Universidad Nacional Mayor de San Marcos
Universidad del Perú. Decana de América
Facultad de Ciencias Biológicas
Escuela Profesional de Genética y Biotecnología

Ensamblaje y análisis comparativo del genoma mitocondrial del cuy doméstico (*Cavia porcellus*) y silvestre (*Cavia tschudii*) de Perú

TESIS

Para optar el Título Profesional de Biólogo Genetista
Biotecnólogo

AUTOR

Francisco ASCUE OROSCO

ASESOR

Abelardo Lenin MATURRANO HERNÁNDEZ

Lima, Perú

2022



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

https://creativecommons.org/licenses/by-nc-sa/4.0/

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Ascue, F. (2022). *Ensamblaje y análisis comparativo del genoma mitocondrial del cuy doméstico (Cavia porcellus) y silvestre (Cavia tschudii) de Perú*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Biológicas, Escuela Profesional de Genética y Biotecnología]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	Francisco Ascue Orosco
Tipo de documento de identidad	DNI
Número de documento de identidad	705484379
URL de ORCID	https://orcid.org/0000-0002-5036-4190
Datos de asesor	
Nombres y apellidos	Abelardo Lenin Maturrano Hernádez
Tipo de documento de identidad	DNI
Número de documento de identidad	15725076
URL de ORCID	https://orcid.org/0000-0001-8819-7335
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	María Gisella Orjeda Fernández
Tipo de documento	DNI
Número de documento de identidad	10839204
Miembro del jurado 1	
Nombres y apellidos	Nadia Felicita Vera Munárriz
Tipo de documento	DNI
Número de documento de identidad	00517731
Miembro del jurado 2	
Nombres y apellidos	Miguel Ángel Neira Gonzáles
Tipo de documento	DNI
Número de documento de identidad	09640619
Datos de investigación	
Línea de investigación	A.1.1.3 Diversidad Genética
Grupo de investigación	No aplica

Agencia de financiamiento	Perú. Instituto Nacional de Innovación Agraria (INIA). Programa financiador: Programa Nacional de Innovación Agraria (PNIA). 171_PI Perú. Universidad Nacional Mayor de San Marcos. Vicerrectorado de Investigación y Posgrado. Programa de Promoción de Tesisde Pregrado. E18030044-PTPGRADO.
Ubicación geográfica de lainvestigación	Lugar: Instituto Nacional de Innovación Agraria. País: Perú Departamento: Lima Provincia: Lima Distrito: La Molina Calle: Av. La Molina 1981 Coordenadas geográficas: 12°4′36″ S, 76°56′43″ W (-12.07664, -76.94252).
Año o rango de años en que se realizó la investigación	Abril 2019 – Marzo 2020
URL de disciplinas OCDE	Genética, Herencia https://purl.org/pe-repo/ocde/ford#1.06.07 Bioinformática https://purl.org/pe-repo/ocde/ford#1.02.03

Universidad Nacional Mayor de San Marcos
(Universidad del Perú, Decana de América)

FACULTAD DE CIENCIAS BIOLÓGICAS

ACTA DE SESIÓN PARA OPTAR AL TÍTULO PROFESIONAL DE BIÓLOGO GENETISTA BIOTECNÓLOGO
(MODALIDAD: SUSTENTACIÓN VIRTUAL DE TESIS)

Siendo las 19:05 horas del 07 de febrero de 2022, en el Salón de Grados Virtual de la Facultad de Ciencias Biológicas cuya dirección electrónica fue https://zoom.us/j/93597467762?pwd=NnRPVFBMMDBFcCtOVVBPUksrTmNsdz09, y en presencia del Jurado formado por los profesores que suscriben, se inició la sesión para optar al **Título Profesional de Biólogo Genetista Biotecnólogo** de **FRANCISCO ASCUE OROSCO.**

Luego de dar lectura y conformidad al expediente **N° UNMSM-20210010687**, el titulando expuso su Tesis: **"ENSAMBLAJE Y ANÁLISIS COMPARATIVO DEL GENOMA MITOCONDRIAL DEL CUY DOMÉSTICO (*Cavia porcellus*) Y SILVESTRE (*Cavia tschudii*) DE PERÚ"**, y el Jurado efectuó las preguntas del caso calificando la exposición con la nota 18, calificativo: Aprobado con mención honrosa**.**

Finalmente, el expediente será enviado a la Escuela Profesional de Genética y Biotecnología y al Consejo de Facultad para que se apruebe otorgar el **Título Profesional de Biólogo Genetista Biotecnólogo** a **FRANCISCO ASCUE OROSCO** y se eleve lo actuado al Rectorado para conferir el respectivo grado, conforme a ley.

Siendo las 21:35 horas se levantó la sesión.

Ciudad Universitaria, 07 de febrero de 2022.

UNMSM Firmado digitalmente por ORJEDA FERNANDEZ Maria Gisella FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 08.02.2022 07:35:29 -05:00

Dra. GISELLA ORJEDA FERNANDEZ
(PRESIDENTA)

UNMSM Firmado digitalmente por MATURRANO HERNANDEZ Abelardo Lenin FAU 20148092282 soft
Motivo: Soy el autor del documento
Fecha: 07.02.2022 22:24:33 -05:00

Dr. LENIN MATURRANO HERNANDEZ
(ASESOR)

Mg. NADIA VERA MUNÁRRIZ
(MIEMBRO)

Blgo. MIGUEL NEIRA GONZALES
(MIEMBRO)

“A todas las personas que se han convertido en mis amigos y me ayudaron en la realización de esta tesis”

AGRADECIMIENTOS

El desarrollo de este trabajo no hubiera sido posible sin la contribución de muchas personas, y quisiera retribuirles de esta manera el tiempo que dedicaron en la realización de mi tesis. Debo aclarar que seguramente mis palabras no reflejaran completamente mi agradecimiento a cada uno de ellos, ya que nos unen varias anécdotas y tratare de ser breve.

Mi tesis no hubiera sido posible sin la labor de Claudia Yalta y el Ing. Eudosio Veli, además de sus sabios consejos y predisposición, su amistad hizo más llevadera mi estadía en El Laboratorio de Biotecnología Animal – INIA, tanto para mí como para las tesistas que estuvieron conmigo durante ese periodo. Agradezco también a la Dra. Dina Gutiérrez, por ser una persona muy amable y a la vez profesional con las personas que trabajamos bajo su dirección.

Quiero agradecer a Juan Daniel Montenegro por su mentoria y sincera disposición para enseñarme. Por su transparencia y trato horizontal con todos los que estuvimos compartiendo experiencias, me deja una sensación de que es un ejemplo a seguir en muchos aspectos, gracias por compartir tu tiempo y te deseo muchos éxitos.

A mi asesor de tesis Dr. Lenin Maturrano, por acompañarme en este proceso, aunque no fue directamente, le agradezco mucho ya que siempre tuvo la disposición en ayudarme, en la documentación y postulaciones de nuevos proyectos.

También quiero agradecer a las personas que pertenecen a la Dirección de Recursos Genéticos y Biotecnología, que de alguna u otra manera compartieron conmigo y permitieron un ambiente de trabajo agradable.

Por último, quisiera dar el reconocimiento a los fondos de financiación del Programa Nacional de Innovación Agraria (PNIA) y el Instituto Nacional de Innovación Agraria (INIA) que permitieron este trabajo, sin el cual muchos jóvenes no hubieran tenido esta experiencia, a mi consideración espero sinceramente que muchos de estos proyectos continúen y existan fondos estatales para impulsar la investigación, y a la par exista la continuidad necesaria en los centros de investigación.

GLOSARIO

DLP : D-Loop + Promotores

ESS : *Effective population size*

MCMC : *Markov Chain Monte Carlo*

NGS : Next Generation Sequencing

PCR : Polymerase chain reaction

PE : Pair End - Libería de secuenciación illumina

SE : Single End - Liberia de secuenciación illumina

Q30 : Phred Quality 30

STR : Short Tandem Repeat

SNP : Single-nucleotide Polymorphism

OLC : *Overlap Layout Consensus*

WGS : Whole Genome Sequencing

INDICE GENERAL

INDICE DE TABLAS

INDICE DE IMÁGENES

RESUMEN

ABSTRACT **Pag.**

1. INTRODUCCIÓN 1

2. MARCO TEÓRICO 2

2.1 *Cavia porcellus*.....2

2.2 Investigación en la genética y mejoramiento de cuy.....3

2.2.1 Diversidad y estructura poblacional del cuy 5

2.2.2 *Cavia porcellus* como modelo biológico 6

2.3 Mitocondria 7

2.3.1 Genoma mitocondrial en mamíferos 9

2.3.2 Genoma mitocondrial en *Cavia porcellus* 10

2.4 Secuenciación de Nueva Generación (NGS) 11

2.4.1 Secuenciación illumina 13

2.4.2 Ensamblaje de genomas 14

2.4.2.1 Gráficos o grafos de De Brujin 15

2.4.2.2 Ensamblaje de genomas a partir de datos illumina 17

2.4.2.3 Control de calidad 18

2.4.2.4 Recorte y filtrado de secuencias 19

2.4.2.5 Análisis de conteo de *K-mers* 19

2.4.2.6 Ensamblado de secuencias 20

2.4.2.7 Validación de ensamblaje 21

2.4.3 Ensamblaje de genoma mitocondrial 23

2.4.4 Análisis filogenómico de mitogenomas 24

3. HIPOTESIS Y OBJETIVOS 25

3.1 Hipótesis 25

3.2 Objetivos 25

3.2.1 Objetivo General 25

3.2.2 Objetivos Específicos 25

4. MATERIALES Y METODOS 26

4.1 Materiales 26

4.2.1 Material Biológico 26

4.2.2 Material de Laboratorio 26

4.2 Métodos 26

4.2.1 Extracción de ADN genómico total 28

4.2.2 Evaluación de la calidad de ADN 28

4.2.3 Secuenciación de genoma completo (WGS) 29

4.2.4 Procesamiento de información 29

4.2.5 Análisis de calidad de secuenciación 30

4.2.5.1 Análisis de calidad 30

4.2.5.2 Análisis de conteo de *k-mers* 30

4.2.6 Filtrado de secuencias por tamaño y baja calidad 31

4.2.7 Alineamiento de secuencias al genoma de referencia Cavpor v3.0 31

4.2.8 Filtrado de secuencias mitocondriales 31

4.2.9 Ensamblaje de genoma mitocondrial 32

4.2.10 Validación de ensamblaje de genomas 32

4.2.11 Ensamblaje de secuencias mitocondriales de la base de datos SRA-NCBI 33

4.2.12 Anotación del genoma mitocondrial 33

4.2.13 Análisis de distancias genéticas 34

4.2.14 Análisis filogenómicos, demográficos y haplotipos 34

5. RESULTADOS 36

5.1 Extracción y evaluación de calidad de ADN genómico total 36

5.2 Secuenciación de genoma completo 37

5.3 Análisis de calidad de secuenciación 37

5.4 Filtrado de secuencias de genoma completo 40

5.5 Alineamiento al genoma de referencia 41

5.6 Filtrado de secuencias mitocondriales 42

5.7 Ensamblaje de genoma mitocondrial 42

5.7.1 Ensamblaje del genoma mitocondrial con SPAdes 45

5.7.2 Ensamblaje del genoma mitocondrial con NOVOplasty 47

5.8 Validación de ensamblaje de genomas 48

5.9 Ensamblaje de secuencias mitocondriales de la base de datos SRA - NCBI 51

5.10 Anotación de genoma mitocondrial 53

5.11 Distancias genéticas de mitogenomas 59

5.12 Filogenómica y análisis demográficos - haplotipicos. 60

6. DISCUSIÓN 63

6.1 Secuenciación y ensamblaje de genoma mitocondrial 64

6.2 Anotación de genomas y distancias genéticas entre *C. porcellus y tshcudii* 69

6.3 Análisis filogenéticos, demográficos y haplotipicos 70

7. CONCLUSIONES 74

8. RECOMENDACIONES 75

9. REFERENCIAS BIBLIOGRAFICAS 76

10. ANEXOS 99

INDICE TABLAS

Tabla 1. Excepciones del código genético estándar en mitocondrias **.....8**

Tabla 2. Cuadro resumen de extracción y calidad de muestras de ADN**.....36**

Tabla 3. Cuadro resumen de secuenciación de genoma completo **.....37**

Tabla 4. Cuadro resumen de lecturas del genoma mitocondrial recuperadas **.....43**

Tabla 5. Cuadro resumen del ensamblaje de genoma mitocondrial con SPAdes**.....46**

Tabla 6. Cuadro resumen del ensamblaje de genoma mitocondrial con NOVOplasty**.47**

Tabla 7. Cuadro resumen de calidad de ensamblaje de genoma mitocondrial con el programa SPAdes. **.....49**

Tabla 8. Cuadro resumen de calidad de ensamblaje de genoma mitocondrial con el programa NOVOplasty. **.....50**

Tabla 9. Cuadro resumen de calidad de secuenciación de genoma completo (SRX4501347), descargado de la base de datos SRA. **.....52**

Tabla 10. Cuadro resumen de ensamblaje y validación de ensamblaje del genoma mitocondrial SRX4501347**.....53**

INDICE IMÁGENES

Figura 1. Distribución geográfica del genero *Cavia* **.....3**

Figura 2. Organización y estructura del genoma mitocondrial en mamíferos **.....9**

Figura 3. Costo del secuenciación de genomas desde el 2001 a 2019**.....11**

Figura 4. Desarrollo cronológico de tecnologías de secuenciación**.....12**

Figura 5. Metodología de secuenciación por síntesis (illumina) **.....13**

Figura 6. Conformación de K-meros y Grafos de *De Brujin* **.....15**

Figura 7. Diferencia de complejidad de análisis entre grafos De Brujin **.....16**

Figura 8. Flujo de trabajo para el ensamblaje de genomas **.....18**

Figura 9. Formato FastQ **.....18**

Figura 10. Análisis de la distribución del conteo de *K-mers***.....20**

Figura 11. Número de publicaciones de genoma mitocondrial.**.....24**

Figura 12. Flujo de trabajo para en ensamblaje y análisis del genoma mitocondrial **...27**

Figura 13. Gel de corrida electroforética de muestras secuenciadas**.....36**

Figura 14. Calidad por cada base de las lecturas *forward* y *reverse*.**.....38**

Figura 15. Calidad de secuencias y distribución de tamaño **.....38**

Figura 16. Contenido de bases por secuencias y distribución de contenido de GC **....39**

Figura 17. Grafica de conteo de *K-mers***.....40**

Figura 18. Grafica de filtrado de secuencias de baja calidad del genoma completo **...41**

Figura 19. Grafica de porcentaje de alineamiento al genoma de referencia Cavpor v3.0

.....42

Figura 20. Grafica de cobertura de genomas mitocondriales de *Cavia porcellus* de la Estación Experimental Andahuaylas – Chumbibamba. **.....44**

Figura 21. Grafica de cobertura de genomas mitocondriales de *Cavia tschudii* (cuyes silvestres)**.....44**

Figura 22. Grafica de cobertura de genomas mitocondriales de *Cavia porcellus* de la Estación Experimental - La Molina. **.....45**

Figura 23. Grafica de BLAST de las secuencias ensambladas con el programa SPAdes al genoma de referencia de *Cavia porcellus* (NC_000884.1)**.....46**

Figura 24. Grafica de BLAST de las secuencias ensambladas con el programa NOVOplasty al genoma de referencia de *Cavia porcellus* (NC_000884.1)**.....47**

Figura 25. Análisis de calidad de secuenciación y filtrado de secuencias de baja calidad de las lecturas de secuenciamiento de la base de datos SRA (SRX4501347). **.....51**

Figura 26. Grafica de BLAST del ensamblaje de SRX4501347 con el programa SPAdes al genoma de referencia de *Cavia porcellus* (NC_000884.1). **.....52**

Figura 27. Anotación de genoma mitocondrial de códigos: SA42911 y SA42912**.....54**

Figura 28. Anotación de genoma mitocondrial de códigos: SA42913 y SA42914**.....55**

Figura 29. Anotación de genoma mitocondrial de códigos: SA42976 y SA42977**.....56**

Figura 30. Anotación de genoma mitocondrial de códigos: SA42978 y SA42979**.....57**

Figura 31. Anotación de genoma mitocondrial de códigos: SA42980 y SA42981**.....58**

Figura 32. Anotación de genoma mitocondrial de la base de datos SRA **.....59**

Figura 33. Comparación de genomas ensamblados con el genoma de referencia y *heatmap* de distancias genéticas **.....60**

Figura 34. Reconstrucción filogenética a partir de mitogenomas ensamblados y disponibles en la base de datos NCBI. **.....61**

Figura 35. Análisis de coalescencia por bayesian skyline a partir de mitogenomas de *C. porcellus y C. tschudii* **.....61**

Figura 36. Median-joining network de haplotipos de *C. porcellus y C. tschudii* a partir de mitogenomas. **.....62**

RESUMEN

Cavia porcellus es un roedor doméstico nativo de los Andes que es parte integral de la cultura y la alimentación tradicional andinas. Pese a su importancia y múltiples usos, los recursos genéticos disponibles para su estudio y mejoramiento son escasos y están sesgados al excluir individuos silvestres y nativos. En este estudio se ha identificado y secuenciado 10 individuos entre mejorados, "criollos" y silvestres del género *Cavia* provenientes de Perú. En total se ha reconstruido el genoma mitocondrial de los 10 individuos secuenciados y se incluyó 1 genoma de la base de datos SRA-NCBI, evaluándose sus relaciones filogenéticas. El individuo de *C. tschudii* más distanciado genéticamente (i.e. Pachachaca, Andahuaylas) se comporta como un grupo basal mientras que otros C. *tschudii* (i.e. Pantanos de Villa, Lima) son indiferenciables de *C. porcellus.* Este trabajo muestra que en *C. porcellus* ha experimentado probablemente una introgresión de *C. tschudii* o viceversa, que pudo generar poblaciones heterogéneas dentro de *C. tschudii*.

Palabras claves*:* *Cavia porcellus*, *Cavia tschudii*, NGS, genoma mitocondrial

ABSTRACT

Cavia porcellus is a domestic rodent native to the Andes that is an integral part of the traditional Andean culture and diet. Despite its importance and multiple uses, the genetic resources available for its study and improvement are scarce and are biased by excluding wild and native individuals. In this study, 10 individuals of the genus *Cavia* from Peru have been identified and sequenced between improved, creole and wild. In total, the mitochondrial genome of these 10 genomes and 01 genome from the SRA-NCBI database were included and their phylogenetic relationships evaluated. The most genetically distanced individual of *C. tschudii* (i.e. Pachachaca, Andahuaylas) performs as a basal group while others (i.e. Pantanos de Villa, Lima) are indistinguishable from *C. porcellus*. This work shows *C. porcellus* could have an introgression from *C. tschudii* or vice versa, that could produce heterogeneous populations within C. tschudii.

keywords: *Cavia porcellus*, *Cavia tschudii*, NGS, mitocondrial genome

1. INTRODUCCIÓN

La especie *Cavia porcellus* es un roedor domesticado, originario de la región andina donde es tradicionalmente consumido (Wagner *et al.* 1976; Chauca, 1997). En la actualidad es una de las fuentes cárnicas de mayor demanda en el mercado peruano, con un incremento en la exportación a otros países según la Dirección de Estudios Económicos e Información Agraria del MINAGRI. Cuenta con una población estimada de 23 millones de individuos en el territorio peruano, y los departamentos donde la crianza de cuyes está más desarrollada son: Ancash, Apurímac, Cajamarca, Cusco, Huánuco, Junín, La libertad, Lima y Puno.

Los cuyes han sido catalogados como una fuente cárnica promisoria (Sánchez-Macías *et al.* 2018), debido a su calidad nutricional. Tiene bajo contenido de grasas y colesterol y alto contenido proteico. A ello se le suma que el ciclo de vida es corto, lo que acelera su producción. Es por ello que, en los últimos 50 años, han sido sometidos a un proceso de mejoramiento, generando líneas o razas en el Instituto Nacional de Innovación Agraria (INIA) como: raza Perú, Andina e Inti. Actualmente, el INIA está incorporando información genética para realizar nuevos estudios de mejoramiento, haciendo uso de marcadores nucleares y mitocondriales. Es en ese sentido, que la presente tesis tiene como finalidad reconstruir la secuencia del genoma mitocondrial de cuyes mejorados, "criollos" y cuyes silvestres (*Cavia tschudii)*, que se conoce como una de las fuentes de domesticación inicial (Spotorno *et al.* 2004; Dunnum *et al.* 2009). La reconstrucción de la secuencia mitocondrial permitirá identificar variantes entre genomas y estudiar la evolución y origen de los cuyes domesticados. Este es el primer reporte de genomas mitocondriales de cuyes de razas mejoradas, "criollos" y nativos de Perú que serán de utilidad para realizar estudios filogenómicos.

2. MARCO TEÓRICO

2.1 *Cavia porcellus*

Cavia porcellus es una especie de roedor de la familia Caviidae (4.1 +/- 1.4 Ma; Opazo, 2005), son criados como mascotas, animales de laboratorio y para consumo humano; siendo una fuente de alimentos en diferentes países (Wagner *et al*., 1976; Chauca, 1997; Kelt *et al.,* 2007; Fitzpatrick *et al.,* 2007; Giovas *et al.,* 2012; LeFebvre y deFrance 2014), *Cavia porcellus* (cuy) está presente en toda Sudamérica con excepción de la Amazonia y el sur de Argentina y Chile, mientras que el grupo polifilético de *Cavia tschudii* (cuy silvestre) está presente en el norte de Argentina, la costa de Chile y los andes de Perú y Bolivia (Dunnum & Salazar-Bravo, 2010, Figura 1.).

