



Universidad Nacional José Faustino Sánchez Carrión

Facultad de Ciencias

Escuela Profesional de Biología con Mención en Biotecnología

**Caracterización y comparación del genoma del cloroplasto de *Opuntia ficus-indica*
y *Opuntia pubescens*, y su relación filogenética con especies del género *Opuntia***

Tesis

Para optar el Título Profesional de Biólogo con Mención en Biotecnología

Autor

Cusy Mao Yupanqui Celestino

Asesora

Dra. Hermila Belba Díaz Pillasca

Univ. Nac. José Faustino Sánchez Carrión
FACULTAD DE CIENCIAS
Dra. HERMILA B. DÍAZ PILLASCA
DIRECTORA EP DE BIOLOGÍA CON MENCIÓN EN BIOTECNOLOGÍA

Huacho – Perú

2026



Reconocimiento – No comercial – Sin derivadas – Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Acreditación: conceder crédito, enlazar a la licencia e identificar modificaciones. Es factible ejecutarlo de manera razonable, pero no de tal manera que requiera el apoyo del licenciante. Sin ánimo de lucro: el material no puede ser empleado de manera comercial. Derivados inexistentes: El material modificado no podrá ser distribuido. Ausencia de restricciones: no se aplica. No existen restricciones jurídicas ni restricciones tecnológicas que restrinjan lo autorizado por la licencia.



UNIVERSIDAD NACIONAL
JOSÉ FAUSTINO SÁNCHEZ CARRIÓN
LICENCIADA

Resolución de Consejo Directivo N° 012-2020-SUNEDU/CD de fecha 27/01/2020

FACULTAD CIENCIAS




ESCUELA PROFESIONAL DE BIOLOGÍA CON MENCIÓN EN BIOTECNOLOGÍA

METADATOS

| DATO DEL AUTO: | | |
|--|------------|---|
| NOMBRES Y APELLIDOS | DNI | FECHA DE SUSTENTACIÓN |
| Cusy Mao Yupanqui Celestino | 75542172 | 02/03/2026 |
| DATOS DEL ASESORA: | | |
| NOMBRES Y APELLIDOS | DNI | CÓDIGO ORCID |
| Dra. Hermila Belba Diaz Pillasca | 15601607 | https://orcid.org/0000-0002-2491-3774 |
| DATOS DE LOS MIEMROS DE JURADO: | | |
| NOMBRES Y APELLIDOS | DNI | CÓDIGO ORCID |
| Dr. William Andrés Guzmán Sánchez | 06015776 | https://orcid.org/0000-0003-1424-4287 |
| Blgo. Cotos Durán Desiderio Elías | 07243334 | https://orcid.org/0000-0001-7456-5379 |
| Dr. Cipriano Bautista Johnny Gregorio | 15732636 | https://orcid.org/0000-0002-7239-4665 |

Cusy Mao Yupanqui Celestino

Caracterización y comparación del genoma del cloroplasto de Opuntia ficus-indica y Opuntia pubescens, y su relación filoge...

-  Quick Submit
-  Quick Submit
-  Facultad de Ciencias

Detalles del documento

Identificador de la entrega
trnoid::1:3459063832

Fecha de entrega
16 ene 2026, 9:39 p.m. GMT-5

Fecha de descarga
16 ene 2026, 9:55 p.m. GMT-5

Nombre del archivo
Borrador_de_tests_-_Mao.pdf

Tamaño del archivo
2.7 MB

90 páginas

21.787 palabras

126.917 caracteres



Página 2 de 98 · Descripción general de integridad

Identificador de la entrega trnoid::1:3459063832

14% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...




Filtrado desde el Informe

- ▶ Bibliografía
- ▶ Coincidencias menores (menos de 10 palabras)

Exclusiones

- ▶ N.º de coincidencias excluidas

Fuentes principales

- 13%  Fuentes de Internet
- 5%  Publicaciones
- 10%  Trabajos entregados (trabajos del estudiante)

Marcas de Integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

DEDICATORIA

A mis padres, Nudig Celestino Miranda y Jesús Marcos Yupanqui de la Cruz: les dedico esta tesis con profunda gratitud por su amor incondicional, su guía constante y su sacrificio con los que forjaron mi camino. Gracias por enseñarme que el esfuerzo y la disciplina son la base de toda conquista y por creer siempre en mis sueños. A mi abuela, Luzmila Miranda Castro gracias por tus lecciones de vida, a ella, que ha sido como una madre y me ha cuidado con ternura desde el año 2000 hasta hoy, y por despertaste mi curiosidad por la ciencia. A mi tía Bety Celestino Miranda, por su confianza y apoyo constante en la búsqueda de mis metas; sin su aliento nunca habría tenido el valor de emprender este camino exigente. Ha Abdías Primos Obregón, por ser mi respaldo desde el primer día, por acompañarme en cada obstáculo y perseverancia todo es posible. A mis amigos y profesores del colegio y la universidad, cuyo entusiasmo, motivación y apoyo inquebrantable me impulsaron a descubrir nuevos horizontes y a superarme continuamente.

Dedico este trabajo a ustedes con todo mi cariño y gratitud.

Mao Yupanqui

AGRADECIMIENTO

Quiero expresar mi más sincero agradecimiento a todas las personas e instituciones que hicieron posible la realización de esta tesis.

A la Dra. Marinoli Rivas Chamorro, mi directora de tesis: gracias por su asesoramiento constante, su paciencia y su exigencia académica. Sus sugerencias metodológicas, la revisión rigurosa de los resultados y su claridad para orientar cada etapa del trabajo fueron decisivas para elevar la calidad de esta investigación.

Al Laboratorio de Genómica y Bioinformática para la Biodiversidad (FCB-UNMSM), mi gratitud por abrirme sus puertas y facilitar el acceso a infraestructuras, reactivos y equipos permitió que los experimentos se realizaran con rigurosidad. Además, valoro el ambiente colaborativo que incentivó el intercambio de ideas y fortaleció mis habilidades en bioinformática.

Al Dr. Sergio Eduardo Contreras-Liza, agradezco su respaldo financiero y su orientación estratégica; su confianza en el proyecto fue fundamental para sostener las etapas más críticas de la investigación.

Ángel David Hernández Amasifuen, mi amigo y mentor por sus consejos prácticos y su enfoque positivo en situaciones difíciles me ayudaron a mantener el rumbo cuando los tiempos se tornaban complejos. Y Allison Esmeralda Lozano Galindo, por su compañía constante y apoyo emocional; su empatía y ánimo fueron un pilar indispensable durante todo el proceso.

No sólo no hubiéramos sido nada sin ustedes, sino con toda la gente que estuvo a nuestro alrededor desde el comienzo; algunos siguen hasta hoy. **¡Gracias.... totales!**

Mao Yupanqui

ÍNDICE

| | |
|--|----|
| INTRODUCCIÓN | 1 |
| CAPITULO I. PLANTEAMIENTO DEL PROBLEMA | 3 |
| 1.1. Descripción de la realidad problemática | 3 |
| 1.2. Formulación del problema | 4 |
| 1.3. Objetivo de la investigación | 5 |
| 1.4. Justificación de la investigación | 6 |
| 1.5. Delimitación del estudio | 7 |
| 1.6. Viabilidad del estudio | 8 |
| CAPITULO II. MARCO TEÓRICO | 8 |
| 2.1. Antecedentes de la investigación | 8 |
| 2.2. Bases teóricas | 13 |
| 2.3. Definición de términos básicos | 25 |
| 2.4. Hipótesis de investigación | 27 |
| 2.5. Operacionalización de variables e indicadores | 28 |
| CAPITULO III. METODOLOGÍA | 28 |
| 3.1. Diseño metodológico | 28 |
| 3.2. Población y muestra | 29 |
| 3.3. Técnicas de recolección de datos | 29 |
| 3.4. Técnicas para el procesamiento de información | 38 |
| 3.5. Matriz de consistencia | 39 |
| CAPITULO IV. RESULTADOS | 40 |
| 4.1. Extracción y purificación de ADN total | 40 |
| 4.2. Análisis de calidad y filtrado de lecturas | 42 |
| 4.3. Ensamblaje y anotación de los genomas de cloroplasto | 43 |
| 4.4. Tipos, estructuras y patrones evolutivos de los genomas de cloroplasto 53 | |
| 4.5. Comparación de los genomas de los cloroplastos | 54 |
| 4.6. Análisis de secuencias repetitivas | 56 |
| 4.6.1. Repeticiones de secuencias simples (SSR) | 56 |
| 4.6.2. Secuencias de repetición largas | 58 |

| | |
|--|-----------|
| 4.7. Análisis filogenético del género <i>Opuntia</i>..... | 60 |
| CAPITULO V. DISCUSIÓN..... | 62 |
| 5.1. Discusión de resultados..... | 62 |
| CAPITULO VI. CONCLUSIÓN Y RECOMENDACIONES..... | 65 |
| 6.1. Conclusión..... | 65 |
| 6.2. Recomendación..... | 65 |
| CAPITULO VII. REFERENCIAS BIBLIOGRÁFICAS | 66 |

INDICE DE FIGURAS

| | |
|--|----|
| Figura 1. Planta entera de <i>Opuntia ficus-indica</i> mostrando las principales características morfológicas..... | 14 |
| Figura 2. Un diagrama que muestra la estructura general de los cloroplastos de plantas superiores..... | 18 |
| Figura 3. Presentación clásica del mapa del genoma cloroplastidial, donde se muestran las cuatro regiones: Dos secuencias de repeticiones invertidas (IR), una secuencia larga de copia única (LSC) y secuencia de corta copia única (SSC)..... | 19 |
| Figura 4. Diagrama de flujo del trabajo del proyecto de investigación..... | 30 |
| Figura 5. (A) Muestras de <i>O. ficus-indica</i> código de acceso PTI 003 INIA. (B) Colecta de <i>O. pubescens</i> | 31 |
| Figura 6. Electroforesis de ADN extraído de <i>Opuntia ficus-indica</i> y <i>Opuntia pubescens</i> en un gel de agarosa 1% (M: marcador de peso molecular de 1kb, OP1: <i>Opuntia ficus-indica</i> , OF1: <i>Opuntia ficus-indica</i>), las muestras OP1B y OF1B son respaldo de muestras OP1 y OF1..... | 41 |
| Figura 7. Calidad de las lecturas forward (A) y reverse (B) del ADN secuenciado de <i>Opuntia ficus-indica</i> | 44 |
| Figura 8. Evaluación de la calidad de las lecturas forward (A) y reverse (B) del ADN secuenciado de <i>Opuntia ficus-indica</i> tras el proceso de filtrado, donde se eliminaron aquellas con un valor de calidad inferior a $Q = 33$ | 45 |
| Figura 9. Calidad de las lecturas forward (A) y reverse (B) del ADN secuenciado de <i>Opuntia pubescens</i> | 46 |
| Figura 10. Evaluación de la calidad de las lecturas forward (A) y reverse (B) del ADN secuenciado de <i>Opuntia pubescens</i> tras el proceso de filtrado, donde se eliminaron aquellas con un valor de calidad inferior a $Q = 33$ | 47 |
| Figura 11. Mapas de versión final de los genomas cloroplastidiales de <i>Opuntia ficus-indica</i> (A) y <i>Opuntia pubescens</i> (B). Se muestra las 4 regiones de genoma: LSC, SSC, IRA e IRB. Los genes ubicados en la parte interna del círculo se transcriben en sentido horario, mientras que los localizados en la parte externa lo hacen en sentido antihorario. Los genes están representados con colores que indican sus respectivos grupos funcionales..... | 49 |
| Figura 12. Configuraciones cuatripartitas de los genomas cloroplastidiales linealizados en el género <i>Opuntia</i> . Se muestran los cuatro tipos (1–4) identificados de arreglo LSC–IRb–SSC–IRa. Cada gen aparece como un recuadro situado por encima o por debajo de la línea central según su dirección de transcripción (+/–). Las regiones invertidas (IRa e IRb) están coloreadas en rojo. El triángulo punteado señala la reubicación del bloque génico trnA-UGC – trnV-GAC en la configuración del Tipo 3..... | 53 |
| Figura 13. Visualización de 18 genomas de los cloroplastos de género <i>Opuntia</i> , mediante mVISTA. Cada banda horizontal representa uno de los genomas analizados. Rojo: regiones no codificantes. Azul: regiones codificantes. Blanco: ausencia de alineación bajo el umbral mínimo de identidad del 70% similitud. Las flechas verdes son genes que se transcriben a derecha (+) y a las izquierdas (-). Amarillo: regiones comparadas con alta diversidad..... | 55 |

Figura 14. Distribución, motivos, tipo de repeticiones de secuencias simples (SSR) en los genomas de cloroplastos de *O. ficus-indica* y *O. pubescens*, así como en otras especies de *Opuntia*. (A) Número de SSR totales por especies. (B) Frecuencia de motivos SSR en diferentes tipos de clases de repetición identificados. (C) Número de tipos de SSR identificados por especies..... 57

Figura 15. Secuencias de repetición largas los genomas de cloroplastos de *O. ficus-indica* y *O. pubescens*, así como en otras especies de *Opuntia*. (A) Número de repeticiones total por especie. (B) Número de diferentes tipos de repeticiones. (C) Longitud de la secuencia de repeticiones..... 59

Figura 16. Árbol filogenético de 39 especies para determinar ubicación *O. ficus-indica* y *O. pubescens*. El árbol fue construido empleando 3 métodos de diferentes: Máxima Verosimilitud (ML), Máxima Parsimonia (MP) e Inferencia Bayesiana (BI). Los asteriscos en negro representan 100 % soporte de bootstrap [BP] e 1.0 probabilidad posterior [PP]..... 61

INDICE DE TABLAS

| | |
|---|----|
| Tabla 1. Tabla de la operacionalización de las variables estudiadas..... | 28 |
| Tabla 2. Ubicación de referencias para <i>Opuntia pubescens</i> | 31 |
| Tabla 3. Números de acceso de GenBank de referencias e incluyendo 2 genomas de <i>Opuntia</i> | 36 |
| Tabla 4. En este punto se aprecia el cuadro de la operacionalización de las variables estudiadas..... | 39 |
| Tabla 5. Volumen, concentración y calidad del ADN extraído de cladodios jóvenes de <i>Opuntia ficus-indica</i> y <i>Opuntia pubescens</i> | 40 |
| Tabla 6. Datos de cantidad y calidad obtenidos por el secuenciamiento de ADN por el laboratorio externo..... | 42 |
| Tabla 7. Principales características del genoma cloroplastidial ensamblado a partir del ADN obtenido de cladodios jóvenes mediante técnicas de extracción y secuenciación de alto rendimiento..... | 50 |
| Tabla 8. Composición genética del genoma cloroplastidial de <i>Opuntia ficus-indica</i> y <i>Opuntia pubescens</i> | 51 |

RESUMEN

El género *Opuntia*, es de gran importancia socioeconómica y ecológica en el Perú, destacando *Opuntia ficus-indica* por su cultivo y la producción de cochinilla, un tinte natural fundamental para la industria. A pesar de su relevancia, los estudios genómicos en *Opuntia* en Perú son escasos y se han enfocado principalmente en caracteres morfológicos, lo cual es insuficiente para delimitar especies y genotipos. Avances en las tecnologías de secuenciación de próxima generación (NGS) permiten analizar genomas de cloroplastos, que son una herramienta valiosa para la conservación y el mejoramiento genético, dado su tamaño relativamente pequeño, estructura simple y alta conservación en plantas. En este contexto, el objetivo general de esta investigación fue caracterizar y comparar el genoma cloroplástico de *Opuntia ficus-indica* y *Opuntia pubescens*, y reconstruir su filogenia en relación con otras especies del género *Opuntia*. Los resultados de la investigación incluyen el ensamblaje y la anotación exitosa de los genomas cloroplásticos de *O. ficus-indica* (152,730 pb) y *O. pubescens* (150,927 pb), ambos con alta calidad y cobertura de secuenciación. Se identificaron 139 genes en *O. ficus-indica* y 137 en *O. pubescens*, que incluyen genes codificantes de proteínas, ARNt y ARNr, así como pseudogenes. El análisis filogenético confirmó que el género *Opuntia* es monofilético con alto soporte, agrupando a *O. pubescens* con especies centroamericanas y a *O. ficus-indica* con otras especies nativas de México. En conclusión, la caracterización de los genomas cloroplásticos de *Opuntia ficus-indica* y *Opuntia pubescens* proporcionó una visión integral de la diversificación del género al identificar sus tipos y estructuras genómicas. Los genomas de cloroplastos ofrecen una base sólida para futuros estudios de diversidad genética y para el desarrollo de programas de mejoramiento en estas especies.

Palabras clave: *Opuntia ficus-indica*, genoma del cloroplasto, reconstrucción filogenética.

ABSTRACT

The *Opuntia* genus, known as prickly pear, is of great socioeconomic and ecological importance in Peru, with *Opuntia ficus-indica* standing out for its cultivation and the production of cochineal, a natural dye that is essential for industry. Despite its relevance, genomic studies on *Opuntia* in Peru are scarce and have focused mainly on morphological characteristics, which is insufficient to delimit species and genotypes. Advances in next-generation sequencing (NGS) technologies allow the analysis of chloroplast genomes, which are a valuable tool for conservation and genetic improvement, given their relatively small size, simple structure, and high conservation in plants. In this context, the overall objective of this research was to characterize and compare the chloroplast genomes of *Opuntia ficus-indica* and *Opuntia pubescens*, and to reconstruct their phylogeny in relation to other species of the genus *Opuntia*. The results of the research include the successful assembly and annotation of the chloroplast genomes of *O. ficus-indica* (152,730 bp) and *O. pubescens* (150,927 bp), both with high quality and sequencing coverage. A total of 139 genes were identified in *O. ficus-indica* and 137 in *O. pubescens*, including protein-coding genes, tRNA, and rRNA genes, as well as pseudogenes. Phylogenetic analysis confirmed that the genus *Opuntia* is monophyletic with high support, grouping *O. pubescens* with Central American species and *O. ficus-indica* with other species native to Mexico. In conclusion, the characterization of the chloroplast genomes of *Opuntia ficus-indica* and *Opuntia pubescens* provided a comprehensive view of the diversification of the genus by identifying its genomic types and structures. Chloroplast genomes provide a solid basis for future studies of genetic diversity and for the development of breeding programs in these species.

Keywords: *Opuntia ficus-indica*, chloroplast genome, phylogenetic reconstruction.

INTRODUCCIÓN

El género *Opuntia* pertenece a la familia Cactácea, es nativo de América y se ha distribuido ampliamente en las regiones de todo el mundo. Debido a su eficiente uso del agua, se adapta perfectamente en regiones áridas y semiáridas. Actualmente, el género *Opuntia* forma parte del sistema agrícola y del entorno paisajístico de muchos países, principalmente de México, Argentina, Bolivia, Perú, Brasil, Estados Unidos, Italia, España, Marruecos, Túnez, Israel y Sudáfrica. En los que se encuentra disponible como: plantaciones comerciales, huertos familiares y comunidades silvestres (Mohamed et al., 1996; Sáenz et al., 2004; Mannai et al., 2024).

En el Perú, los matorrales de *Opuntia* forman parte de uno de los socioecosistemas más importantes de la costa y valles andinos, proporcionando productos diversos como frutas, compuestos químicos farmacéuticos, forrajes suculentos para alimento de animales y materia prima para la fabricación de trabajos ornamentales. Históricamente, los matorrales de *Opuntia* han sido de gran importancia para la crianza de *Dactylopius coccus* (cochinilla del carmín) para la obtención de ácido carmínico, un tinte natural, empleado en la industria alimentaria, farmacéutica y textil (Guevara et al., 1996; Rodríguez & Pascual, 2004; Yanqui et al., 2024). Actualmente, Perú es el mayor productor comercial del mundo de cochinilla del carmín, empleando entre 60% y 70% de la superficie de los matorrales de *Opuntia* para satisfacer la demanda mundial (Almora et al., 2018; Berhe et al., 2023).

En las dos últimas décadas, se han visto avances revolucionarios en las tecnologías de secuenciación masiva de ADN, sobre todo, respecto a la aparición de las técnicas de secuenciación de próxima generación (NGS: Next-Generation Sequencing), permitiendo secuenciar millones de bases de ADN en una sola ronda, ya que, cada vez estas plataformas NGS se vuelven más óptimas y asequibles (Egan et al., 2012). A medida

que los costos y capacidades de estas tecnologías continúan mejorando, se seguirá facilitando la aparición y el avance de nuevos campos de estudio, resolviendo diversas cuestiones que nunca habrían sido posibles de responder antes, para así generar conocimiento en el área de la Biología, pero especialmente de la genómica (Sanger et al., 1977).

Las herramientas NGS disponibles en mercado son lo suficientemente capaces para proporcionar análisis de alta resolución de genomas de plantas. Esta herramienta ha permitido generar grandes cantidades de datos secuenciados, facilitando la elaboración del perfil genómico (Singh et al., 2015; Nguyen et al., 2019). Hasta la fecha, se ha secuenciado aproximadamente 800 genomas completos y más de 4 235 genomas de cloroplastos completos en plantas, los cuales están publicados oficialmente y son de acceso público. La información genómica varía entre especies, poblaciones e incluso individuos (Zhou et al., 2021; Hong et al., 2023; Song et al., 2023). Por esta razón, la abundante información sobre los genomas de cloroplastos de plantas obtenida, a lo largo del tiempo, mediante el uso de herramientas de NGS ha permitido alcanzar una comprensión detallada de la estructura del genoma de los cloroplastos, así como de su composición genética, funcionalidad y procesos evolutivos en diversas especies. Este avance también ha facilitado el progreso en ingeniería genética en plantas superiores, lo que ha llevado a la obtención de cultivos genéticamente mejorados. Esta información crucial ha sido fundamental para impulsar la investigación y el desarrollo en la agricultura (Green, 2011; L. Gao et al., 2010; Bansal & Saha, 2012; Llorente et al., 2021).

CAPITULO I. PLANTEAMIENTO DEL PROBLEMA

1.1. Descripción de la realidad problemática

El género *Opuntia* cuenta con más de 300 especies existentes, destacando en el Perú *Opuntia macbridei*, *Opuntia ficus-indica*, *opuntia pubescens* y *opuntia quitensis*, siendo *ficus-indica* una de las especies con mayor presencia, ventajas comparativas y competitivas de importancia socioeconómica para los departamentos productores como Ayacucho, Cusco, Huancavelica, Apurímac y Lima. Además. Esta especies cuenta con cinco variedades: blanca, amarilla, morada, colorada y forrajera. Durante el año 2020, la producción de los frutos de *Opuntia ficus-indica* (tuna) superó las 62 000 toneladas, provenientes de una superficie cosechada de 12 000 hectáreas, con un rendimiento promedio de 5.3 toneladas por hectárea. Por otro lado, la producción de tuna en el Perú está destinada principalmente al consumo local, ya que más del 98% de la producción total se destina al mercado fresco, siendo Lima el principal mercado, donde las variedades más demandadas son la blanca, la amarilla y la morada (Arenas et al., 2017; Almora et al., 2018; MIDAGRI, 2022).

A pesar de su importancia, la Opuntia en el Perú ha sido poco estudiada. Hasta la fecha de elaboración de este texto no existen investigaciones sobre la filogenia, la delimitación de especies ni la diversidad genética de Opuntia que empleen códigos de barras de ADN. No obstante, se han realizado estudios parciales, como el libro Opuntia del Perú de Ostolaza (2019), que describe las características morfológicas de diversas especies, y la investigación de Tineo (2019) sobre caracterización morfológica y diversidad genética, en la que se identificaron 26 morfotipos entre las accesiones de la Estación de Canaán del INIA (Ayacucho). El escaso número de estudios constituye un obstáculo para abordar investigaciones de mayor envergadura, por ejemplo, genética de poblaciones e mejoramiento; por ello, resulta prioritario generar al menos un genoma de

referencia del cloroplasto que permita diseñar marcadores moleculares para estudios de filogenia y delimitación de especies.

Actualmente, las investigaciones donde se ensambla y se caracteriza nuevos genomas de cloroplasto amplían aún más las fronteras de la investigación sobre biología del cloroplasto, analizando a profundidad las funciones de las secuencias del genoma del cloroplasto para mejorar nuestra comprensión de la transferencia de genes intracelulares, conservación, diversidad y mejorar los rasgos agronómicos de las plantas. Por estas razones ensamblar, analizar y comparar el genoma de cloroplasto de especies del género *Opuntia* es crucial para desarrollar programas de mejoramiento genético y conservación. Tradicionalmente se consideraba que los genomas de cloroplasto estaban altamente conservados en estructura, pero estudios recientes en Cactácea han revelado una notable variación estructural del genoma, incluyendo reordenamientos estructurales, pérdidas de genes y cambios en la estructura de las regiones repetidas (Köhler et al., 2023). En este contexto, el proyecto tiene el propósito de caracterizar el genoma del cloroplasto de la especie cultivada *O. ficus-indica* y la especie silvestre *O. pubescens*, para proporcionar información crucial para entender la historia evolutiva de genoma cloroplasto y la filogenia del género *Opuntia*.

1.2. Formulación del problema

1.2.1. Formulación del problema general

¿Qué aportes para la filogenia y la historia evolutiva del género *Opuntia* puede proporcionar la caracterización detallada del genoma cloroplástico de *O. ficus-indica* y *O. pubescens*?

