.* **61**: 1-12.
- Yang, D., Oyaizu, Y., Oyaizu, H., Olsen, G.J. y Woese, C.R. (1985). Mitochondrial origins. *Proc. Natl. Acad. Sci.* **82**: 4443-4447.
- Zewdie, Y., Havey M.J., Prince, J.P., y Jenderek, M.M. (2005). The first genetic linkages among expressed regions of the garlic genome. *J. American Soc. Horticultural Science* **130**: 569-574.
- Zerbino, D.R. y Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**: 821-829.
- Zeth, K. y Thein, M. (2010). Porins in prokaryotes and eukaryotes: common themes and variations. *Biochem. J.* **431**: 13-22.
- Zimorski, V., Ku, C., Martin, W.F. y Gould, S.B. (2014). Endosymbiotic theory for organelle origins. *Curr. Opin. Microbiol.* **22**: 38-48.