Los cuyes formaron parte de la economía familiar de poblaciones precolombinas, acumulando variantes por selección artificial. Existen restos arqueológicos en Telarmachay (Perú) que datan su presencia por lo menos hace 9000 años, y su consumo y usos espirituales entre 7000 y 4500 años (Wing, 1986). En el estudio de Dunnum y Salazar-Bravo del 2010, se sugiere que *Cavia porcellus* es producto de un único evento de domesticación inicial, criados actualmente como variedad de “criollos” en varios países andinos (Spotorno *et al.* 2006). Se estima que el tiempo de divergencia de las subespecies de *C. tschudii* se encuentra entre el plioceno tardío e inicio del pleistoceno (1.5 - 3.5 Ma; Dunnum & Salazar-Bravo, 2010; Upham & Patterson, 2012). Estudios previos señalan además que una subespecie de *C. tschudii* sea probablemente el antecesor de *C. porcellus* (A. E. Spotorno *et al*., 2006). A partir de estudios con citocromo b en especies del género *Cavia* se reportó que los especímenes de *C. tschudii tschudii* de la localidad de Ica-Perú son posiblemente los parientes silvestres más cercanos de *C. porcellus* debido a su menor distancia genética con los *C. porcellus* criados en la actualidad (Dunnum & Salazar-Bravo, 2010).

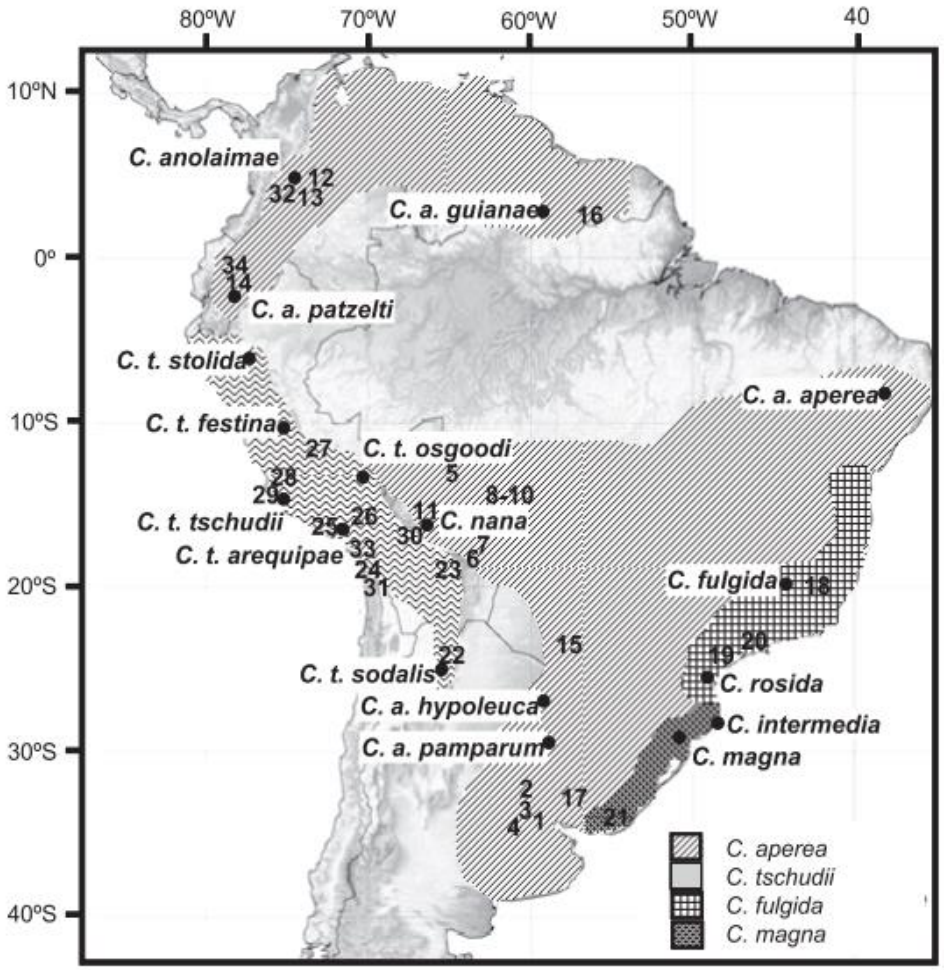


Figura 1. Distribución Geográfica en Sudamérica del género *Cavia*

Fuente: Dunnum y Salazar-Bravo 2010.

2.2 Investigación en la genética y mejoramiento de cuy.

Los trabajos de investigación de la genética en cuyes se iniciaron en Perú en la década del 60, seguidos de trabajos en Colombia y Ecuador en el 70, Bolivia en el 80 y Venezuela en el 90 (Chauca F., 1997). Las mejoras genéticas para esta especie en el Perú se llevaron a cabo por selección artificial, diferenciando dos genotipos de cuyes: "criollo" o nativo y mejorado, conocido también como "peruano" en otros países andinos (Chauca, 1995). Los cuyes denominados "criollos" se encuentran predominantemente en áreas rurales, se desarrollan con poca exigencia en la alimentación y con bajos rendimientos productivos, tiene un bajo desarrollo muscular y se clasifican por el color

de pelaje, siendo predominantes los pelajes claros como blanco, bayo o alazán. Por su tipo de crianza presentan alta consanguinidad, tienen una alimentación herbívora y son seleccionados negativamente por la separación de animales de mayor tamaño para consumo o para venta. Los cuyes mejorados conocidos como "peruanos" fueron seleccionados en la Estación Central La Molina del INIA (Instituto Nacional de Innovación Agraria) a partir de cuyes "criollos" provenientes de una colecta a nivel nacional por el INIA; creándose 3 líneas: Perú, Inti y Andina con características distintivas tanto fenotípicas como productivas (Chauca, 2005).

En la actualidad los estudios sobre el mejoramiento de cuyes siguen en curso a nivel nacional, entre los centros que lo realizan se encuentran las Estaciones Experimentales del INIA, donde encontramos trabajos como: línea Inka (Estación Experimental Baños del Inca - INIA), Mantaro (Estación Experimental Santa Ana - INIA), Saños (Estación Experimental Santa Ana - INIA), Cori (Estación Experimental Santa Ana - INIA), etc. Otros trabajos de mejoramiento que podemos citar son realizados por Instituto Veterinario de Investigaciones Tropicales y de Altura (IVITA) de la Facultad de Medicina Veterinaria de la Universidad Nacional Mayor de San Marcos, en la Estación Experimental IVITA El Mantaro, ubicada en la provincia de Jauja, Junín. Donde se desarrollan líneas paternas ("precoz", "cárnica") y maternas ("prolífica" y "lechera") para generar cuyes de alta productividad denominados cuyes "G" y "RG" (Yamada *et al*., 2018).

Actualmente los estudios de mejoramiento en estos animales se han referido a características de peso y número de la camada (Dillard *et al*., 1972; Chávez, 1979. INIA; Quijandría *et al*., 1983 INIA; Rodríguez L *et al.,* 2013; Camino *et al.,* 2014) estableciendo correlaciones genéticas y fenotípicas, el grado de heredabilidad (Dillard *et al.,* 1972; Quijandría *et al.,* 1983) y repetitividad (Balbin N.R. 1990).

2.2.1 Diversidad y estructura poblacional de cuy.

Dentro de los estudios de diversidad y estructura poblacional se tienen los siguientes reportes: el estudio de Spotorno *et al.* en el 2006, donde se realizaron análisis morfométricos y moleculares usando como marcador el gen mitocondrial de citocromo b, en este estudio se puso a prueba la hipótesis de 3 etapas de domesticación de *Cavia porcellus,* que diferencia los cuyes de domesticación antigua, mascotas/animales de laboratorio y cuyes mejorados para producción de carne.

La caracterización molecular por AFLP realizado en Colombia por Solarte *et al*. en el 2007, encontró una disminución de la diversidad genética dentro de líneas "criollas" y mejoradas.

Se estudió también el posible origen continental de restos de *cavia porcellus* en las islas caribeñas evaluado por Kimura *et al.* en el 2016 con marcadores mitocondriales de citocromo b, atribuyendo a Colombia como la región de origen, relacionando además estos hallazgos con rutas migratorias humanas.

Fueron realizadas tambien comparaciones citogenéticas entre *Cavia porcellus* y *Cavia tschudii*, encontrándose el mismo número de cromosomas, pero con diferencias en cuatro cromosomas a través de bandeo G y perdida de heterocromatina de *Cavia porcellus* por bandeo C (Walker *et al.,* 2014).

Con 6 marcadores microsatélites en cuyes de Colombia, se encontró reducida variabilidad alélica y ausencia de estructuración poblacional en las líneas mejoradas y "nativas" (también denominadas "criollas"), aunque en esta última como una mayor riqueza de alelos, exceptuando los cuyes criados en el departamento de Nariño - Colombia (Burgos-Paz *et al.,* 2011).

Se ha reportado también caracterización genética usando 20 marcadores dinucleótido de microsatélites entre diferentes grupos de cuyes de: Perú, Ecuador, Colombia y Bolivia, encontrando un alto polimorfismo en las poblaciones (alta variabilidad genética) pero sin estructuración poblacional, no se encontraron diferencias entre cuyes

autóctonos de Ecuador y los cuyes introducidos de Perú y otros países (Avilés *et al.,* 2015).

En un estudio de *C. porcellus* de sitios arqueológicos de Perú, Bolivia, Ecuador y Colombia se sugiere además centros alternativos de domesticación de *Cavias* en otras locaciones antes desconocidas, como el caso de muestras antiguas de Colombia, donde se encontraron animales domesticados a partir de *Cavia aperea*, pero que fueron reemplazados por linajes actuales descendientes de *C. tschudii* (E. Lord *et al.,* 2020), esto sugiere que sucesivas migraciones en el pasado permitieron la dispersión de *C. porcellus* (Kimura et al., 2016; Edana Lord *et al.,* 2018).

2.2.2 Cuy como modelo biológico.

Otra área donde *Cavia porcellus* tiene importancia es como modelo biológico para diferentes áreas de investigación, estos animales de laboratorio fueron exportados desde Sudamérica a diferentes países de Europa y Estados Unidos (Wagner JE y PJ Manning, 1976), donde son usados actualmente para estudios como: metabolismo de lípidos (Tugwood *et al.,* 1998; Fernández *et al.,* 2006; Leite *et al.,* 2010; Yang *et al.,* 2016), arteriosclerosis y factores de riesgo cardiovasculares (Volek *et al.,* 2004; Leite et al., 2009), embriología y diferenciación sexual (Wang *et al.,* 2017), desarrollo óseo fetal y postnatal (Finch *et al.,* 2010; Yin, Z. *et al.,* 2014), modelo para estudio de encefalopatía espongiforme bovina (Furuoka *et al.,* 2011), modelo celular y molecular de mecanismos inmunológicos y enfermedades infecciosas (Padilla-Carlin*, et al.* 2008), enfermedades oculares (Ostrin *et al.,* 2016; Howlett *et al.,* 2006), modelo para estudio de asma (Smith *et al.,* 2007; Turner *et al.,* 2017), cambios en redes neuronales inducidos por procesos epileptógenos (Curtis, M *et al.,* 2016) y secuenciación del genoma mitocondrial de *Cavia porcellus* para modelos de osteoartritis (Cui *et al.,* 2016).

2.3 Mitocondria.

La mitocondria es un organelo que se encuentra en la mayoría de los organismos eucariotas. Se encarga de la energética celular, produciendo ATP (adenosin trifosfato) como fuente de energía química, producto de un proceso denominado fosforilación oxidativa (Alberts *et al.,* 2002). Además de suministrar energía, la mitocondria está involucrada en otros procesos como: diferenciación y muerte celular (Birsoy *et al.,* 2015), señalización celular, envejecimiento (De Grey, 2004), β oxidación de ácidos grasos (McBride *et al.,* 2006) y síntesis de aminoácidos y esteroides (Pagliarini y Rutter, 2013). Las mitocondrias producto de su doble envoltura están divididas en 4 compartimientos: membrana externa, espacio intermembrana, membrana interna y matriz. La membrana externa es la parte más externa del organelo, y es el encargado de filtrar las proteínas y compuestos que entran al interior (Hoogenboom *et al.,* 2007). La membrana externa contiene enzimas que están involucradas en diferentes procesos como el alargamiento de ácidos grasos, oxidación de epinefrina y la degradación del triptófano (Alberts *et al.,* 2002). El espacio intermembrana es la transición entre la membrana externa e interna, la composición de moléculas pequeñas como iones y azúcares es igual al citosol, por la alta permeabilidad de la membrana externa, siendo el contenido proteico el que difiere. La membrana interna contiene proteínas de las reacciones de oxidación y reducción de la fosforilación oxidativa, están compuestas por los complejos I, II, III y IV y ATPasa, además se encuentran proteínas de fusión y fisión mitocondrial (Alberts *et al.,* 2002). La membrana interna es rica en el fosfolípido cardiolipina, que contiene 4 ácidos grasos en lugar de 2, ayudando a la impermeabilidad de la membrana interna (McMillin y Dowhan., 2002). Por último, la matriz es el compartimiento clave para el desarrollo de procesos que incluyen: la oxidación de ácidos grasos, oxidación del piruvato y el ciclo de ácido cítrico (Alberts *et al.,* 2002). La matriz contiene además varias copias de su propio material genético, que está empaquetado en nucleoides, estos pueden estar entre 1 a 10 copias (Cole, 2016), se observa además la presencia de ribosomas que traducen

proteínas codificadas por el propio material genético de la mitocondria, con una variante del código genético estándar (Tabla 1. Alberts *et al.,* 2002).

Tabla 1. Excepciones del código genético estándar en mitocondrias

Organismo	Codón	Estándar	Mitocondria
Mamíferos	AGA, AGG	Arginina	Codón de Stop
Invertebrados	AGA, AGG	Serina	Codón de Stop
Fungí	CUA	Leucina	Treonina
Todos los de arriba	AUA	Isoleucina	Metionina
	UGA	Codón de Stop	Triptófano

Fuente: Alberts *et al.,* 2002

Aunque se grafica independientemente, las mitocondrias forman una red altamente dinámica, que se regula entre sí a través de fisiones y fusiones, el conjunto de las mitocondrias presente en una célula constituye el condrioma (Logan, 2010). El tamaño del condrioma en células animales está en el orden de 1000 a 2000 mitocondrias, que varía sustancialmente en tamaño y forma según el tejido y la demanda energética de las células (Fernández-Vizarra *et al.,* 2011). Las mitocondrias forman una compleja red en 3D asociada al citoesqueleto, esta unión además determina su forma y puede verse involucrada en su función (Kuznetsov *et al.,* 2013; Bartolák-Suki *et al.,* 2017). Se puede encontrar a las mitocondrias ancladas a las miofibrillas en los músculos o envueltas alrededor del flagelo del espermatozoide como casos excepcionales (Alberts *et al.* 2002).

2.3.1 Genoma Mitocondrial en mamíferos.

El genoma mitocondrial en animales es relativamente pequeño (~16 Kb), circular y de herencia matrilineal (Alberts *et al.,* 2002), pero esta acepción se debe en gran parte a los trabajos de investigación centrados en animales bilaterios, que no estima las diferentes variantes estructurales del genoma mitocondrial en el reino animal (Ladoukakis y Zouros, 2017). La comparación de genomas mitocondriales entre mamíferos en cambio indica que son notablemente similares, siendo pequeños (~16 Kb) y circulares, conteniendo el mismo set de genes (Lavrov y Pett, 2016).

Las dos cadenas del ADN mitocondrial difieren en composición de bases, siendo una de ellos rico en guanina, que permite separar las cadena H (pesada) y L (ligera) con centrifugación en gradiente CsCl2 (Bignell *et al.,* 1996). Podemos encontrar orígenes de replicación (OHR y OLR) y promotores para transcripción policistrónica (Falkenberg, 2018).

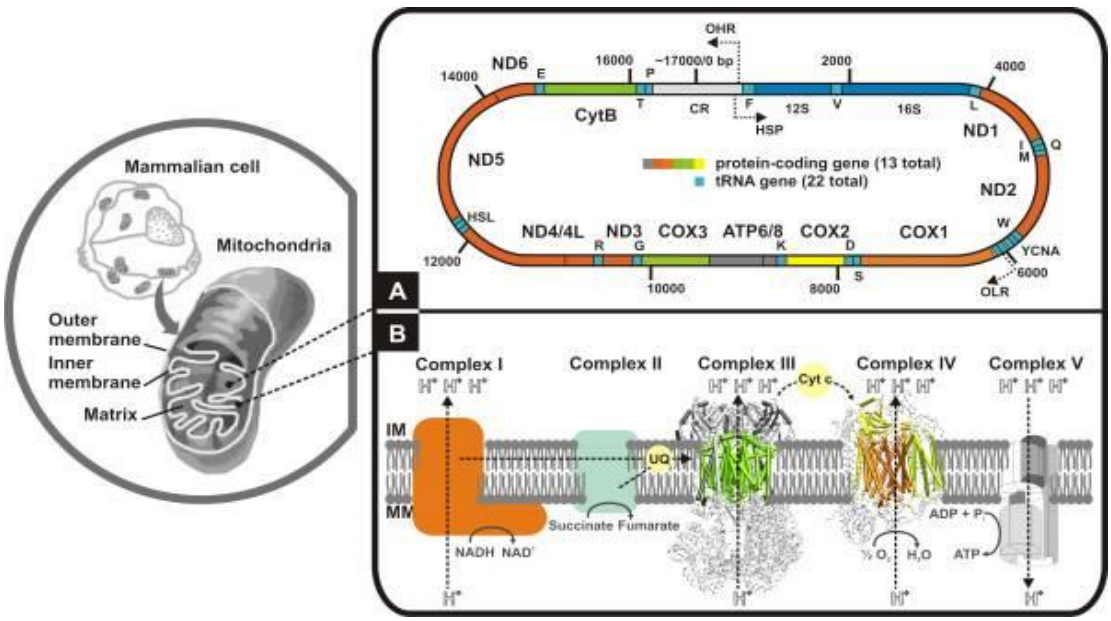


Figura 2. Organización y estructura del genoma mitocondrial de mamíferos.

Fuente: da Fonseca *et al.,* 2008.

El ADNmt codifica 13 genes para proteínas involucradas en la maquinaria de la fosforilación oxidativa: 7 subunidades corresponden a NADH deshidrogenasa (ND: ND 1, 2, 3, 4, 4L, 5 y 6), 1 subunidad del citocromo b, del complejo de citocromo bc (CytB), 3 subunidades de citocromo oxidasa C (COX), y dos subunidades de ATPasa (ATP 6 y ATP8), que son complementadas con genes del genoma nuclear (Alberts *et al.,* 2002). Para la traducción de sus propias proteínas, el ADNmt codifica también 22 ARN de transferencia y 2 ARN ribosomales (Chan, 2006).

El ADNmt contiene una región no codificante llamada región control, esta se confunde con la región D-Loop, la región control considera además de la región D-Loop a las regiones promotoras de transcripción adyacentes, por esta razón se puede llamar a la región control con el acrónimo DLP (D-Loop and associated Promoters) (Falkenberg, 2018).

2.3.2 Genoma mitocondrial de *Cavia porcellus*

El genoma mitocondrial de *Cavia porcellus* fue secuenciado inicialmente por D´erchia *et al.* en 1996, presente en la base de datos del GenBank - NCBI como genoma de referencia (NC_ 000884, actualmente cuenta con la versión NC_ 000884.1), con una extensión de 16 801 pb, codificando 37 genes: 13 genes codificantes de proteínas, 22 de ARNt y 2 genes de ARNr, y una región no codificante (DLP). Además de este genoma se ha reportado también un genoma mitocondrial para modelo de osteoartritis (Cui *et al.*, 2016) y en cuanto a los genomas de especies relacionadas, se cuenta únicamente con un genoma mitocondrial de *Cavia aperea* (Cao and Xia, 2015) depositado en la base de datos de Genbank (NC_469667.1), hasta la actualidad no se cuenta con genomas mitocondriales de *Cavia tschudii*.

2.4 Secuenciación de nueva generación (NGS)

El inicio de la era de secuenciación se inició con el método de secuenciamiento por síntesis propuesto por Sanger (Sanger *et al.,* 1977), método referencial que fue ampliamente usado durante 40 años, este método permite secuenciar fragmentos de tamaño mediano (1000 pb), útiles para identificación de SNP, STR o identificación de variantes. Este método sin embargo tiene un elevado precio y es utilizado en la secuenciación de bajo rendimiento. A pesar de estas limitaciones fue usado para el secuenciamiento de genoma Humano (Ackerman, 1999), impulsando y dando paso a las tecnologías actuales.

Las tecnologías de nueva generación desarrolladas posteriormente, son utilizadas en la secuenciación de alto rendimiento, capaces de secuenciar genomas completos en días u horas (Slatko *et al.,* 2018), con una disminución de los costos en los últimos años, incluso desviándose de la Ley de Moore (Figura 3), estas técnicas permiten el secuenciamiento en paralelo de varias moléculas de ADN ahorrando tiempo y costos.