1.2.2. Formulación del problema específico

¿Es posible obtener genoma de cloroplasto mediante ensamblaje de novo y anotación de *O. ficus-indica* y *O. pubescens* adecuados para análisis comparativos?

¿Qué variaciones estructurales presentan el genoma de cloroplasto de *O. ficus-indica* y *O. pubescens* respecto a otras especies de *Opuntia*?

¿Qué tipos y abundancia de secuencias repetitivas contienen el genoma de cloroplasto de *O. ficus-indica* y *O. pubescens* y cómo difieren de otras especies de *Opuntia*?

¿Qué topología filogenética se obtiene al utilizar genomas de cloroplasto de *O. ficus-indica*, *O. pubescens* y otras especies del género *Opuntia*?

1.3. Objetivo de la investigación

1.3.1. Objetivo general

Caracterizar el genoma del cloroplasto de la especie cultivada *O. ficus-indica* y la especie silvestre *O. pubescens*, para entender la historia evolutiva de genoma cloroplasto y la filogenia del género *Opuntia*.

1.3.2. Objetivo específico

Ensamblar de novo y anotar el genoma del cloroplasto de *O. ficus-indica* y *O. pubescens*.

Comparar las variaciones estructurales del genoma del cloroplasto de *O. ficus-indica* y *O. pubescens*, así como en otras especies de *Opuntia*.

Identificar y caracterizar secuencias repetitivas en los genomas del cloroplasto de *Opuntia ficus-indica* y *Opuntia pubescens*, así como en otras especies de *Opuntia*,

Realizar una reconstrucción filogenética utilizando secuencias del genoma de cloroplasto de *O. ficus-indica*, *O. pubescens* y todas las especies de *Opuntia* disponibles en el NCBI.

1.4. Justificación de la investigación

La importancia económica y ecológica que tiene el género *Opuntia*, radica a sus múltiples usos y beneficios; además, sus metabolitos han demostrado diversas actividades farmacológicas, entre las que se incluyen efectos antimicrobianos, antioxidantes, antiinflamatorios, antitumorales, neuroprotectores, hepatoprotectores e hipotensores (López et al., 2014; Shirazinia et al., 2019) ha llamado la atención de su investigación en diferentes partes de mundo. Sin embargo, en Perú aún son escasos los estudios genéticos y genómicos dentro de este grupo. Tanto así que existen investigaciones de análisis de la variabilidad genética en *Opuntia* que solamente se basan en caracterización morfológica (Tineo, 2019), cuando en otras partes del mundo se usan marcadores moleculares de ADN. Además, en el Perú algunas especies endémicas de género *Opuntia* se encuentran categorizadas como vulnerables y otras en peligro de extinción (Ostolaza, 2019). No se sabe el origen de este problema, pero se tiene algunas hipótesis como factores abióticos debido al cambio climático y factores bióticos como la presencia de insectos fitófagos y enfermedades.

Los estudios sobre diversidad genética en el género *Opuntia* son escasos y aún no se han realizado estudios genómicos de este género en Perú, según la literatura disponible. La investigación existente se ha centrado principalmente en la taxonomía y la variabilidad fenotípica de *O. ficus-indica* (tuna), con algunos trabajos que exploran la diversidad genética intraespecífica entre las diferentes variedades principalmente por características morfológico (Sejuro, 1988; Tineo, 2019; Ostolaza, 2019). Pero no obstante Köhler *et al.*,

(2023), ha empezado a investigar y publicar investigaciones en genoma cloroplastidiales en los de la tribu Opuntieae. La falta de interés en comprender este grupo de plantas, que es de gran importancia para el Perú, ha impedido el desarrollo de estrategias adecuadas para su conservación y mejora genética. Esta falta de atención podría resultar en la pérdida de recursos genéticos en las especies silvestres en el género de *Opuntia*.

La información genómica de cloroplasto, puede ser una herramienta valiosa para la conservación de especies vegetales. El genoma del cloroplasto es relativamente pequeño entre 20 a 25 Kb y presenta una estructura simple cuatripartita conformada por la región larga única (LSC) la región corta única (SSC), y las dos regiones repetidas invertidas (IRa y IRb), así mismo se encuentran en múltiples copias en cada célula vegetal, lo que facilita su secuenciación y análisis. A partir de esta información, se pueden desarrollar estrategias de conservación y el rescate genético (Caycho, 2023). Además, la información del genoma del cloroplasto permite realizar investigaciones en diversidad genética de poblaciones, genotipado, fenotipado e ingeniería genética, todas ellos importantes para desarrollar fitomejoramiento (Daniell et al., 2016).

Este proyecto se tiene propósito de caracterizar el genoma del cloroplasto de la especie cultivada *O. ficus-indica* (tuna) y la especie silvestre *O. pubescens*, para proporcionar información crucial para entender la historia evolutiva de genoma cloroplasto y la filogenia del género *Opuntia*.

1.5. Delimitación del estudio

El estudio se centrará en describir y comparar la información del genoma cloroplástidial de *O. ficus-indica* y *O. pubescens*, y en reconstruir sus relaciones filogenéticas con otras especies del género *Opuntia*. Para ello se emplearán herramientas bioinformáticas de ensamblaje de *novo*, anotación genómica, detección de variaciones estructurales, análisis de secuencias repetitivas y análisis filogenético.

Este proyecto se realizará el Laboratorio de Genómica y Bioinformática para la Biodiversidad (FCB), en la Universidad Nacional Mayor de San Marcos, Lima – Perú. Donde se realizará las extracciones de ADN y los análisis bioinformáticos. Así mismo se tendrá soporte de otros laboratorios como el Laboratorio de Biotecnología de la Producción, de la Facultad Ingeniería Agraria, Industrias Alimentarias y Ambiental de la Universidad Nacional José Faustino Sánchez Carrión, Huacho - Perú.

1.6. Viabilidad del estudio

Este estudio es viable dado que el Instituto Nacional de Innovación Agraria (INIA) cuenta con un banco de germoplasma de *Opuntia ficus-indica*, y se cuenta con las coordenadas de colecta para *Opuntia pubescens*. Asimismo, se cuenta con las experiencias del grupo de investigación del Laboratorio de Genómica y Bioinformática para la Biodiversidad (LGBB) de la Facultad de Ciencias Biológicas, UNMSM. Por otra parte, este proyecto cuenta con financiamiento para el secuenciamiento del genoma de dos especies y los kits de extracción de ADN por parte de proyecto titulado Caracterización y análisis comparativo del genoma de cloroplasto de la tuna (*Opuntia ficus-indica*) y su pariente silvestre (*Opuntia pubescens*). Los gastos faltantes serán cubiertos por el Laboratorio de Genómica y Bioinformática para la Biodiversidad (LGBB).

CAPITULO II. MARCO TEÓRICO

2.1. Antecedentes de la investigación

2.1.1. Antecedentes internacionales

En la actualidad existe diversas investigaciones relacionadas con caracterización genómica de cloroplasto en género *Opuntia*, pero muy pocas en *O. ficus-indica*, y ninguna en *O. pubescens*.

La caracterización de los genomas organelares ha permitido comprender mejor la evolución y plasticidad genética en Cactaceae. *Opuntia cochenillifera*, una especie de reproducción clonal con relevancia agrícola e industrial, ha sido objeto de estudios recientes que han revelado una gran dinámica en sus genomas mitocondrial y cloroplástidial. Liu et al. (2024) secuenciaron y ensamblaron el genoma mitocondrial de esta especie, identificando su compleja organización y la presencia de fragmentos homólogos compartidos entre ambos organelos. Estos hallazgos resaltan la importancia de la variabilidad estructural en la evolución de los genomas organelares y su potencial aplicación en programas de mejoramiento genético. En este contexto, el análisis del genoma cloroplastidial de *O. ficus-indica* y *O. pubescens* permitirá ampliar el conocimiento sobre la evolución estructural de los genomas cloroplastidiales en *Opuntia* y su impacto en la filogenia del grupo.

Matias Köhler *et al.*, demostró que los genomas cloroplastidiales suelen ser altamente conservados en plantas, pero en algunas líneas evolutivas, como Cactaceae, se han identificado modificaciones estructurales significativas. Estudios recientes han demostrado que los genomas cloroplastidiales en *Opuntia*, presentan una notable variabilidad en tamaño, estructura y contenido génico, con eventos de expansión y contracción de la región de repeticiones invertidas (IR), así como la pérdida o pseudogenización de ciertos genes. Köhler et al. (2023) analizaron los genomas cloroplastidiales de 43 especies de *Opuntieae*, los cuales fueron agrupados en base a su estructura en nueve tipos distintos, lo que evidencia la dinámica evolución de estos genomas en el grupo. A pesar de que los datos cloroplastidiales han permitido esclarecer relaciones filogenéticas dentro de la tribu, algunos nodos continúan siendo conflictivos y requieren más estudios.

Majure *et al.*, (2023) estudiaron la biogeografía del clado Xerocarpa de *Opuntia*, compuesto por especies adaptadas a ambientes desérticos del oeste de América del Norte. Este grupo se originó en el Plioceno medio-tardío, probablemente en el altiplano del Colorado, y posteriormente se dispersó hacia los desiertos de Mojave, Sonora y Chihuahua, así como regiones montañosas de California. El clado presenta adaptaciones morfológicas notables que han facilitado su éxito ecológico: *O. basilaris* desarrolló una epidermis pubescente para tolerar condiciones más cálidas y áridas, mientras que *O. polyacantha*, con tallos cubiertos de espinas, logró expandirse hacia ambientes más fríos del norte. Asimismo, la sinapomorfía de frutos secos influyó en la capacidad de dispersión de las especies, limitando a unas como *O. basilaris* y favoreciendo a otras como *O. polyacantha*. Estos cambios reflejan cómo el género *Opuntia* ha evolucionado para adaptarse a hábitats contrastantes, mostrando patrones complejos de diversificación influenciados por la geografía y el clima.

Majure *et al.*, (2012), sostienen que género *Opuntia*. constituye un clado bien sustentado que se originó en el suroeste de América del Sur, diversificándose rápidamente tras una migración hacia el norte o un evento de dispersión a larga distancia hacia las zonas áridas del oeste de América del Norte. En este proceso, la evolución reticulada y la poliploidización han desempeñado un papel crucial, al generar fenotipos novedosos e incrementar la riqueza específica. La presencia de múltiples niveles de ploidía entre especies ha dificultado la resolución de relaciones filogenéticas dentro del grupo, resaltando la necesidad de determinar el nivel de ploidía de todos los taxones para reconstruir con precisión su historia evolutiva. En este contexto, la caracterización el genoma del cloroplasto de la especie cultivada *O. ficus-indica* (tuna) y la especie silvestre *O. pubescens* proporcionará información crucial para entender la historia evolutiva de genoma cloroplasto y la filogenia del género *Opuntia*.

2.1.2. Antecedentes nacionales

En la actualidad al momento que se escribe el presente trabajo, no existe trabajos en sobre caracterización genómica usando el genoma cloroplastial para especies nativas del género *Opuntia* en Perú. Así mismo tampoco existen trabajos empleando técnica de Código de barras de ADN para identificación de especies nativas de *Opuntia*.

En el contexto de la diversidad de cactáceas del sur del Perú, Pauca & Quipuscoa (2020) abordaron la taxonomía del género *Cumulopuntia* en el departamento de Arequipa, realizando una revisión sistemática de 12 especies registradas en esta región. En este estudio, se designaron neotipos para varias especies originalmente descritas bajo el género *Opuntia*, como *O. dimorpha*, *O. unguispina*, *O. ignota* y *O. rauppiana*, lo que evidencia la importancia histórica y taxonómica del género *Opuntia* en la región andina. Las redefiniciones nomenclaturales y tipificaciones efectuadas contribuyen significativamente a la estabilización del conocimiento sistemático de este grupo, aportando a la correcta identificación y clasificación dentro de la subfamilia Opuntioideae.

Por otra parte, existe un trabajo en caracterización morfológica y análisis de la variabilidad genética de la colección nacional de germoplasma de tuna (*Opuntia spp*) del Perú. Tineo (2016) realizó la caracterización de 183 accesiones del género *Opuntia* o de *O. ficus-indica* (tuna) en la colección nacional de germoplasma del Perú, identificando 26 morfotipos de fruto y 20 morfotipos de planta mediante el análisis de caracteres morfológicos. Mediante análisis de componentes principales, se determinó que variables como el peso y tamaño del fruto, así como el diámetro de la semilla, son altamente discriminantes en la variabilidad genética del germoplasma. Además, se identificaron accesiones promisorias con características comerciales destacadas, lo que resalta la

importancia de la caracterización genética para el mejoramiento y conservación de las especies del género *Opuntia*.

En esta línea, García-Godos (2015) evaluó la variabilidad genética de *Opuntia ficus-indica* en 40 accesiones conservadas en el banco de germoplasma del Instituto Nacional de Innovación Agraria (INIA) de Ayacucho, utilizando marcadores moleculares AFLP. El análisis molecular reveló un alto nivel de polimorfismo, con 208 loci polimórficos sobre un total de 233 bandas generadas por cinco combinaciones de primers. La variación genética promedio fue de 0,276, y el índice de información de Shannon alcanzó 0,426, indicando una considerable diversidad genética. El agrupamiento de las accesiones mediante análisis de distancia con el coeficiente Simple Matching permitió identificar cuatro grupos claramente diferenciados, sin evidenciar duplicados, lo que resalta la riqueza genética del material conservado. Estos resultados demuestran la utilidad de los marcadores moleculares en la caracterización genética y ofrecen criterios valiosos para programas de mejoramiento y conservación del germoplasma.

Rafael-Ramírez *et al.*, (2012) describe detalladamente la morfología *O. ficus-indica*, señalando que presenta un sistema radical extenso y superficial, alcanzando profundidades de hasta 80 cm. Sus tallos modificados, denominados cladodios o pencas, tienen forma ovoide, elíptica u oblonga, y pueden medir entre 33-60 cm de largo por 18-25 cm de ancho, con un grosor de 1.8-2.3 cm. Estos órganos fotosintéticos pueden ser espinosos o no, dependiendo de la variedad, y poseen areolas con espinas grandes (hojas modificadas) y gloquidios. Las flores, hermafroditas y sésiles, miden entre 6 y 7 cm y se desarrollan en la parte apical del cladodio. Además, se resalta la fisiología CAM de la planta, que le permite optimizar el uso del agua en condiciones áridas mediante un intercambio gaseoso nocturno.

No obstante, estos estudios se han enfocado principalmente en características morfológicas, que resulta ser insuficiente para delimitar especies y genotipos. En este contexto, la caracterización del genoma cloroplástico de *O. ficus-indica* y *O. pubescens* complementará dichas investigaciones al proporcionar información relevante sobre la arquitectura genómica, los patrones evolutivos y la filogenia del género *Opuntia*, aportando nuevas perspectivas para su conservación y mejoramiento genético.

2.2. Bases teóricas

2.2.1. Generalidades de las dos especies perteneciente al género *Opuntia*

a. *Opuntia ficus-indica* (L.) Mill.

O. ficus-indica conocido como tuna es la especie de mayor importancia agronómica histórica, debido al sabor dulce del fruto, pero también debido a su tallos, que se utiliza como forraje para el ganado y en la producción de insecto escamoso *Dactylopius sp.* para obtención de carmín (Inglese, 2017).

El cultivo de *O. ficus-indica* históricamente está relacionado con las antiguas civilizaciones mesoamericanas, específicamente en las regiones semiáridas y llevado a otras partes del mundo mediante el cultivo y el comercio como a África, Europa y Asia (Sidney Novoa, 2006; Yahia & Sáenz, 2011). En los sistemas agropastoriles de los Andes peruanos esta especie representa un recurso de gran importancia agronómica. Esta cactácea se encuentra ampliamente distribuida en los valles interandinos, donde ha logrado establecerse debido a las condiciones favorables. Su principal uso se centra en la cría de un insecto llamado cochinilla, cuyo pigmento natural, conocido como carmín, es ampliamente utilizado en diversas industrias, incluyendo la alimentaria, farmacéutica y textil (Rodríguez et al., 2006; Novoa, 2006). El Perú continúa siendo uno de los principales productores mundiales de carmín, satisfaciendo alrededor del 80% de la demanda global de este pigmento. La actividad productiva se sustenta en la acopio artesanal de cochinilla productores de carmín, principalmente en los departamentos de

Apurímac , Ayacucho, Cusco y Huancavelica (Rodríguez & Pascual, 2004; AGRARIA.PE, 2022).

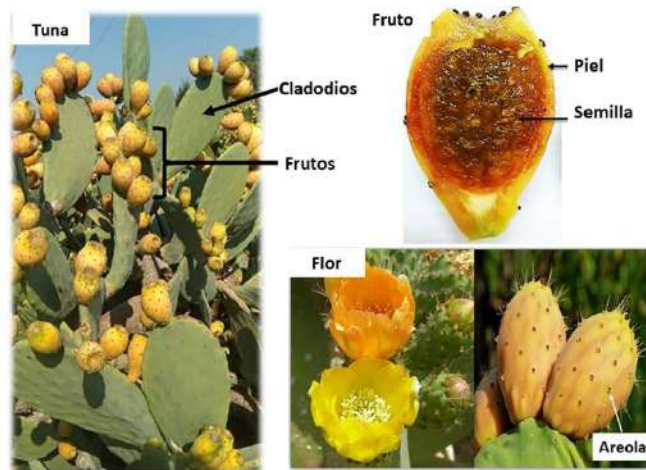


Figura 1. Plata entera de *Opuntia ficus-indica* mostrando las principales características morfológicas.

(Abbas et al., 2022)

O. ficus-indica (L.) Mill. es un arbusto erguido y extendido que alcanza una altura de 3 a 5 metros. Sus cladodios, que son las estructuras similares a las hojas, son suculentos y típicamente tienen forma elíptica, específicamente espatulado-oblongos, con dimensiones que varían entre 25 y 60 cm de largo y 20 a 40 cm de ancho (Figura 1). Las areolas son ovaladas, se encuentran a 2 mm por debajo de la superficie de la piel y se distribuyen en forma helicoidal. Pueden desarrollar espinas cortas parecidas a pelos llamadas gloquidias. Por lo general, las espinas se encuentran en las primeras etapas del crecimiento de los cladodios, y el aumentar la temperatura hace que, la mayoría se caigan. Las espinas tienen una superficie rugosa, mientras que los gloquidias la tienen lisa. Las espinas son blancas y están presentes como una o dos espinas largas acompañadas de dos más pequeñas, mientras que los gloquidias se encuentran en grupos de 7 a 12. Las flores son hermafroditas y actinomorfo con pétalos de color amarillo brillante, aunque en ocasiones también pueden tener colores naranjas, rosa, morado, rojo o blanco (Figura 1).

Las frutas tienen diferentes formas, tamaños y colores (verde, amarillo, naranja, rojo y morado) según la variedad, además tienen forma de barril de 5 a 10 cm de largo y de 4 a 9 cm de diámetro, mostrando areolas que pueden tener espinas, pero siempre tienen gloquidias. Las semillas están incrustadas en la pulpa carnosa y tienen formas redondeadas de color marrón pálido con 5 mm de diámetro (Galati et al., 2003; Abbas et al., 2022; F. J. Barba et al., 2022).

b. *Opuntia pubescens*

Opuntia pubescens (Cariofilales: Cactáceas; Opuntia: Opuntia) es una planta subarborescente y suberecta que puede alcanzar hasta 30 cm de altura, caracterizada por tener una epidermis pubescente. Sus cladodios son verdes y de forma cilíndrica a ligeramente aplanada, con dimensiones que oscilan entre 10 y 26 cm de largo y de 1 a 2.7 cm de ancho, con una superficie tuberculada que se desprende fácilmente. Presenta gloquidios amarillos con espinas de 1 a 4 barbas. Las flores son amarillo-verdosas y tienen una longitud de 3 a 5 cm, con un diámetro proporcional. Los frutos son de color verde con un ligero tinte rojizo y tienen un tamaño aproximado de 2 a 2.5 cm de largo (Ostolaza, 2019).

2.2.2. Importancia económica y ecológica de *Opuntia* en el Perú

La *Opuntia* comúnmente conocida como tuna, tiene un notable valor económico y ecológico en el Perú. Su cultivo y comercialización generan ingresos importantes para los agricultores, particularmente en regiones áridas donde otros cultivos tienen dificultades para crecer. La *Opuntia* sirve como ingrediente valioso en diversos productos alimenticios, medicinales y cosméticos, contribuyendo a la diversificación económica y la creación de empleo. La producción de tuna en el Perú ha superado las 62 000 toneladas, con una superficie cosechada de 12 000 hectáreas y un rendimiento promedio de 5,3 toneladas por hectárea. Si bien la tendencia ha sido negativa en los últimos años,

es evidente una recuperación, con rendimientos superiores al promedio histórico (Novoa, 2006; Arenas et al., 2017; MIDAGRI, 2022).

Ecológicamente, *Opuntia* juega un papel crucial en la conservación del suelo y la preservación de la biodiversidad y riqueza genética. Sus raíces profundas ayudan a prevenir la erosión del suelo, mientras que su capacidad para prosperar en condiciones áridas contribuye a estabilizar los ecosistemas afectados por la desertificación. Además, *Opuntia* proporciona alimento y refugio a una amplia gama de especies silvestre, lo que ayuda a mantener la diversidad biológica en los hábitats donde está presente (Berhe et al., 2023; Ostolaza, 2019; Yanqui et al., 2024).

2.2.3. Los cloroplastos

a. Origen y funciones de los cloroplastos

Los cloroplastos son orgánulos únicamente de las plantas y no se hallan de forma natural en otros grupos eucariotas desde su aparición. Su origen se remonta hace unos 3.500 millones de años, cuando surgió a partir de una endosimbiosis entre una célula ancestral y una cianobacteria. Esta cianobacteria que poseía una forma ancestral del fotosistema (fotosistema anoxigénico), las cuales no generan oxígeno durante el proceso fotosintético. Necesito unos mil millones de años para que sucede una reelaboración del nuevo aparato fotosintético con la capacidad de utilizar el agua como fuente de electrones, dando lugar así a la fotosíntesis que libera oxígeno (Rascio, 2013). Posteriormente estas procariotas similares a las cianobacterias fueron engullidas por un protoeucariota heterótrofo aeróbico, esta endosimbiosis dio origen a los cloroplastos, los organoides en las células vegetales responsables de la fotosíntesis.

Los cloroplastos realizan muchas funciones metabólicas importantes para la planta, participando en fotocinesis, síntesis de carbohidratos, respiración, síntesis de lípidos y otras rutas metabólicas. Así mismo, los cloroplastos preservan su propio genoma

altamente reducido, con genes propios que no se localizan en el genoma nuclear. El genoma del cloroplasto opera de manera armonizada con el genoma nuclear para la biogénesis y función de los cloroplastos (Liebers et al., 2022; H. Gao et al., 2023).

b. Estructura del cloroplasto

La estructura de los cloroplastos tiene una serie de estructuras clave, que permiten su funcionalidad. La estructura está definida por tres sistemas: las membranas envolventes exterior e interior y el sistema de tilacoides (Figura 2) (Staehelein, 1986; Kirchhoff, 2019). La membrana externa es una capa envolvente suave que constituye una bicapa de fosfolípidos y permeable a pequeñas moléculas orgánicas. Dentro de la membrana externa existe una membrana interna más compleja y menos permeable a estas mismas moléculas, pero que atraviesan a través de las proteínas transportadoras de membrana específicos (Krupinska et al., 2013). Al interior del cloroplasto se encuentra el estroma, espacio donde se encuentra el sistema genético del cloroplasto y una variedad de enzimas metabólicas, asimismo alberga reacciones metabólicas independientes de la luz e incluidas las responsables de la conversión crítica de CO₂ en carbohidratos mediante la fotosíntesis (Cooper, 2000; Nevo et al., 2012; Kirchhoff, 2019).

Los tilacoides se encuentran en el estroma del cloroplasto, como se muestra en la Figura 2. Estos tilacoides están organizados en pilas, formando sacos aplanados que pueden estar aislados o conectados entre sí. En estos tilacoides se lleva a cabo la fotosíntesis, gracias a la presencia de complejos multiproteicos y clorofilas involucradas en las reacciones luminosas. Además de la fotosíntesis, en los tilacoides también tienen lugar otros procesos metabólicos propios del cloroplasto (Rascio, 2013; Pribil et al., 2014). El ADN de cloroplasto está organizado en conglomerados de ADN y proteínas llamados nucleoides de cloroplasto, que se encuentran dentro del estroma del cloroplasto (Figura 2). Los nucleoides presentan 8 numerosas copias del genoma del cloroplasto y se

creo que los nucleoides de los cloroplastos son la unidad funcional de diversos procesos, incluida la herencia, recombinación, replicación, reparación y transcripción del ADN (Kuroiwa, 1991; Sakai et al., 2004; Kamimura et al., 2018).

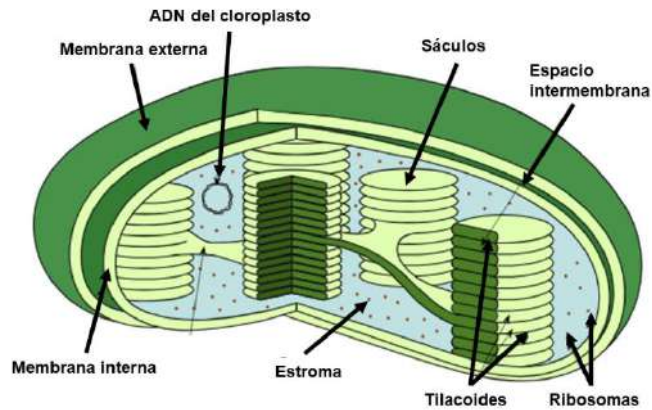


Figura 2. Un diagrama que muestra la estructura general de los cloroplastos de plantas superiores.
(Kirchhoff, 2019)

c. Estructura del genoma cloroplastial

El genoma del cloroplasto (cpDNA) es una molécula de ADN bicatenaria en forma circular y su longitud varía aproximadamente entre 120 y 200 Kb. Los cloroplastos presentan su sistema genético reducido lo que evidencia sus orígenes evolutivos a partir de bacterias fotosintéticas, pero preservado genes indispensables para la fotosíntesis y la biogénesis de los cloroplastos e plastidios (Wicke et al., 2011; Kamimura et al., 2018; Turudić et al., 2021). En la estructura del cpDNA está mayormente constituido por cuatro partes (Figura 3): una región larga de copia única (LSC), una región corta de copia única (SSC) y dos regiones repetidas invertidas (IR). Existen dos regiones IR que están separadas por la región LSC y una sección corta de SSC (Schwarz et al., 2015; Morley et al., 2019; Wang & Lanfear, 2019; Turudić et al., 2022).

Las regiones IR están altamente conservadas en diferentes especies de plantas y tienen una longitud aproximadamente de 20 000 a 25 000 pb, además la recombinación homóloga (flip-flop) originó la aparición de dos haplotipos estructurales, manifestándose como la invertida la orientación de la secuencias (Palmer, 1983; Sloan et al., 2012; Wang & Lanfear, 2019). Se ha informado la existencia de especies que han perdido las regiones IRs o una copia de la repetición invertida en los genomas de sus cloroplastos, estos caso se encuentran en leguminosas, coníferas, Cactaceae, Arecaceae, y la mayoría de las plantas parásitas y no fotosintéticas de la familia Orobanchaceae (Tangphatsornruang et al., 2010; Wicke et al., 2011; Wu & Chaw, 2014; Barrett et al., 2016).

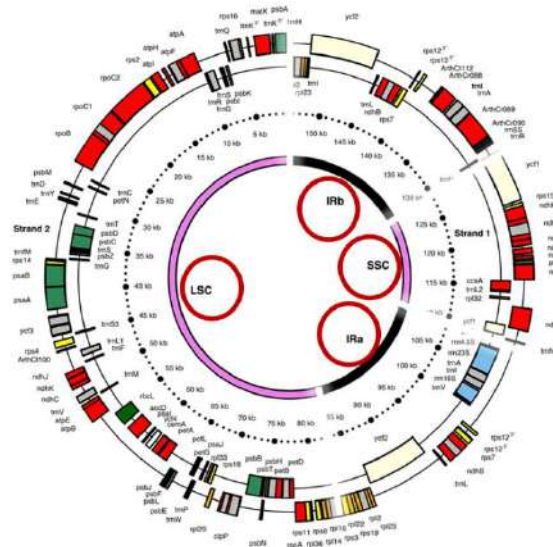


Figura 3. Presentación clásica del mapa del genoma cloroplastal, donde se muestran las cuatro regiones: Dos secuencias de repeticiones invertidas (IR), una secuencia larga de copia única (LSC) y secuencia de corta copia única (SSC).

(Dobrogojski et al., 2020)

El genoma del cloroplasto es rico en adenina (A) y timina (T) y una concentración más baja de guanina (G) y citosina (C) en comparación con el ADN nuclear. El cpDNA tiene regiones codificantes y no codificantes. Estas regiones codificantes son

responsables de la síntesis de proteínas indispensables para el funcionamiento del cloroplasto. El número de genes codificados en cpDNA se encuentra aproximadamente en 9 100 genes (Wicke et al., 2011), y en número de genes codificados que varía entre diferentes especies de plantas puede ir de 0 hasta 315 según Dobrogojski et al., (2020). Además, en el genoma del cloroplasto se encuentran codificados diversos tipos de ARN, entre ellos genes que contienen información para la síntesis de ARN de transferencia (ARNt) y ARN ribosomal (ARNr).

2.2.4. Secuenciación de ADN de última generación

La secuenciación de próxima generación (NGS) son un grupo de tecnologías que permite secuenciación de ADN y ARN. NGS ofrecen una gama diversa de métodos, lo que brinda a los investigadores la capacidad de abordar prácticamente cualquier pregunta relacionada con el genoma, transcriptoma o epigenoma de cualquier organismo. Las técnicas de secuenciación varían principalmente en cómo se adquieren las muestras de ADN o ARN, así como en las opciones de análisis de datos empleadas. Aunque el número de métodos de NGS está en constante crecimiento, se presenta aquí un breve resumen de los métodos más comunes.

2.2.4.1. Tecnología de primera generación: Sanger y Maxam & Gilbert

Los primeros esfuerzos para secuenciar el ADN y el ARN se basaron en técnicas de degradación química de las moléculas, lo que permitía generar fragmentos que podían ser analizados individualmente. Uno de los hitos importantes en este proceso fue el logro de Robert Holley en 1964, cuando secuenció una molécula de ácido nucleico, el ARNt de alanina, utilizando la ribonucleasa de *S. cerevisiae* (Holley et al., 1965). Y la secuenció el ARNt de alanina usando ribonucleasa. Luego Walter Gilbert y Allan Maxam desarrollaron una técnica de degradación química para secuenciar PhiX174. Sin embargo, el verdadero avance llegó con la introducción del método de secuenciación basado en

terminación de cadena por parte de Frederick Sanger. Este método, conocido como secuenciación de Sanger, utilizaba didesoxinucleótidos (ddNTP) que interrumpían el alargamiento de la cadena de ADN durante la replicación, permitiendo la obtención de producción de secuencia de hasta unos pocos cientos de nucleótidos de longitud (Sanger et al., 1977).

La técnica de Sanger fue ampliamente adoptada y revolucionó el campo de la biología molecular al permitir la secuenciación rápida y precisa de ADN y ARN. En 1987, se lanzó en los Estados Unidos la primera máquina comercial de secuenciación automatizada denominado Applied Biosystems ABI 370. Esta máquina utilizaba ddNTP marcados con fluorescencia y electroforesis capilar para automatizar el método de secuenciación de Sanger, aumentando significativamente la velocidad y precisión de la secuenciación del ADN (Schuster, 2008; M. Barba et al., 2014). Esta tecnología proporciono importantes descubrimientos y avances en nuestra comprensión de la genética y la biología molecular, así como en la diagnosis y tratamiento de enfermedades genéticas y otras condiciones médicas.

2.2.4.2. Tecnología de secuenciación de segunda generación

a. Tecnología Illumina

Las plataformas de secuenciación de Illumina utilizan la fluorescencia para secuenciar moléculas individuales de ADN, tras una amplificación clonal en un soporte sólido. La secuenciación ocurre en una celda de flujo con múltiples canales que pueden contener varias muestras. El ADN se une a la celda y se amplifica en puentes clonales (Płoski, 2016). Después, la secuenciación procede mediante síntesis con ciclos de polimerización y detección de fluorescencia. Se eliminan las cadenas complementarias y se procede a secuenciar el otro extremo del ADN. Esta técnica es altamente confiable, con un margen de error inferior al 0.1%, y rápida. Sin embargo, puede enfrentar

dificultades al secuenciar regiones con secuencias repetidas de un solo nucleótido, conocidas como regiones homopoliméricas. La tecnología de Illumina ha sido fundamental en la secuenciación de próxima generación (NGS), revolucionando la genética y la biología molecular. Su capacidad para secuenciar ADN y ARN ha permitido avances significativos en la investigación científica, medicina y diagnóstico genético (Adey et al., 2010; Płoski, 2016).

b. Tecnología Ion Torrent

La metodología de Ion Torrent basadas en semiconductores emplea la detección de cambios de pH durante la síntesis de ADN para la secuenciación. Este proceso, conocido como PCR en emulsión, amplifica moléculas de ADN individuales (Dressman et al., 2003; Rothberg et al., 2011). La preparación de la biblioteca comienza con la fragmentación del ADN y la ligación de adaptadores, seguida de PCR en emulsión, donde la biblioteca se mezcla con microperlas en una emulsión de aceite y solución acuosa. Después de la PCR, las perlas se separan, se enriquecen y se preparan para la secuenciación. Las perlas se colocan en pozos de un chip especializado (Ion Chip) y se someten a ciclos de síntesis de ADN, donde la incorporación de nucleótidos libera iones H⁺, detectados por sensores. La secuenciación es rápida, con una sola ejecución que toma aproximadamente 2 horas y altamente confiable con error (<0.1%). El sistema Ion Torrent fue el primero en utilizar esta tecnología en 2010, capaz de secuenciar exomas (Płoski, 2016).

2.2.4.3. Tecnología de Secuenciación de tercera generación

a. Tecnología de secuenciación SMRT de Pacific Biosciences

La técnica más conocida de SMS es la secuenciación en tiempo real de molécula única (SMRT) de (PacBio). SMRT utiliza una química de secuenciación por síntesis similar a la de la segunda generación, pero con menos material inicial y sin PCR, lo que

reduce los errores y produce lecturas largas en menos tiempo. Durante el proceso de secuenciación, está la preparación de la biblioteca de ADN, síntesis de ADN a partir ADNss circular, después las señales de fluorescencia se activan mediante un láser tan pronto como se incorpora un dNTP marcado al ADN. Luego, un sistema de cámara registra el color y la duración de la luz emitida en tiempo real en la celda de flujo equipada con guías de onda de modo cero (ZMV) (Xiao & Zhou, 2020; Athanasopoulou et al., 2022). Todas las plataformas de PacBio se basan en la tecnología de secuenciación en tiempo real de molécula única. Por ejemplo, el último secuenciador de SMRTbell template (SMRTbell) puede producir 4 millones de lecturas con más del 99% de precisión en solo 30 horas (Garrido-Cardenas et al., 2017; Slatko et al., 2018). Todas plataformas ofrecen lecturas contiguas, largas y rápidas, pero presenta una tasa de error (>5%) y además costos elevados.

b. Tecnología de secuenciación de nanoporos de Oxford Technologies

La tecnología de Oxford Nanopore es utilizado para secuenciación de ADN en tiempo real desde una sola molécula, empleando nanosensores que forman estructuras de canales y transportan la muestra a un sensor que permite la detección de cada residuo nucleotídico presente en la cadena de ADN. Esta tecnología se basa en creación de poros para separar dos compartimentos. La molécula de ADN cruza el poro gracias a la acción de una proteína motora, produciendo una alteración corriente iónica entre ambos lados del poro, lo que permite su identificación por la proteína lectora (Wick et al., 2019). Esta plataforma no solo detecta la composición de una plantilla de ADN, sino que también puede interpretar secuencias k-mer. Esta tecnología utiliza ADN no modificado, lo que ofrece resultados rápidos a partir de cantidades mínimas de muestra. Sin embargo, a pesar del potencial prometedor de estas plataformas de secuenciación de moléculas simples en

tiempo real, aún tienen la gran desventaja de baja precisión (Y. Yang et al., 2013; Garrido-Cardenas et al., 2017).

2.2.5. Metodologías de ensamblaje y anotación de genoma de cloroplasto

2.2.5.1. Método de ensamblaje

Para el ensamblaje del genoma cloroplastidial, existen diversas metodologías, como señala Turudić et al., (2021), identifican tres enfoques distintos. Sin embargo, estos se pueden agrupar en dos categorías principales: Las rutas bioinformáticas, también conocidas como pipelines, y el uso de ensambladores especializados. En este trabajo uso ensambladores especializados, dado que el primer método implica una combinación extensa de varios programas que realizan funciones diversas en el proceso, lo que resulta en una menor flexibilidad en la entrada de datos.

La metodología que utiliza ensambladores especializados para este propósito surgió hace unos años. Freudenthal et al. (2020), llevaron a cabo una comparación exhaustiva de siete programas especializados en el ensamblaje de genomas cloroplásticos, además menciona que estas herramientas extraen datos rápidos correspondiente al genoma cloroplastidial y permitiendo ensamblar y solucionar el genoma circular del cloroplasto. GetOrganelle (<https://github.com/Kinggerm/GetOrganelle>) es un grupo de herramientas que permite ensamblar genomas de orgánulos precisos a partir de datos NGS. Además, superó significativamente a todos los demás ensambladores especializados en consistencia, precisión y tasa de éxito (Freudenthal et al., 2020; Jin et al., 2020). Este conjunto de herramientas recluta lecturas utilizando un enfoque modificado de cebo y mapeo iterativo, realiza un ensamblaje de *novo*, filtra y desenreda el gráfico de ensamblaje y produce todas las configuraciones posibles de genomas de orgánulos circulares. Además, GetOrganelle pueden producir configuraciones flip-flop u

otros isómeros de ensamblaje mediados por repeticiones (Guo et al., 2014; Qu et al., 2017; Tsumura et al., 2000).

2.2.5.2. Método de anotación

Para la anotación de los genomas cloroplastidiales, la herramienta más utilizada es GeSeq (<https://clorobox.mpimp-golm.mpg.de/geseq.html>) para la anotación rápida y precisa de secuencias de genomas de organelos, en particular genomas de cloroplastos. A diferencia de las herramientas existentes, GeSeq combina el procesamiento por lotes con una selección de secuencias de referencia totalmente personalizable de registros del genoma organelar del NCBI y/o referencias cargadas por el usuario. Por otra parte, GeSeq permite identificar genes u otras regiones que codifican características mediante búsquedas de homología basadas en BLAT y búsquedas de perfil HMM de genes codificantes de proteínas y ARNr (Tillich et al., 2017). Debido a su capacidad para manejar diferentes tipos de genes, esta herramienta se ha convertido en la más popular para la anotación de genomas cloroplásticos en los últimos 5 años.

2.3. Definición de términos básicos

ADN genómico: El material genético que contiene la información hereditaria de un organismo. Se encuentra en el núcleo de las células eucariotas y en el citoplasma de las células procariotas. El ADN genómico incluye todos los genes y secuencias no codificantes que conforman el genoma de un organismo. Su estructura y organización varían entre diferentes organismos, pero su función principal es almacenar y transmitir la información genética necesaria para el desarrollo, funcionamiento y reproducción del organismo.

Anotación del genoma: La anotación del genoma es el proceso de identificar y marcar los elementos funcionales dentro de una secuencia de ADN. Esto incluye la identificación de genes, exones, intrones, regiones regulatorias, y otros elementos

relevantes. La anotación del genoma es crucial para comprender las funciones de diferentes partes del genoma y cómo contribuyen al organismo en su totalidad. Este proceso generalmente implica el uso de software bioinformático y bases de datos genéticas para predecir y categorizar las características genómicas.

Análisis bioinformático: Disciplina que combina la biología, la informática y la estadística para analizar y comprender datos biológicos. Este campo implica el uso de herramientas y técnicas computacionales para procesar grandes volúmenes de datos biológicos, como secuencias de ADN, proteínas y metabolitos. El objetivo del análisis bioinformático es extraer información valiosa y generar hipótesis sobre la estructura, función y evolución de los organismos. Las aplicaciones típicas incluyen la anotación de genes, el estudio de interacciones proteína-proteína, la reconstrucción de árboles filogenéticos, y la identificación de biomarcadores.

Las secuencias repetitivas simples (SSR): También conocidos como microsatélites, son secuencias cortas de ADN que se repiten de manera consecutiva en el genoma. Cada repetición suele tener una longitud de 1 a 6 pares de bases. Por ejemplo, una secuencia de repetición podría ser "ACACACAC", donde "AC" se repite cuatro veces.

Análisis genómico: El análisis genómico es el estudio exhaustivo del genoma completo de un organismo, lo que incluye la secuenciación, la anotación y la interpretación de todas las secuencias de ADN. Este análisis permite identificar los genes y sus funciones, así como las variaciones genéticas que pueden estar asociadas con enfermedades o rasgos específicos.

Filogenia: Estudio de las relaciones evolutivas entre diferentes especies o grupos de organismos. Este campo de estudio se centra en la reconstrucción de los árboles

filogenéticos, que representan las hipótesis sobre las relaciones de parentesco basadas en características heredadas. Los análisis filogenéticos utilizan datos genéticos, morfológicos y moleculares para entender cómo las especies están relacionadas y cómo han evolucionado a lo largo del tiempo.

2.4. Hipótesis de investigación

2.4.1. Hipótesis general

Los genomas de cloroplasto de *O. ficus-indica* y *O. pubescens* presentan diferencias en su estructura, contenido génico y secuencias repetitivas, las cuales reflejan su historia evolutiva y permiten esclarecer las relaciones filogenéticas dentro del género *Opuntia*.

2.4.2. Hipótesis específico

Es posible ensamblar y anotar el genoma de cloroplasto de *O. ficus-indica* y *O. pubescens* mediante metodologías bioinformáticas actuales.

Los genomas de cloroplasto de *O. ficus-indica* y *O. pubescens* presentan expansiones/contracciones de IR, pérdidas/duplicaciones génicas detectables en comparación con otras especies del género *Opuntia*.

Lo genoma de cloroplasto de *O. ficus-indica* y *O. pubescens* difieren en la composición y abundancia de secuencias repetitivas con otros opuntias

La reconstrucción filogenética basada en genomas completos de cloroplasto generará topologías más resolutivas y ubicará de forma consistente a *O. ficus-indica* y *O. pubescens* dentro del género *Opuntia*.

2.5. Operacionalización de variables e indicadores

Tabla 1. Tabla de la operacionalización de las variables estudiadas.

| Variables | Definición Operacional | Indicadores | Escala de Medición |
|--|--------------------------|--|--|
| Variables: - Datos de secuenciación de ADN de lecturas cortas (151 pb). - Genoma cloroplástico ensamblado y anotado. - Secuencias de cloroplasto propias y otras especies relacionadas. - Identificación y caracterización de variaciones estructurales. - Secuencias repetitivas identificadas. - Relaciones filogenéticas entre las especies de <i>Opuntia</i> | Ensambla y anotar | <ul style="list-style-type: none"> - N° de contigs. - Tamaño total (~120–160 kb). - Cobertura media (> 50×). - N° de k-mers. - Número total de genes anotados (esperado: ~110–130). - Presencia y correcta ubicación de regiones IR, LSC y SSC. - Concordancia con genomas cloroplásticos de referencia (>80% identidad). - Anotación funcional completa y sin errores (genes codificantes, tRNAs, rRNAs). | <ul style="list-style-type: none"> - N° de contigs - Tamaño - Cobertura - k-mers - N° genes - Concordancia |
| | Analizar variaciones | <ul style="list-style-type: none"> - Existencia de inversiones, deleciones, duplicaciones, translocaciones. - Tamaño medio de cada evento (bp). - Ubicación (LSC/SSC/IR; génica/intergénica) - N° de genes afectados por inversiones o reordenamientos | <ul style="list-style-type: none"> - Existencia de eventos - Tamaño - % genoma afectado - N° genes afectados |
| | Identificar repeticiones | <ul style="list-style-type: none"> - Número total de secuencias repetitivas. - Porcentaje del genoma ocupado por repeticiones. - Tipos de repeticiones detectadas (palindrómicas, tándem, dispersas, microsátélites). - Comparación del patrón de repeticiones entre las especies analizadas. | <ul style="list-style-type: none"> - N° total - % ocupado - Densidad - Tipos |
| | Reconstruir filogenia | <ul style="list-style-type: none"> - Soporte estadístico de las ramas (valores bootstrap o posterior probabilities $\geq 70\%$) - Número total de especies incluidas en el árbol. - Concordancia con clasificaciones taxonómicas previas. - Topología coherente con datos moleculares previos por literatura. | <ul style="list-style-type: none"> - Soporte - N° especies - Coherencia topológica |

Fuente propia

CAPITULO III. METODOLOGÍA

3.1. Diseño metodológico

El tipo de investigación fue descriptivo y no experimental, lo que implicó observar y describir el fenómeno del genoma del cloroplasto de *O. ficus-indica* y *O. pubescens*, sin influir sobre ellos de ninguna manera.

Esta investigación se enfocó en caracterizar la información generada a partir de la secuenciación masiva de ADN, el ensamblaje y la anotación de los genomas de

cloroplasto de *O. ficus-indica* y *O. pubescens*; asimismo, estos genomas de cloroplasto obtenidos fueron comparados con genomas de taxones filogenéticamente relacionados.

3.2. Población y muestra

3.2.1. Población

La población está constituida por un espécimen para cada especie (*O. ficus-indica* y *O. pubescens*). La muestra de cladodios de *O. ficus-indica* se realiza fueron obtenidos de la Estación Experimental Canaán del INIA (Instituto Nacional de Innovación Agraria). La muestra de cladodios de *O. pubescens* se colectó en las afueras de Santa Eulalia, Lima con coordenadas -11.83330°S, -76.62657°O, a 1517 m s. n. m. Antes de la colecta se llevó a cabo un reconocimiento morfológico detallado para garantizar la correcta identificación de la especie y evitar posibles confusiones.

3.2.2. Muestra

Las muestras para la extracción de ADN estuvieron constituidas por epidermis de cladodios secos de *O. ficus-indica* y *O. pubescens*. Para su obtención, se separó la epidermis de los cladodios y luego se desecó a 50 °C empleando perlas de gel de sílice.

3.3. Técnicas de recolección de datos

Se muestra el diagrama de flujo de la investigación propuesta (Figura 4).

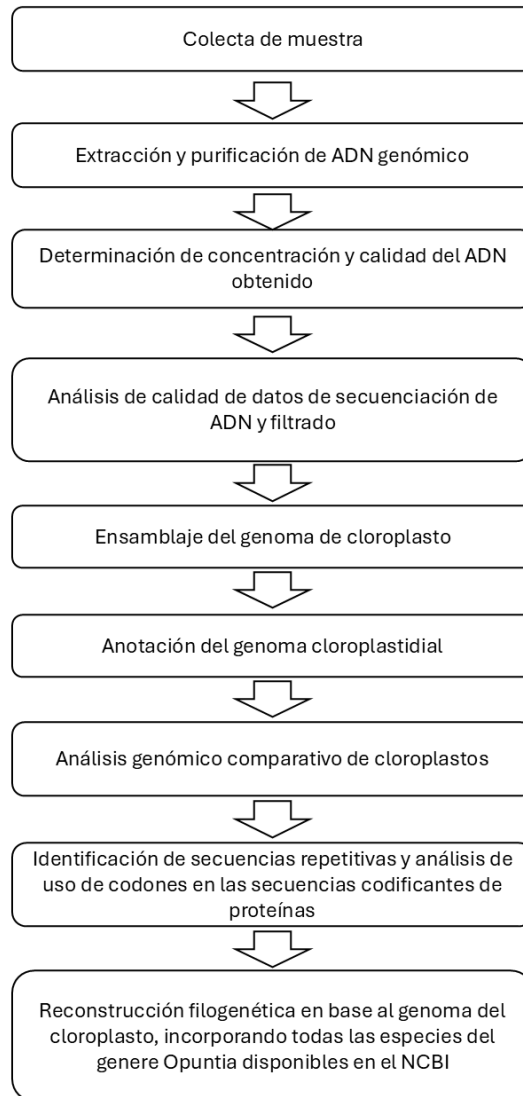


Figura 4. Diagrama de flujo del trabajo del proyecto de investigación.

Fuente propia

3.3.1. Métodos de campo

Se llevó a cabo la búsqueda de las especies *Opuntia pubescens*, en la que se indicaron las ubicaciones geográficas de referencias sobre el avistamiento de estas especies (Tabla 2). Una vez encontradas, se realizó el reconocimiento morfológico de especies para evitar algún tipo de confusión, como se muestra en la Figura 5. Luego se extrajeron cladodios jóvenes de cada uno de los individuos identificados y se guardaron con perlas de silicagel. Por otra parte, las muestras biológicas de tunas (*Opuntia ficus-indica*) fueron compradas en la Estación Experimental Canaán del INIA, Ayacucho.

Posteriormente, los cladodios obtenidos fueron transportados a las instalaciones del Laboratorio de Biotecnología de la Producción de la Universidad Nacional José Faustino Sánchez Carrión, donde fueron disecados obteniéndose la epidermis y secando 50 °C en estufa.