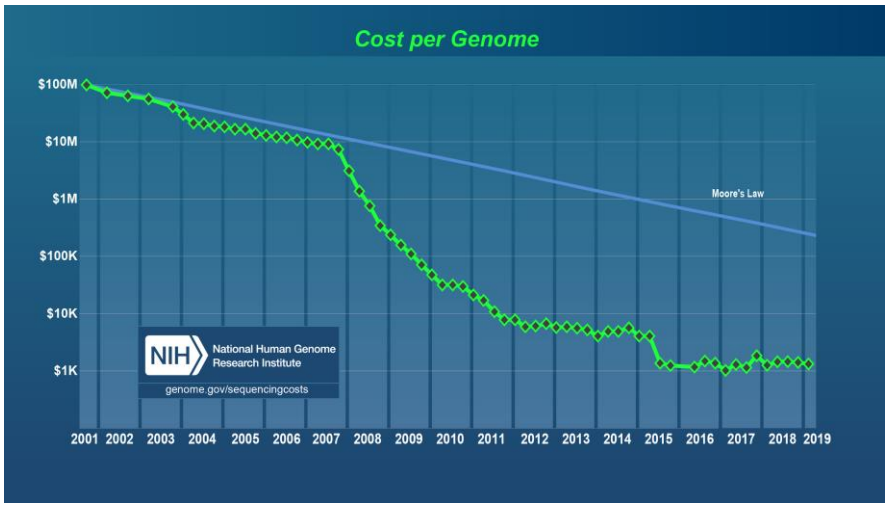


Figura 3. Costo de secuenciación de genoma desde el 2001 al 2019

Fuente: National Institutes of Health (NIH) - EEUU

Dentro de las tecnologías de nueva generación (denominadas segunda generación actualmente), se encuentran: Secuenciación de protones (Ion Torrent), pirosecuenciamiento (Roche 454 *sequencing*) y secuenciación por síntesis (Illumina) (Heather y Chain, 2016), el tiempo y el número de pares de bases que se secuencian varía entre las tecnologías (Figura 4), pero es común entre ellas la producción de lecturas de pequeño tamaño denominadas *Short reads* (150 a 450 pb).

Las tecnologías de secuenciación que actualmente están en desarrollo son denominadas de tercera generación, dentro se encuentran: secuenciación por tecnología SMRT de la empresa Pacific Biosciences (PACBIO) y Oxford Nanopore (Slatko *et al.,* 2018), estas tecnologías producen lecturas de gran tamaño, mayores de 5000 pb denominadas como *Long reads*.

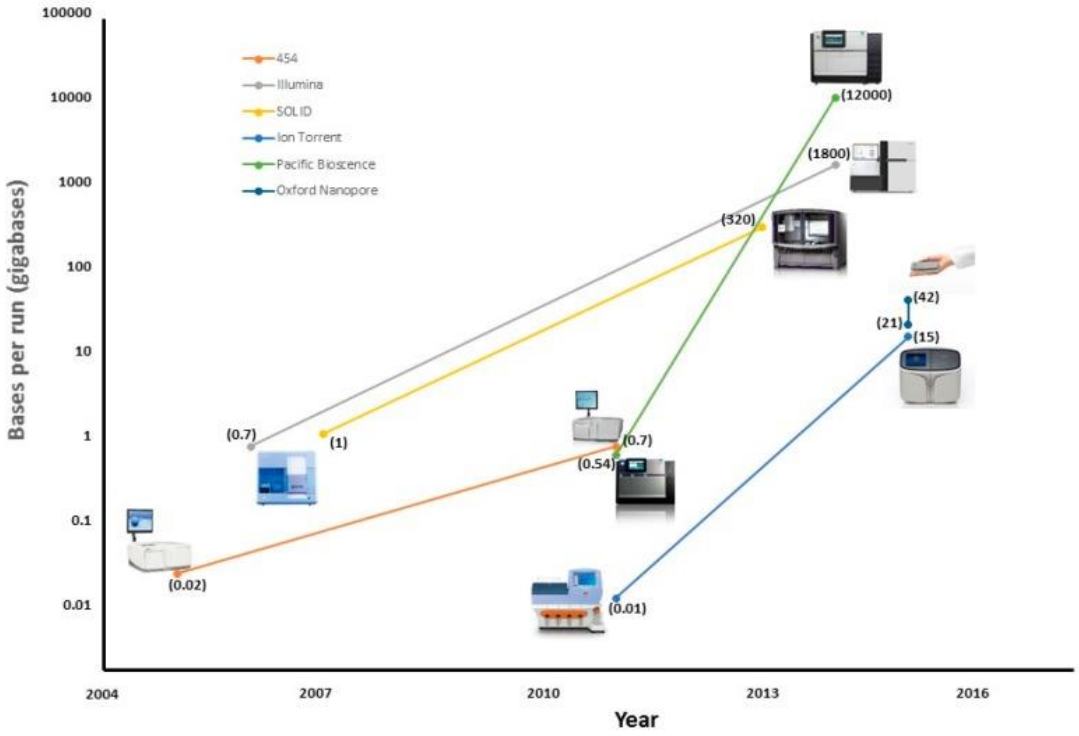


Figura 4. Desarrollo cronológico de tecnologías de secuenciación

Fuente: Garrido-Cardenas *et al.,* 2017

2.4.1 Secuenciación por Illumina

El secuenciamiento illumina, es una de las técnicas de mayor aplicación en los últimos años, inicialmente introducida al mercado con el nombre de Solexa (Bentley, 2006), marcó un antes y un después en las técnicas de secuenciación, tanto que en la actualidad por la cantidad de aplicaciones que tiene, es una de las técnicas más requeridas para la secuenciación de alto rendimiento (McCombie *et al.,* 2018). Las diferencia con las tecnologías predecesoras son: la preparación de librerías, estas son producidas por métodos físicos o químicos ligados a adaptadores y cebadores universales sin la necesidad de introducirlos en bacterias u hongos; la unión de análogos de nucleótidos marcados con fluorescencia que actúan como terminadores reversibles de amplificación, esta idea es similar a la técnica de secuenciación por Sanger solo que la terminación en la PCR es reversible y es realizada a escala microscópica; multiplicación clonal de secuencias por medio de una PCR en puente (*Bridge pcr,* Figura 5.) que permite tener varias copias de una misma molécula para amplificar la fluorescencia emitida (Bentley *et al.,* 2009).

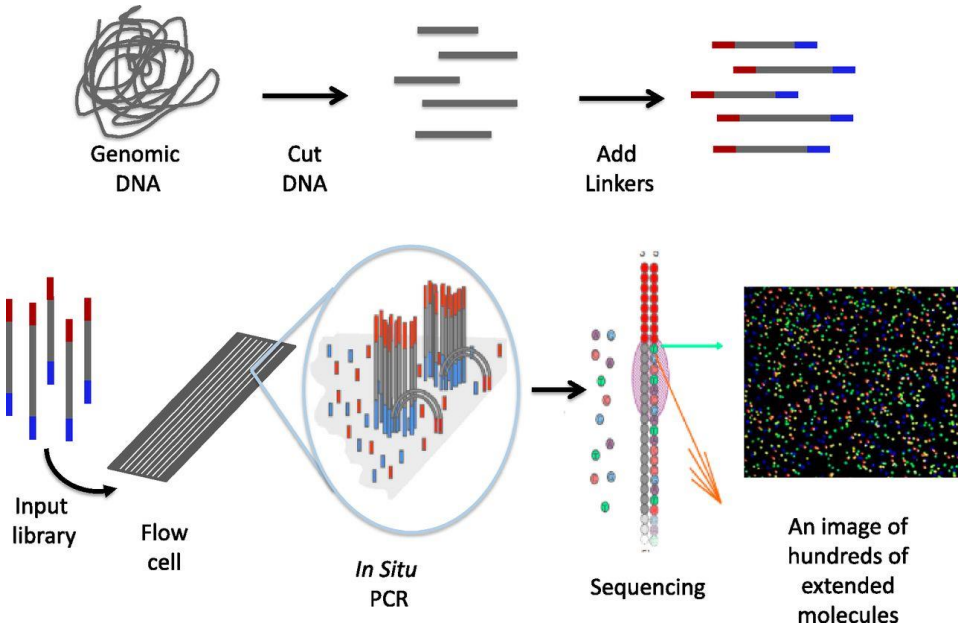


Figura 5. Metodología de secuenciación por illumina

Fuente: Johnsen *et al.,* 2013

2.4.2 Ensamblaje de genomas

El ensamblaje de genomas requiere de alta demanda computacional, dependiendo del tamaño del genoma que se requiere procesar, el número de individuos y el método usado para la secuenciación (Dominguez Del Angel *et al.,* 2018).

El ensamblaje de genomas, proviene de la idea de superposición lecturas producidas por un alineamiento, por un grafo (*eulerian / hamiltonian graphs*) o una versión mixta, estas son trasladas a algoritmos informáticos, entre ellos SCSP (*Shortest Common Superstring Problem)*, OLC (*Overlap Layout Consensus*) y grafos de *De Bruijin*. En el inicio de la era genómica, se utilizaron ensambladores como: Phrap (http://www.phrap.org/), Celera assembler (Myers *et al.,* 2000), ARACHNE (Batzoglou *et al.,* 2002), Atlas (Havlak *et al.,* 2004) que utilizaron datos de secuenciamiento Sanger. Sin embargo, actualmente tendrían problemas para ensamblar secuencias más cortas provenientes de NGS, como illumina. La secuenciación por tecnología NGS genera una gran cantidad de datos que implica algunas demandas computacionales, como alta capacidad de almacenamiento de disco (Domínguez Del Angel *et al.,* 2018), alta capacidad de memoria RAM y la necesidad de adecuados algoritmos (Sohn y Nam, 2016). Es por ello que el ensamblaje de genomas a partir de estos datos trata de minimizar costos computacionales sin comprometer el correcto ensamblaje, los ensambladores de genomas a partir de datos NGS utilizan métodos de grafos eulerianos y hamiltonianos que son un tipo especial de grafos denominados de *De Bruijn* (Sohn y Nam, 2016).

2.4.2.1 Gráficos o Grafos de *De Bruijn*

Los gráficos o grafos de *De Bruijn* son un tipo especial de gráficos dirigidos, que une nodos que representan superposiciones entre secuencias (Compeau *et al.,* 2011), en el contexto de genómica computacional estas secuencias corresponden a nucleótidos, la idea detrás del ensamblaje de genomas a través de grafos, es la capacidad de realizar una secuencia ininterrumpida de lecturas, de modo que permita hallar la secuencia de inicio a fin de un genoma (Compeau *et al.,* 2011), esta metodología es utilizada en la mayoría de los ensambladores con lecturas provenientes de NGS. Para realizar los gráficos de *De Bruijn* computacionalmente y que sea viable un algoritmo, se utiliza subsecuencias de un tamaño K que se obtiene de las lecturas, estas secuencias son obtenidas dividendo la secuencia en agrupaciones de un mismo tamaño continuo (Figura 6. A), denominados en bioinformática como *K-mers* (Sohn y Nam, 2016), por ejemplo una lectura de tamaño L es cortada en un total de (L - K + 1) *k-mers*, que luego serán utilizados para elaborar los grafos de *De Bruijn* (Figura 6. B y C).

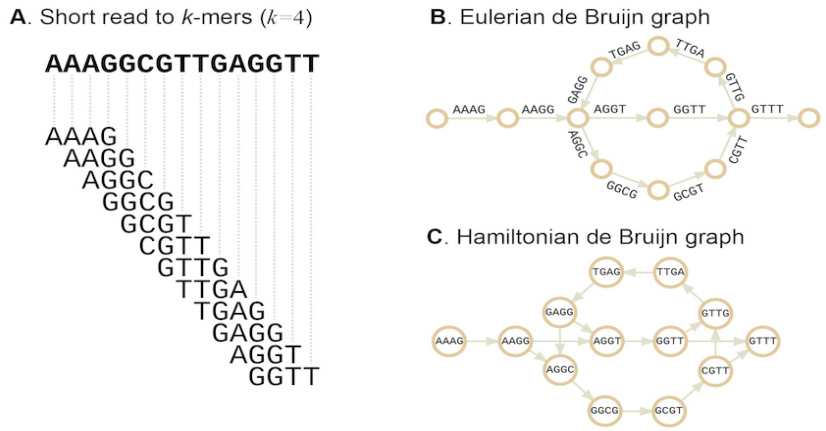


Figura 6. A. Conformación de *K-mers* **B.** Grafico de *De Bruijn* Euleriano **C.** Grafico de *De Bruijn* Hamiltoniano. Fuente: Sohn y Nam, 2016.

Los grafos de *De Bruijn* son clasificados en 2 tipos: Grafos de *De Bruijn* Hamiltoniano y Euleriano, en los grafos Hamiltoniano las secuencia de *K-mers* representan los nodos,

y las secuencias sufijo (Secuencia final del *K-mer* anterior) y prefijo (Secuencia inicial del K-mer) que no llegan a solaparse corresponden a las aristas o arcos, en cambio en los grafos de *De Bruijn* Eulerianos se considera de manera opuesta, la secuencia del *K-mers* interno son los vértices y los sufijos y prefijos de los *K-mers* consecutivos corresponden a los nodos (Sohn y Nam, 2016).

Algunos ejemplos de ensambladores con grafos de *De Bruijn* hamiltonianos son: SOAPdenovo (Luo *et al.,* 2012), SGA (Simpson y Durbin, 2012), ABySS (Simpson *et al.,* 2009) y Velvet (Zerbino y Birney, 2008); en estos programas a medida que el tamaño del genoma o la complejidad topológica aumenta requieren mayor esfuerzo computacional, con el siguiente patrón exponencial O ($m \,.\, 2^n$) donde: m es el número total de nodos y n el número de ramificaciones de nodos (Sohn y Nam, 2016, Figura 7.). Algunos programas simplifican la complejidad computacional removiendo las ramificaciones de los nodos (Luo *et al.,* 2012).

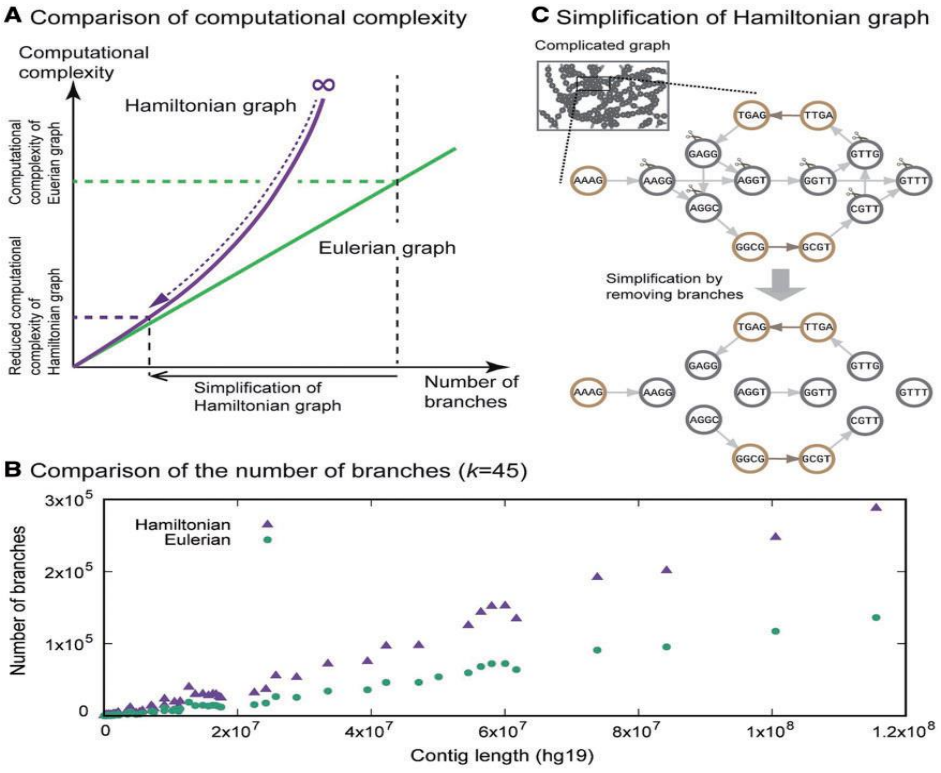


Figura 7. A. Complejidad computacional entre grafos Hamiltoniano y Euleriano **B.** Comparación del número de ramas formadas entre grafos Hamiltonianos y Eulerianos **C.** Grafo de *De Brujin* Hamiltoniano. Fuente: Sohn y Nam, 2016.

Los grafos de *De Bruijn* Eulerianos en cambio, resuelven la complicación topológica del grafo con un solo recorrido de una vía Euleriana, reduciendo la complejidad computacional O (n^2) (Sohn y Nam, 2016). Algunos programas que utilicen este método son: EULER (Pevzner *et al.* 2001), SPAdes (Bankevich *et al.,* 2012), ALLPATHS-LG (Gnerre *et al.,* 2010) y MaSurCA (Zimin *et al.,* 2013).

2.4.2.2 Ensamblado de secuencias a partir de datos illumina.

Los secuenciadores illumina proveen secuencias o lecturas en el formato FastQ (homólogo al formato FASTA, pero con datos de calidad) (Cock *et al.,* 2010). El formato FastQ asocia datos de calidad de secuenciación a cada base reportada, los datos sin procesar (*raw data*) especifican además las coordenadas de la lectura de bases, el número de la celda de flujo de secuenciación y de ser el caso de lecturas PE (*Pair End*) se especifican la lectura par (*File Format Guide-SRA-NCBI*). Estas secuencias sin procesar reportadas en el formato FastQ, contienen gran cantidad de bases con diferente calidad, además de las secuencias de adaptadores que deben ser removidos antes de proceder al ensamblaje del genoma (Dominguez Del Angel *et al.,* 2018). El procedimiento para el ensamblaje de genomas se observa en la Figura 8, donde se incluyen las observaciones discutidas arriba y otras que se desarrollaran más adelante.

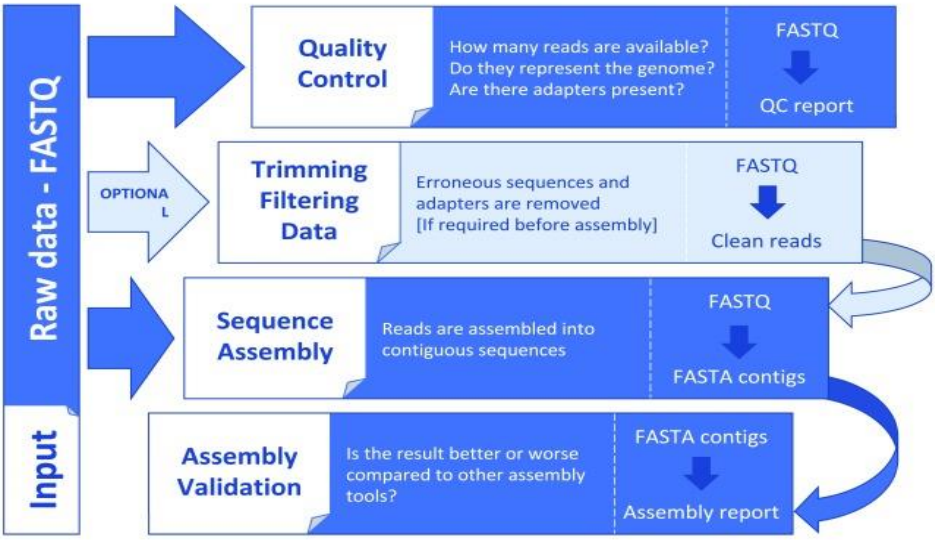


Figura 8. Flujo de trabajo para el ensamblaje de genomas
Fuente: Dominguez Del Angel *et al.,* 2018

2.4.2.3 Control de calidad.

La calidad de las lecturas, mide la probabilidad de que una base sea llamada incorrectamente, esto se mide con los puntajes de calidad denominados *Phred score o Qscore* (Cock *et al.,* 2010).

Probabilidad de error: $P = 10^{-Q/10}$

Q score: $Q = -10 \log_{10} (P)$

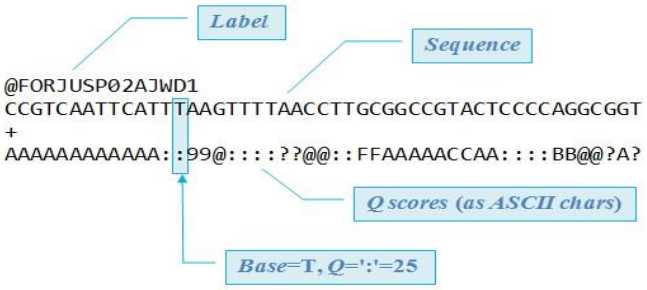


Figura 9. Formato FastQ.

Los puntajes *Phrep* son codificados con caracteres ASCII (*American Standard Code for Information Interchange*) para disminuir costos de almacenamiento. El *Qscore* está codificado con caracteres únicos que pueden ser almacenados en 7 bits de información (Ewing *et al.,* 1998). Para el cálculo de estadísticas de la calidad de secuenciamiento

con gráficas de resumen, se utilizan programas como FastQC (Andrews, 2010), MultiQC (Ewels *et al.,* 2016), AfterQC (Chen *et al.,* 2017), etc.

2.4.2.4 Recorte y filtrado de Secuencias.

Las lecturas secuenciadas contienen adaptadores y cebadores para la amplificación del PCR (Trivedi *et al.,* 2014). Estas secuencias al igual que las lecturas que no tienen adecuada calidad son filtradas, ya que la presencia de estas secuencias puede alterar posteriormente el ensamblaje de genomas. Este proceso es realizado con programas como: Trimmomatic (Bolger *et al.,* 2014), PRINSEQ (Schmieder y Edwards, 2011) FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), CD-HIT (Fu *et al.,* 2012), etc.

2.4.2.5 Análisis de conteo de *K-mers*.

Uno de los principales retos para el ensamblaje de genomas es la corrección de errores de secuenciación. En los secuenciadores illumina existen errores en regiones que tienen elevado contenido de “GC” o “AT”, como el caso de la heterocromatina, incluidos centrómeros, telómeros o regiones altamente repetitivos (Pfeiffer *et al.,* 2018). La tasa de error aleatorio en secuenciadores illumina es < 1 % y tienden a acumularse en el extremo 3` (Pfeiffer *et al.,* 2018). La presencia de estos errores acumulados aumenta la complejidad de los grafos de *De Brujin* (Sohn y Nam, 2016.).

El método de conteo de *K-mers* permite realizar la corrección asumiendo que los *K-mers* con baja profundidad corresponde a errores de secuenciamiento, esto debido que la frecuencia de errores no es muy elevada y por lo general no ocurren en el mismo sitio más de dos veces seguidas (Alic *et al.,* 2016). En un caso ideal donde el secuenciamiento de un genoma ocurra sin errores se esperaría una distribución normal de *K-mers*, en un secuenciamiento con muchos errores se esperaría una curva de distribución exponencial decreciente (Sohn y Nam, 2016). El análisis de secuencias que corresponden a individuos diploides heterocigotos, tienen una disminución de cobertura

D´/2, donde D´ es la cobertura promedio de *K-mers* (Figura 10). La presencia de regiones repetidas en cambio produce una alta frecuencia de *K-mers*, donde los errores son más difíciles de detectar (Manekar y Sathe, 2018).

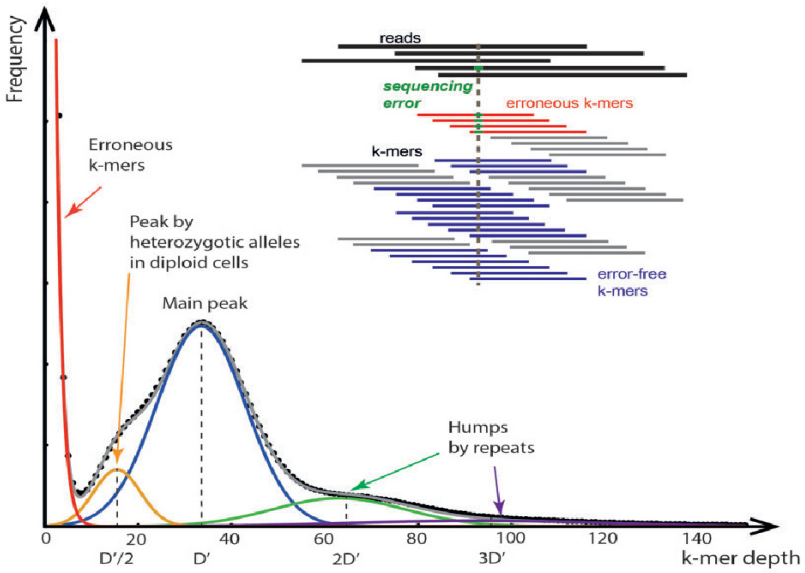


Figura 10. Análisis de la distribución del conteo de *K-mers*
Fuente: Sohn y Nam, 2016.

2.4.2.6 Ensamblado de secuencias.

La reconstrucción de un genoma fue en el pasado una tarea ardua y costosa, el producto de un esfuerzo aunado entre varios países permitió decodificar el primer genoma humano (Ackerman, 1999). Un poco más 1000 pb constituían las lecturas producto de los primeros secuenciadores Sanger, sin embargo estas lecturas no cubrían la totalidad del genoma y no estaban exentas de errores, es por ello que se necesitó más de 8 copias de cada pieza del genoma, aproximadamente 30 millones de piezas, para reconstruir el genoma, una tarea muy complicada de resolver en su momento (Baker, 2012).

En la actualidad los conceptos como cobertura y profundidad heredados de los primeros estudios de ensamblaje son tomados en cuenta para ensamblar genomas a partir de lecturas de NGS (Dominguez Del Angel *et al.,* 2018). El ensamblaje de *novo* con datos NGS (illumina), debe tener una profundidad mínima de 30X (un promedio de 30 copias

para cada posición del genoma) (Ekblom y Wolf, 2014). El proceso de ensamblaje puede ser resumido en tres pasos i) Ensamblaje de *contigs* ii) Ensamblaje del *Scaffold* y iii) Relleno de vacíos (*gaps*) (Sohn y Nam, 2016).

En el primer paso, las lecturas son ensambladas a una secuencia consenso más larga denominada "*contig*". Después los c*ontigs* son ordenados y conectados en un andamiaje más grande denominado "*Scaffold*", que en un escenario ideal debería representar la totalidad de un cromosoma, sin embargo aún existen limitaciones que no permiten llenar los vacíos que existen entre s*caffolds,* algunas limitaciones como: algoritmos utilizados (Baker, 2012) , librerías (Fragmentos de ADN) y el tamaño de lecturas producidas, que se refiere al tamaño de los fragmentos de ADN que son secuenciados, es por ello que se requieren de otras tecnologías de secuenciación independiente para rellenar estos vacíos, que incluyen la secuenciación por Sanger y otras tecnologías denominadas *long read sequencing* o *third generation sequencing* (Sohn y Nam, 2016).

2.4.2.7 Validación del ensamblaje

El ensamblaje de genomas es un proceso con posibles errores, como se explicó líneas arriba, ya sea por las limitaciones bioinformáticas (inadecuados algoritmos) o la tecnología utilizada (errores de secuenciamiento), por ello es necesario corroborar los resultados de ensamblaje con estadísticas y referencias externas del genoma ensamblado (Salzberg *et al.,* 2012).

Para estos fines se han desarrollado algunos programas, que permiten realizar estadísticas de validación como CAGE (Salzberg *et al.,* 2012), QUAST (Gurevich *et al.,* 2013), LAP (Ghodsi *et al.,* 2013), REAPR (Hunt *et al.,* 2013), etc. Para la validación del ensamblaje se puede utilizar el genoma de referencia o someter a otros análisis el genoma que se está ensamblando. Las secuencias de genes ensamblados se pueden contrastar con bases de datos que contienen información experimental, para ello se utilizan los siguientes programas: BUSCO (Simão *et al.,* 2015), tRNAscan-SE v2.0

(Lowe y Chan., 2016), GenomeQC (Manchanda *et al.,* 2020), etc. Para constatar el correcto ensamblaje del genoma, habitualmente se realizán verificaciones con las siguientes estadísticas (Earl *et al.,* 2011):

N50: Definido como la longitud del *contig*, que corresponde a la mediana de los *contigs* obtenidos en el ensamblaje.

NG50: Este parámetro es similar a N50, solo que considera el tamaño del genoma de referencia.

L50: Este parámetro es muy similar al N50, solo que se define en términos de número de *contigs*.

A estas métricas se deben sumarse algunas más, que se han desarrollado a lo largo de los años y que clasificaremos en las siguientes métricas:

1. **Métricas de Contigüidad**

   Las métricas de contigüidad, consiste en estadísticas básicas que muestran el grado de fragmentación del ensamblaje, dentro de estas estadísticas se incluyen: Tamaño total de ensamblaje, *contig* más largo, fracción del genoma, N50, NG50 y L50 (Earl *et al.,* 2011).

2. **Métricas de Coherencia**

   Las métricas de coherencia, permite comprobar la fidelidad del ensamblaje, realizando un realineamiento (mapeo) de las lecturas que se utilizaron para el ensamblaje al genoma obtenido, esperándose un alto porcentaje de alineamiento. Las estadísticas que se reportan para esta métrica son: Cobertura de alineamiento, Porcentaje de alineamiento (%), Alineamiento propiamente apareados y Alineamiento sin pares (*Singletons*), etc. (Gurevich *et al.,* 2013).

3. **Métricas de Integridad**

   Las métricas de integridad permiten contrastar el ensamblaje con bases de datos de genes, de tal manera que los genes ensamblados coincidan con los genes reportados anteriormente. En estas bases de datos se contrastan los genes codificantes de proteínas (Simão *et al.,* 2015), genes de ARN de transferencia y ARN ribosomal (Lowe y Chan., 2016).

2.4.3 Ensamblaje de genoma mitocondrial

Antes de la tecnologías de NGS, el secuenciamiento y ensamblaje de los genomas mitocondriales fueron realizados a partir de productos de PCR, clonación de fragmentos de PCR o a través de la separación de genoma mitocondrial por procesos químicos (Dierckxsens *et al.* 2017). Actualmente la secuenciación del genoma mitocondrial se puede llevar a cabo a partir de ADN genómico total mediante NGS, y las metodologías para el ensamblaje del genoma mitocondrial que se puede implementar son: separación de secuencias mitocondriales de las lecturas nucleares mediante un mapeo de las lecturas al genoma completo (Flicek y Birney, 2009); método de extensión de "semilla", que toma un gen o un genoma mitocondrial/plastidial de referencia, para filtrar por alineamientos sucesivos lecturas del genoma, esto es realizado por el programa NOVOplasty (Dierckxsens *et al.,* 2017) y método de separación por *k-mers* de alta frecuencia, que filtra secuencias de alta repetitividad, realizado por el programa Norgal (Al-Nakeeb *et al.,* 2017). El ensamblaje de genomas mitocondriales sigue el mismo procedimiento explicado anteriormente para genomas, con la modificación de filtrado de las lecturas mitocondriales a partir de un WGS, además no necesita tener una alta profundidad para poder realizar el ensamblaje de genomas mitocondriales por la cantidad de mitogenomas presentes en el ADN total, reduciendo aún más los costos (Rustagi *et al.,* 2017).

2.4.4 Análisis filogenómico de mitogenomas.

Debido a la cantidad de mitogenomas disponibles en las bases de datos y a su constante crecimiento (Smith, 2016), y gracias a los avances de la tecnologías de NGS los estudios de filogenómica comparativa han cobrado relevancia en los últimos años (Chan y Ragan, 2013), dentro de los programas más utilizados para estudios filogenéticos, podemos nombrar a aquellos que se basan en el método de máxima verosimilitud (Maximun likelihood) como RaxML (Stamatakis, 2014) o PhyML (Guindon *et al.,* 2010.); y análisis filogenéticos con el método de inferencia bayesiana como Beast (Suchard *et al.,* 2018) o MrBayes (Ronquist F. *et al.,* 2012).

Los mitogenomas también son utilizados para análisis de patrones de domesticación a partir de inferencia bayesiana de coalescencia (Drummond *et al.,* 2005) en animales que han sufrido de una selección intensiva, estudios que son complementados con los estudios de diversidad de linajes matrilineales (Sun *et al.,* 2020).

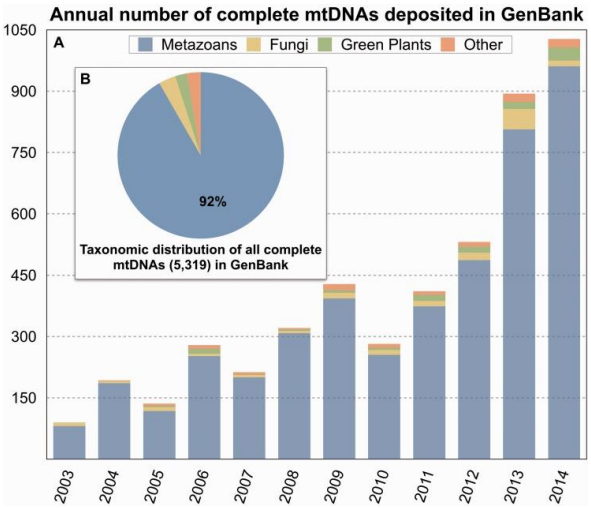


Figura 11. Genomas mitocondriales depositados anualmente en GenBank. Fuente: Smith, 2016

3. HIPÓTESIS Y OBJETIVOS

3.1 Hipótesis

Existen variantes genéticas entre los genomas mitocondriales en *C. porcellus* y *C. tschudii* debido a procesos de selección que permite reconstruir su grado de relacionamiento filogenético.

3.2 Objetivos

3.2.1 Objetivo general.

- Secuenciar, ensamblar y analizar los genomas mitocondriales de *C. porcellus* ("criollos" y mejorados) y *C. tschudii* provenientes de Perú para establecer su relación filogenética.

3.2.2 Objetivos específicos.

- Realizar una adecuada extracción de ADN para la secuenciación de genoma completo de *C. porcellus* y *C. tschudii.*
- Ensamblar y anotar los genomas mitocondriales de *C. porcellus* y *C. tschudii*.
- Establecer las distancias genéticas, así como las relaciones filogenéticas entre *C. porcellus* y *C. tschudii* a partir de los mitogenomas ensamblados.
- Establecer la red de haplotipos, al igual que la dinámica poblacional de *C. porcellus* y *C. tschudii* a partir de los mitogenomas ensamblados.

4. MATERIALES Y MÉTODOS.

4.1 Materiales

4.1.1 Material Biológico.

Se utilizaron muestras de tejido de *C. porcellus* “Criollos” provenientes de la Estación Experimental Andahuaylas-INIA sede Chumbibamba, *C. porcellus* mejorados de raza Perú de la Sede Central La Molina-INIA, *C. tschudii* provenientes del refugio silvestre Los Pantanos de Villa y muestras de folículo piloso de cuyes silvestres provenientes de la región de “Pachachaca”- Abancay Apurímac. Para colecta de los animales se formalizaron permisos para colecta de material genético de cuyes (permiso de colecta N° 005-2019-SERNANP-JEF y Memorándum N°815-2019-MINAGRI.INIA-DGIA/SDRIA) (ver **Anexo 1**).

4.1.2 Material de Laboratorio.

Los materiales utilizados para esta tesis fueron proporcionados por el Laboratorio de Biología Molecular y Genómica del INIA (LBMG-INIA), que cuenta con áreas de extracción de ADN, análisis de calidad de ADN y electroforesis. Los equipos utilizados son mencionados pertinentemente en el desarrollo de los métodos.

4.2 Métodos

El procedimiento utilizado en todo el proceso de la elaboración de esta tesis es resumido en flujo de trabajo de la Figura 12.

Ensamblaje y Análisis del Genoma Mitocondrial

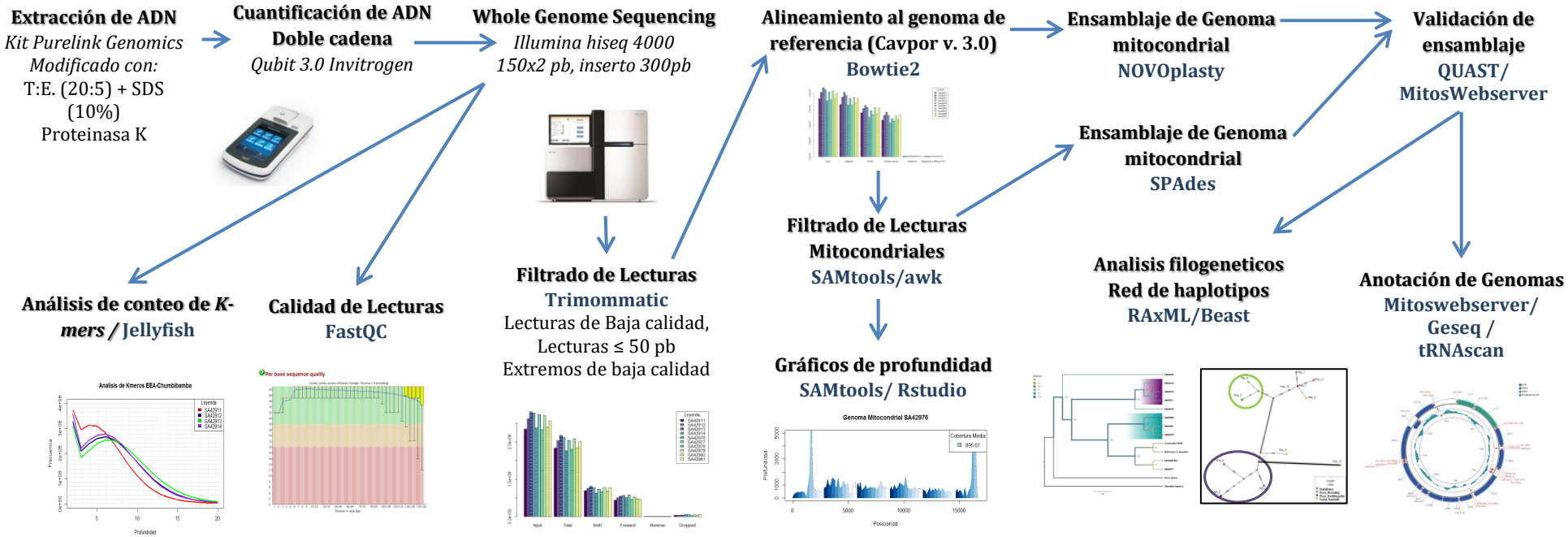


Figura 12. Flujo de trabajo desde extraccion de ADN hasta el esamblaje de genomas.

4.2.1 Extracción de ADN genómico total.

La extracción de las muestras de ADN fue realizada con el kit de extracción *Purelink ™ Genomic mini kit de* la marca Thermo Fisher Scientific, con una modificación previa. Para la extracción se cortaron los tejidos con ayuda de un bisturí, posteriormente se añadió 750 µl del buffer TE (20:5), 75 µl SDS (10%) y 15 µl de Proteinasa K (10 ng/µl*)* y se incubaron a 55°C por 3 horas hasta obtener una lisis completa; las muestras se enfriaron a temperatura ambiente, luego se añadió 20 µl de RNAasa (*PureLink™ Genomic Mini Kit*) y se incubó por 2 minutos a temperatura ambiente según las indicaciones de la enzima. Después de la incubación se añadió 500 µl de Acetato de potasio 1M y se refrigeró por 10 minutos a – 20°C, posteriormente se centrifugó por 10 minutos a 14 000 RPM y se recuperó el sobrenadante en un tubo de microcentrifuga de 2.0 ml que contenía 800 µl de isopropanol helado, se centrifugó por 10 minutos a 14 000 RPM y se eliminó el sobrenadante. A partir de este punto se procedió con el manual de extracción de Kit *PureLink™ Genomic* (**Anexo 2**).

4.2.2 Evaluación de la calidad del ADN.

Se realizó electroforesis en geles de agarosa al 1%, para verificar la calidad del ADN, con las siguientes condiciones: Se utilizó cámaras electroforéticas horizontales, buffer TBE 1X (Tris-HCl, ácido bórico, EDTA; pH 8,0), buffer de carga compuesta por Sal B 2X (azul de bromofenol, xilencianol, naranja G, sucrosa) y Gel Red, voltaje de corrida de 100 V y tiempo de corrida electroforética de 20 minutos. Los geles se registraron utilizando un transiluminador UV Chemi XR (BioRad) y se editaron mediante el programa Quantity One v. 4.6.3 (BioRad). Las concentraciones de ADN y los parámetros de absorbancia (260/280 y 260/230) se calcularon con el espectrofotómetro NanoDrop™ 8000 (Thermo Scientific), se evaluó también la concentración de ADN de doble cadena con el lector de fluorescencia Qubit® 3.0 (Thermo Scientific™).

Para el envío de las muestras de ADN, se tomó en cuenta los siguientes requisitos:

● Cantidad de ADN > 200 ng

● Alto peso molecular, verificable en geles de agarosa.

● Concentración de ADN > 20 ng/ul, con una concentración óptima entre: 20 ng/ul a 500 ng/ul en Tris - EDTA (10:1).

● Pureza del ADN con los siguientes parámetros de absorbancia: 260nm/280nm: 1.8 - 2.0, 260nm/230nm ≥ 1.8.

● Muestras libres de ARN.

4.2.3 Secuenciación de genoma completo (WGS).

Las muestras de ADN genómico total fueron enviados para la secuenciación del genoma completo (*Whole Genome Sequencing*, WGS), a la empresa QuickBiology - USA. Que fue realizado a través de la tecnología illumina utilizando el secuenciador $HiSeq^{TM}\ 4000$ con un requerimiento de 100 Millones de lecturas, tamaño de librería de 300 pb y fragmentos de lecturas *Pair-end* (150x2).

4.2.4 Equipos de procesamiento de información.

Para el procesamiento de la información proveniente del secuenciamiento del genoma completo, se trabajó en un servidor de la marca Hewlett Packard Enterprise (HPE), memoria RAM 160 GB, memoria en disco 2x2.4 TB SAS- 1x240 GB SATA SSD, sistema operativo Linux Ubuntu server 18.04.3 LTS y un dispositivo conectado a red NAS (*Network Attached Storage*) de 10 TB de capacidad de almacenamiento.

4.2.5 Análisis de calidad de Secuenciación.

El análisis de calidad de secuenciación se dividió en dos partes, concernientes a la calidad *Phred* de las lecturas resultantes y al análisis de *k-mers*:

4.2.5.1 Análisis de calidad con FastQC.

Para el análisis de la calidad de secuenciación se utilizó el programa FastQC (Andrews., 2010), se analizaron las secuencias *forward* y *reverse* de cada individuo en formato comprimido y se consideró para el análisis: Calidad de las muestras con una puntuación de *phred score* superior a 30 para cada posición de las bases, calidad de secuencias para identificar desviaciones de calidad (grupos de lecturas con diferente calidad), contenido de las bases con diferencias máxima de %5, distribución normal de contenido de GC, presencia de bases N, distribución homogénea del tamaño de lecturas, presencia de duplicación de secuencias y presencia de secuencias sobrerrepresentadas.

4.2.5.2 Análisis de conteo de *K-mers*.

Se realizó el análisis de conteo de *k-mers* con el programa Jellyfish (Marçais y Kingsford., 2011), se consideró para el análisis: tamaño de *K-mer*s de 31 pb, tabla *hash* de 10 Gb, límite superior conteo de *k-mers* de 1000 e inferior de 0. A partir de los datos obtenidos de conteo se realizó un histograma con el mismo programa y las gráficas se realizaron en el lenguaje de programación R versión 4.0

Para el análisis se consideró la presencia de un pico de baja profundidad y alta frecuencia atribuida a errores de secuenciación, presencia de picos de heterocigosidad y alta frecuencia de número de copias de regiones repetitivas.

4.2.6 Filtrado de Secuencias por tamaño y baja calidad.

Se utilizó el programa Trimmomatic (Bolger *et al.,* 2014) para filtrar los datos de secuenciación con las siguientes características: se indicó al programa la librería *Pair End* (PE), *score* de calidad *Phred 33*, librería de adaptadores Truseq3-PE-2 para su identificación y remoción, recorte de extremos de baja calidad y un tamaño mínimo de secuencias de 50 pb.