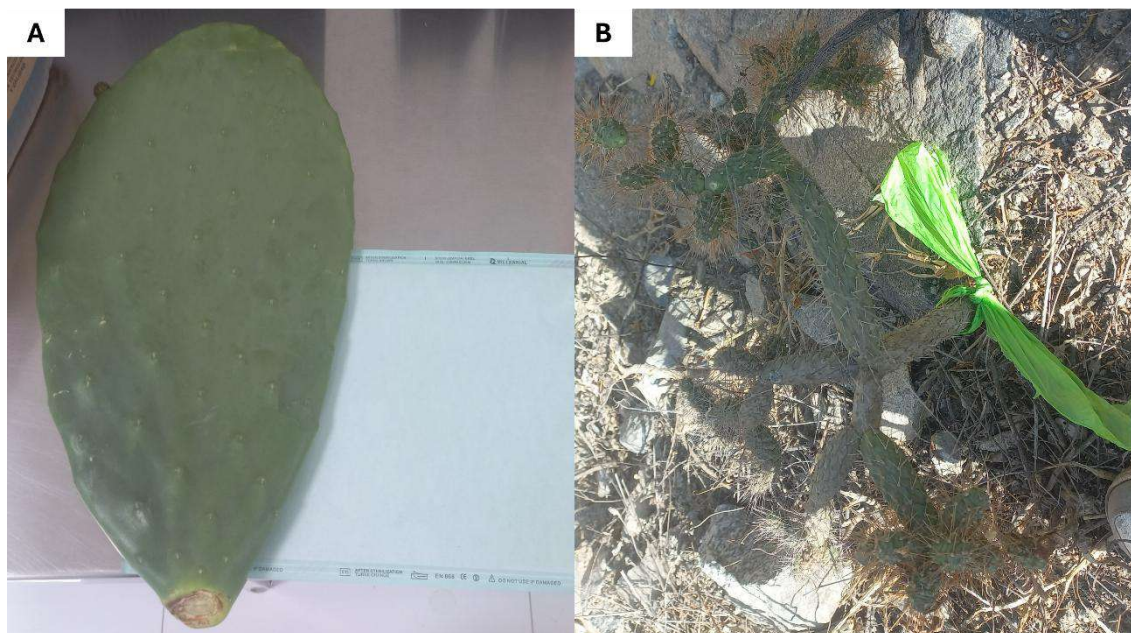


Figura 5. (A) Muestras de *O. ficus-indica* código de acceso PTI 003 INIA. (B) Colecta de *O. pubescens*.

Fuente propia

Tabla 2. Ubicación de referencias para *Opuntia pubescens*.

| Especies | Lugar | Referencias geográficas |
|--------------------------|---------------|---|
| <i>Opuntia pubescens</i> | Santa Eulalia | -11.83330°S, -76.62657°O, 1517 m s. n. m. |

Fuente propia

3.3.2. Método de laboratorio

3.3.2.1. Extracción y purificación de ADN genómico

La extracción del ADN se realizó en las instalaciones del Laboratorio de Genómica y Bioinformática para la Biodiversidad (FCB-UNMSM), para ello se empleó

el kit Plant Genomic DNA (Tiangen, China), siguiendo el protocolo GDP305-Plant Genomic DNA Kit-240514 (https://en.tiangen.com/upload/file/20240712/20240712135931_23071.pdf), con una modificación en el Paso 1, en el que se añadieron 20 µl de proteinasa K y 4 µl de RNase A.

El ADN total purificado se almacenó entre $-25\text{ }^{\circ}\text{C}$ y $-50\text{ }^{\circ}\text{C}$ durante los días siguientes, hasta que se evaluó su calidad y concentración.

3.3.2.2. Determinación de concentración y calidad del ADN obtenido

La calidad y cantidad del ADN extraído se evaluó utilizando el espectrofotómetro Nanodrop™ One C de Thermo Scientific y el fluorómetro Qubit™ 4 de Invitrogen empleando el kit Qubit™ dsDNA Quantification Assay. Para ambas mediciones se utilizaron 2 µL de ADN por muestra. Además, la integridad del ADN fue verificada mediante electroforesis en gel de agarosa al 1 %, utilizando el tinte de ácido nucleico Diamond™ de Promega. La visualización de los resultados se realizó mediante iluminación UV con un foto-documentador.

La secuenciación del ADN se realizó utilizando la plataforma NovaSeq 6000 de Illumina, a través de los servicios del laboratorio externo Macrogen (Macrogen Inc., Seúl, República de Corea). Para la preparación de la biblioteca se empleó el kit TruSeq DNA PCR-Free de Illumina. Durante el proceso de secuenciación se generaron lecturas paired-end en ambas direcciones (forward 5'–3' y reverse 3'–5'), con una longitud de 151 pb.

Los reportes de control de calidad asociados al proceso fueron los siguientes:

- [Evaluación de la muestra original – Reporte 1](#)
- [Evaluación de la muestra original – Reporte 2](#)

- [Control de calidad de la biblioteca NGS](#)
- [Reporte final de secuenciamiento NGS](#)

3.3.3. Método bioinformático

3.3.3.1. Análisis de calidad de datos de secuenciación de ADN y filtrado

A partir de este punto, se llevaron a cabo análisis bioinformáticos utilizando una Workstation PC con las siguientes características: Intel Xeon Gold 5115 (20) @ 3.200 GHz, 3038 MiB / 127427 MiB de memoria y sistema operativo Ubuntu 22.04.4 LTS x86_64.

Para cada una de las especies, se realizó por separado un análisis de calidad de los datos de secuenciación y un proceso de filtrado. La evaluación de la calidad de los datos obtenidos a través de la secuenciación por Illumina se llevó a cabo utilizando FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Posteriormente, se escanearon las secuencias con una ventana de desplazamiento de 30 nucleótidos, eliminando aquellas con una calidad promedio inferior a $Q = 33$; este proceso de eliminación se llevó a cabo utilizando Trimmomatic (<http://www.usadellab.org/cms/?page=trimmomatic>). Una vez completado el proceso de filtrado, se procedió a realizar una nueva evaluación de la calidad de las lecturas utilizando FastQC, con el objetivo de asegurar que los datos filtrados cumplieran con los criterios de calidad establecidos previamente (Bolger et al., 2014).

3.3.3.2. Ensamblaje del genoma cloroplastidial de *O. ficus-indica* y *O. pubescens*

La secuencia consenso del genoma de cloroplasto de *Opuntia ficus-indica* y *Opuntia pubescens* se obtuvo por separado para cada especie.

El ensamblaje *de novo* se llevó a cabo utilizando el software GetOrganelle, disponible en el enlace <https://github.com/Kinggerm/GetOrganelle>. Este software integró varias herramientas, incluyendo Bowtie2 (Langmead & Salzberg, 2012), BLAST (Camacho et al., 2009), SPAdes (Bankevich et al., 2012), así como librerías de Python. Su función principal fue identificar las secuencias relacionadas con el genoma cloroplastidial (cpDNA) y ensamblarlas.

El flujo de trabajo de GetOrganelle comenzó con la identificación de una secuencia semilla en sus librerías, la cual se mapeó con las lecturas proporcionadas por Bowtie2, actuando como “carnadas”. Luego, SPAdes buscó lecturas superpuestas, extendiendo la secuencia hasta completar el ensamblaje de *contigs*. Estos *contigs* se guardaron como archivos de gráfico de ensamblaje (GFA). Posteriormente, se eliminaron los *contigs* ensamblados que no estaban relacionados con la semilla inicial, y se etiquetaron las regiones del ensamblaje según la base de datos de etiquetas de GetOrganelle.

El archivo filtrado y etiquetado se empleó para obtener los mejores resultados de un genoma circular, los cuales fueron sometidos a análisis posterior. Se especificó en GetOrganelle el uso de los datos filtrados previamente (lecturas *forward* y *reverse*), se emplearon 15 rondas de iteraciones de extensión y se buscó un plastidio de planta. Los demás parámetros se mantuvieron por defecto.

El archivo gráfico de ensamblaje obtenido se visualizó utilizando Bandage (<https://rrwick.github.io/Bandage/>) para comprender la estructura del cpDNA y determinar la profundidad de la secuenciación (Wick et al., 2015).

A partir del ensamblaje realizado por GetOrganelle se generaron archivos FASTA por cada especie. Las secuencias obtenidas en el caso de *O. ficus-indica* fueron analizadas

empleando EMBOSS <http://emboss.open-bio.org/>, con la opción *stretcher*, para determinar la secuencia con mayor identidad a los genomas de referencia existentes de esta misma especie mediante alineamiento global (Rice et al., 2000).

3.3.3.3. Anotación del genoma cloroplastidial

La anotación del genoma del cloroplastidial se realizó por separado para cada especie, empleando la herramienta GeSeq (Tillich et al., 2017), siguiendo flujograma desarrollada por Caycho (2023) con algunas modificaciones. GeSeq, disponible en el servidor web [Chlorobox](#), permitió realizar la anotación a partir de los archivos FASTA obtenidos en el ensamblaje anterior, utilizando las opciones por defecto del software.

El software generó dos resultados de anotación, los cuales fueron comparados con las secuencias génicas de los genomas de *Arabidopsis thaliana* (NC_000932), y *Opuntia macbridei* (NC_085762), disponibles en [NCBI](#). Esta comparación permitió determinar cuál de los dos resultados era más adecuado para los análisis posteriores.

Una vez seleccionada una de las dos versiones del ensamblaje, se especificó que la secuencia correspondía a un genoma de cloroplastidial de plantas terrestres. Se procedió a la anotación de diversas características, incluyendo las repeticiones invertidas (IR), las secuencias codificantes de proteínas (CDS), ARN ribosomales (ARNr), ARN de transferencia (ARNt) y las secuencias génicas interespaciadas *rps12*. Para la anotación, se establecieron umbrales del 25 % de identidad para proteínas y del 85 % para ARN.

Se empleó tRNAscan-SE como segundo anotador de ARNt y se utilizó el anotador externo Chloë v0.1.0 (<https://github.com/ian-small/chloe>), disponible en el mismo servidor web (Chan & Lowe, 2019). El servidor devolvió archivos en formato GenBank con los genes identificados.

Posteriormente, se realizó una corrección manual de cada gen identificado, comparándolos con sus homólogos en otros cpDNA de especies del género *Opuntia*, con el fin de completar la anotación de manera precisa y exhaustiva.

Del mismo modo, el reconocimiento y la anotación de las regiones LSC, SSC, IRa e IRb se llevaron a cabo utilizando Geneious Prime® v1.3, en función de los resultados obtenidos con GeSeq.

3.3.3.4. Análisis estructural del genoma cloroplastidial

Para el análisis del cpDNA, se utilizaron los datos obtenidos de la anotación y otros genomas de *Opuntia* como se muestra en la Tabla 3. Para el análisis se empleó Geneious Prime® (<https://www.geneious.com/>), lo cual permitió extraer el orden de los genes y clasificarlos por tipo, en función de las configuraciones cuatripartitas de los genomas (Nhat et al., 2023).

Tabla 3. Números de acceso de GenBank de referencias e incluyendo nuestros 2 genomas de *Opuntia*.

| Especies | Tamaño del genoma pb | GenBank |
|---|----------------------|----------|
| <i>Opuntia ficus-indica</i> (L.) Mill. | 150,927 | PV849967 |
| <i>Opuntia pachyrrhiza</i> H.M.Hern., Gómez-Hin. & Bárcenas | 152,690 | OQ613397 |
| <i>Opuntia stenopetala</i> Engelm. | 152,717 | OQ613406 |
| <i>Opuntia pycnantha</i> Engelm. | 152,441 | OQ613399 |
| <i>Opuntia quimilo</i> K.Schum. | 150,374 | OQ613400 |
| <i>Opuntia arechavaletae</i> Speg. | 149,275 | OQ613377 |
| <i>Opuntia colubrina</i> A.Cast. | 147,853 | OQ613385 |
| <i>Opuntia retrorsa</i> Speg. | 147,640 | OQ613402 |
| <i>Opuntia macbridei</i> Britton & Rose | 148,031 | OQ613392 |
| <i>Opuntia quitensis</i> F.A.C.Weber | 148,114 | OQ613401 |
| <i>Opuntia auberi</i> Pfeiff. | 138,379 | OQ613378 |

| | | |
|--|---------|----------|
| <i>Opuntia dejecta</i> Salm-Dyck | 137,452 | OQ613387 |
| <i>Opuntia gaumeri</i> (Britton & Rose) R.Puente & Majure | 137,818 | OQ613389 |
| <i>Opuntia caracassana</i> Salm-Dyck | 138,141 | OQ613382 |
| <i>Opuntia jamaicensis</i> Britton & Harris | 137,596 | OQ613391 |
| <i>Opuntia guatemalensis</i> Britton & Rose | 137,494 | OQ613390 |
| <i>Opuntia scheeri</i> F.A.C.Weber | 151,001 | OQ613404 |
| <i>Opuntia pubescens</i> H.L.Wendl. ex Pfeiff. | 152,730 | PV849966 |

Fuente propia

3.3.3.5. Análisis genómico comparativo del genoma cloroplastidial

Los genomas de anotados de *O. ficus-indica* y *O. pubescens* se compararon con los de dieciséis especies del género *Opuntia* (Tabla 3) mediante el programa mVISTA (<https://genome.lbl.gov/vista/index.shtml>), empleando el modo Shuffle-LAGAN (Frazer et al., 2004), identificando regiones divergentes.

3.3.3.6. Identificación de secuencias repetitivas

Para la identificación de secuencias repetitivas, se empleó varios programas. Se identificó secuencias SSR (Short Sequence Repeat) empleando el programa software MISA (Beier et al., 2017), aplicando los siguientes parámetros de búsqueda: diez repeticiones para mononucleótidos, ocho para dinucleótidos, cuatro para trinucleótidos y tetranucleótidos, y tres para pentanucleótidos y hexanucleótidos.

Para identificar las repeticiones en tándem (forward, palindromic, reverse y complementarias) en estas especies, utilizamos REPuter (Kurtz & Schleiermacher, 1999) con los siguientes parámetros: distancia de Hamming igual a 3, tamaño mínimo de repetición establecido en 30 pb y tamaño máximo de repetición calculado en 300 pb.

3.3.3.7. Reconstrucción filogenética

Para la reconstrucción filogenética se emplearon los genomas cloroplastidiales de *O. ficus-indica* y *O. pubescens* anotados en el presente trabajo y 39 genomas

cloroplastidiales del género *Opuntia* disponibles en el NCBI (Anexo 2). Previamente se realizó un alineamiento múltiple utilizando MAFFT v7 (Kato & Standley, 2013). Luego se empleó trimAl v1.5 para realizar el recorte de las alineaciones con un umbral de cobertura global del 80 % para eliminar posiciones no conservadas (Capella-Gutiérrez *et al.*, 2009).

La reconstrucción filogenética se llevó a cabo mediante tres enfoques: Máxima Verosimilitud (ML) usando el software IQ-TREE (Nguyen *et al.*, 2015; Minh *et al.*, 2020), empleando 10,000 réplicas de Bootstrap; Máxima Parsimonia (MP) usando MPBoot v.1.1.1 (Hoang *et al.*, 2018) empleando 10,000 réplicas de Bootstrap; e Inferencia Bayesiana (BI) usando MrBayes v.3.2.7 (Ronquist *et al.*, 2012), empleando algoritmo Markov Chain Monte Carlo (MCMC) para 200,000 generaciones, con una frecuencia de muestreo de árboles cada 200,000 generaciones y un burn-in del 20% , utilizando el modelo de sustitución GTR+F+I. El algoritmo Markov Chain Monte Carlo (MCMC) se calculó para 200,000 generaciones, y el árbol de muestreo para cada 10,000 generaciones. El 20% de las generaciones se descartó como burn-in.

3.4. Técnicas para el procesamiento de información

Para el procesamiento de la información, se emplearon técnicas de ensamblaje de novo mediante GetOrganelle, seguido de la anotación del genoma cloroplastidial con *GeSeq*. Las variaciones estructurales se analizaron mediante alineamientos múltiples en *MAFFT* y Geneious Prime ®, para la identificación de SSR se empleará software MISA. Para la reconstrucción filogenética, se utilizaron herramientas de inferencia bayesiana y máxima verosimilitud (MrBayes, MPBoot y IQ-TREE), permitiendo establecer relaciones evolutivas entre *O. ficus-indica*, *O. pubescens* y otras especies del género *Opuntia* disponibles en NCBI.

3.5. Matriz de consistencia

Tabla 4. En este punto se aprecia el cuadro de la operacionalización de las variables estudiadas.

| PROBLEMA | OBJETIVOS | HIPÓTESIS | VARIABLE | DIMENSIONES | INDICADORES | DISEÑO DE INVESTIGACIÓN |
|---|---|---|--|---------------------------------|--|--|
| <p>Problema general</p> <p>¿Qué aportes para la filogenia y la historia evolutiva del género <i>Opuntia</i> puede proporcionar la caracterización detallada del genoma cloroplástico de <i>O. ficus-indica</i> y <i>O. pubescens</i>?</p> | <p>Objetivo general</p> <p>Caracterizar y comparar el genoma cloroplástico de <i>Opuntia ficus-indica</i> y <i>Opuntia pubescens</i>, y reconstruir su filogenia en relación con otras especies del género <i>Opuntia</i></p> | <p>Hipótesis general</p> <p>Los genomas de cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i> presentan diferencias en su estructura, contenido génico y secuencias repetitivas, las cuales reflejan su historia evolutiva y permiten esclarecer las relaciones filogenéticas dentro del género <i>Opuntia</i>.</p> | <ol style="list-style-type: none"> Datos de secuenciación de ADN de lecturas cortas (151 pb). Genoma cloroplástico ensamblado y anotado. Secuencias de cloroplasto propias y otras especies relacionadas. Identificación y caracterización de variaciones estructurales. Secuencias repetitivas identificadas. Relaciones filogenéticas entre las especies de <i>Opuntia</i> | <p>Ensambla y anotar</p> | <ul style="list-style-type: none"> - N° de contigs. - Tamaño total (~120–160 kb). - Cobertura media (> 50×). - N° de k-mers. - Número total de genes anotados (esperado: ~110–130). - Presencia y correcta ubicación de regiones IR, LSC y SSC. - Concordancia con genomas cloroplásticos de referencia (>80% identidad). - Anotación funcional completa y sin errores (genes codificantes, tRNAs, rRNAs). | <p>INVESTIGACIÓN Descriptiva</p> <p>TIPO DE INVESTIGACIÓN Básica</p> <p>DISEÑO No experimental</p> |
| <p>Problema específico</p> <ul style="list-style-type: none"> - ¿Es posible obtener genoma de cloroplasto mediante ensamblaje de novo y anotación de <i>O. ficus-indica</i> y <i>O. pubescens</i> adecuados para análisis comparativos? - ¿Qué variaciones estructurales presentan el genoma de cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i> respecto a otras especies de <i>Opuntia</i>? - ¿Qué tipos y abundancia de secuencias repetitivas contienen el genoma de cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i> y cómo difieren de otras especies de <i>Opuntia</i>? - ¿Qué topología filogenética se obtiene al utilizar genomas de cloroplasto de <i>O. ficus-indica</i>, <i>O. pubescens</i> y otras especies del género <i>Opuntia</i>? | <p>Objetivos específicos</p> <ul style="list-style-type: none"> - Ensamblar de novo y anotar el genoma del cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i>. - Comparar las variaciones estructurales del genoma del cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i>, así como en otras especies de <i>Opuntia</i>. - Identificar y caracterizar secuencias repetitivas en los genomas del cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i>, así como en otras especies de <i>Opuntia</i>. - Realizar una reconstrucción filogenética utilizando secuencias del genoma de cloroplasto de <i>O. ficus-indica</i>, <i>O. pubescens</i> y todas las especies de <i>Opuntia</i> disponibles en el NCBI. | <p>Hipótesis específico</p> <ul style="list-style-type: none"> - Es posible ensamblar y anotar el genoma de cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i> mediante metodologías bioinformáticas actuales. - Los genomas de cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i> presentan expansiones/contracciones de IR, pérdidas/duplicaciones génicas detectables en comparación con otras especies del género <i>Opuntia</i>. - Lo genoma de cloroplasto de <i>O. ficus-indica</i> y <i>O. pubescens</i> difieren en la composición y abundancia de secuencias repetitivas con otros opuntias - La reconstrucción filogenética basada en genomas completos de cloroplasto generará topologías más resolutivas y ubicará de forma consistente a <i>O. ficus-indica</i> y <i>O. pubescens</i> dentro del género <i>Opuntia</i>. | <ol style="list-style-type: none"> Identificación y caracterización de variaciones estructurales. Relaciones filogenéticas entre las especies de <i>Opuntia</i> | <p>Analizar variaciones</p> | <ul style="list-style-type: none"> - Existencia de inversiones, deleciones, duplicaciones, translocaciones. - Tamaño medio de cada evento (bp). - Ubicación (LSC/SSC/IR; génica/intergénica) - N° de genes afectados por inversiones o reordenamientos | |
| | | | | <p>Identificar repeticiones</p> | <ul style="list-style-type: none"> - Número total de secuencias repetitivas. - Porcentaje del genoma ocupado por repeticiones. - Tipos de repeticiones detectadas (palindrómicas, tándem, dispersas, microsátélites). - Comparación del patrón de repeticiones entre las especies analizadas. | |
| | | | | <p>Reconstruir filogenia</p> | <ul style="list-style-type: none"> - Soporte estadístico de las ramas (valores bootstrap o posterior probabilities $\geq 70\%$) - Número total de especies incluidas en el árbol. - Concordancia con clasificaciones taxonómicas previas. - Topología coherente con datos moleculares previos por literatura. | |

CAPITULO IV. RESULTADOS

4.1. Extracción y purificación de ADN total

La extracción de ADN total se realizó con el kti Plant Genomic DNA el cual está basado en el protocolo publicado por Doyle y Doyle (1990). Para la extracción se necesitó 0.1 g de cladodios jóvenes deshidratados y triturado. Al terminar la extracción de ADN se realizó la cuantificación con Nandodrop para determinar las concentraciones de muestras extraídas, si las concentraciones eran muy bajas se concentraban con el centrifugador Concentrador Plus (Eppendorf) se juntaba en un solo tubo de 1.5 ml, y se volvía a cuantificar por Nandodrop y fluorometro Qubit. Los resultados de esta evaluación mostraron que se obtuvo ADN en concentraciones y calidades ($A_{260}/280 \geq 1.80$) en valor adecuadas para ser secuenciadas (Tabla 5). También se corrió un gel de agarosa 1% con 2 μ L de cada muestra, encontrando que el ADN se encontraba mayormente integro como una sola banda, con un tamaño mayor de 10 000 pb como se muestra en la (Figura 6).

Tabla 5. Concentración y calidad del ADN extraído de cladodios jóvenes de *Opuntia ficus-indica* y *Opuntia pubescens*.

| Código de muestra | Concentra en Nanododrop (ng/ μ L) | Concentra en Qubit (ng/ μ L) | A260/280 |
|-------------------|---------------------------------------|----------------------------------|----------|
| OF1 | 177.0 | 81.1 | 1.9 |
| OF1B | 69.5 | 65.8 | 1.8 |
| OP1 | 63.0 | 53.2 | 1.9 |
| OP1B | 130.0 | 65.8 | 1.9 |

A260/A280: relación entre los valores de absorbancia a 260 nm y 280 nm.

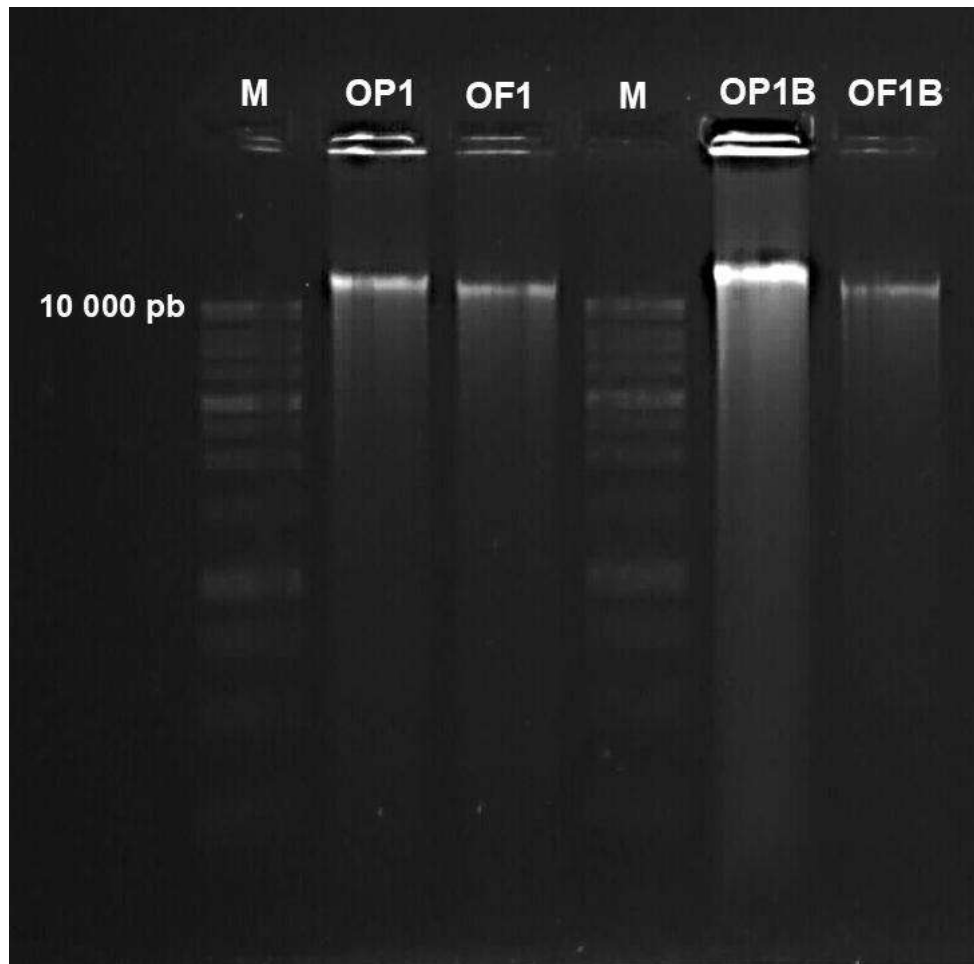


Figura 6. Electroforesis de ADN extraído de *O. ficus-indica* y *O. pubescens* en un gel de agarosa 1% (M: marcador de peso molecular de 1kb, OP1: *O. pubescens*, OF1: *O. ficus-indica*), la muestras OP1B y OF1B son respaldo de muestras OP1 y OF1.

Los mejores valores de DIN (Número de integridad del ADN) fueron las muestras OP1 y OF1B. así mismo las muestras OF1 y OP1B tienen una mayor fragmentación de ADN genómico las cuales fueron obtuvieron mediante la concentración por centrifugación empleando Concentrador Plus a partir de varias muestras extraídas de bajas concentraciones de ADN (Anexo 1).

La secuenciación del ADN produjo un total de 593909978 lecturas (83.56 Gb) para *O. ficus-indica* y 592026904 lecturas (83.33 Gb) para *O. pubescens*, con una longitud

de lectura de 151 pares de bases (pb). Los datos exactos de la distribución de base y calidad de lecturas obtenidos por secuenciación de ADN se encuentran en la Tabla 6.

Tabla 6. Datos de cantidad y calidad obtenidos por el secuenciamiento de ADN por el laboratorio externo.

| Código de muestra | Total de pares de bases (bp) | Total de lecturas | GC% | AT% | Q20% | Q30% |
|-------------------|------------------------------|-------------------|------|------|------|------|
| OF1B | 89 680 406 678 | 593 909 978 | 37.8 | 62.2 | 96.2 | 90.6 |
| OP1 | 89 396 062 504 | 592 026 904 | 38.3 | 61.7 | 96.1 | 90.4 |

4.2. Análisis de calidad y filtrado de lecturas

Los datos obtenidos de la secuenciación se descargaron en una estación de trabajo (computadora de alto rendimiento), la cual se encuentra en el LGBB de la Facultad de Ciencias Biológicas – UNMSM. Los datos fueron analizados empelando FastQC tanto las lecturas forward como reverse de cada uno de las especies, presentado una alta calidad de secuenciamiento (Figura 7 y 9). Posterior se realizó un proceso de filtrado de lecturas utilizando Trimmomatic, estableciendo un umbral de calidad mínimo de Q = 33 en la escala Phred para eliminar nucleótido o lectura cuya calidad estuviera por debajo de este valor.

Después del filtrado, se realizó una nueva evaluación de la calidad, confirmando se encuentran pro encima de Q = 33. Las lecturas totales que pasaron el filtro de calidad en *O. ficus-indica* son 191 787 444, lo que representa el 33.71% de las lecturas originales y las lecturas eliminadas son 200 205 936, equivalente al 33.71% del total inicial (Figura 8). Para *O. pubescens*, el número de lecturas que pasaron el filtro de calidad ascendió a 294,579,560 que equivale 49.76% del total inicial, mientras que 121 211 444 lecturas

fueron descartadas, lo que corresponde al 20.47% de las lecturas originales de esta especie (Figura 10).

4.3. Ensamblaje y anotación de los genomas de cloroplasto

Los datos filtrados se utilizaron para realizar el ensamblaje de novo del ADN de cloroplasto de ambas especies. Para ello, se empleó el programa GetOrganelle, configurado con parámetros específicos para plantas y ejecutado con 15 rondas de iteración, según las recomendaciones del desarrollador.

En el caso del ensamblaje de *O. ficus-indica*, se obtuvieron seis rutas posibles de genomas candidatos. Para seleccionar el ensamblaje más adecuado, se utilizó el paquete EMBOSS, aplicando funciones de comparación con un genoma de referencia (Accesión: NC_085762). Finalmente, se seleccionó la ruta que mostró mayor similitud con el genoma de referencia, la cual presentó una profundidad promedio de cobertura de 550X.

Para *O. pubescens* se obtuvieron dos rutas posibles de ensamblaje, que diferían principalmente en la región SSC (Copia Única Corta) del genoma cloroplastídico. Ambas rutas presentaron una profundidad de cobertura aproximada de 545X. Tras la comparación, se seleccionó la segunda opción por mostrar mayor coherencia estructural con genomas de cloroplasto en *Opuntia*.

Los genomas seleccionados para cada especie fueron invertidos y convertidos a su reverso complementario, ya que el ensamblaje original se encontraba en esa orientación. Posteriormente, se asignaron los nombres genome_OF_12 para *O. ficus-indica* y genome_OP_2 para *O. pubescens*. El tamaño de los genomas ensamblados fue de 152,730 pb y 150,927 pb, respectivamente. Estos genomas correspondieron a la versión final, sobre la cual se procedió a realizar la anotación genética (Figura 11).

Opuntia ficus-indica

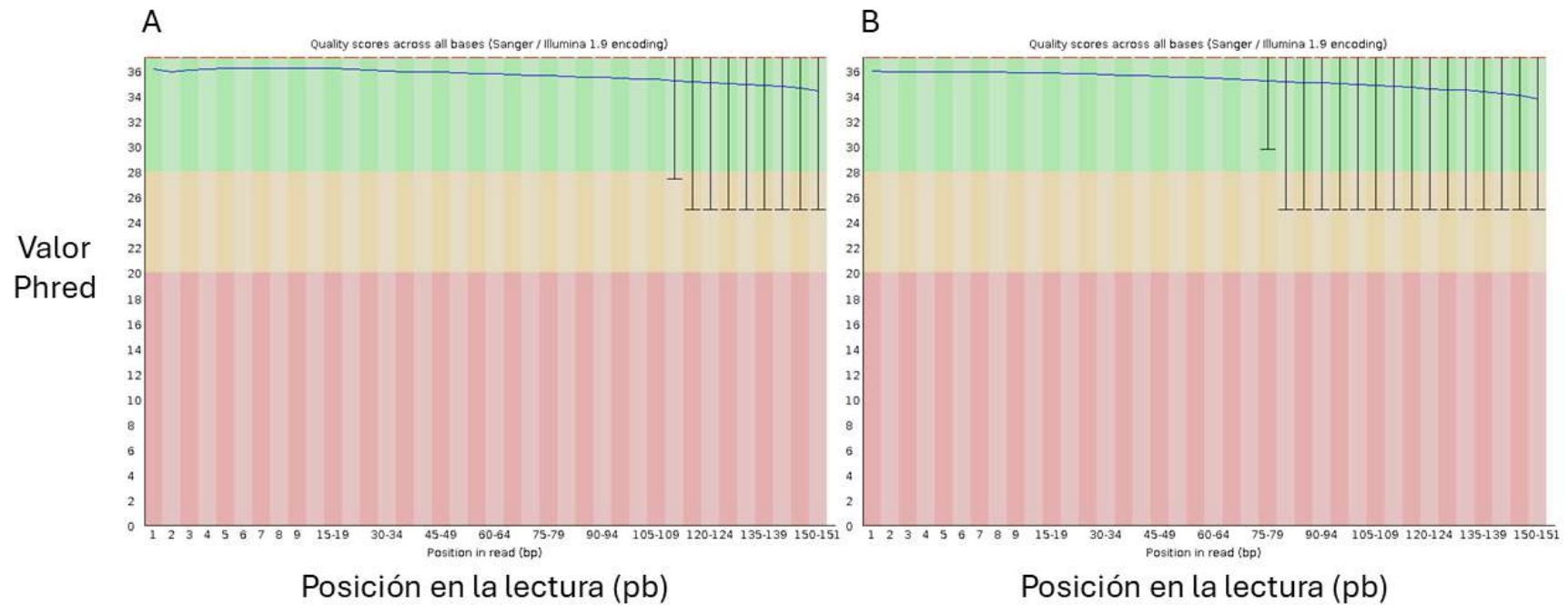


Figura 7. Calidad de las lecturas forward (A) y reverse (B) del ADN secuenciado de *Opuntia ficus-indica*.

Opuntia ficus-indica

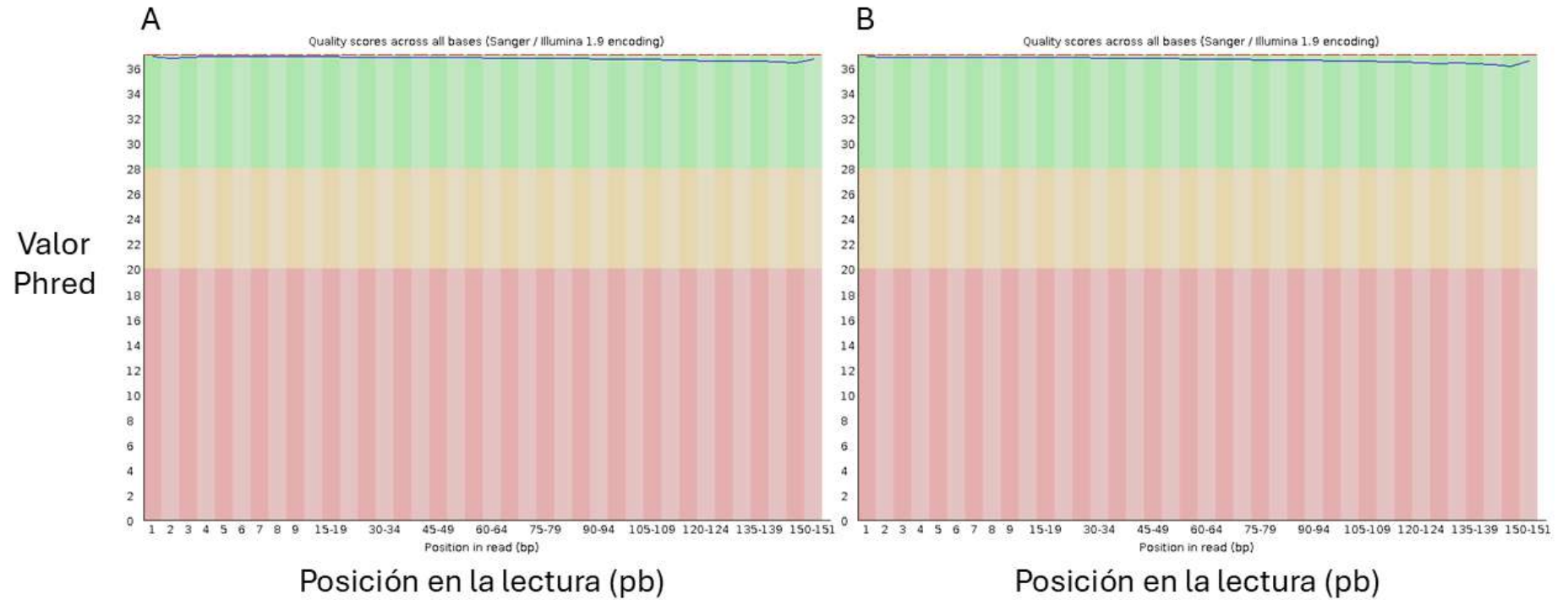


Figura 8. Evaluación de la calidad de las lecturas forward (A) y reverse (B) del ADN secuenciado de *Opuntia ficus-indica* tras el proceso de filtrado, donde se eliminaron aquellas con un valor de calidad inferior a $Q = 33$.

Opuntia pubescens

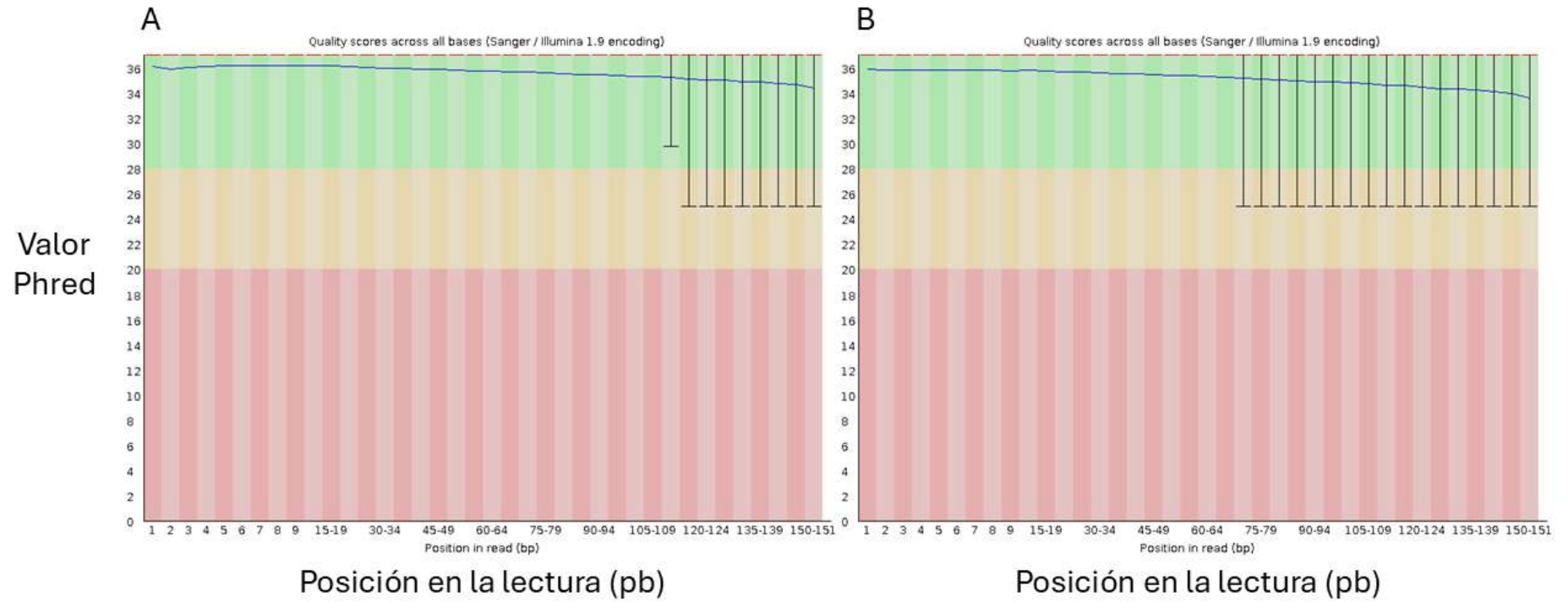


Figura 9. Calidad de las lecturas forward (A) y reverse (B) del ADN secuenciado de *Opuntia pubescens*.

Opuntia pubescens

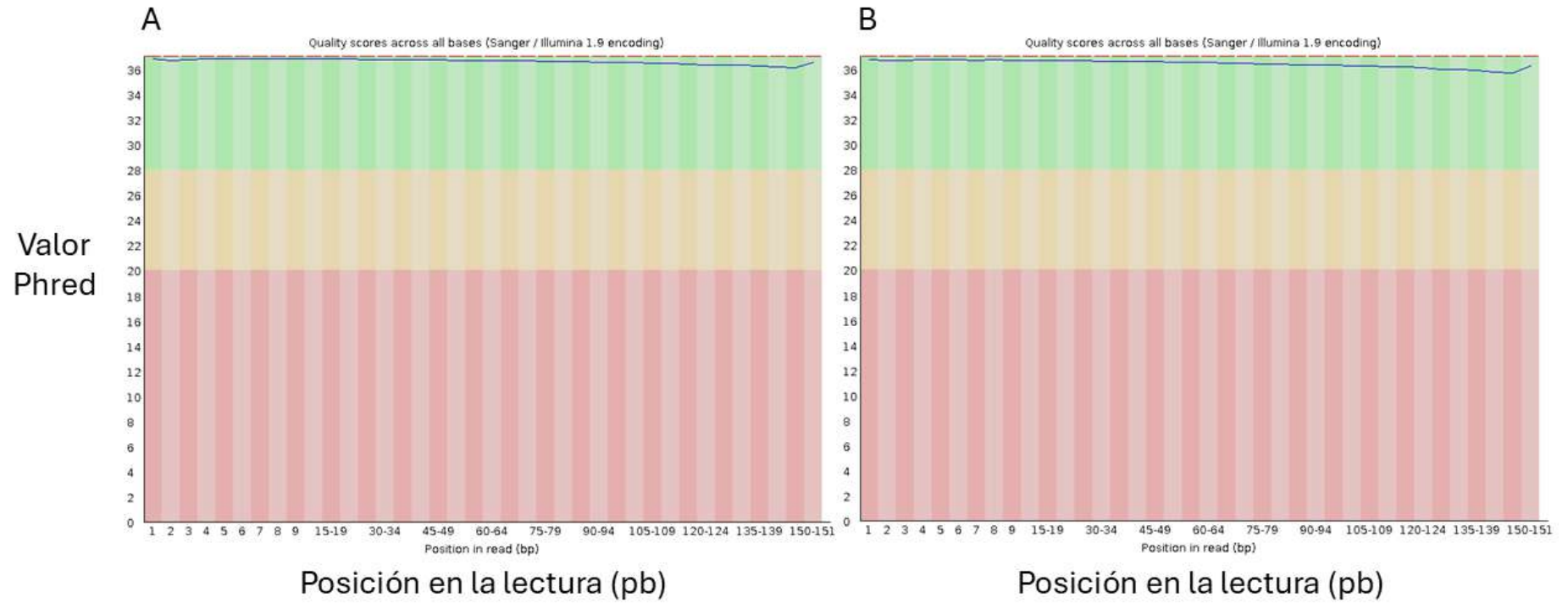


Figura 10. Evaluación de la calidad de las lecturas forward (A) y reverse (B) del ADN secuenciado de *Opuntia pubescens* tras el proceso de filtrado, donde se eliminaron aquellas con un valor de calidad inferior a $Q = 33$.

En *O. ficus-indica* se identificaron un total de 139 genes, de los cuales 25 corresponden a genes duplicados. Este conjunto incluye 89 genes codificantes de proteínas, que abarcan una longitud total de 67 188 pb, 37 genes de ARN de transferencia (ARNt), con una longitud combinada de 2 786 pb, y 8 genes de ARN ribosomal (ARNr), que suman 9 108 pb (Tabla 7).

En *O. pubescens* se identificaron 137 genes totales, de cuales procese 22 genes duplicados. Se identificaron 87 genes codificantes de proteína, que corresponde a 67 086 pb de longitud, 37 genes de ARN de transferencia (ARNt), con una longitud combinada de 2786 pb, y 8 genes de ARN ribosomal (ARNr) a 9 106 pb (Tabla 7).

En *O. ficus-indica* y *O. pubescens* se identificaron los siguientes genes funcionales: 15 genes (*psb*) relacionados con el Fotosistema II, 5 genes (*psa*) del Fotosistema I, 6 genes (*atp*) codificantes de ATP sintasa, 17 genes (*ndh*) asociados a la NADH deshidrogenasa, 6 genes (*pet*) del complejo citocromo b6/f, 1 gen (*rbcL*) correspondiente a la Rubisco, 12 genes (*rpl*) en *O. ficus-indica* y 11 genes (*rpl*) en *O. pubescens*, los cuales codifican para las subunidades grandes de las proteínas ribosomales, 4 genes (*rpo*) asociado polimerasa de ARN dependiente de ADN, 16 genes (*rps*) en *O. ficus-indica* y 15 genes (*rps*) en *O. pubescens* asociados a las subunidades pequeñas de proteínas ribosomales. También se hallaron genes asociados a funciones específicas como el de la subunidad de la Acetil-CoA carboxilasa (*accD*), el gen de síntesis del citocromo tipo c (*ccsA*), la proteína de membrana del sobre del cloroplasto (*cemA*), la subunidad proteolítica de la proteasa Clp (*clpP*), la maturasa K (*matK*), el factor de inicio de la traducción 1 (*infA*), y seis genes (*ycf*) codificantes de proteínas hipotéticas de función desconocida. Asimismo, se observó que *O. pubescens* presenta una copia adicional de los genes *rpl2*, *rpl23* y *rps19*, y cuatro pseudogenes, en comparación con *O. ficus-indica*, que posee cinco pseudogenes (Tabla 8).

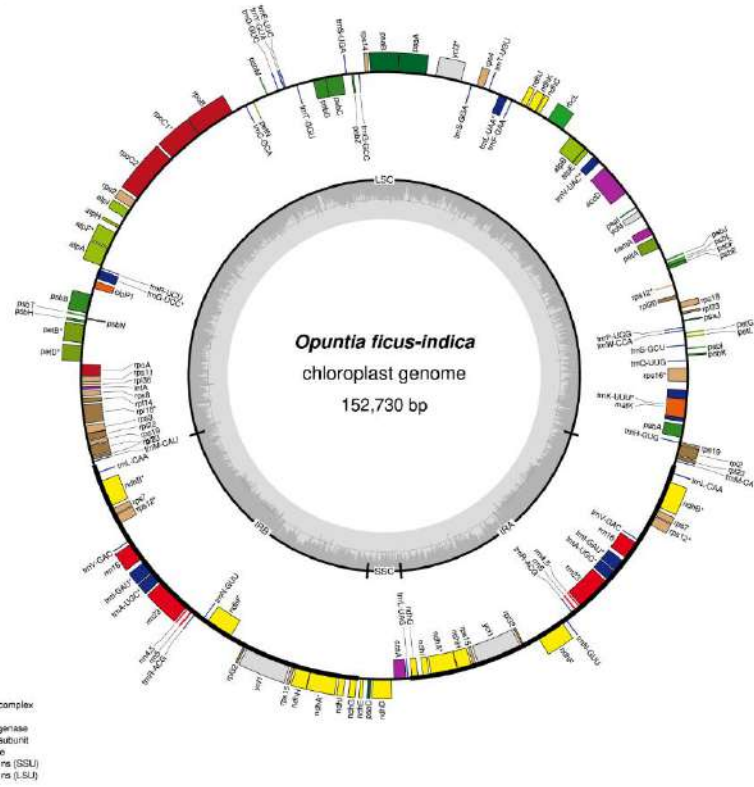
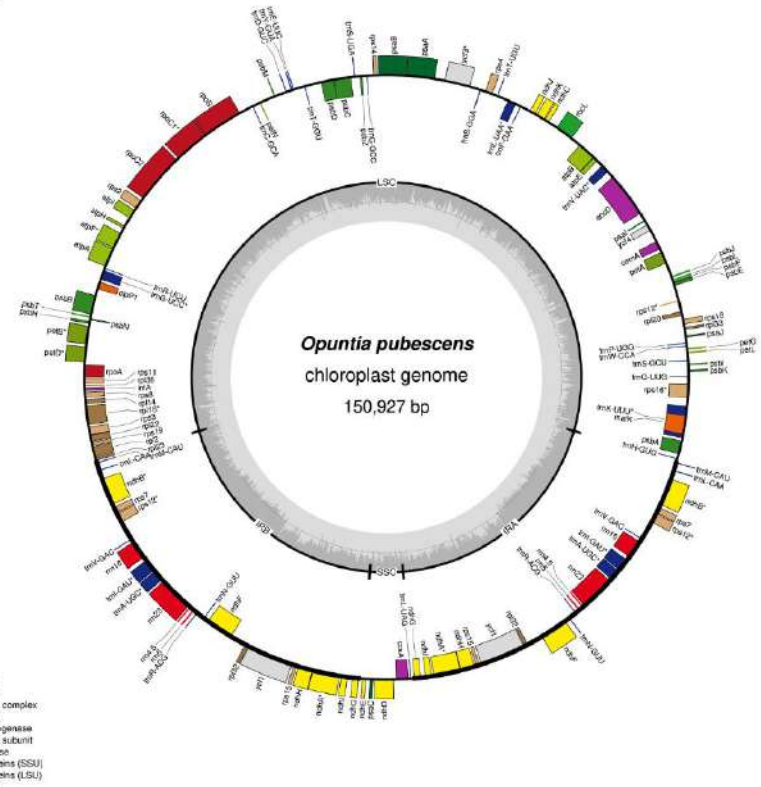
A**B**

Figura 11. Mapas de versión final de los genomas cloroplastidiales de *Opuntia ficus-indica* (A) y *Opuntia pubescens* (B). Se muestra las 4 regiones de genoma: LSC, SSC, IRA e IRB. Los genes ubicados en la parte interna del círculo se transcriben en sentido horario, mientras que los localizados en la parte externa lo hacen en sentido antihorario. Los genes están representados con colores que indican sus respectivos grupos funcionales.