4.2.7 Alineamiento de secuencias al genoma de referencia Cavpor v3.0.

Se realizó alineamientos de las lecturas filtradas del secuenciación de cada individuo con el genoma completo de referencia de *Cavia porcellus* Cavpor v3.0 (Numero de accesión del NCBI: GCF_000151735.1) con el programa Bowtie2 (Langmead y Salzberg., 2012), para el alineamiento se utilizó la opción de alineamiento sensible de extremo a extremo `(--end-to-end` *mode*). Para reducir la cantidad de espacio utilizado se convirtió y ordeno el archivo de alineamiento "SAM" a un formato binario "BAM", con el programa SAMtools (Li *et al.,* 2009).

4.2.8 Filtrado de secuencias mitocondriales.

Las lecturas mitocondriales fueron filtradas como sugieren Flicek y Birney (2009). Del archivo BAM del alineamiento del genoma completo (Cavpor v3.0) se filtraron las lecturas que corresponden a las secuencias del genoma mitocondrial con el programa SAMtools (Li *et al.,* 2009), para esto se identificó el código del genoma mitocondrial (NC_00884.1) presente en la referencia Cavpor v3.0, y se extrajeron secuencias *forward* y *reverse* propiamente mapeadas, además secuencias *forward* y *reverse* que no estuvieron emparejadas, estas secuencias luego fueron emparejadas haciendo una búsqueda en los archivos filtrados por calidad (sección 4.2.6). Adicionalmente se realizaron análisis de profundidad de las secuencias alineadas al genoma mitocondrial con el programa SAMtools (Li *et al.,* 2009), y se realizaron gráficas de cobertura con el lenguaje de programación R versión4.0. Para

realizar el ensamblaje se convirtieron las secuencias de alineamiento al formato *fastq* con el programa seqtk (Shen *et al.,* 2016) y se filtraron aleatoriamente para reducir profundidad con el mismo programa para garantizar el correcto ensamblaje debido a que se identificaron secuencias con una cobertura mayor a 1000X.

4.2.9 Ensamblaje del genoma mitocondrial.

El ensamblaje del genoma mitocondrial fue realizado por 2 metodologías:

1. Para el ensamblaje por el método de grafos de De Brujin (Grafos Eulerianos), se utilizó el programa SPAdes (Bankevich *et al.,* 2012). Para ello se tomaron las secuencias convertidas al formato *fastq* de la sección 4.2.8, y se realizó un ensamblaje *de novo* considerando *K-mers* de 51, 71 y 91 pb.
2. Para el ensamblaje por el método basado por extensión de semilla (*seed-extend based assembler*) se utilizó el programa NOVOplasty (Dierckxsens, Mardulyn, y Smits, 2016), este ensamblaje fue realizado con los datos filtrados por calidad en la sección 4.2.6 que incluye secuencias del genoma total, se indicó en el programa la opción de ensamblaje para mitocondrias y para realizar la extensión por "semilla" (seed), se descargó una secuencia del gen citocromo *b* de *Cavia porcellus* (número de accesión Genbank: HM447187.1).

4.2.10 Validación de ensamblaje de genomas

Para validar los ensamblajes de los 2 métodos utilizados se tomaron en cuenta las siguientes Métricas:

- Métricas de contigüidad; con el programa QUAST (Gurevich, *et al.,* 2013), donde se evaluó los siguientes parámetros: tamaño total del ensamblaje, N50, L50, NG50, c*onting* más largo.

- Métricas de coherencia; fueron realizados con QUAST (Gurevich*, et al.,* 2013), se alineó las secuencias utilizadas como *input* el genoma ensamblado y se evaluaron los siguientes parámetros: Tamaño total del alineamiento, porcentaje de alineamiento, secuencias propiamente pareadas, porcentaje de secuencias sin par (*Singletons*) y porcentaje de alineamiento con cobertura mayor a 1.
- Métricas de Integridad (*Completenesss*); fue realizado con el servidor web MITOwebserver (Bernt *et al.,* 2013) con parámetros estándar, se evaluó la presencia de regiones: codificantes de proteínas (CDS), genes de ARN de transferencia y genes de ARN ribosomales.

4.2.11 Ensamblaje de secuencias mitocondriales de la base de datos SRA NCBI

Se realizó el ensamblaje de genoma mitocondrial de un WGS disponible en la base de datos SRA (código de accesión: SRX4501347), las lecturas contaban con una librería PE (*Pair end*) de 2x250 y tamaño de librería de 450pb, la secuenciación fue realizado con un equipo Illumina Hiseq 2500 a una profundidad de 34.8X. El procesamiento de estas secuencias fue realizando siguiendo las mismas pautas descritas en las secciones 4.2.5 al 4.2.10.

4.2.12 Anotación de genoma mitocondrial.

Para la anotación de las secuencias mitocondriales se utilizó el servidor web MITOwebserver (Bernt *et al.,* 2013); junto con los programas Geseq (Tillich *et al.,* 2017) y tRNAscan-SE v2.0 (Lowe y Chan., 2016) dentro de la plataforma web MPI-PI CHOLOROBOX. Se generaron archivos de anotación en formato GFF y Genbank, estas anotaciones fueron editadas manualmente para cruzar los resultados de los análisis con el programa UGENE (Okonechnikov, Golosova, y Fursov, 2012), los anotaciones fueron luego visualizadas con el programa Cgview (Grant y Stothard, 2008)

4.2.13 Análisis de distancias genéticas.

Para los análisis de distancia genéticas se utilizaron los genomas ensamblados en la sección 4.2.9 con el programa SPAdes (Bankevich *et al.,* 2012), el genoma ensamblado de la base de datos SRA (SRX_4501347); y de la base de datos de Genbank los genomas con código: NC_000884.1 (refseq), KP100656.1 y MT_0017565.1.

Estas secuencias fueron alineadas con el programa clustalw (Thompson, Higgins, y Gibson, 1998) y recortadas para el análisis, se realizaron análisis de distancia genética con el modelo de K80 con el paquete de R (R *core* version: 4.0) adegenet (Jombart, 2008).

4.2.14 Análisis filogenómicos, demograficos y haplotipos.

Para los análisis filogenéticos también se analizaron las secuencias de la sección 4.2.13, se añadió el genoma de referencia de *Cavia aperea* (NC_469667.1) y se consideró como grupo externo el genoma mitocondrial de referencia de *Chinchilla lanigera* (NC_21386.1).

Para los análisis filogenéticos se escogió el modelo de sustitución nucleotídica GTR+G+I con MEGA-X (Kumar *et al.,* 2018). Los análisis de máxima verosimilitud fueron realizados con RAxML-HPC v8.2.11 (Stamatakis, 2014) con el modelo GTRGAMMAI, 1000 iteraciones de bootstrap con un análisis rápido (*rapid Bootstrap analysis*). Para los análisis de máxima parsimonia se utilizaron los paquetes *Phangorn* (Schliep, 2011) *y ape* (Paradis & Schliep, 2019) de R (R core version: 4.0). Los análisis de inferencia bayesiana y demográficos por coalescencia, se realizaron con el programa BEAST v. 1.10.4 (Suchard *et al.,* 2018). Para el análisis de inferencia bayesiana se utilizó el modelo de sustitución nucleotídica GTR+G+I, reloj molecular estricto, el análisis de MCMC fue realizado con 10,000,000 iteraciones con muestreo posterior cada 1000 iteraciones y se seleccionó el modelo *birth-death* para el "*tree prior*". Se evaluó el tamaño efectivo de muestreo con Tracer v1.7 (Rambaut *et al.,* 2018) y se extrajo el árbol filogenético consenso usando TreeAnnotator v1.10.4 con un 10 % de burn-in.

Para los análisis de coalescencia (*bayesian skyline*) se cambió el modelo de "*tree prior*" por *bayesian skyline* (Drummond *et al.,* 2005) manteniendo los demás parámetros similares. La reconstrucción bayesiana skyline fue realizada con Tracer v1.7 (Rambaut *et al.,* 2018) verificando un ESS (*Effective population size*) mayor de 200. Los árboles generados por máxima verosimilitud, máxima parsimonia e inferencia bayesiana fueron editados y visualizados con FigTree v1.4.3. Para análisis de haplotipos de mitogenomas se calcularon con DnaSP6 (Rozas *et al.,* 2017) y se reconstruyó una red de haplotipos median-joining network en popART (Leigh & Bryant, 2015).

5. RESULTADOS

5.1 Extracción y Evaluación de calidad de ADN genómico total

Se extrajo ADN de 7 muestras de *Cavia porcellus* y 3 muestras de *Cavia tschudii*, la Tabla 2 muestra el resumen, indicando el tejido, la procedencia y la evaluación de la extracción de ADN. Se obtuvo por espectrofotometría (NanoDrop) una concentración de ADN superior a 20 ng/µl, y las tasas de absorbancia 260/280 y 260/230 se mantuvieron dentro de los parámetros esperados.

Tabla 2. Cuadro resumen de extracción y calidad de muestras ADN.

	Código	Tejido	Procedencia	Concentración (ng/µl) NanoDrop	260/280	260/230	Concentración (ng/µl) Qubit
C. porcellus " Mejorados"	CP-1135	Músculo	EEA La Molina	59.24	1,84	2,15	58.4
	CP-1136	Músculo	EEA La Molina	79.63	1,89	2,31	67
	CP-1137	Músculo	EEA La Molina	32.00	1,9	2,02	30.1
C. porcellus " Criollos"	CP-1144	Músculo	EEA Chumbibamba	98.57	1,87	2,19	98.6
	CP-1145	Músculo	EEA Chumbibamba	59.25	1,84	2,42	60
	CP-1146	Músculo	EEA Chumbibamba	137.1	1,9	2,37	132
	CP-1068	Cartílago	EEA Chumbibamba	314.6	1.91	1.9	120
C. tschudii " Silvestres"	CT-034	Folículo Piloso	Pachachaca-Abancay Apurímac	30,09	1,84	2,07	24.80
	CT-035	Músculo	Refugio de vida Silvestre Pantanos de Villa	37.81	1,88	2,27	33,9
	CT-036	Músculo	Refugio de vida Silvestre Pantanos de Villa	69.16	1,86	2,35	56,2

EEA: Estación Experimental Agraria.

Se alcanzó una concentración superior a 20 ng/µl con el lector de fluorescencia (Qubit) sin una caída significativa en comparación de las lecturas por espectrofotometría y buena integridad de las muestras de ADN en el gel de electroforesis (Figura13).

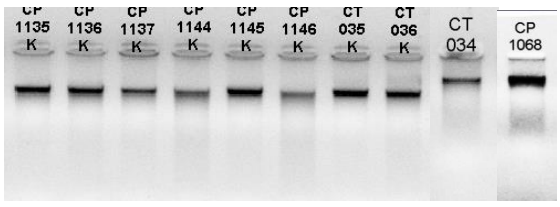


Figura 13. Gel de corrida electroforética de muestras

5.2 Secuenciación del Genoma completo

La secuenciación de las muestras supera en todos los casos el requerimiento de 100 millones de lecturas. La Tabla 3 muestra el tamaño de la librería, los códigos asignados en la secuenciación, la cantidad total de secuencias generadas y la calidad promedio de las lecturas con *Phred score* mayor o igual a 30. Las secuencias obtenidas tienen una media Q30 mayor del 90 %.

Tabla 3. Cuadro resumen de secuenciación de genoma completo.

Código	Bioanalizador tamaño (bp)	ID de muestra (*)	Secuencias Totales (**)	Media % Q30
CP-1135	464	SA42979	135.152.047	93,91
CP-1136	440	SA42980	119.965.953	93,95
CP-1137	482	SA42981	128.122.069	92,83
CP-1144	455	SA42911	117.340.314	93,6
CP-1145	449	SA42912	130.787.116	94,1
CP-1146	448	SA42913	140.368.353	94,22
CP-1068	426	SA42914	138.641.244	93,71
CT-034	446	SA42978	116.493.452	93,62
CT-035	456	SA42976	119.162.747	93,56
CT-036	441	SA42977	136.615.724	93,63

* Los códigos de identificación para el secuenciamiento son utilizado para realizar los análisis posteriores.
** Las secuencias totales hace referencia al total de insertos generados en la secuenciación.

5.3 Análisis de calidad de secuenciación

Análisis de calidad con FastQC

Se realizó el análisis de la calidad de la secuenciación de 10 individuos. En general con buena calidad de secuencias en los gráficos resumen del programa FastQC. La figura 14 muestra la calidad por cada posición de bases, el diagrama de cajas muestra una calidad dentro de la región verde (*very good quality*) conservando su promedio dentro de la misma área (Figura

14), pero se observa una diferencia de calidad de secuenciación entre las secuencias *forward* y *reverse*, en las últimas posiciones de la secuencia.

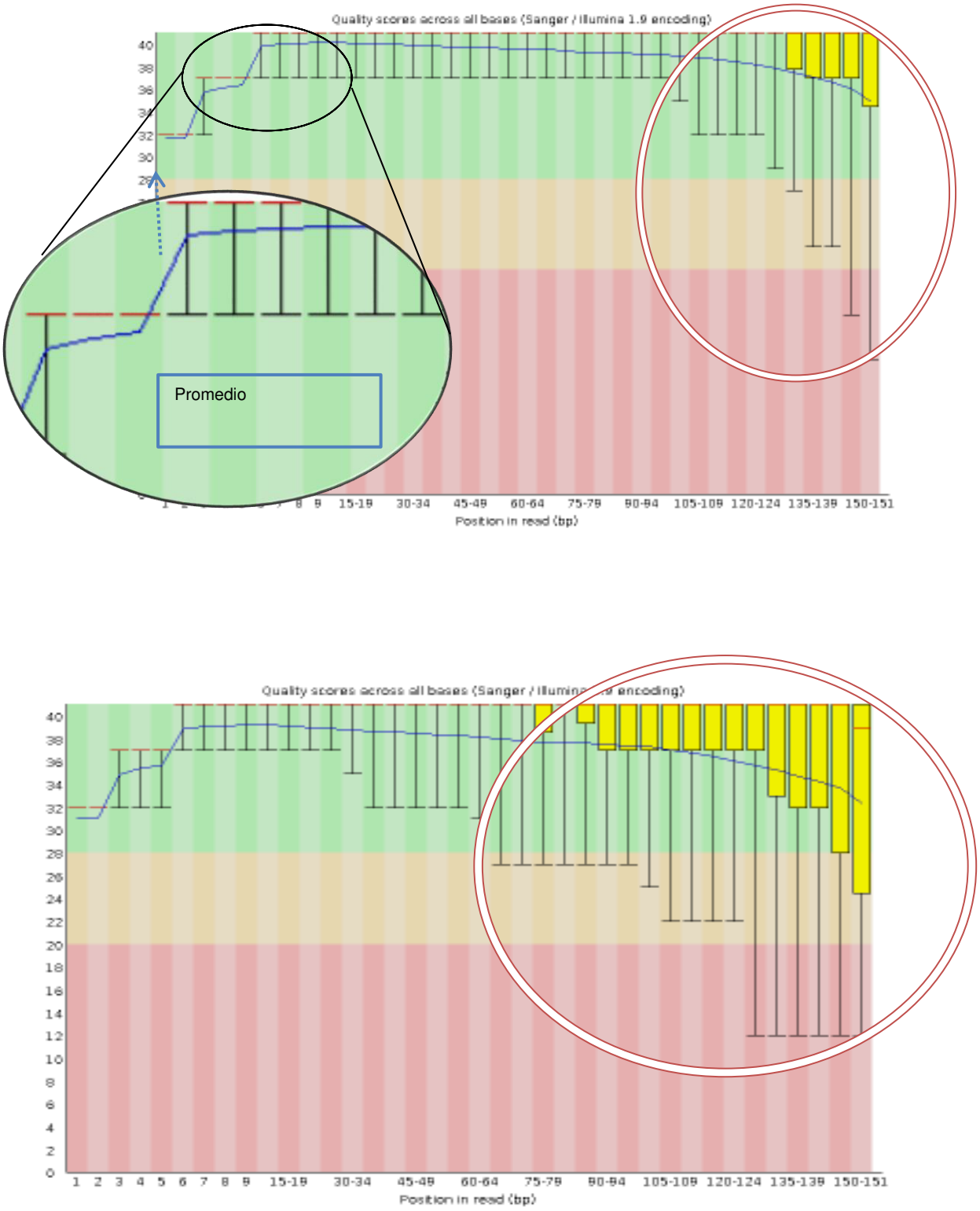


Figura 14. Calidad por cada base de las secuencias forward (izquierda), reverse (derecha).

No se observa un subconjunto de secuencias de diferente calidad en las gráficas de calidad de secuencias, manteniendo las secuencias en conjunto una buena calidad (Figura 16 A), una distribución de calidad con pico de tamaño en 151 pb (Figura 16 B).

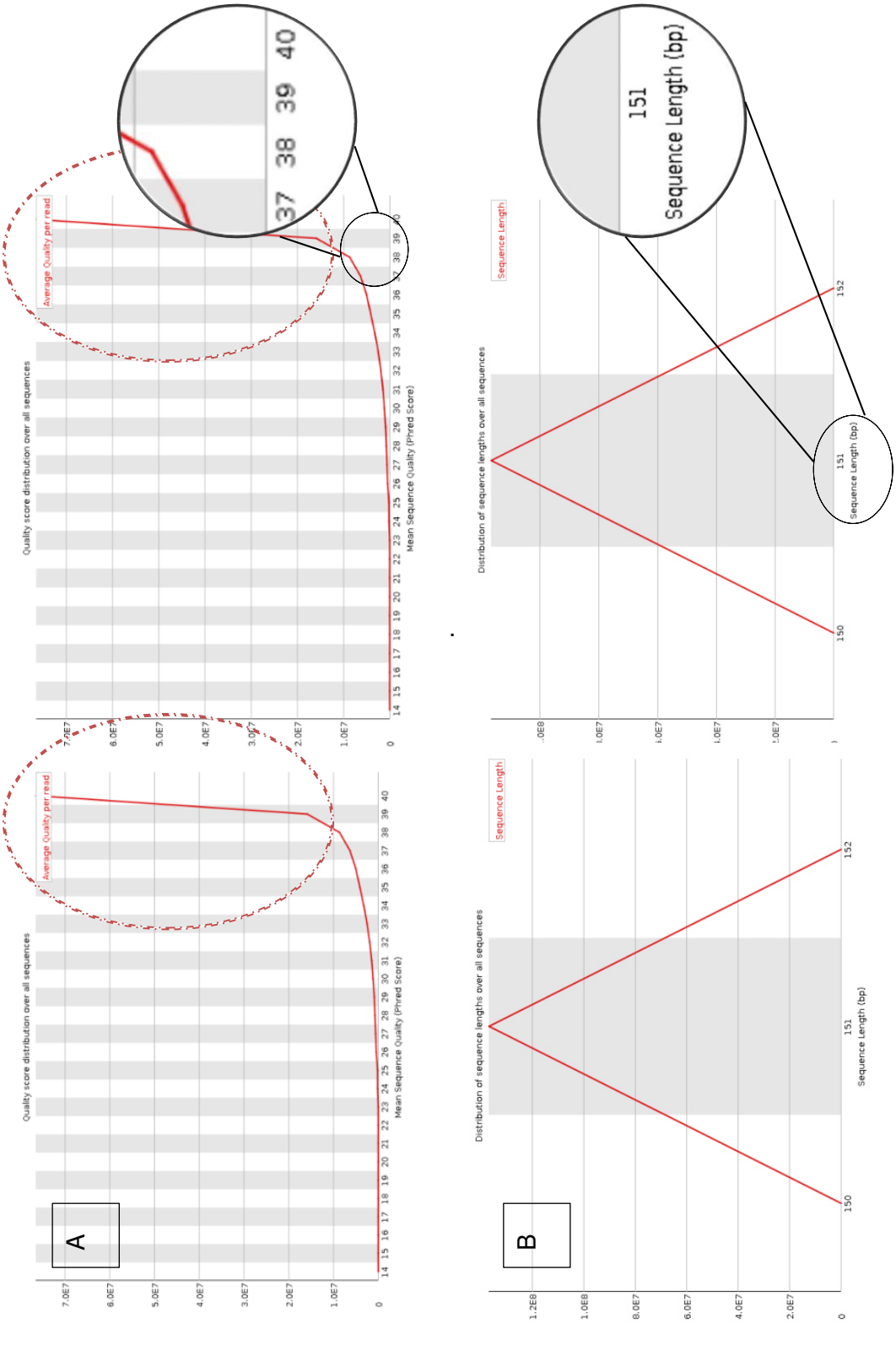


Figura 15. A. *Score* de calidad por secuencias. **B**. Distribución de longitud Forward (izquierda), reverse (derecha).

Se aprecia una variación del contenido de bases en las primeras posiciones de las secuencias, regularizándose a partir de la posición 15; sin embargo, se observa en las últimas posiciones desviaciones del contenido de bases, siendo más evidente en las secuencias *reverse* (Figura 16. A). El contenido GC (rojo) de la figura 16. B, sigue una distribución equiparable a la distribución normal de referencia (azul), sin desviaciones significativas (**Anexo 5**). Los reportes de calidad resumen de todas las muestras secuenciadas se puede revisar a mayor detalle en el **Anexo 5**.