Fuente propia

Tabla 7. Principales características del genoma cloroplastidial de *O. ficus-indica* y *O. pubescens* ensamblado a partir del ADN obtenido de cladodios jóvenes mediante secuenciación de alto rendimiento.

| Características | Valor | Valor |
|---|--------------------------|-----------------------------|
| Especie | <i>Opuntia pubescens</i> | <i>Opuntia ficus-indica</i> |
| Origen del genoma | Cloroplasto | Cloroplasto |
| Tamaño (pb) | 150927 | 152730 |
| Contenido de GC total (%) | 36.6 | 36.6 |
| Tamaño LSC (pb) | 88850 | 87373 |
| Tamaño SSC (pb) | 4108 | 4109 |
| Tamaño IR (pb) | 28985 | 30673 |
| Regiones codificantes de proteínas (pb) | 67086 | 67188 |
| Regiones codificantes de ARNt (pb) | 2786 | 2786 |
| Regiones codificantes de ARNr (pb) | 9106 | 9108 |
| Número de genes | 137 | 139 |
| Número de genes codificantes de proteínas | 87 | 89 |
| Número de ARNr | 8 | 8 |
| Número de ARNt | 37 | 37 |
| Genes con intrones | 23 | 23 |
| Seudogenes | 5 | 5 |

Fuente propia

Tabla 8. Composición genética del genoma cloroplastidial de *O. ficus-indica* y *O. pubescens*.

| Categoría de genes | Grupo funcional | Genes de <i>Opuntia pubescens</i> | Genes de <i>Opuntia ficus-indica</i> |
|---|-------------------------|---|---|
| Fotosíntesis | Fotosistema II | psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ | psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ |
| | Fotosistema I | psaA, psaB, psaC, psaI, psaJ | psaA, psaB, psaC, psaI, psaJ |
| | ATP sintasa | atpA, atpB, atpE, atpF ^b , atpH, atpI | atpA, atpB, atpE, atpF ^b , atpH, atpI |
| | NADH deshidrogenasa | ndhA ^{ab} , ndhB ^{ab} , ndhC, ndhD, ndhE, ndhF ^a , ndhG ^a , ndhH ^a , ndhI ^a , ndhJ, ndhK | ndhA ^{ab} , ndhB ^{ab} , ndhC, ndhD, ndhE, ndhF ^a , ndhG ^a , ndhH ^a , ndhI ^a , ndhJ, ndhK |
| | Complejo Citocromo b6/f | petA, petB ^b , petD ^b , petG, petL, petN | petA, petB ^b , petD ^b , petG, petL, petN |
| | Rubisco | rbcL | rbcL |
| Síntesis de proteínas y replicación del ADN | ARN de transferencia | trnA-UGC ^{ab} , trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-GCC, trnG-UCC ^b , trnH-GUG, trnI-GAU ^{ab} , trnI-CAU ^a , trnK-UUU ^b , trnL-CAA ^a , trnL-UAA ^b , trnL-UAG, trnfM-CAU, trnM-CAU, trnN-GUU ^a , trnP-UGG, trnQ-UUG, trnR-ACG ^a , trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC ^a , trnV-UAC ^b , trnW-CCA, trnY-GUA | trnA-UGC ^{ab} , trnC-GCA, trnD-GUC, trnE-UUC, trnF-GAA, trnG-GCC, trnG-UCC ^b , trnH-GUG, trnI-GAU ^{ab} , trnI-CAU ^a , trnK-UUU ^b , trnL-CAA ^a , trnL-UAA ^b , trnL-UAG, trnfM-CAU, trnM-CAU, trnN-GUU ^a , trnP-UGG, trnQ-UUG, trnR-ACG ^a , trnR-UCU, trnS-GCU, trnS-GGA, trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC ^a , trnV-UAC ^b , trnW-CCA, trnY-GUA |

| Categoría de genes | Grupo funcional | Genes de <i>Opuntia pubescens</i> | Genes de <i>Opuntia ficus-indica</i> |
|---|---|---|--|
| Síntesis de proteínas y replicación del ADN | ARN ribosomal | <i>rrn4.5^a, rrn5^a, rrn16^a, rrn23^a</i> | <i>rrn4.5^a, rrn5^a, rrn16^a, rrn23^a</i> |
| | Subunidades grandes de proteínas ribosomales | <i>rpl2, rpl14, rpl16^b, rpl20, rpl22, rpl23, rpl23^ψ, rpl32^a, rpl33, rpl36</i> | <i>rpl2^a, rpl14, rpl16^b, rpl20, rpl22, rpl23^a, rpl32^a, rpl33, rpl36</i> |
| | Polimerasa de ARN dependiente de ADN | <i>rpoA, rpoB, rpoC1^b, rpoC2</i> | <i>rpoA, rpoB, rpoC1^b, rpoC2</i> |
| | Subunidades pequeñas de proteínas ribosomales | <i>rps2, rps3, rps4, rps7^a, rps8, rps11, rps12^{ab}, rps14, rps15^a, rps16^b, rps18, rps19</i> | <i>rps2, rps3, rps4, rps7^a, rps8, rps11, rps12^{ab}, rps14, rps15^a, rps16^b, rps18, rps19^ψ, rps19</i> |
| Otras funciones | Subunidad Acetil-CoA Carboxilato | <i>accD</i> | <i>accD</i> |
| | Gen de síntesis del Citocromo tipo C | <i>ccsA</i> | <i>ccsA</i> |
| | Proteína de envoltura de membrana | <i>cemA</i> | <i>cemA</i> |
| | Proteasa | <i>clpP</i> | <i>clpP</i> |
| | Maturasa K | <i>matK</i> | <i>matK</i> |
| | Traducción | <i>infA</i> | <i>infA</i> |
| Desconocida | Marcos de lectura conservados | <i>ycf1^{ψa}, ycf2^{ψa}, ycf3^b, ycf4</i> | <i>ycf1^{ψa}, ycf2^{ψa}, ycf3^b, ycf4</i> |

^a genes duplicados, ^b genes con intrón(es), ^ψ pseudogenes.

4.4. Tipos, estructuras y patrones evolutivos de los genomas de cloroplasto

Para analizar los tipos, estructuras y patrones evolutivos de los genomas cloroplastidiales en el género *Opuntia*, seleccionamos aquellos genomas que presentan la típica estructura cuatripartita (LSC, SSC e IR). Estos genomas se detallan en la Tabla 3, la cual incluye tanto nuestras anotaciones como las de otras especies del género. El análisis comparativo de su arquitectura reveló cuatro configuraciones cuatripartitas diferentes (Tipos 1–4), tal como se muestra en la Figura 12. El tamaño de los cloroplastos varió entre 137 y 162 kb y contenía de 126 a 139 genes en total: 83–89 codificantes de proteínas, 32–37 de ARNt y 7–8 de ARNr.

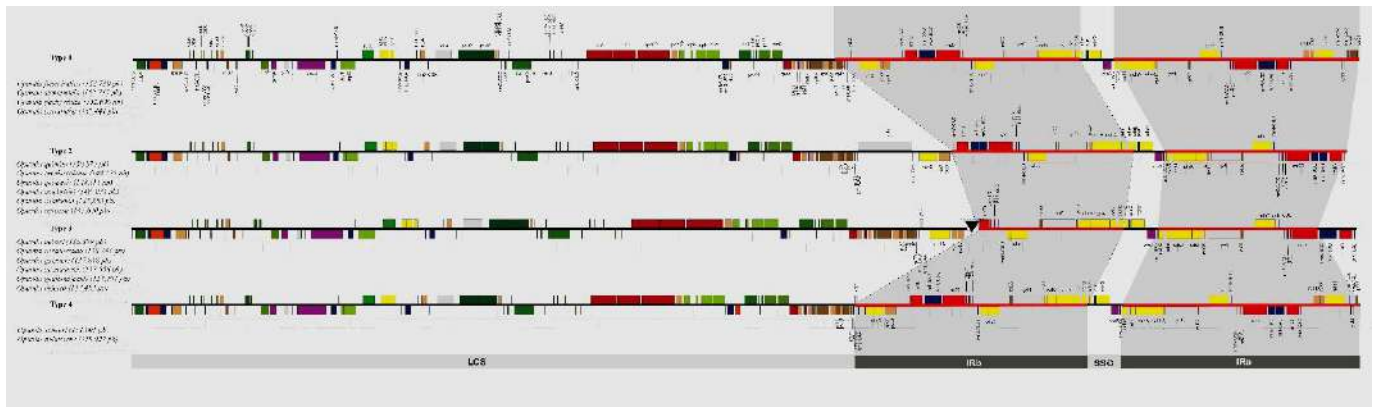


Figura 12. Configuraciones cuatripartitas de los genomas cloroplastidiales linealizados en el género *Opuntia*. Se muestran los cuatro tipos (1–4) identificados de arreglo LSC–IRb–SSC–IRa. Cada gen aparece como un recuadro situado por encima o por debajo de la línea central según su dirección de transcripción (+/–). Las regiones invertidas (IRa e IRb) están coloreadas en rojo. El triángulo punteado señala la reubicación del bloque génico *trnA-UGC* – *trnV-GAC* en la configuración del Tipo 3.

El Tipo 1 es el más completo, con 139 genes totales, y exhibe la LSC extendida desde *trnH-GUG* hasta *rpl22*. Las regiones invertidas IRb e IRa abarcan simétricamente el bloque génico *rps19-ndhG*, flanqueando una SSC de *trnL-UAG* a *ndhE*. Esta simetría

sugiere una alta conservación estructural y funcional, probablemente asociada a una estabilidad replicativa y reparativa óptima de las IR.

El Tipo 2, de tamaño moderado (130 genes totales), muestra una LSC algo más acotada, que va de *trnH-GUG* a *rps12*. Las IR parten de *trnV-GAC* y terminan en *ndhG*, mientras que la SSC mantiene el mismo rango de *trnL-UAG* a *ndhE*. La compresión del límite IR podrían reflejar procesos de contracción selectiva en regiones no críticas para la expresión génica esencial con respecto a genomas de tipo 1 y 4.

El Tipo 3 se caracteriza por su minimalismo genético, con sólo 126 genes totales, y por una reubicación notable de un conjunto de ARNt y ARNr: el bloque *trnA-UGC/trnI-GAU/rrn16/trnV-GAC*, originalmente en la IRs, migra a la LSC. Su LSC tiene un intervalo *trnH-GUG-rpl22*, las IR abarcan de *rrn23* a *ndhG*, y la SSC, de *ndhE* a *trnL-UAG*.

El Tipo 4, intermedio (137 genes totales), presenta LSC de *trnH-GUG* a *rpl2*, IR delimitadas por *rpl23* y *ndhG*, y la tradicional SSC entre *ndhE* y *trnL-UAG*.

4.5. Comparación de los genomas de los cloroplastos

Se examinaron dieciocho genomas cloroplastidiales dentro del género *Opuntia* (*O. ficus-indica*, *O. pachyrrhiza*, *O. stenopetala*, *O. pycnantha*, *O. quimilo*, *O. arechavaletae*, *O. colubrina*, *O. retrorsa*, *O. macbridei*, *O. quitensis*, *O. auberi*, *O. dejecta*, *O. gaumeri*, *O. caracasana*, *O. jamaicensis*, *O. guatemalensis*, *O. scheeri* y *O. pubescens*) utilizando mVISTA para identificar regiones divergentes dentro los genomas como se muestra en la Figura 13. El análisis reveló una baja sintenia en las regiones IR,

acompañada de una notable reorganización génica, lo que sugiere una divergencia evolutiva significativa a nivel estructural del genoma, evidenciada en las Figuras 12 y 13.

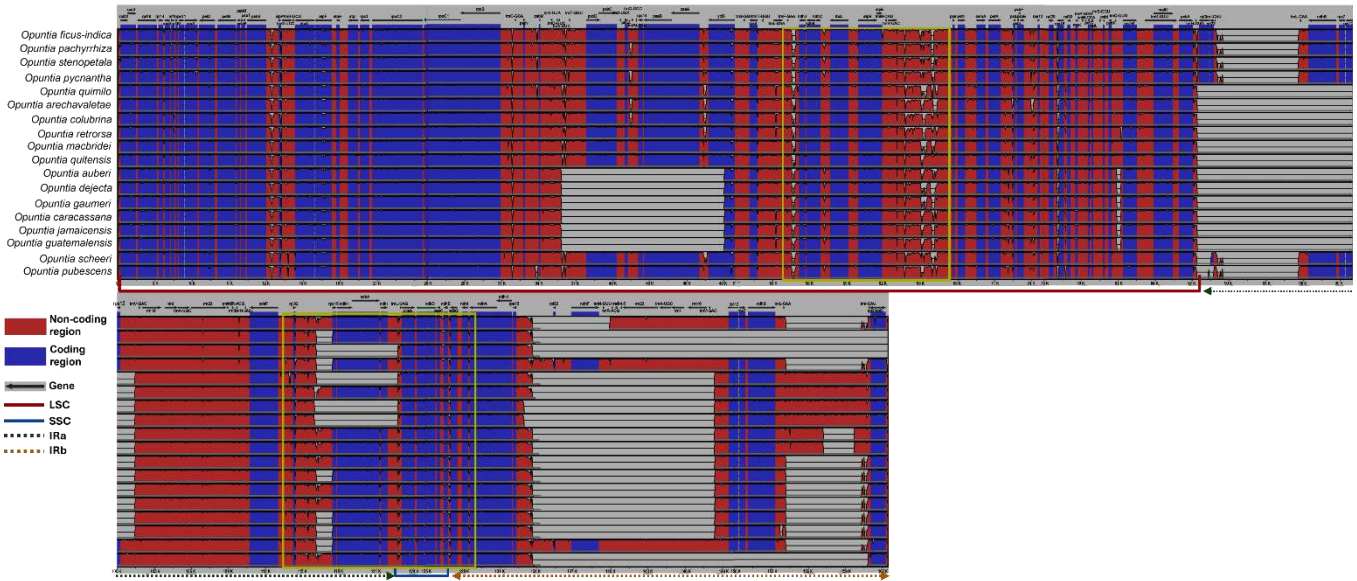


Figura 13. Visualización de 18 genomas de los cloroplastos de género *Opuntia*, mediante mVISTA. Cada banda horizontal representa uno de los genomas analizados. Rojo: regiones no codificantes. Azul: regiones codificantes. Blanco: ausencia de alineación bajo el umbral mínimo de identidad del 70% similitud. Las flechas verdes son genes que se transcriben en sentido 5' a 3' (+) y en sentido 3' a 5' (-). Amarillo: regiones comparadas con alta diversidad.

En general, los resultados revelaron una mayor divergencia en las regiones no codificantes en comparación con las codificantes. Las regiones IRa e IRb mostraron una baja identidad de secuencia e incluso ausencia de alineamiento en algunos casos, lo que sugiere una posible erosión estructural en ciertas especies del género *Opuntia*. Entre las regiones codificantes con mayor variabilidad se identificaron los genes *infA*, *rpoA*, *petD*, *clpP*, *atpF*, *yef3*, *rps18*, *rpl32*, *rps15*, *ndhH*, *ndhA*, *ndhI*, *ccsA* y *ndhE*. Por otro lado, la mayor diversidad en las regiones no codificantes se concentró en varios intervalos

intergénicos, destacando los segmentos comprendidos entre *psbB-clpP*, *rpoB-psbD*, *rps4-ndhJ*, *atpE-psaI*, *psbE-rps12* y *rpl32-rps15*.

Por otro lado, la región SSC resultó ser la más conservada entre los genomas analizados, seguida por la región LSC. Sin embargo, incluso dentro de estas zonas conservadas se detectaron focos de diversidad, particularmente en regiones intergénicas. En la región LSC, el tramo comprendido entre *psbD* y *ycf3* presentó una baja identidad de secuencia en aquellas especies que exhiben la configuración estructural Tipo 3, como se evidencia en la Figura 13.

4.6. Análisis de secuencias repetitivas

4.6.1. Repeticiones de secuencias simples (SSR)

Considerando el papel de secuencias repeticiones simples de genoma cloroplastidial (cpSSRs) como marcadores filogenéticos y recursos para rastrear la evolución genómica, examinamos y cuantificamos seis tipos de repeticiones (mono- a hexanucleotídicas), que engloban dieciocho clases de motivos (Figura 14). El número total de SSR identificados osciló entre 88 en *O. ficus-indica* con valor máximo y 69 en *O. arechavaletae* con el valor mínimo como muestra en la Figura 14A.

Todas las especies compartieron repeticiones mononucleotídicas A/T, seguidas de dinucleotídicas AT/AT y trinucleotídicas AAT/ATT. Entre las pentanucleotídicas, los motivos AAAAC/GTTTT, AAAGT/ACTTT y ACTAG/AGTCT estuvieron presentes en todas las muestras. El motivo C/G también fue casi universal, con la excepción en *O. ficus-indica* y *O. quimilo*, mientras que AATAT/ATATT estuvo ausente en *O. arechavaletae*. Por su parte, AAAT/ATTT resultó ser exclusivo de *O. quitensis* (Figura 14B).

entre los cuales destacan por su abundancia A/T con 59 repeticiones y AAT/ATT con 4 repeticiones. A diferencia de *O. ficus-indica*, *O. pubescens* carece de los motivos C/G, AAAAT/ATTTT y AAATT/AATTT, y no presenta repeticiones hexanucleotídicas.

4.6.2. Secuencias de repetición largas

Las secuencias de repetición largas se detectaron con REPuter, identificando un rango de 95 a 300 repeticiones entre los 18 genomas de *Opuntia* (Figura 15A), donde *O. ficus-indica* alcanzó el valor máximo de 300 y *O. colubrina* el mínimo de 95. Al clasificar los tipos de repeticiones, observamos que las repeticiones directas (Forward) son las más abundantes, que varía desde 178 en *O. pachyrrhiza* hasta 70 en *O. dejecta* (Figura 15B). El segundo grupo en orden de frecuencia lo constituyen las repeticiones Palindrómicas, encabezadas por *O. pachyrrhiza* con 105 repeticiones y con el menor número en *O. retrorsa* 16 repeticiones.

En cuanto a la longitud de la secuencia de repetición, las secuencias de 31–40 pb fueron las más comunes, variando entre 105 en *O. ficus-indica* y 31 repeticiones en *O. arechavaletae* como muestra en la Figura 15C. Les siguieron los intervalos de 0–30 pb y 41–60 pb. Asimismo, se detectó un elevado número de repeticiones mayores a 90 pb en prácticamente todos los genomas analizados, con la excepción de *O. dejecta*, que presentó solo tres de este tamaño. Esta alta densidad de repeticiones > 90 pb sugiere una trayectoria evolutiva en la que los genomas de cloroplasto de *Opuntia* han experimentado episodios recurrentes de duplicación interna o inserciones mediadas por mecanismos recombinantes.

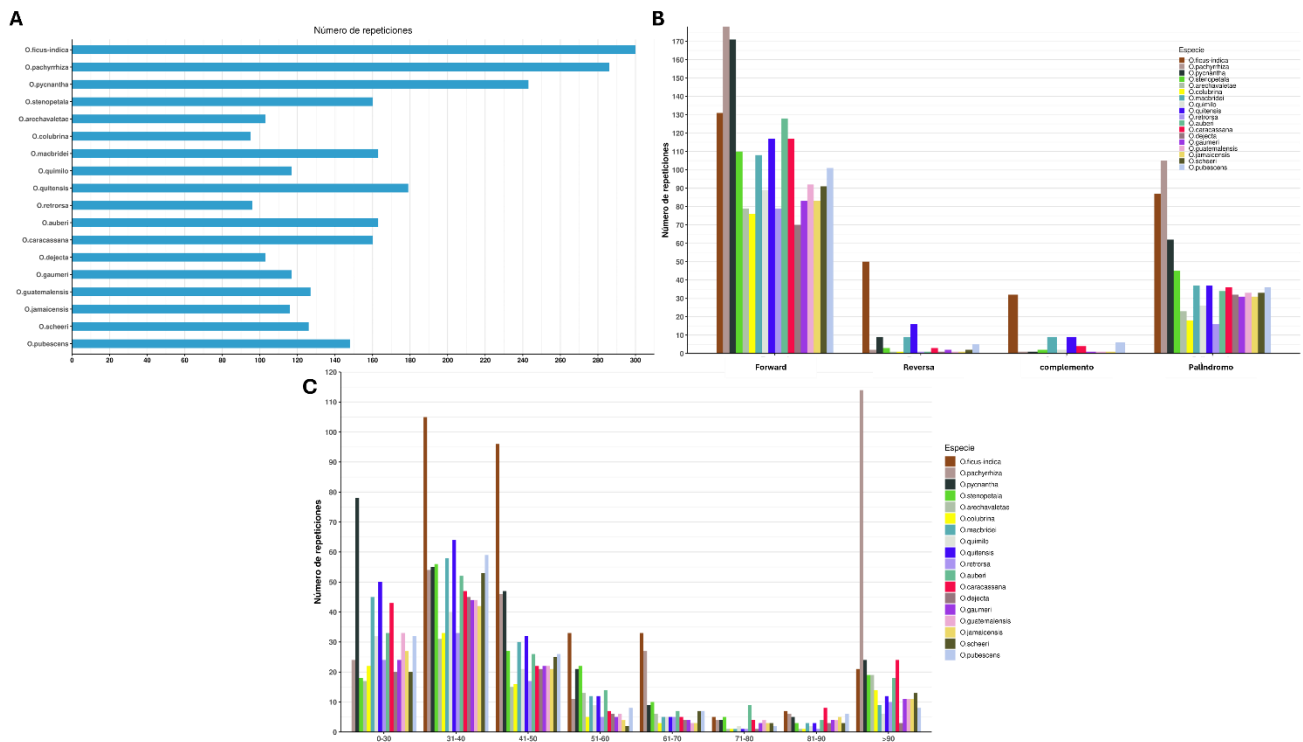


Figura 15. Secuencias de repetición largas los genomas de cloroplastos de *O. ficus-indica* y *O. pubescens*, así como en otras especies de *Opuntia*. (A) Número de repeticiones total por especie. (B) Número de diferentes tipos de repeticiones. (C) Longitud de la secuencia de repeticiones.

En *O. ficus-indica* se detectaron 300 repeticiones largas en el genoma de cloroplasto, frente a 148 en *O. pubescens*. Al clasificarlas por tipo, las repeticiones directas (forward) fueron las más frecuentes (131 en *O. ficus-indica* y 101 en *O. pubescens*), seguidas de las palindrómicas (87 y 36), las reversas (50 y 5) y las de complemento (32 y 6) (Figura 15B).

Respecto a la longitud, el intervalo de 31–40 pb concentró la mayor proporción de repeticiones en ambas especies con 105 en *O. ficus-indica* frente y 59 en *O. pubescens* (Figura 15C). Esta diferencia cuantitativa se alinea con el hecho de que *O. ficus-indica* posee un plastoma de mayor tamaño y un número superior de genes, tal como se observa

en la Fig. 12 y la Tabla 3, lo cual podría favorecer la acumulación de secuencias repetitivas largas.

4.7. Análisis filogenético del género *Opuntia*

Los resultados de reconstrucción filogenética se hicieron utilizando los genomas de cloroplastos de 39 especies previamente publicadas para determinar la posición de *O. ficus-indica* y *O. pubescens*. Para esto se crearon árboles de Máxima Verosimilitud (ML), Máxima Parsimonia (MP) e Inferencia Bayesiana (BI). Los resultados muestran que el género *Opuntia* es un clado monofilético con valores altos de soporte bootstrap del 100% en ML y MP, y probabilidad posterior bayesiana de 1.0 en IB (Fig. 16). Los grupos externos identificados en el árbol corresponden a los géneros *Airamboa*, *Brasiliopuntia* y *Salmonopuntia*.

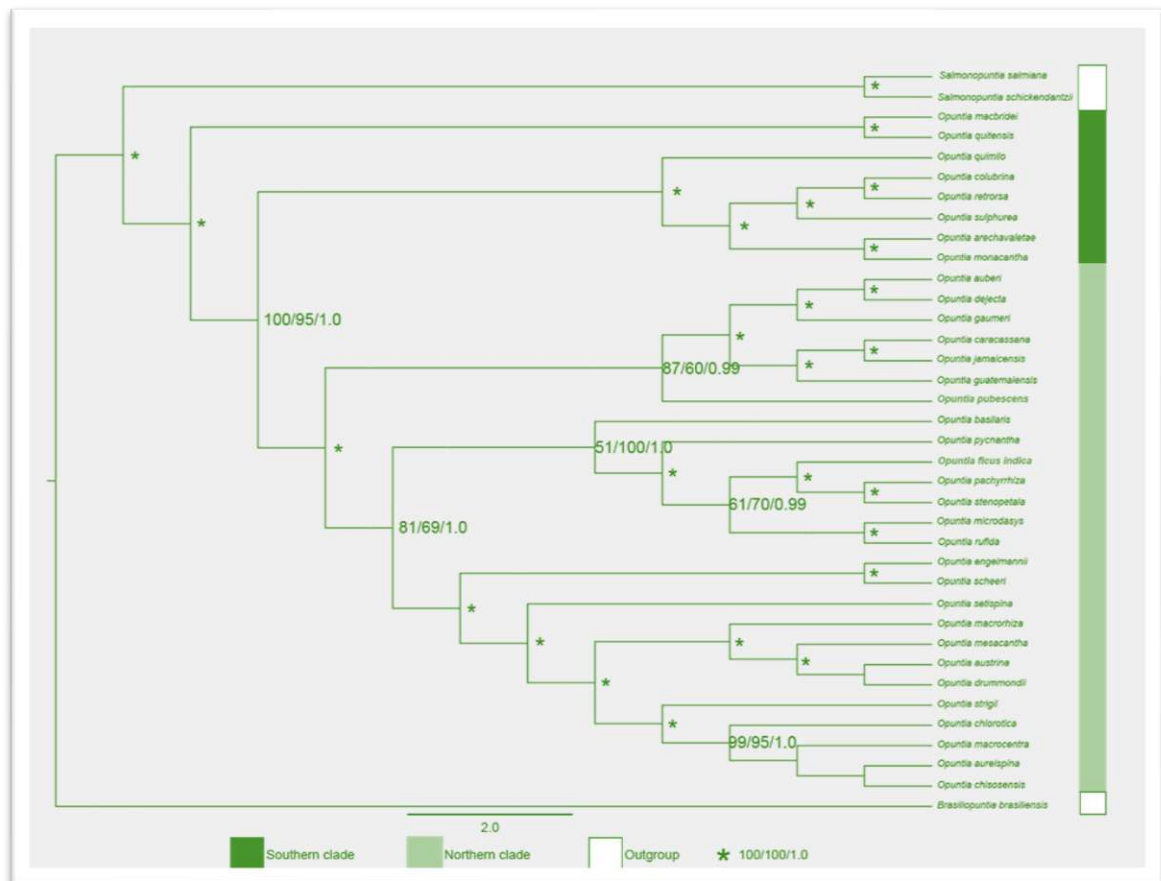


Figura 16. Árbol filogenético de 39 especies para determinar ubicación *O. ficus-indica* y *O. pubescens*. El árbol fue construido empleando 3 métodos de diferentes: Máxima Verosimilitud (ML), Máxima Parsimonia (MP) e Inferencia Bayesiana (BI). Los asteriscos en negro representan 100 % soporte de bootstrap [BP] y 1.0 de probabilidad posterior [PP].