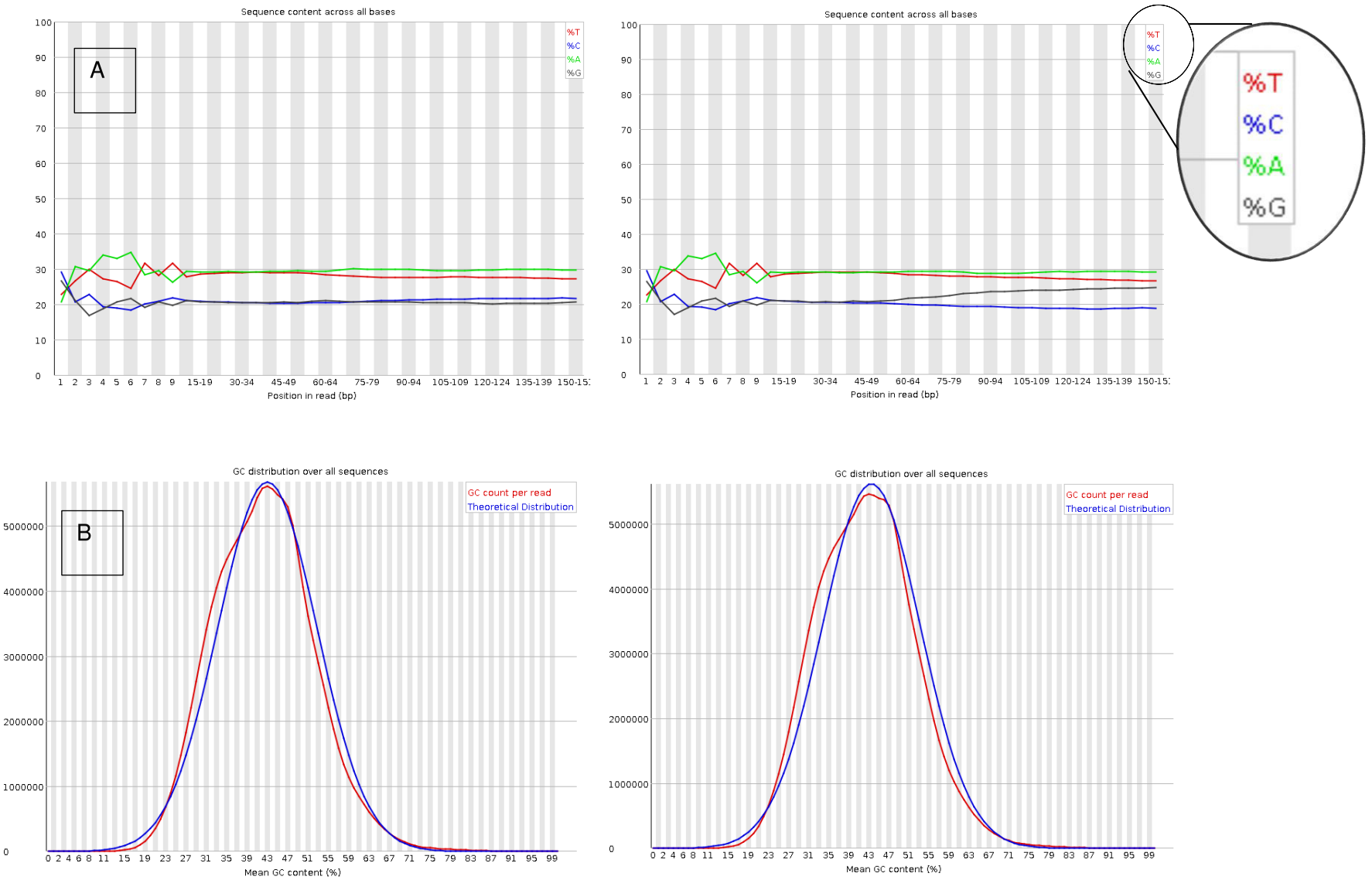


Figura 16. A. Contenido de bases por secuencias **B**. Distribución del contenido de GC. Forward (izquierda), reverse (derecha).

Análisis de conteo de *K-mers*

En el análisis de conteo de *k-mers*, se obtuvieron 2 picos de alta frecuencia de *k-mers,* uno correspondiente a errores intrínsecos de secuenciación illumina normalmente a una profundidad de 1x, y otro pico de la distribución promedio de *k-mers* correspondientes a las secuencias unicas del genoma, no se observan picos de regiones repetitivas y heterocigosidad sobrerrepresentadas (Figura 17).

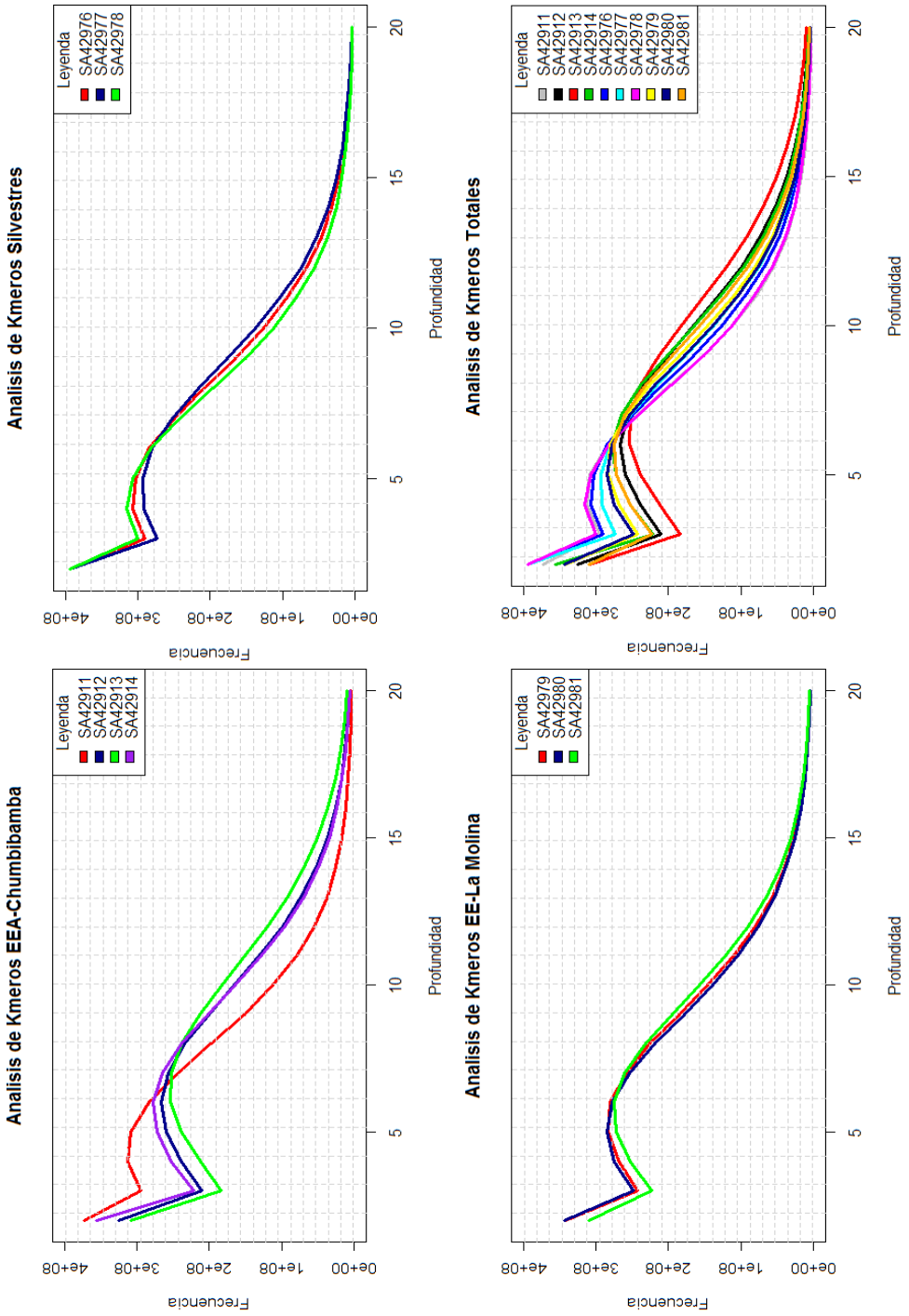


Figura 17. Grafica de conteo de *K-mers*

5.4 Filtrado de secuencias de genoma completo

Se recuperaron en promedio 76,8 % de lecturas, estas secuencias estuvieron compuestas en promedio por un 56,9 % de lecturas pares, 39.6% de lecturas no apareadas *Forward*, se obtuvo poco porcentaje de lecturas *Reverse* no apareados (0.29 %) y un 3.21% en promedio de lecturas removidas (Figura 18). El número de lecturas recuperadas varía entre los individuos según la cantidad total de lecturas (*input*) generadas en el secuenciamiento, estos porcentajes se resumen en la tabla del **Anexo 3**.

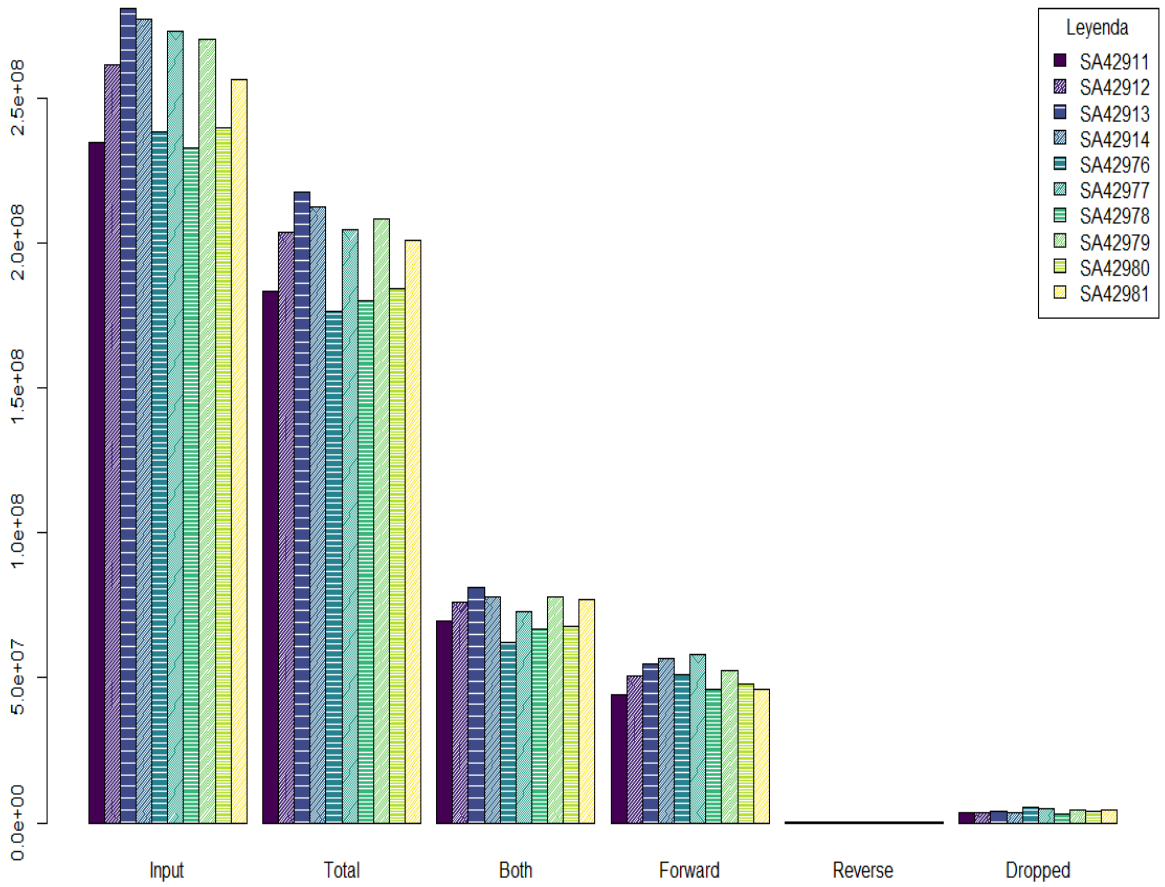


Figura 18. Grafica de filtrado de secuencias de baja calidad del genoma completo, **Input**: Secuencias antes del filtro; **Total**: Secuencias después del filtro; **Both:** Secuencias con pares (*forward* y *reverse*); ***Forward, Reverse*:** Secuencias sin pares después del filtro; **Dropped:** Secuencias descartadas después del filtro.

5.5 Alineamiento al genoma de referencia

Se obtuvo en promedio 91,6 % de lecturas totales alineadas al genoma de referencia, de las lecturas pares un 84.8 % de secuencias fueron propiamente alineados, se evidencio bajo porcentaje de lecturas sin par (*Singletons*) y lecturas alineadas a otra región diferente (Figura 19). Las estadísticas alineamiento al genoma de referencia Cavpor v3.0 están detalladas en el **Anexo 4.**

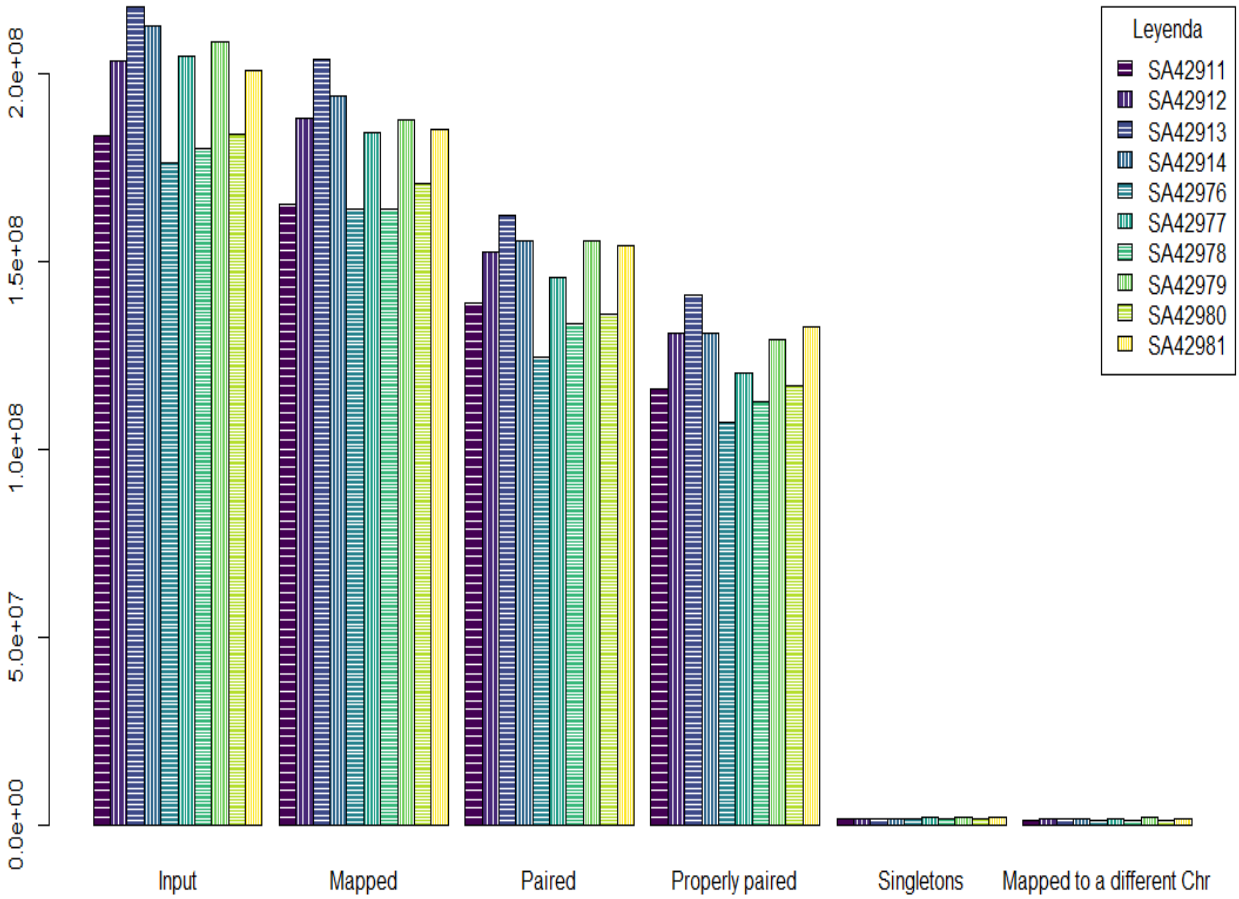


Figura 19. Grafica de porcentaje de alineamiento al genoma de referencia Cavpor v3.0, **Input**: Secuencias antes del alineamiento; **Mapped**: Secuencias alienadas al genoma Cavpor v3.0; **Paired:** Secuencias con pares (*forward* y *reverse*); **Properly paired:** Secuencias con pares (*forward* y *reverse*) alienadas en el mismo cromosoma o región; **Singletons:** Secuencias sin pares; **Mapped to different Chr:** Secuencias mapeadas a diferentes cromosomas o regiones que su par alineado.

5.6 Filtrado de secuencias mitocondriales

Se recuperaron lecturas mitocondriales de 10 individuos (Tabla 4), la mayor profundidad en promedio de lecturas alineadas al genoma mitocondrial correspondiente a muestras de tejido muscular, seguido de la muestra de bulbo piloso y finalmente las muestras con menor profundidad fueron los de cartílago.

Tabla 4. Cuadro resumen de lecturas del genoma mitocondrial recuperadas.

	Código	Total de secuencias	Provenientes de Alineamiento	Propiamente Mapeados	Singletons	Alineados a diferente región (*)
C. porcellus " Criollos"	SA42911	55526	55385 (99.75%)	41772	141	1519
	SA42912	64613	64433 (99.72%)	44962	180	2150
	SA42913	43849	43723 (99.71%)	31690	126	1393
	SA42914	9635	9598 (99.62%)	6772	37	335
C.tschudii	SA42976	115812	115447(99.68%)	70962	365	3363
	SA42977	29190	29056 (99.54%)	17064	134	1362
	SA42978	22910	22753 (99.31%)	16780	157	546
C. porcellus " Mejorados	SA42979	56962	56708 (99.55%)	39242	254	2054
	SA42980	107979	107594 (99.64%)	76940	385	3088
	SA42981	251406	250587 (99.67%)	189676	819	6204

* Lecturas recuperadas del filtrado de la sección 5.4

En los gráficos de cobertura están ausentes lecturas que alinean a la región repetitiva (RR) del genoma mitocondrial de las muestras de *C. porcellus* "Criollos" (Figura 20), a diferencia de los demás alineamientos (Figura 21 y 22). Se aprecia regiones del genoma mitocondrial con una cobertura sobrerrepresentada de secuencias, estas regiones corresponden a: la secuencia del gen de ARN ribosomal 16S, la región de origen de replicación de la cadena ligera (ORL) y la región repetitiva dentro de la región D-Loop. La región correspondiente al gen de ARN ribosomal de 16S con alta profundidad únicamente está ausente el individuo del código SA42977, la región correspondiente al origen de replicación de la cadena ligera en cambio no se encuentra con picos de profundidad pronunciada en los individuos de los códigos: SA42914, SA42976, SA42977, SA42978 y SA42979. Otra región con un pico de profundidad aparente se ubica en el gen NADH deshidrogenasa subunidad 4L, pero esta

solo puede ser apreciada en la mitad de los individuos de códigos: SA42911, SA42912, SA42913, SA42976 y SA42981 (Se señala con flechas, ver figuras 20 al 22).

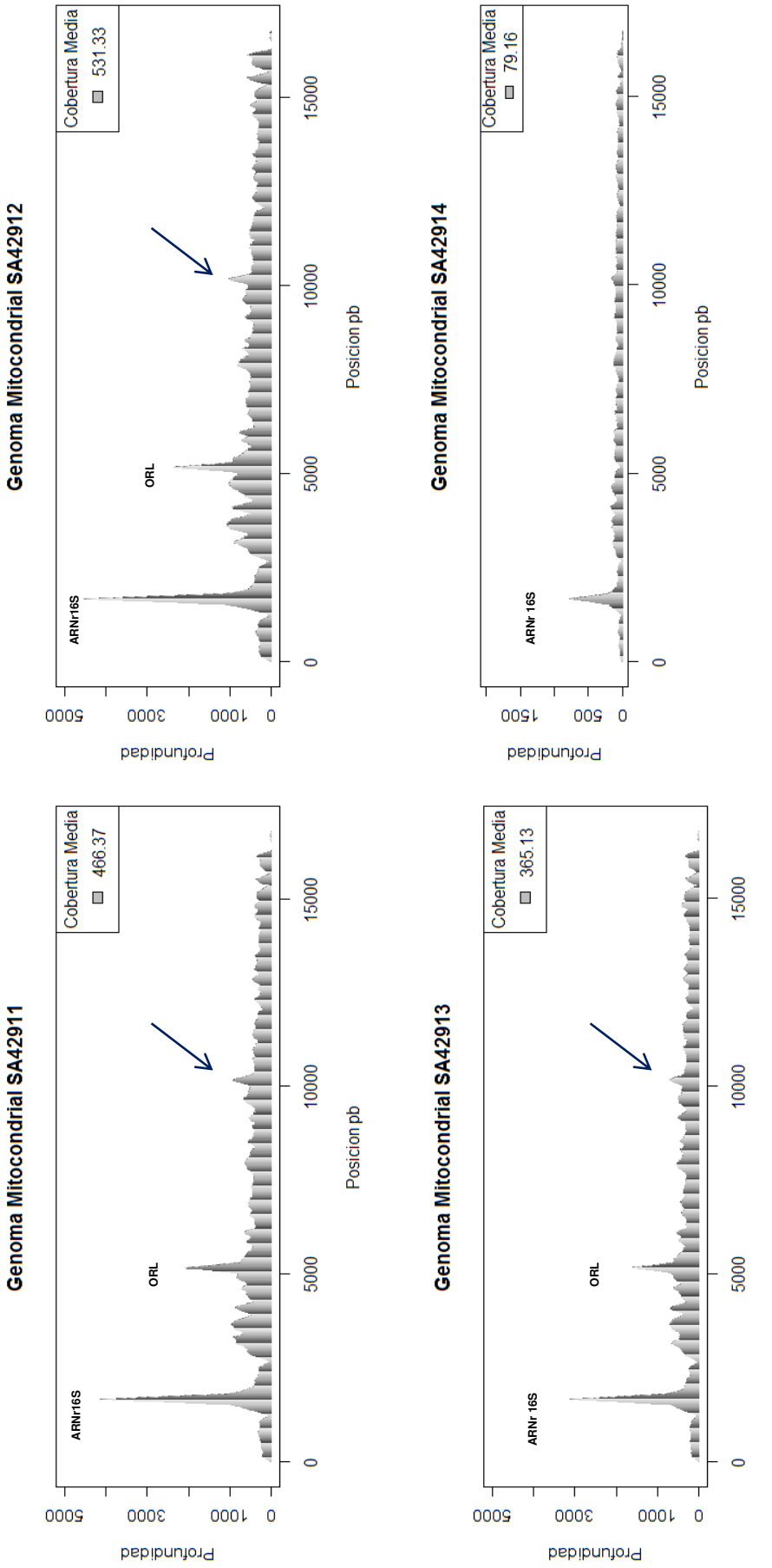


Figura 20. Grafica de cobertura de genoma mitocondrial de *Cavia porcellus* "Criollos" de la Estación Experimental Andahuaylas – Chumbibamba, se indican con siglas ARN ribosomal 16S (**ARNs16s**), origen de replicación de la cadena ligera (**ORL**) y con una flecha se indica la región proxima del gen NADH deshidrogenasa subunidad 4L.