Opuntia pubescens se agrupa con *O. dejecta*, *O. jamaicensis* y *O. guatemalensis* en un clado respaldado por valores altos de soporte (BP=88 y 98 para ML y MP respectivamente; y BI =1.0), mientras que *O. ficus-indica* forma un clado con *O. stenopetala*, *O. pachyrrhiza* y *O. pycnantha*, con soporte de (BP=92 y 94 para ML y MP respectivamente; y BI =1.0). La posición más derivada de *O. ficus-indica* refleja su

historia de domesticación y selección humana tras la divergencia de un ancestro común próximo a *O. pubescens*.

CAPITULO V. DISCUSIÓN

5.1. Discusión de resultados

La investigación genómica empleando tecnología NGS se ha masificado rápidamente permitiendo la secuenciación eficiente de genomas de cloroplasto completo (Moore et al., 2006). Los genomas de cloroplasto tienen varias ventajas sobre el genoma nuclear, como estructura haploide única, herencia materna y baja tasa de evolución (Vadim et al., 2003); además permite diferenciar moléculamente especies de un mismo género y poblaciones o variedades dentro de una misma especie (Yang et al., 2013).

En nuestra investigación, el genoma cloroplastidial de *O. ficus-indica* presenta una longitud de 152 730 pares de bases (pb), ligeramente superior a la reportada por Silva (2021), cuyo ensamblaje alcanza los 152 387 pb, eso se podría deber a heteroplasmia o limitaciones del ensamblaje. Por su parte, el genoma de *O. pubescens* posee 150 927 pb, siendo considerablemente mayor que los tamaños observados en especies relacionadas como *O. dejecta* (137 452 pb), *O. jamaicensis* (137 596 pb) y *O. guatemalensis* (137 494 pb) (Köhler et al., 2023).

En cuanto al contenido génico, el genoma de *O. ficus-indica* incluye un total de 139 genes, ligeramente superior los 137 genes reportados por Silva (2021). Esta diferencia se explica por la presencia de 89 genes codificantes de proteínas, 37 genes de ARN de transferencia (ARNt), 8 genes de ARN ribosomal (ARNr) y cinco pseudogenes, dos más que los identificados en el estudio comparativo. Además, se detectaron dos copias del gen *ycf2* en nuestra anotación. Por otra parte, *O. pubescens* presenta un total de 137 genes, cantidad que coincide con la reportado para *O. scheer*. No obstante, existen

diferencias en la composición génica: *O. pubescens* posee 87 genes codificantes, mientras que *O. scheer* cuenta con 85 genes codificantes. Asimismo, se observa una variación en el número de pseudogenes, ya que *O. scheer* presenta 7 pseudogenes, en contraste con los 5 pseudogenes identificados en *O. pubescens* (Köhler et al., 2023). Aunque estas dos especies comparten la misma cantidad de genes y comparten la estructura del genoma cloroplasto no se encuentra emparentadas, esto se puede deber a que la adaptación convergente a ambientes extremos puede producir similitudes en la composición de los genes de especies no emparentadas (Xu et al., 2020).

El análisis de las configuraciones cuatripartitas de los genomas cloroplastidiales en *Opuntia* muestra una clara concordancia con los tipos estructurales descritos por Köhler et al., (2023). Nuestra estructura tipo 1 se corresponde con su tipo 1; el tipo 2 con su tipo 3; el tipo 3 con su tipo 6; y el tipo 4 con su tipo 9 (Figura 12). En este marco, *O. ficus-indica* se ubica dentro del tipo 1, mientras que *O. pubescens* se asocia al tipo 9.

Un total de 69 a 88 repeticiones de secuencia simple del genoma cloroplastidial (cpSSRs) fueron encontradas en las configuraciones cuatripartitas en género *Opuntia* (Figura 14), menores que los reportado por Silva (2021), el cual identifico entre 253 y 240 SSRs para *O. ficus-indica*, *O. monacantha* y *B. brasiliensis*. En cuanto a la composición de los motivos, los más abundantes en nuestros resultados corresponden al tipo A/T, lo cual coincide con lo observado por Silva (2021). Esta diferencia en la cantidad de repeticiones de secuencia simple podría deberse a que se utilizaron distinto parámetro a compresión de nosotros que utilizamos los parámetros estándares de búsqueda: diez repeticiones para mononucleótidos, ocho para dinucleótidos, cuatro para trinucleótidos y tetranucleótidos, y tres para pentanucleótidos y hexanucleótidos (Contreras-Díaz et al., 2024).

El género *Opuntia* resultó monofilético lo cual concuerda con estudios previos Köhler et al. (2023). Los análisis basados en secuencias completas de plastoma han demostrado un elevado poder resolutivo para inferir las relaciones filogenéticas en este grupo. En nuestro estudio, recuperamos diez clados principales, cifra que difiere de los ocho clados delimitados por Köhler et al. (2023) y de los ocho clados descritos por Majure et al. (2023a), quienes integraron datos de plastoma y caracteres morfológicos. Sin embargo, en un análisis posterior, Majure et al. (2023b) determinaron diez clados principales, lo cual sugiere que las diferencias en el muestreo taxonómico y en los criterios de delimitación pueden influir significativamente en la resolución filogenética.

En nuestros análisis filogenéticos se identifican dos clados con distribución estrictamente sudamericana: el clado Quitensis, integrado por *O. macbridei* y *O. quitensis*; y el clado Elatae, que agrupa a *O. sulphurea*, *O. colubrina*, *O. monacantha*, *O. retrorsa*, *O. arechavaletae*, *O. quimilo* y *O. caracasana* (excluyendo *O. gaumeri* y *O. auberi*) (Majure et al., 2012). El resto de los clados está conformado por especies originarias de Centro y Norteamérica.

Opuntia pubescens presenta una distribución que abarca desde Norteamérica hasta Sudamérica. Filogenéticamente se asocia con especies centroamericanas como *O. jamaicensis*, *O. guatemalensis* y *O. dejecta*. Además, en la Figura 16 nuestro estudio sitúa a *O. pubescens* en un ancestro común con *O. engelmannii* y *O. basilaris*, hallazgo que coincide con lo reportado por Boudjouan et al. (2022).

Por otro lado, *O. ficus-indica*, especie nativa de México, se agrupa en un clado junto a *O. stenopetala*, *O. pachyrrhiza* y *O. pycnantha* las cuales son nativas de norte América. Este patrón filogenético está respaldado por los resultados de Boudjouan et al. (2022), quienes también ubican a *O. ficus-indica* y *O. pycnantha* en el mismo clado.

CAPITULO VI. CONCLUSIÓN Y RECOMENDACIONES

6.1. Conclusión

Presentamos el ensamblaje y la caracterización de los genomas de cloroplasto de *Opuntia ficus-indica* y *Opuntia pubescens*, aportando datos completos sobre su estructura, contenido génico y patrones de repetición. Estos resultados amplían el conocimiento sobre la diversidad estructural de los plastomas en *Opuntia* y proporcionan recursos moleculares útiles para estudios comparativos y taxonómicos.

Las señales filogenéticas derivadas de los genomas de cloroplasto sitúan a *O. ficus-indica* y *O. pubescens* dentro del clado norteamericano de *Opuntia*, lo que refuerza relaciones evolutivas previamente sugeridas. La reconstrucción filogenómica basada en los plastomas constituye un marco sólido para futuros estudios sobre biogeografía, diversidad genética y conservación.

6.2. Recomendación

Para asegurar un desarrollo fluido de la tesis y obtener resultados robustos, proponemos las siguientes recomendaciones:

- Definir fases claras: muestreo, extracción de ADN, control de calidad, ensamblaje del genoma, anotación y análisis comparativo.
- Estandarización de la extracción de ADN: Diseñar el protocolo definitivo, evaluar de forma comparativa varios métodos o kits comerciales (CTAB modificado, columnas de sílice, kits de perlas magnéticas) con cual te permita obtener ADN (≥ 50 ng/ μ L). Y realizar pruebas piloto midiendo rendimiento (μ g de ADN por g de tejido), pureza (260/280 y 260/230) e integridad (electroforesis en gel).
- Verificar la fragmentación y contaminación del ADN con espectrofotómetro y TapeStation/Bioanalyzer.

CAPITULO VII. REFERENCIAS BIBLIOGRÁFICAS

- Abbas, E., Ezzat, M., El-Hefnawy, H., & Abdel-sattar, E. (2022). An overview and update on the chemical composition and potential health benefits of *Opuntia ficus-indica* (L.) Miller. *Journal of Food Biochemistry*, 46. <https://doi.org/10.1111/jfbc.14310>
- Adey, A., Morrison, H. G., Asan, Xun, X., Kitzman, J. O., Turner, E. H., Stackhouse, B., MacKenzie, A. P., Caruccio, N. C., Zhang, X., & Shendure, J. (2010). Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biology*, 11(12), R119. <https://doi.org/10.1186/gb-2010-11-12-r119>
- Agraria.pe. (2022). Perú es el principal productor mundial de cochinilla. Agencia Agraria de Noticias. <https://agraria.pe/noticias/peru-es-el-principal-productor-mundial-de-cochinilla-27137>
- Alamro, H., Alzamel, M., Iliopoulos, C. S., Pissis, S. P., & Watts, S. (2021). IUPACpal: Efficient identification of inverted repeats in IUPAC-encoded DNA sequences. *BMC Bioinformatics*, 22(1), 51. <https://doi.org/10.1186/s12859-021-03983-2>
- Almora, F. C., Retamozo, E. C., & Malpartida, R. P. (2018). Análisis económico de la cadena de valor de tuna (*Opuntia ficus indica*) en Ayacucho. *Investigación*, 26(1), 29-35. <https://doi.org/10.51440/unsch.revistainvestigacion.2018.1.55>
- Arenas, A. A., Cepero, V. C., & Otiniano, A. J. (2017). Caracterización de fincas productoras de tuna (*Opuntia ficus indica*) para la producción de cochinilla del carmín (*Dactylopius coccus*) en La Joya (Arequipa, Perú). *Aporte Santiaguino*, ág. 245-258. <https://doi.org/10.32911/as.2017.v10.n2.167>
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, 19(5), 455-477. <https://doi.org/10.1089/cmb.2012.0021>
- Bansal, K. C., & Saha, D. (2012). Chloroplast Genomics and Genetic Engineering for Crop Improvement. *Agricultural Research*, 1(1), 53-66. <https://doi.org/10.1007/s40003-011-0010-6>
- Barba, F. J., Garcia, C., Fessard, A., Munekata, P. E. S., Lorenzo, J. M., Aboudia, A., Ouadia, A., & Remize, F. (2022). *Opuntia Ficus Indica* Edible Parts: A Food and Nutritional Security Perspective. *Food Reviews International*, 38(5), 930-952. <https://doi.org/10.1080/87559129.2020.1756844>
- Barba, M., Czosnek, H., & Hadidi, A. (2014). Historical Perspective, Development and Applications of Next-Generation Sequencing in Plant Virology. *Viruses*, 6(1), Article 1. <https://doi.org/10.3390/v6010106>
- Barrett, C. F., Baker, W. J., Comer, J. R., Conran, J. G., Lahmeyer, S. C., Leebens-Mack, J. H., Li, J., Lim, G. S., Mayfield-Jones, D. R., Perez, L., Medina, J., Pires, J. C., Santos, C., Wm. Stevenson, D., Zomlefer, W. B., & Davis, J. I. (2016). Plastid

- genomes reveal support for deep phylogenetic relationships and extensive rate variation among palms and other commelinid monocots. *New Phytologist*, 209(2), 855-870. <https://doi.org/10.1111/nph.13617>
- Berhe, Y. K., Portillo, L., & Vázquez-Bolaños, M. (2023). Testing *Opuntia ficus-indica* genotypes for resistance against *Dactylopius coccus* (Hemiptera: Dactylopiidae). *Revista Colombiana de Entomología*, 49(2), Article 2. <https://doi.org/10.25100/socolen.v49i2.12760>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Boudjouan, F., Zeghib, W., Carneiro, J., Silva, R., Morais, J., Vasconcelos, V., & Lopes, G. (2022). Comparison Study on Wild and Cultivated *Opuntia* sp.: Chemical, Taxonomic, and Antioxidant Evaluations. *Agriculture*, 12(11), 1755. <https://doi.org/10.3390/agriculture12111755>
- García-Godos Alcázar, P. . (2019). Diversidad genética de la colección de *Opuntia ficus indica* “tuna” del INIA. *Investigación*, 27(1), 93 –. <https://doi.org/10.51440/unsch.revistainvestigacion.2019.1.108>
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10(1), 421. <https://doi.org/10.1186/1471-2105-10-421>
- Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* (Oxford, England), 25(15), 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>
- Carlos Ostolaza Nano. (2019). Todos los Cactus del Perú. *Sociedad Peruana de Cactáceas y Suculentas*, 572. <https://www.librosperuanos.com/libros/detalle/20173/Todos-los-Cactus-del-Peru>
- Caruso, M., Currò, S., Las Casas, G., La Malfa, S., & Gentile, A. (2010). Microsatellite markers help to assess genetic diversity among *Opuntia ficus indica* cultivated genotypes and their relation with related species. *Plant Systematics and Evolution*, 290(1), 85-97. <https://doi.org/10.1007/s00606-010-0351-9>
- Caycho Ortiz, E. N. (2023). Ensamblaje, anotación y análisis del genoma cloroplastidial del árbol de algarrobo *Neltuma pallida* (subfamilia: Caesalpinioideae). *Universidad Nacional Mayor de San Marcos*. <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/19768>
- Chan, P. P., & Lowe, T. M. (2019). tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods in molecular biology* (Clifton, N.J.), 1962, 1-14. https://doi.org/10.1007/978-1-4939-9173-0_1
- Contreras-Díaz, R., Carevic, F. S., van den Brink, L., Huanca-Mamani, W., & Jung, P. (2024). Structure, gene composition, divergence time and phylogeny analysis of the woody desert species *Neltuma alba*, *Neltuma chilensis* and *Strombocarpa*

- strombulifera. *Scientific reports*, 14(1), 13604. <https://doi.org/10.1038/s41598-024-64287-y>
- Cooper, G. M. (2000). Chloroplasts and Other Plastids. En *The Cell: A Molecular Approach*. 2nd edition. Sinauer Associates. <https://www.ncbi.nlm.nih.gov/books/NBK9905/>
- Daniell, H., Lin, C.-S., Yu, M., & Chang, W.-J. (2016). Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biology*, 17, 134. <https://doi.org/10.1186/s13059-016-1004-2>
- Dobrogojski, J., Adamiec, M., & Luciński, R. (2020). The chloroplast genome: A review. *Acta Physiologiae Plantarum*, 42(6), 98. <https://doi.org/10.1007/s11738-020-03089-x>
- Paolo Inglese, C. S. (2017). *Crop ecology, cultivation and uses of cactus pear*. FAO. <https://www.fao.org/documents/card/es?details=6cb6bff1-66f3-4d2e-92cc-c58130c1f9dc/>
- Egan, A. N., Schlueter, J., & Spooner, D. M. (2012). Applications of next-generation sequencing in plant biology. *American Journal of Botany*, 99(2), 175-185. <https://doi.org/10.3732/ajb.1200020>
- Felker, P., Rodriguez, S. D. C., Casoliba, R. M., Filippini, R., Medina, D., & Zapata, R. (2005). Comparison of *Opuntia ficus indica* varieties of Mexican and Argentine origin for fruit yield and quality in Argentina. *Journal of Arid Environments*, 60(3), 405-422. Scopus. <https://doi.org/10.1016/j.jaridenv.2004.06.003>
- Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., & Dubchak, I. (2004). VISTA: Computational tools for comparative genomics. *Nucleic Acids Research*, 32(Web Server issue), W273-279. <https://doi.org/10.1093/nar/gkh458>
- Freudenthal, J. A., Pfaff, S., Terhoeven, N., Korte, A., Ankenbrand, M. J., & Förster, F. (2020). A systematic comparison of chloroplast genome assembly tools. *Genome Biology*, 21(1), 254. <https://doi.org/10.1186/s13059-020-02153-6>
- Galati, E. M., Mondello, M. R., Giuffrida, D., Dugo, G., Miceli, N., Pergolizzi, S., & Taviano, M. F. (2003). Chemical characterization and biological effects of Sicilian *Opuntia ficus indica* (L.) mill. Fruit juice: Antioxidant and antiulcerogenic activity. *Journal of Agricultural and Food Chemistry*, 51(17), 4903-4908. <https://doi.org/10.1021/jf030123d>
- Ganopoulos, I., Kalivas, A., Kavroulakis, N., Xanthopoulou, A., Mastrogianni, A., Koubouris, G., & Madesis, P. (2015). Diversidad genética de la colección de higos de Berbería (*Opuntia ficus-indica*) en Grecia con marcadores moleculares ISSR. *Plant Gene*, 2, 29-33. <https://doi.org/10.1016/j.plgene.2015.04.001>
- Gao, H., McCormick, A. J., Roston, R. L., & Lu, Y. (2023). Editorial: Structure and function of chloroplasts, Volume III. *Frontiers in Plant Science*, 14. <https://doi.org/10.3389/fpls.2023.1145680>

- Gao, L., Su, Y.-J., & Wang, T. (2010). Plastid genome sequencing, comparative genomics, and phylogenomics: Current status and prospects. *Journal of Systematics and Evolution*, 48(2), 77-93. <https://doi.org/10.1111/j.1759-6831.2010.00071.x>
- Green, B. R. (2011). Chloroplast genomes of photosynthetic eukaryotes. *The Plant Journal*, 66(1), 34-44. <https://doi.org/10.1111/j.1365-313X.2011.04541.x>
- Guevara, J. C., Estevez, O. R., Stasi, C. R., & Monge, A. S. (1996). Botanical composition of the seasonal diet of cattle in the rangelands of the Monte Desert of Mendoza, Argentina. *Journal of Arid Environments*, 32(4), 387-394. <https://doi.org/10.1006/jare.1996.0032>
- Guo, W., Grewe, F., Cobo-Clark, A., Fan, W., Duan, Z., Adams, R. P., Schwarzbach, A. E., & Mower, J. P. (2014). Predominant and substoichiometric isomers of the plastid genome coexist within Juniperus plants and have shifted multiple times during cupressophyte evolution. *Genome Biology and Evolution*, 6(3), 580-590. <https://doi.org/10.1093/gbe/evu046>
- Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, 22(2), 160-174. <https://doi.org/10.1007/BF02101694>
- Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., & Zamir, A. (1965). Structure of a Ribonucleic Acid. *Science*, 147(3664), 1462-1465. <https://doi.org/10.1126/science.147.3664.1462>
- Hong, K., Radian, Y., Manda, T., Xu, H., & Luo, Y. (2023). The Development of Plant Genome Sequencing Technology and Its Conservation and Application in Endangered Gymnosperms. *Plants*, 12(23), Article 23. <https://doi.org/10.3390/plants12234006>
- Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., dePamphilis, C. W., Yi, T.-S., & Li, D.-Z. (2020). GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology*, 21(1), 241. <https://doi.org/10.1186/s13059-020-02154-5>
- Kamimura, Y., Tanaka, H., Kobayashi, Y., Shikanai, T., & Nishimura, Y. (2018). Chloroplast nucleoids as a transformable network revealed by live imaging with a microfluidic device. *Communications Biology*, 1(1), 1-7. <https://doi.org/10.1038/s42003-018-0055-1>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772-780. <https://doi.org/10.1093/molbev/mst010>
- Kim, S.-C., Ha, Y.-H., Park, B. K., Jang, J. E., Kang, E. S., Kim, Y.-S., Kimspe, T.-H., & Kim, H.-J. (2023). Comparative analysis of the complete chloroplast genome of Papaveraceae to identify rearrangements within the Corydalis chloroplast genome. *PLOS ONE*, 18(9), e0289625. <https://doi.org/10.1371/journal.pone.0289625>

- Kirchhoff, H. (2019). Chloroplast ultrastructure in plants. *New Phytologist*, 223(2), 565-574. <https://doi.org/10.1111/nph.15730>
- Köhler, M., Reginato, M., Jin, J.-J., & Majure, L. C. (2023). More than a spiny morphology: Plastome variation in the prickly pear cacti (Opuntieae). *Annals of Botany*, 132(4), 771-786. <https://doi.org/10.1093/aob/mcad098>
- Krupinska, K., Melonek, J., & Krause, K. (2013). New insights into plastid nucleoid structure and functionality. *Planta*, 237(3), 653-664. <https://doi.org/10.1007/s00425-012-1817-5>
- Kuroiwa, T. (1991). The Replication, Differentiation, and Inheritance of Plastids with Emphasis on the Concept of Organelle Nuclei. En K. W. Jeon & M. Friedlander (Eds.), *International Review of Cytology* (Vol. 128, pp. 1-62). Academic Press. [https://doi.org/10.1016/S0074-7696\(08\)60496-9](https://doi.org/10.1016/S0074-7696(08)60496-9)
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4), 357-359. <https://doi.org/10.1038/nmeth.1923>
- Liebers, M., Cozzi, C., Uecker, F., Chambon, L., Blanvillain, R., & Pfannschmidt, T. (2022). Biogenic signals from plastids and their role in chloroplast development. *Journal of Experimental Botany*, 73(21), 7105-7125. <https://doi.org/10.1093/jxb/erac344>
- Llorente, B., Segretin, M. E., Giannini, E., Lobais, C., Juárez, M. E., Paulsen, I. T., & Blanco, N. E. (2021). Homecoming: Rewinding the reductive evolution of the chloroplast genome for increasing crop yields. *Nature Communications*, 12(1), 6734. <https://doi.org/10.1038/s41467-021-26975-5>
- López-Romero, P., Pichardo-Ontiveros, E., Avila-Nava, A., Vázquez-Manjarrez, N., Tovar, A. R., Pedraza-Chaverri, J., & Torres, N. (2014). The Effect of Nopal (*Opuntia Ficus Indica*) on Postprandial Blood Glucose, Incretins, and Antioxidant Activity in Mexican Patients with Type 2 Diabetes after Consumption of Two Different Composition Breakfasts. *Journal of the Academy of Nutrition and Dietetics*, 114(11), 1811-1818. <https://doi.org/10.1016/j.jand.2014.06.352>
- Liu J, Feng Y, Chen C, Yan J, Bai X, Li H, Lin C, Xiang Y, Tian W, Qi Z, Yu J and Yan X (2024) Genomic insights into the clonal reproductive *Opuntia cochenillifera*: mitochondrial and chloroplast genomes of the cochineal cactus for enhanced understanding of structural dynamics and evolutionary implications. *Front. Plant Sci.* 15:1347945. doi: 10.3389/fpls.2024.1347945
- Mannai, F., Elhleli, H., Ben Mosbah, M., Khiari, R., Nacer, S. N., Belgacem, M. N., & Moussaoui, Y. (2024). Comparative study of conventional and combined ultrasound-assisted methods on the quality of mucilage extracted from *Opuntia ficus-indica* cladodes. *Industrial Crops and Products*, 214, 118566. <https://doi.org/10.1016/j.indcrop.2024.118566>
- Majure, L. C., Achá, S., Baker, M. A., Puente-Martínez, R., Köhler, M., & Fehlberg, S. (2023). Phylogenomics of One of the World's Most Intriguing Groups of CAM Plants, the Opuntioids (Opuntioideae: Cactaceae): Adaptation to Tropical Dry

- Forests Helped Drive Prominent Morphological Features in the Clade. *Diversity*, 15(4), 570. <https://doi.org/10.3390/d15040570>
- Majure, L. C., Murphy, T. H., Köhler, M., Puente, R., & Hodgson, W. C. (2023). Evolution of the Xerocarpa clade (Opuntia; Opuntieae): Evidence for the Role of the Grand Canyon in the Biogeographic History of the Iconic Beavertail Cactus and Relatives. *Plants*, 12(14), 2677. <https://doi.org/10.3390/plants12142677>
- Majure, L.C., Puente, R., Griffith, M.P., Judd, W.S., Soltis, P.S. and Soltis, D.E. (2012), Phylogeny of *Opuntia* s.s. (Cactaceae): Clade delineation, geographic origins, and reticulate evolution†. *American Journal of Botany*, 99: 847-864. <https://doi.org/10.3732/ajb.1100375>
- Michael, D., Gurusaran, M., Santhosh, R., Hussain, M. K., Satheesh, S. N., Suhan, S., Sivaranjan, P., Jaiswal, A., & Sekar, K. (2019). RepEx: A web server to extract sequence repeats from protein and DNA sequences. *Computational Biology and Chemistry*, 78, 424-430. <https://doi.org/10.1016/j.compbiolchem.2018.12.015>
- MIDAGRI. (2022). *Análisis de Mercado—Tuna 2015—2021*. <https://www.gob.pe/institucion/sse/informes-publicaciones/2765089-analisis-de-mercado-tuna-2015-2021>
- Mock, T., Otilar, R. P., Strauss, J., McMullan, M., Paajanen, P., Schmutz, J., Salamov, A., Sanges, R., Toseland, A., Ward, B. J., Allen, A. E., Dupont, C. L., Frickenhaus, S., Maumus, F., Veluchamy, A., Wu, T., Barry, K. W., Falciatore, A., Ferrante, M. I., ... Grigoriev, I. V. (2017). Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature*, 541(7638), 536-540. <https://doi.org/10.1038/nature20803>
- Mohamed–Yasseen, Y., Barringer, S. A., & Splittstoesser, W. E. (1996). A note on the uses of *Opuntia* spp. In Central/North America. *Journal of Arid Environments*, 32(3), 347-353. <https://doi.org/10.1006/jare.1996.0028>
- Moore, M.J., Dhingra, A., Soltis, P.S. (2006). Rapid and accurate pyrosequencing of angiosperm plastid genomes. *BMC Plant Biol* 6, 17. <https://doi.org/10.1186/1471-2229-6-17>
- Morley, S. A., Peralta-Castro, A., Brieba, L. G., Miller, J., Ong, K. L., Ridge, P. G., Oliphant, A., Aldous, S., & Nielsen, B. L. (2019). *Arabidopsis thaliana* organelles mimic the T7 phage DNA replisome with specific interactions between Twinkle protein and DNA polymerases Pol1A and Pol1B. *BMC Plant Biology*, 19(1), 241. <https://doi.org/10.1186/s12870-019-1854-3>
- Nevo, R., Charuvi, D., Tsabari, O., & Reich, Z. (2012). Composition, architecture and dynamics of the photosynthetic apparatus in higher plants. *The Plant Journal: For Cell and Molecular Biology*, 70(1), 157-176. <https://doi.org/10.1111/j.1365-313X.2011.04876.x>
- Nguyen, K. L., Grondin, A., Courtois, B., & Gantet, P. (2019). Next-Generation Sequencing Accelerates Crop Gene Discovery. *Trends in Plant Science*, 24(3), 263-274. <https://doi.org/10.1016/j.tplants.2018.11.008>

- Nhat Nam, N., Hoang Danh, N., Minh Thiet, V., & Do, H. D. K. (2023). New Insights Into The Evolution of Chloroplast Genomes in *Ochna* Species (Ochnaceae, Malpighiales). *Evolutionary Bioinformatics Online*, 19, 11769343231210756. <https://doi.org/10.1177/11769343231210756>
- Novoa, S. (2006). *Sobre el Origen de la Tuna en el Perú Algunos alcances*. <https://www.semanticscholar.org/paper/Sobre-el-Origen-de-la-Tuna-en-el-Per%C3%BA-Algunos-Novoa/526bbaf72d1af973c2fe5c30c7b6baa80748620e>
- Palmer, J. D. (1983). Chloroplast DNA exists in two orientations. *Nature*, 301(5895), 92-93. <https://doi.org/10.1038/301092a0>
- Pauca-Tanco, G. A. & V. Quipuscoa. 2020. El género *Cumulopuntia* (Cactaceae, Opuntioideae) en el Departamento de Arequipa, Perú. *Darwiniana*, nueva serie 8(1): 337-371. <https://doi.org/10.14522/darwiniana.2020.81.853>
- Płoski, R. (2016). Chapter 1—Next Generation Sequencing—General Information about the Technology, Possibilities, and Limitations. En U. Demkow & R. Płoski (Eds.), *Clinical Applications for Next-Generation Sequencing* (pp. 1-18). Academic Press. <https://doi.org/10.1016/B978-0-12-801739-5.00001-5>
- Pribil, M., Labs, M., & Leister, D. (2014). Structure and dynamics of thylakoids in land plants. *Journal of Experimental Botany*, 65(8), 1955-1972. <https://doi.org/10.1093/jxb/eru090>
- Qu, X.-J., Wu, C.-S., Chaw, S.-M., & Yi, T.-S. (2017). Insights into the Existence of Isomeric Plastomes in Cupressoideae (Cupressaceae). *Genome Biology and Evolution*, 9(4), 1110-1119. <https://doi.org/10.1093/gbe/evx071>
- Rafael-Ramírez Héctor., Espino-Rodríguez, Edgar., Tingal-Infante, Isolina. (2012). Morfología de tuna (*Opuntia ficus indica* Mill.) <http://hdl.handle.net/20.500.14074/2853>
- Rascio, N. (2013). Chloroplasts. En W. J. Lennarz & M. D. Lane (Eds.), *Encyclopedia of Biological Chemistry (Second Edition)* (pp. 506-510). Academic Press. <https://doi.org/10.1016/B978-0-12-378630-2.00141-9>
- Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: The European Molecular Biology Open Software Suite [Software]. European Bioinformatics Institute. <https://emboss.sourceforge.net/>
- Rodríguez, L. C., Pascual, U., & Niemeyer, H. M. (2006). Local identification and valuation of ecosystem goods and services from *Opuntia* scrublands of Ayacucho, Peru. *Ecological Economics*, 57(1), 30-44. <https://doi.org/10.1016/j.ecolecon.2005.03.022>
- Rodríguez, L. C., & Pascual, U. (2004). Land clearance and social capital in mountain agro-ecosystems: The case of *Opuntia* scrubland in Ayacucho, Peru. *Ecological Economics*, 49(2), 243-252. <https://doi.org/10.1016/j.ecolecon.2004.03.023>

- Ruiz Pérez-Cacho, M. P., Galán-Soldevilla, H., Corrales García, J., & Hernández Montes, A. (2006). Sensory characterization of nopalitos (*Opuntia* spp.). *Food Research International*, 39(3), 285-293. <https://doi.org/10.1016/j.foodres.2005.08.001>
- Sáenz, C., Sepúlveda, E., & Matsuhira, B. (2004). *Opuntia* spp mucilage's: A functional component with industrial perspectives. *Journal of Arid Environments*, 57(3), 275-290. [https://doi.org/10.1016/S0140-1963\(03\)00106-X](https://doi.org/10.1016/S0140-1963(03)00106-X)
- Sakai, A., Takano, H., & Kuroiwa, T. (2004). Organelle Nuclei in Higher Plants: Structure, Composition, Function, and Evolution. En *International Review of Cytology* (Vol. 238, pp. 59-118). Academic Press. [https://doi.org/10.1016/S0074-7696\(04\)38002-2](https://doi.org/10.1016/S0074-7696(04)38002-2)
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74(12), 5463-5467. <https://doi.org/10.1073/pnas.74.12.5463>
- Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nature Methods*, 5(1), 16-18. <https://doi.org/10.1038/nmeth1156>
- Schwarz, E. N., Ruhlman, T. A., Sabir, J. S. M., Hajrah, N. H., Alharbi, N. S., Al-Malki, A. L., Bailey, C. D., & Jansen, R. K. (2015). Plastid genome sequences of legumes reveal parallel inversions and multiple losses of rps16 in papilionoids. *Journal of Systematics and Evolution*, 53(5), 458-468. <https://doi.org/10.1111/jse.12179>
- Sejuro Nanetti, O. (1988). *Plantas medicinales utilizadas por los curanderos de Nasca*. CONCYTEC.
- Shirazinia, R., Rahimi, V. B., Kehkhaie, A. R., Sahebkar, A., Rakhshandeh, H., & Askari, V. R. (2019). *Opuntia dillenii*: A Forgotten Plant with Promising Pharmacological Properties. *Journal of Pharmacopuncture*, 22(1), 16-27. <https://doi.org/10.3831/KPI.2019.22.002>
- Silva, M. C. (2021). Aspectos evolutivos de los plástidos de la subfamilia Opuntioideae (Cactaceae). 2021. 120 págs. Tesis (Maestría en Fisiología Vegetal) - Universidad Federal de Viçosa.
- Singh, V. K., Singh, A. K., Singh, S., & Singh, B. D. (2015). Next-Generation Sequencing (NGS) Tools and Impact in Plant Breeding. *Advances in Plant Breeding Strategies: Breeding, Biotechnology and Molecular Tools* (pp. 563-612). Springer International Publishing. https://doi.org/10.1007/978-3-319-22521-0_20
- Sloan, D. B., Alverson, A. J., Wu, M., Palmer, J. D., & Taylor, D. R. (2012). Recent Acceleration of Plastid Sequence and Structural Evolution Coincides with Extreme Mitochondrial Divergence in the Angiosperm Genus *Silene*. *Genome Biology and Evolution*, 4(3), 294-306. <https://doi.org/10.1093/gbe/evs006>
- Song, B., Ning, W., Wei, D., Jiang, M., Zhu, K., Wang, X., Edwards, D., Odeny, D. A., & Cheng, S. (2023). Plant genome resequencing and population genomics: Current status and future prospects. *Molecular Plant*, 16(8), 1252-1268. <https://doi.org/10.1016/j.molp.2023.07.009>

- Staehelin, L. A. (1986). Chloroplast Structure and Supramolecular Organization of Photosynthetic Membranes. En L. A. Staehelin & C. J. Arntzen (Eds.), *Photosynthesis III: Photosynthetic Membranes and Light Harvesting Systems* (pp. 1-84). Springer. https://doi.org/10.1007/978-3-642-70936-4_1
- Stefan Kurtz. (2017). The Vmatch large scale sequence analysis software. <http://www.vmatch.de/>
- Tangphatsornruang, S., Sangsrakru, D., Chanprasert, J., Uthaipaisanwong, P., Yoocha, T., Jomchai, N., & Tragoonrung, S. (2010). The Chloroplast Genome Sequence of Mungbean (*Vigna radiata*) Determined by High-throughput Pyrosequencing: Structural Organization and Phylogenetic Relationships. *DNA Research*, 17(1), 11-22. <https://doi.org/10.1093/dnares/dsp025>
- Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., & Greiner, S. (2017). GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research*, 45(W1), W6-W11. <https://doi.org/10.1093/nar/gkx391>
- Tineo Canchari, J. I. (2019). Caracterización morfológica y análisis de la variabilidad genética de la colección nacional de germoplasma de tuna (*Opuntia* sp.) del Perú. *Instituto Nacional de Innovación Agraria*. <https://repositorio.inia.gob.pe/handle/20.500.12955/1766>
- Tsumura, Y., Suyama, Y., & Yoshimura, K. (2000). Chloroplast DNA inversion polymorphism in populations of *Abies* and *Tsuga*. *Molecular Biology and Evolution*, 17(9), 1302-1312. <https://doi.org/10.1093/oxfordjournals.molbev.a026414>
- Turudić, A., Liber, Z., Grdiša, M., Jakše, J., Varga, F., & Šatović, Z. (2021). Towards the Well-Tempered Chloroplast DNA Sequences. *Plants (Basel, Switzerland)*, 10(7), 1360. <https://doi.org/10.3390/plants10071360>
- Turudić, A., Liber, Z., Grdiša, M., Jakše, J., Varga, F., & Šatović, Z. (2022). Chloroplast Genome Annotation Tools: Prolegomena to the Identification of Inverted Repeats. *International Journal of Molecular Sciences*, 23(18), Article 18. <https://doi.org/10.3390/ijms231810804>
- Vadim V. G., Hirsch-Ernst, K. I., Stefan-Wöfl., Hellwig F. H. (2003). Analysis of the *Amborella trichopoda* Chloroplast Genome Sequence Suggests That *Amborella* Is Not a Basal Angiosperm, *Molecular Biology and Evolution*, Volume 20, Issue 9, Pages 1499–1505, <https://doi.org/10.1093/molbev/msg159>
- Wang, W., & Lanfear, R. (2019). Long-Reads Reveal That the Chloroplast Genome Exists in Two Distinct Versions in Most Plants. *Genome Biology and Evolution*, 11(12), 3372-3381. <https://doi.org/10.1093/gbe/evz256>
- Wick, R. R., Schultz, M. B., Zobel, J., & Holt, K. E. (2015). Bandage: Interactive visualization of de novo genome assemblies. *Bioinformatics (Oxford, England)*, 31(20), 3350-3352. <https://doi.org/10.1093/bioinformatics/btv383>

- Wicke, S., Schneeweiss, G. M., dePamphilis, C. W., Müller, K. F., & Quandt, D. (2011). The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Molecular Biology*, 76(3-5), 273-297. <https://doi.org/10.1007/s11103-011-9762-4>
- Wu, C.-S., & Chaw, S.-M. (2014). Highly rearranged and size-variable chloroplast genomes in conifers II clade (cupressophytes): Evolution towards shorter intergenic spacers. *Plant Biotechnology Journal*, 12(3), 344-353. <https://doi.org/10.1111/pbi.12141>
- Xia, X. (2013). DAMBE5: A Comprehensive Software Package for Data Analysis in Molecular Biology and Evolution. *Molecular Biology and Evolution*, 30(7), 1720-1728. <https://doi.org/10.1093/molbev/mst064>
- Xu, S., Wang, J., Guo, Z., He, Z., & Shi, S. (2020). Genomic Convergence in the Adaptation to Extreme Environments. *Plant communications*, 1(6), 100117. <https://doi.org/10.1016/j.xplc.2020.100117>
- Yahia, E. M., & Sáenz, C. (2011). Cactus pear (*Opuntia* species). En E. M. Yahia (Ed.), *Postharvest Biology and Technology of Tropical and Subtropical Fruits* (pp. 290-331e). Woodhead Publishing. <https://doi.org/10.1533/9780857092762.290>
- Yang, JB., Tang, M., Zhi-Rong Z., De-Zhu L. (2013) Complete chloroplast genome of the genus *Cymbidium*: lights into the species identification, phylogenetic implications and population genetic analyses. *BMC Evol Biol* 13, 84. <https://doi.org/10.1186/1471-2148-13-84>
- Yang, Z. (1994). Estimating the pattern of nucleotide substitution. *Journal of Molecular Evolution*, 39(1), 105-111. <https://doi.org/10.1007/BF00178256>
- Yanqui, W. T., Ramos, J. B. C., Garcia, A. O., Nano, C. O., Moat, J., & Whaley, O. (2024). Distribución geográfica y estado de conservación del cactus endémico *AustrocylindrOpuntia pachypus* y nuevos registros para el sur del Perú. *Revista Peruana de Biología*, 31(1), Article 1. <https://doi.org/10.15381/rpb.v31i1.25847>
- Zhou, J., Zhang, S., Wang, J., Shen, H., Ai, B., Gao, W., Zhang, C., Fei, Q., Yuan, D., Wu, Z., Tembrock, L. R., Li, S., Gu, C., & Liao, X. (2021). Chloroplast genomes in *Populus* (Salicaceae): Comparisons from an intensively sampled genus reveal dynamic patterns of evolution. *Scientific Reports*, 11(1), 9471. <https://doi.org/10.1038/s41598-021-88160-4>



INSTITUTO NACIONAL DE
INNOVACIÓN AGRARIA

**ESTACION EXPERIMENTAL AGRARIA
SANTA ANA-JUNIN**
CAR. SAÑOS GRANDE / HUALAHYOY KM. 8 FND. SANTA ANA
EL TAMBO - HUANCAYO - JUNIN
Teléfono : 064 -246206
E-Mail : santaana@inia.gob.pe santaana_admin@inia.gob.pe
Web : www.inia.gob.pe

R.U.C. 20568503253

BOLETA DE VENTA

ELECTRÓNICA

Nº : B006-2364

Cliente : **RUA RUA CLEIDE**

D.N.I. : 28251029

Dirección : AYACUCHO

Forma de Pago : Contado – Efectivo ventas menores

Moneda : SOLES

Fecha Emisión : 12/11/2024

| Cant. | U/M | Descripción | Precio Unitario | Importe Total |
|-------|-------|------------------------------------|-----------------|---------------|
| 1.00 | UNID. | PENCA DE TUNA BLANCA FINA S/E | 1.00 | 1.00 |
| 1.00 | UNID. | PENCA DE TUNA MORADA CORRIENTE S/E | 1.00 | 1.00 |



PAGADO

Información adicional:

| | | |
|-----------------------|------|------|
| Venta Gravada | : S/ | 1.69 |
| Total I.G.V. | : S/ | 0.31 |
| Total Precio de Venta | : S/ | 2.00 |

SON: DOS Y 00/100 SOLES

"SIRVASE ABONAR EN NUESTRAS CUENTAS"

| BANCO | MONEDA | CUENTA | CCI |
|-------|--------|--------|-----|
| | | | |

"GRACIAS POR SU PREFERENCIA"

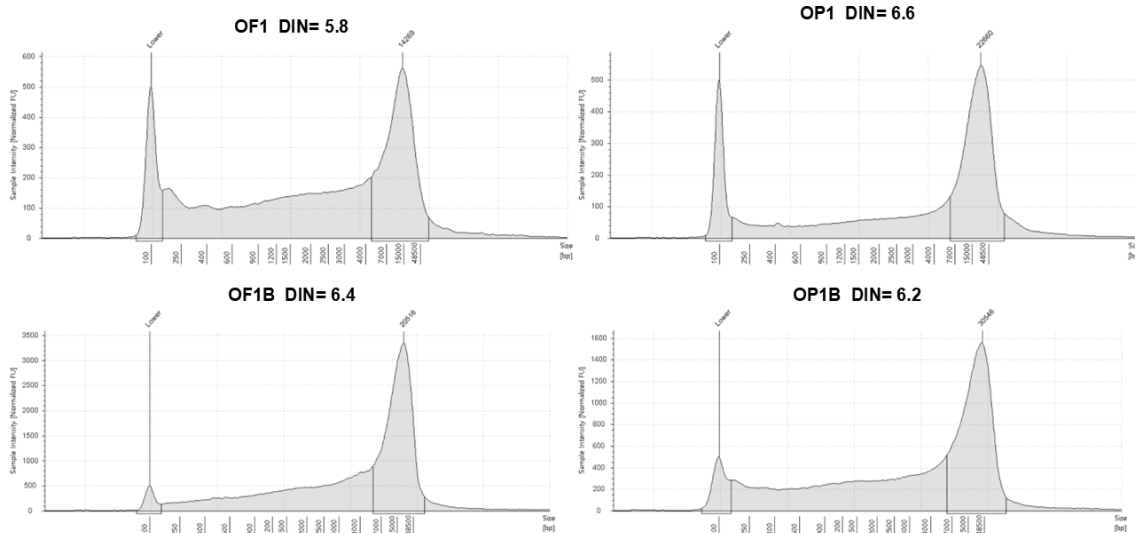


Para visualizar el presente documento ingrese a: cpe.telemovil.global o en <http://www.sunat.gob.pe> con su clave SOL.
Representación impresa de la **BOLETA DE VENTA ELECTRÓNICA**. Autorizado mediante la Resolución: N° 034-030-0000103/SUNAT
www.facturaonline.pe

Factura Online

Anexo 1. Boleto de compra de muestra de *Opuntia ficus-indica* (tuna).

Fuente propia



Anexo 2. Evaluación del DIN (Número de integridad) de ADN genómico de las muestras mediante TapeStation por el laboratorio externo.

Fuente propia

Anexo 3. Lista de secuencias empleadas para realizar la reconstrucción filogenética en genero *Opuntia*.

| Especies | Cobertura base promedio | Tamaño del genoma pb | GC% | GenBank |
|---|--------------------------------|-----------------------------|------------|----------------|
| <i>Opuntia arechavaletae</i> Speg. | 551.6 | 149,275 | 36.6 | OQ613377 |
| <i>Opuntia auberi</i> Pfeiff. | 664.7 | 138,379 | 35.9 | OQ613378 |
| <i>Opuntia aureispina</i> (S.Brack & K.D.Heil) Pinkava & B.D. Parfitt | 595.2 | 124,016 | 36.1 | OQ613379 |
| <i>Opuntia austrina</i> Small | 198.9 | 123,091 | 36.1 | OQ613380 |
| <i>Opuntia basilaris</i> Engelm. & Bigelow | 538.8 | 123,500 | 36.1 | OQ613381 |
| <i>Opuntia caracassana</i> Salm-Dyck | 660.0 | 138,141 | 35.9 | OQ613382 |
| <i>Opuntia chisosensis</i> (Anthony) D.J.Ferguson | 584.5 | 123,987 | 36.1 | OQ613383 |
| <i>Opuntia chlorotica</i> Engelm. & Bigelow | 367.0 | 123,748 | 36.1 | OQ613384 |
| <i>Opuntia colubrina</i> A.Cast. | 319.8 | 147,853 | 36.7 | OQ613385 |
| <i>Opuntia engelmannii</i> var. cuija | 687.1 | 123,275 | 36.1 | OQ613386 |
| <i>Opuntia dejecta</i> Salm-Dyck | 166.2 | 137,452 | 35.9 | OQ613387 |
| <i>Opuntia drummondii</i> Salm-Dyck | 672.5 | 123,611 | 36.1 | OQ613388 |

| | | | | |
|---|-------|---------|------|----------|
| <i>Opuntia gaumeri</i> (Britton & Rose) R.Puente & Majure | 424.7 | 137,818 | 35.9 | OQ613389 |
| <i>Opuntia guatemalensis</i> Britton & Rose | 210.1 | 137,494 | 35.9 | OQ613390 |
| <i>Opuntia jamaicensis</i> Britton & Harris | 114.3 | 137,596 | 35.9 | OQ613391 |
| <i>Opuntia macbridei</i> Britton & Rose | 284.1 | 148,031 | 36.6 | OQ613392 |
| <i>Opuntia macrocentra</i> Engelm. | 596.5 | 123,988 | 36.1 | OQ613393 |
| <i>Opuntia macrorhiza</i> Engelm. | 235.0 | 123,881 | 36.1 | OQ613394 |
| <i>Opuntia mesacantha</i> Raf. | 679.2 | 123,784 | 36.1 | OQ613395 |
| <i>Opuntia microdasys</i> (Lehm.) Pfeiff. | 591.8 | 124,337 | 36.2 | OQ613396 |
| <i>Opuntia pachyrrhiza</i> H.M.Hern., Gómez-Hin. & Bárcenas | 196.0 | 152,690 | 36.6 | OQ613397 |
| <i>Opuntia pycnantha</i> Engelm. | 500.8 | 152,441 | 36.6 | OQ613399 |
| <i>Opuntia quimilo</i> K.Schum. | 301.6 | 150,374 | 36.6 | OQ613400 |
| <i>Opuntia quitensis</i> F.A.C.Weber | 376.8 | 148,114 | 36.6 | OQ613401 |
| <i>Opuntia retrorsa</i> Speg. | 517.3 | 147,640 | 36.7 | OQ613402 |
| <i>Opuntia rufida</i> Engelm. | 577.8 | 123,145 | 36.1 | OQ613403 |
| <i>Opuntia scheeri</i> F.A.C.Weber | 210.4 | 151,001 | 36.7 | OQ613404 |
| <i>Opuntia setispina</i> Engelm. in Salm-Dyck | 133.1 | 123,480 | 36.1 | OQ613405 |
| <i>Opuntia stenopetala</i> Engelm. | 477.8 | 152,717 | 36.7 | OQ613406 |
| <i>Opuntia strigil</i> Engelm. | 600.7 | 123,994 | 36.1 | OQ613407 |
| <i>Opuntia sulphurea</i> Gillies ex Salm-Dyck | - | 147,442 | - | PV464448 |
| <i>Opuntia monacantha</i> | - | 149,076 | 36.6 | MZ579523 |
| <i>Opuntia pubescens</i> H.L.Wendl. ex Pfeiff. | 545.5 | 152,730 | 36.6 | PV849966 |
| <i>Opuntia ficus-indica</i> (L.) Mill. | 550.0 | 150,927 | 36.6 | PV849967 |
| <i>Airampoa erectoclada</i> (Backeb.) Doweld | 165.3 | 145,686 | 36.6 | OQ613412 |
| <i>Airampoa soehrensii</i> (Britton & Rose) Lodé | 404.9 | 156,992 | 36.8 | OQ613413 |
| <i>Salmonopuntia salmiana</i> (Pfeiff.) P.V.Heath | 308.1 | 147,474 | 36.7 | OQ613408 |
| <i>Salmonopuntia schickendantii</i> (F.A.C.Weber) Font & M.Köhler | 229.3 | 148,035 | 36.7 | OQ613409 |
| <i>Brasiliopuntia schulzii</i> (A.Cast. & H.V.Lelong) Backeb. | 80.1 | 162,564 | 36.7 | OQ613372 |

Fuente propia