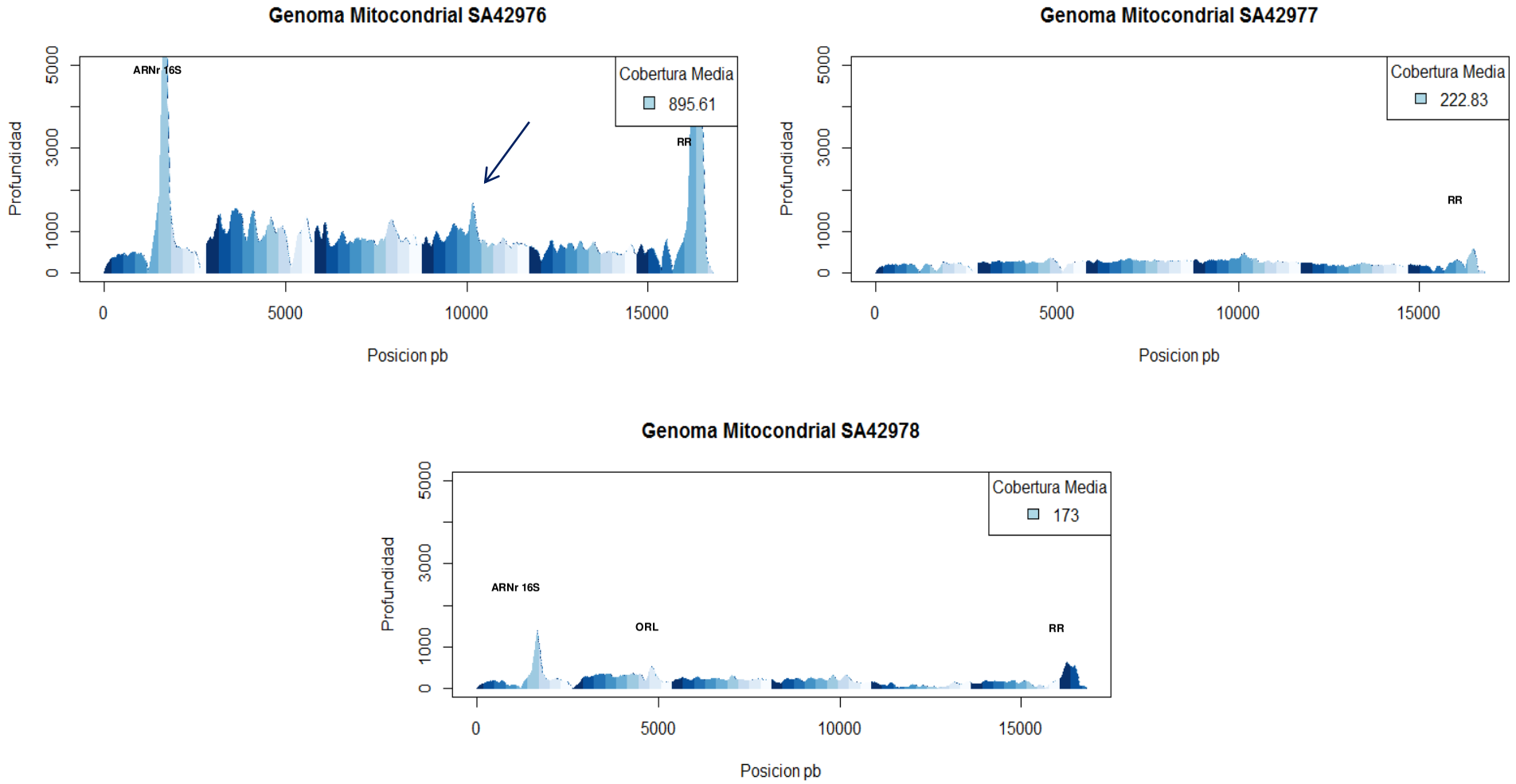


Figura 21. Grafica de cobertura de genoma mitocondrial de *Cavia tschudii* (cuyes silvestres), se indican con siglas ARN ribosomal 16S (**ARNs16s**), origen de replicación de la cadena ligera (**ORL**), region repetitiva de DLP (**RR**) y con una flecha se indica la región proxima del gen NADH deshidrogenasa subunidad 4L.

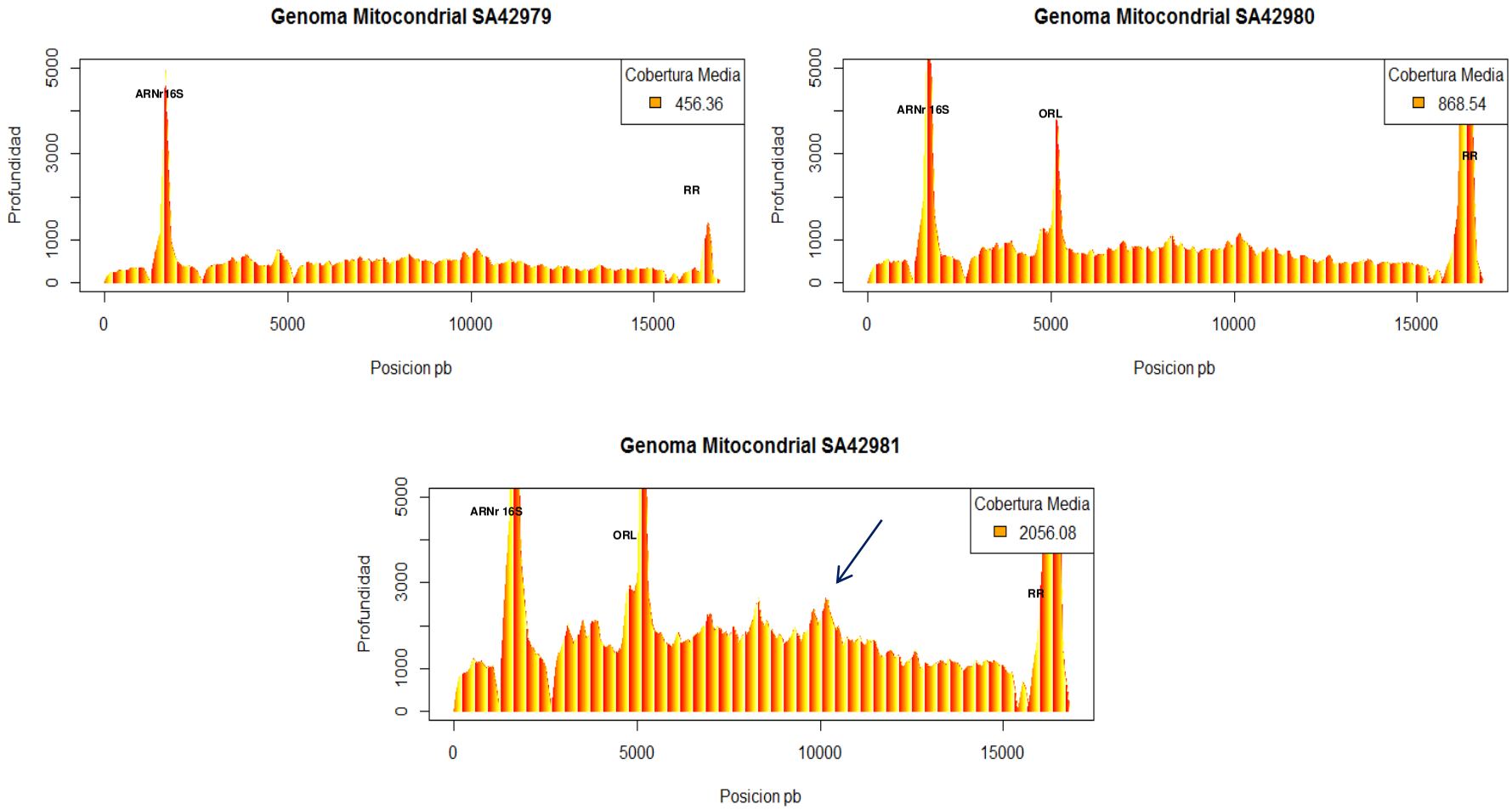


Figura 22. Grafica de cobertura de genoma mitocondrial de *Cavia porcellus* "mejorados" de la Estación Experimental - La Molina. se indican con siglas ARN ribosomal 16S (**ARNs16s**), origen de replicación de la cadena ligera (**ORL**), region repetitiva de DLP (**RR**) y con una flecha se indica la región proxima del gen NADH deshidrogenasa subunidad 4L.

Contrario a las regiones sobrerrepresentadas en el alineamiento, observamos que los extremos del genoma sufren una disminución de profundidad, alternativamente regiones que evidencia baja profunidad se ubican en los siguientes genes y regiones mitocondriales: Gen de ARN de transferencia de Valina (ARNt-Val), Gen de ARN de transferencia de Leucina (ARNt-Leu) y la region D-Loop (regiones flanqueantes a la repetición).

5.7 Ensamblaje de genoma mitocondrial

5.7.1 Ensamblaje de genoma mitocondrial con SPAdes.

Se obtuvo ensamblajes de 10 genomas mitocondriales no circularizados, sin resolución de la región repetitiva de la región D-Loop (recuadro de la Figura 23), obteniendo 2 *contigs* como máximo de ensamblaje, el *contig* de mayor longitud fue de 16747 pb (SA42979) y de menor longitud fue de 681pb (SA42914). Y se obtuvieron mejores ensamblajes después de filtrar subconjuntos de lecturas en las muestras SA42979 y SA42981.

Tabla 5. Cuadro resumen de ensamblaje genoma mitocondrial.

	Código	Tamaño total de ensamblaje (pb)	Numero de *Contigs*	*Contig* más Largo (pb)	*Contig* más Corto (pb)	Total de Lecturas (*)
C. porcellus " Criollos"	SA42911	16533	1	16533	-	42054
	SA42912	16335	1	16335	-	45322
	SA42913	16432	1	16432	-	31942
	SA42914	16335	2	15696	681	21196
C.tschudii	SA42976	16393	1	16393	-	71692
	SA42977	16541	2	15265	1276	17332
	SA42978	16575	2	11915	4660	17094
C. porcellus " Mejorados	SA42979	16747	1	16747	-	9000(**)
	SA42980	16414	1	16414	-	77710
	SA42981	16402	1	16402	-	20000(**)

* Secuencias filtradas en la sección 5.6

** Subconjunto de las secuencias filtradas de la sección 5.6

El ensamblaje de genomas mitocondriales presenta regiones no definidas (presencias de bases "N"), estas regiones se observan en el alineamiento con BLAST y corresponden con

las regiones de baja profundidad, del gen de ARNt–Leucina e inicio de la región de D-Loop (Flechas en la Figura 23).

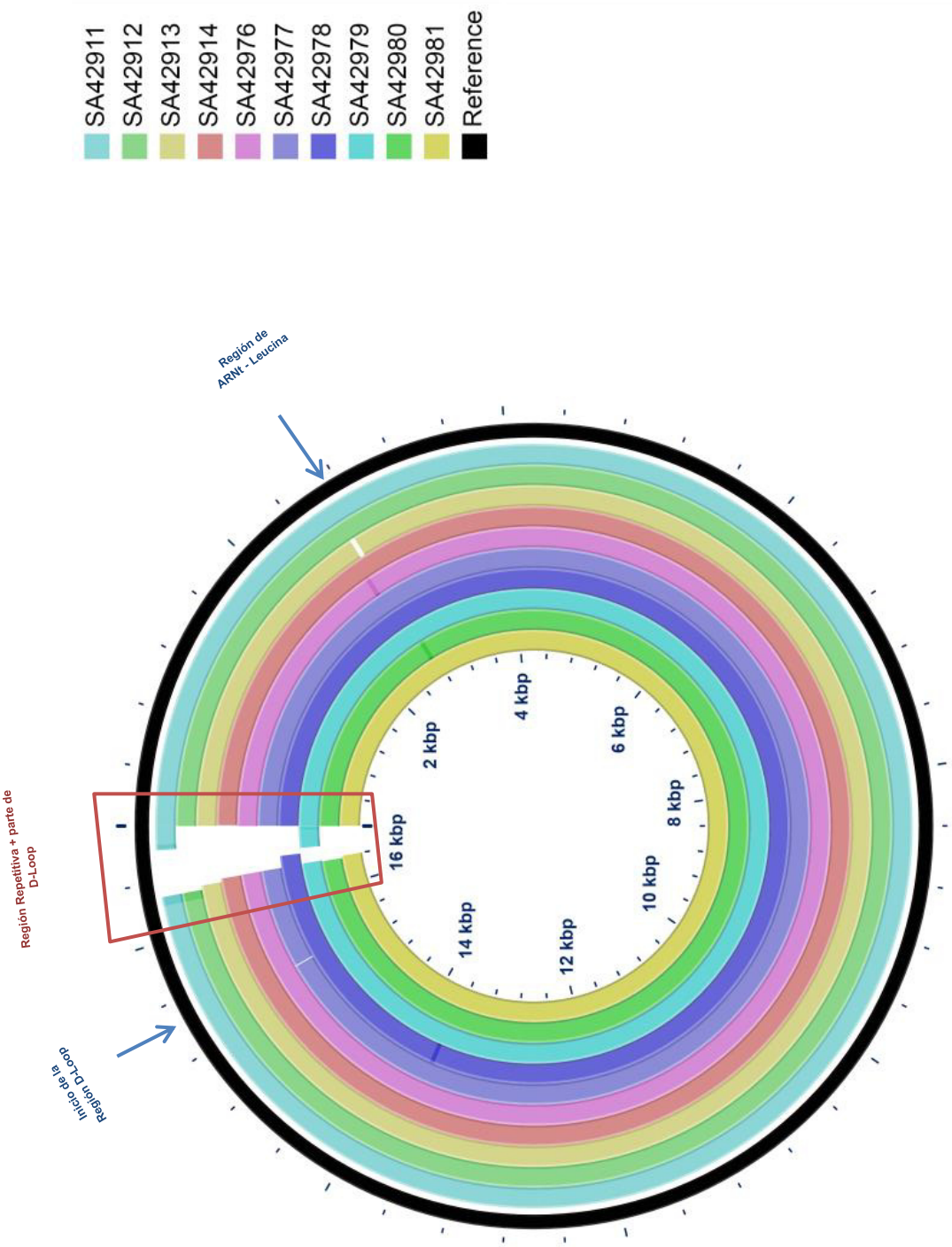


Figura 23. Grafica de BLAST al genoma de referencia de *Cavia porcellus* (NC_000884.1) de los genomas ensamblados con el programa SPAdes.

5.7.2 Ensamblaje de genoma mitocondrial con NOVOplasty

Se obtuvieron 9 genomas mitocondriales no circularizados con 2 *contigs* por individuo con un tamaño total de ensamblaje mayor al genoma de referencia (16801 pb) en su mayoría, con la excepción de los genomas SA42978, SA42980 y SA42981 (Figura 24).

Tabla 6. Cuadro resumen de ensamblaje de genoma mitocondrial.

	Código	Tamaño total de ensamblaje (pb)	Numero de *Scaffolds*	*Contig* más Largo (pb)	*Contigs* más Corto (pb)	Total de Lecturas (*)
C. porcellus " Criollos"	SA42911	16826	2	12223	4603	373288
	SA42912	16810	2	12195	4615	412958
	SA42913	17025	2	10885	6140	117590
	SA42914	17763	1	17763	-	61328
C.tschudii	SA42976	16806	2	12402	4404	584572
	SA42977	16855	2	12193	4662	242694
	SA42978	16710	2	12533	4177	123218
C. porcellus " Mejorados	SA42979	16919	2	11919	5000	444176
	SA42980	16705	2	12181	4624	688352
	SA42981	16789	2	12369	4420	1650278

* Secuencias filtradas al alineamiento de la secuencia semilla de CytB

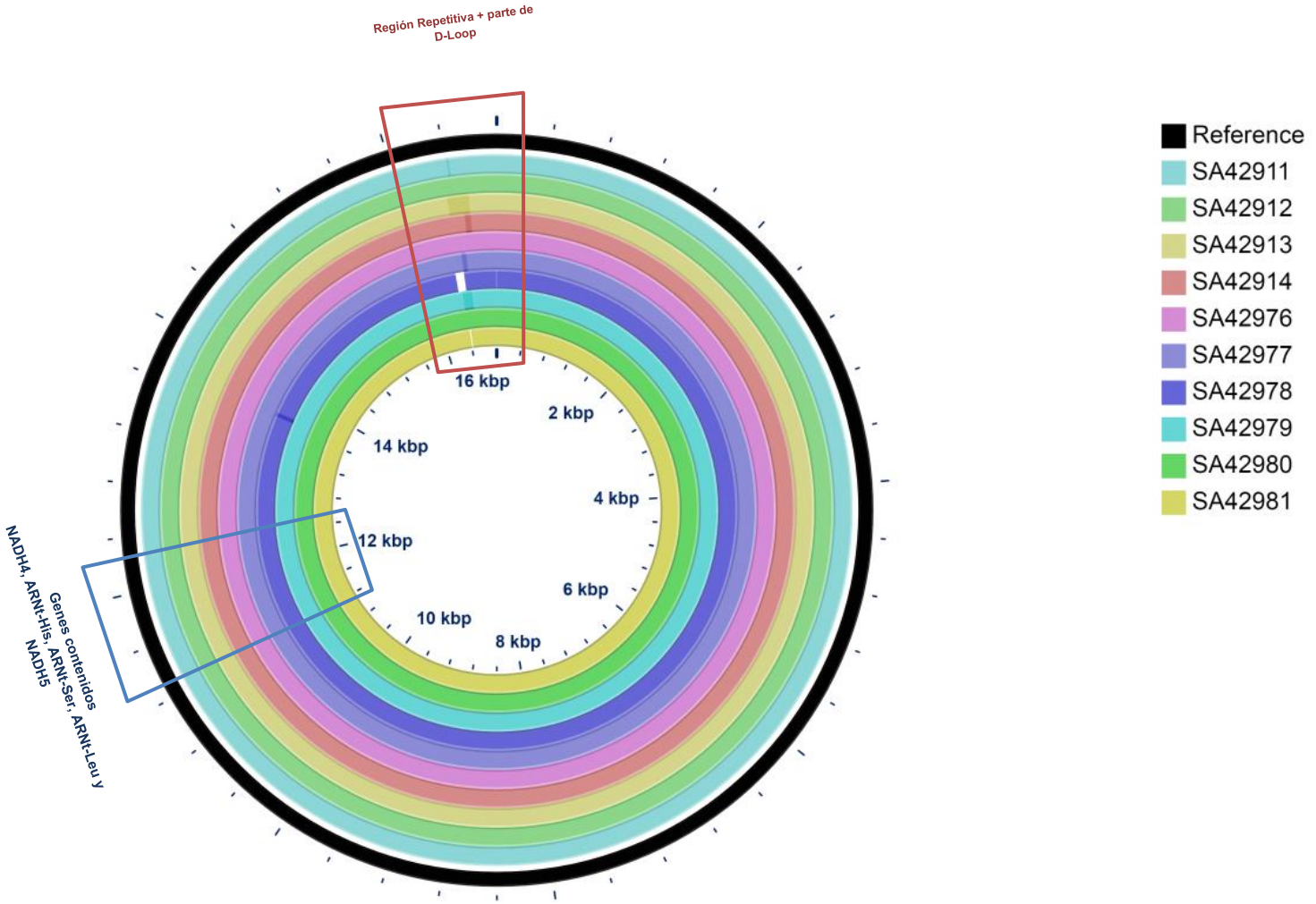


Figura 24. Grafica de BLAST al genoma de referencia de *Cavia porcellus* (NC_000884.1) de los genomas ensamblados con el programa NOVOplasty.

Adicionalmente se obtuvo un genoma circularizado correspondiente al individuo SA42914, con el *contig* de mayor tamaño (17763 pb). El *contig* de menor longitud fue de 4177 pb (SA42978) y se observa que los *contigs* de mayor tamaño de los 9 genomas no circularizados fueron alienados al genoma de referencia entre la posición 1 y el rango de: 10460 - 12158 del genoma mitocondrial de referencia, que contiene a los genes NADH4, 3 genes de ARNt (His, Ser y Leu) y NADH5).

5.8 Validación de ensamblaje de genomas

Se realizó la evaluación de los ensamblajes para cada metodología, en ambos ensamblajes se obtuvieron buenas métricas de contigüidad y coherencia, existiendo diferencia en las métricas de Integridad (*Completeness*) y la resolución de la región repetitiva (Tabla 7 y 8).

Validación de ensamblaje de genomas del programa SPAdes.

Tabla 7. Cuadro resumen de calidad de secuenciación de genoma completo con el programa SPAdes.

	Código	Métricas de contigüidad				Métricas de coherencia					Métricas de Integridad		
		Fracción del genoma (%)	N50	NG50	LG50	Tamaño total de alineamiento (pb)	Alineamiento (%)	Propiamente pareados	*Singletons* (%)	Cobertura ≥ 1 (%)	CDS Obs/Esp	ARNt Obs/Esp	ARNr Obs/Esp
EEA - Chumbibamba	SA42911	96.851	16533	16533	1	16285	99.92	99.22	0.08	99.94	13/13	22/22	2/2
	SA42912	96.851	16335	16335	1	16285	99.91	99.07	0.08	99.76	13/13	22/22	2/2
	SA42913	96.839	16432	16432	1	16363	99.89	98.93	0.09	99.88	13/13	**21/22 ***	2/2
	SA42914	97.268	15696	15696	1	16361	99.73	99.06	0.27	99.88	13/13	22/22	2/2
Silvestres	SA42976	97.399	16393	16393	1	16393	99.7	98.27	0.29	99.94	13/13	**21/22 ***	2/2
	SA42977	97.393	15265	15265	1	16502	99.68	98.17	0.28	99.88	13/13	22/22	2/2
	SA42978	98.048	11915	11915	1	16575	99.63	98.92	0.34	99.88	13/13	22/22	2/2
EE-La Molina	SA42979	98.81	16747	16747	1	16606	99.78	97.63	0.22	99.3	13/13	22/22	2/2
	SA42980	97.399	16414	16414	1	16398	99.5	98.05	0.47	99.85	13/13	22/22	2/2
	SA42981	97.56	16402	16402	1	16402	99.48	98.08	0.52	99.94	13/13	22/22	2/2

* Ausencia del gen de *ARNt-Leucina*

Los ensamblajes en promedio de las secuencias representan más del 96 % de genoma de referencia, se observa ausencia de ensamblajes pequeños (parámetro de N50/NG50 y L50), las secuencias mitocondriales presentaron entre 1 a 2 *contigs*. Las métricas de coherencia resultaron en un alto porcentaje de alineamiento de las secuencias al resultado del ensamblaje, superiores en todos los casos al 99.5 %, más de 98 % de secuencias propiamente pareados, bajo porcentaje de secuencias sin pares (*singletons*) y una cobertura mayor a 1 en más del 99% de las bases alineadas. En las métricas de integridad (*Completeness*) con el programa MITOS, se encontraron el 100% de los genes esperados en 8 individuos; los individuos con ausencia de un gen de ARNt-Leucina fueron SA42913 y SA42976. La secuencia de la región D-Loop se encuentra parcialmente ensamblada y en su mayoría la secuencia de la región repetitiva no presenta resolución (Excepción del individuo SA42979).

Validación de ensamblaje de genomas del programa NOVOplasty.

Tabla 8. Cuadro resumen de calidad de secuenciación de genoma completo con el programa NOVOplasty.

	Código	Métricas de contigüidad				Métricas de coherencia					Métricas de Integridad		
		Fracción del genoma (%)	N50	NG50	LG50	Tamaño total de alineamiento (pb)	Alineamiento (%)	Propiamente pareados	*Singletons* (%)	Cobertura ≥ 1 (%)	CDS Obs/Esp	ARNt Obs/Esp	ARNr Obs/Esp
EEA - Chumbibamba	SA42911	100,15	12223	12223	1	16815	98.56	95.06	1.44	100	13/13	22/22	2/2
	SA42912	100,05	12195	12195	1	16799	98.49	94.82	1.51	100	13/13	22/22	2/2
	SA42913	101,33	10885	10885	1	16984	98.43	94.68	1.57	99.98	13/13	22/22	**3/2***
	SA42914	105,73	17763	17763	1	16874	98.94	96.96	1.06	100	13/13	22/22	**3/2***
Silvestres	SA42976	100,03	12402	12402	1	16799	98.48	94.57	0.7	100	13/13	22/22	2/2
	SA42977	100,32	12193	12193	1	16855	98.39	94.51	1.61	100	13/13	22/22	2/2
	SA42978	99,46	12533	12533	1	16702	98.33	94.50	1.67	100	13/13	22/22	2/2
EE- La Molina	SA42979	100,70	11919	11919	1	16915	98.36	94.58	1.64	100	13/13	22/22	2/2
	SA42980	99,43	12181	12181	1	16801	98.49	94.79	1.51	100	13/13	22/22	2/2
	SA42981	99,93	12369	12369	1	16785	98.33	94.57	0.89	100	13/13	22/22	2/2

* Duplicación del gen ARNr 12S

Se obtuvieron 7 genomas con tamaño de ensamblaje mayor al genoma de referencia, superando el tamaño de representación del genoma de referencia, no se observaron ensamblajes pequeños (presencia de 1 a 2 *contigs*). Las métricas de coherencia mostraron más un 98 % de lecturas alineadas al resultado del ensamblaje, las secuencias propiamente apareados fueron mayor a 94 %, se observó menos de 1% de secuencias sin pares (*Singletons*) y una cobertura mayor a 1 en cerca del 100% de las secuencias alineadas. Las métricas de coherencia de los ensamblajes indicaron el 100% de genes esperados del genoma mitocondrial, sin embargo, se obtuvo 1 gen adicional de ARNr 12S en dos individuos (SA42913 y SA42976). Se observa además que la región repetitiva en los individuos se encuentra expandida, especialmente en la secuencia del individuo SA42979.

5.9 Ensamblaje de secuencias mitocondriales de la base de datos SRA del NCBI.

Las secuencias descargadas de la base de datos de SRA contiene 91.7 G de bases, se observaron parámetros de buena calidad de secuenciación (calidad *Phred score* y conteo de *k-mers*) y se realizaron filtrado de las secuencias de baja calidad (Figura 25).

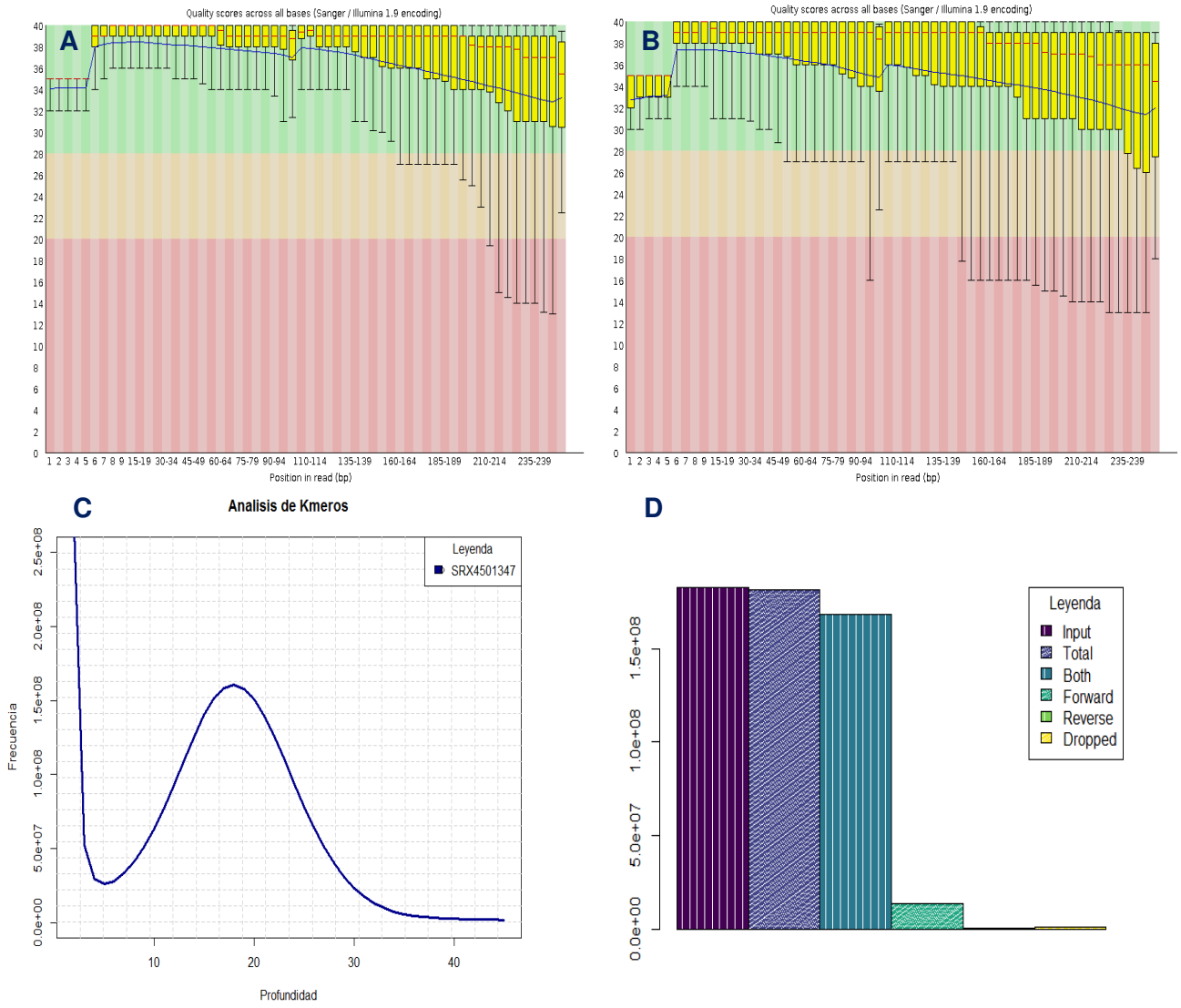


Figura 25. Análisis de calidad de secuenciación y filtrado de secuencias de baja calidad de las lecturas de secuenciación de la base de datos SRA (SRX4501347). **A.** Análisis de calidad de secuencias por cada posición de bases Forward. **B.** Análisis de calidad de secuencias por cada posición de bases Reverse. **C.** Análisis de conteo de *K-mers* de secuenciación de genoma completo. **D.** Grafica de barras del filtrado de secuencias de baja calidad.

Las secuencias filtradas, obtuvieron un alto porcentaje de alineamiento (92.58%) al genoma referencia Cavpor v3.0, y se separaron 14973 lecturas del genoma mitocondrial (Tabla 9).

Tabla 9. Cuadro resumen de calidad de secuenciación de genoma completo (SRX4501347 – SRA).

Alineamiento		Genoma mitocondrial Filtrado	
Input	350544882 (100%)	Total	14973
Alineados	92.58%	Alineados	14898 (99.50%)
Lecturas pares Alineados	96.31%	Propiamente Pareados	14350
Propiamente Pareados	89.67%	*Singletons*	75
Singletons	1.65%	Alineados a diferente región	1
Alineados a un diferente Chr.	0.12%		

Se obtuvo un genoma no circularizado después del ensamblaje, con un total de 2 *contigs* resultantes sin solapamiento (Figura 26). No se logró resolución de la región D-Loop y parte de la región correspondiente al gen de ARNt de Prolina,

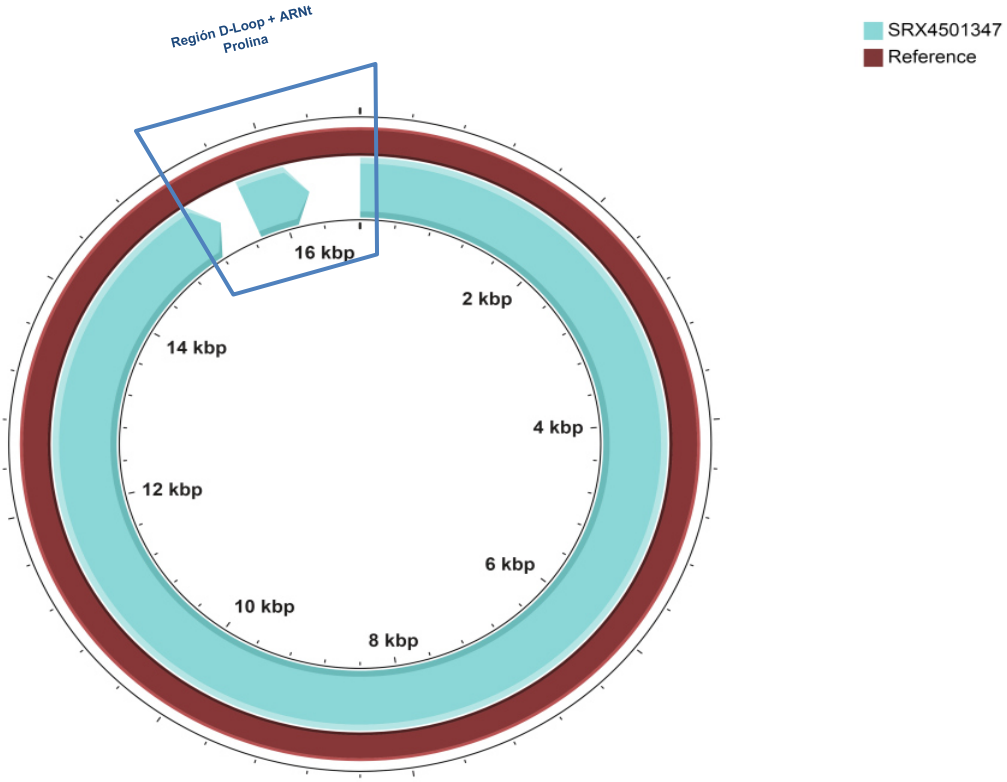


Figura 26. Grafica de BLAST del ensamblaje de SRX4501347 con el programa SPAdes al genoma de referencia de *Cavia porcellus* (NC_000884.1).

El *contig* de mayor tamaño fue de 15 391 pb y el de menor tamaño de 631 pb representando un 95.26 % del genoma de referencia. Las métricas de coherencia indican un 99.62% de alineamiento de las lecturas filtradas.

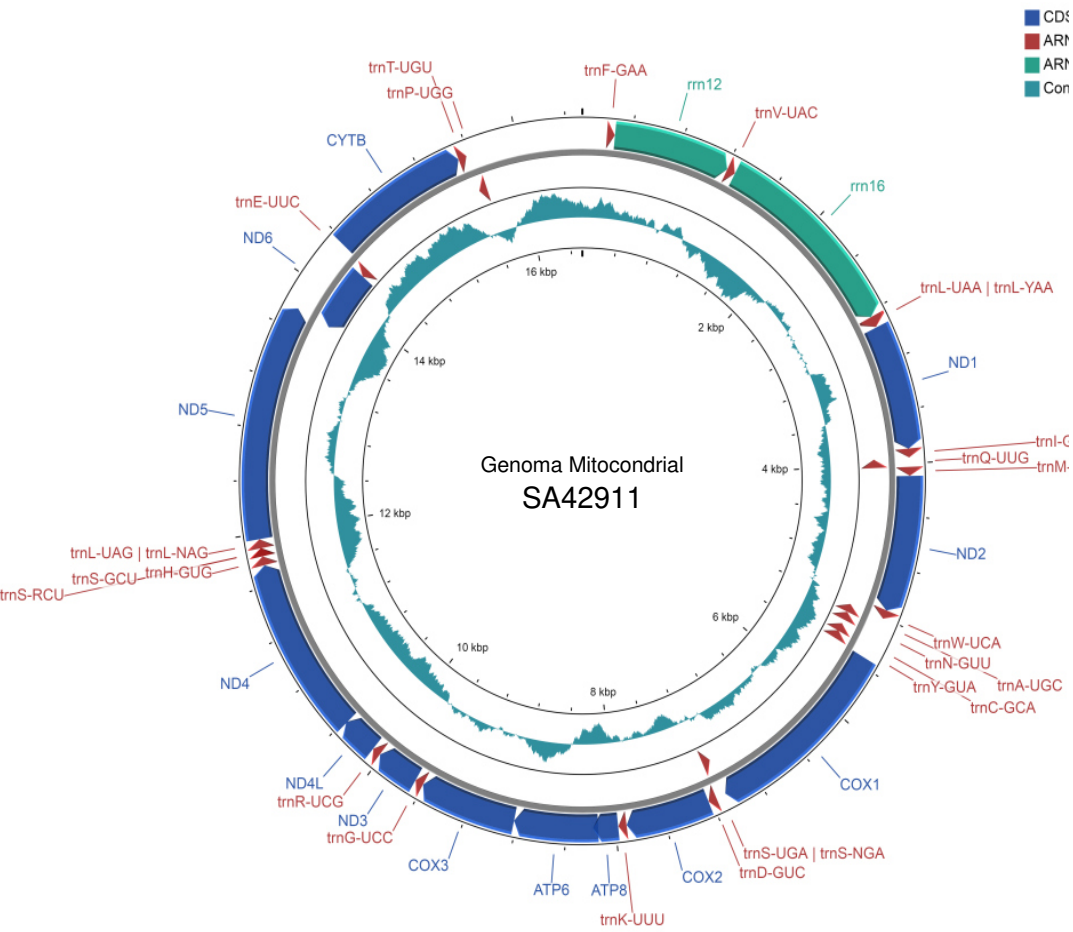
Para las métricas de integridad se encontraron 36 genes de los 37 esperados para el genoma mitocondrial, debido a la ausencia del gen de ARNt de prolina por baja resolución (Tabla 10).

Tabla 10. Cuadro resumen de ensamblaje y validación de ensamblaje del genoma mitocondrial SRX4501347.

Resultados del ensamblaje		**Métricas de Contigüidad**		**Métricas de Coherencia**		**Métricas de Integridad (*Completeness*)**	
Tamaño total de ensamblaje	16022	Fracción del genoma	95.26%	Tamaño total de alineamiento	16014 pb	CDS Obs/Esp	13/13
Numero de *Scaffolds*	2	N50	15391	Alineamiento (%)	99.62 %	ARNt Obs/Esp	21/22*
Contig más Largo	15391	NG50	15391	Propiamente pareados	98.99%	ARNr Obs/Esp	2/2
Contig más Corto	631	LG50	1	*Singletons*	0.38 %	* Ausencia del gen de ARNt de Prolina	
Total de Lecturas	14350			Cobertura ≥ 1 (%)	99.94		

5.10 Anotación de genoma mitocondrial

Se realizó la anotación de 11 genomas mitocondriales: 10 genomas secuenciados para la presente tesis (Figura 27 – 31) y 1 genoma descargado de la base de datos SRA (Figura 32). De los datos cruzados entre la base de datos se observó que únicamente el individuo del código SA42913 carece de un gen de ARNt Leucina.



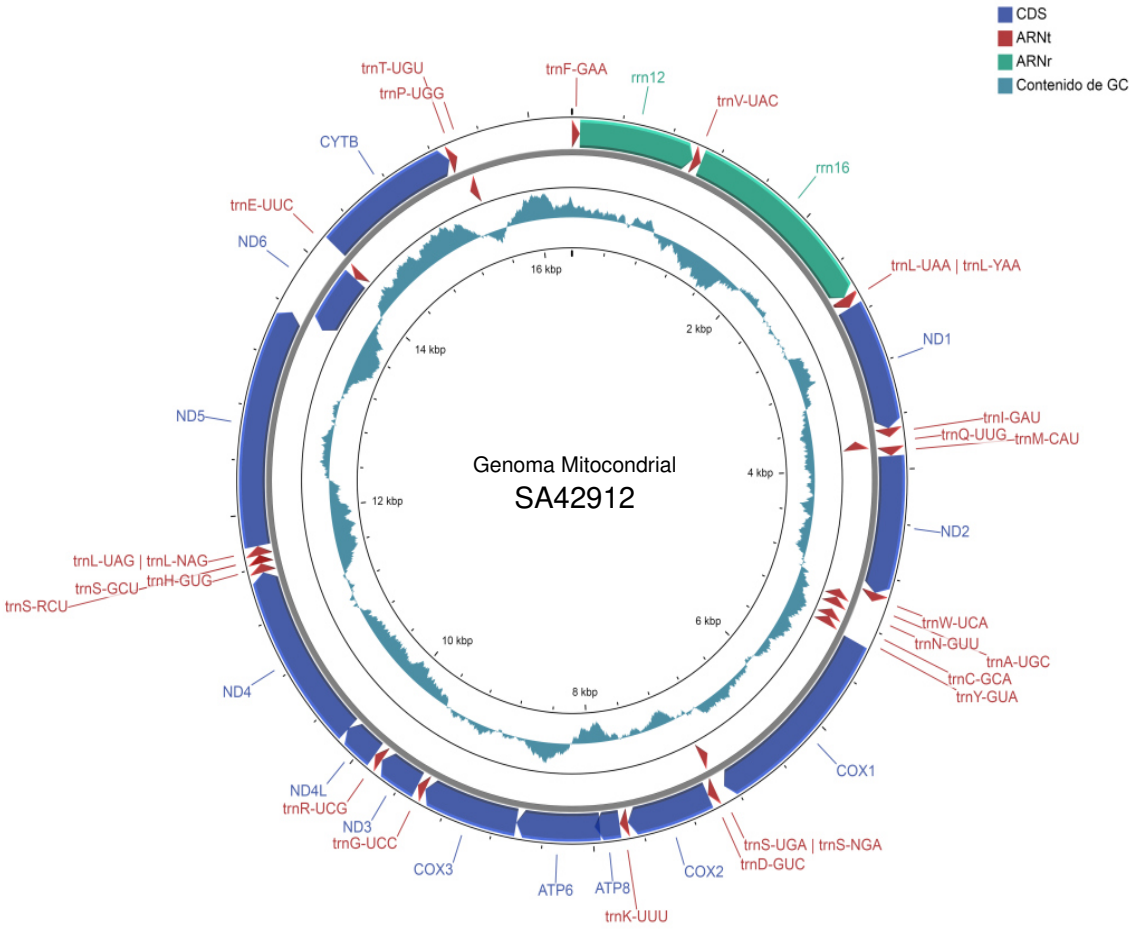


Figura 27. Anotación de genoma mitocondrial de códigos: SA42911 y SA42912. Se codifican por colores genes codificantes de proteínas (13 CDS), genes de ARN ribosomal (2 ARNr) y genes de ARN de transferencia (22 ARNt), se graficaron los genes en sentido horario en el exterior del circulo para genes de la cadena pesada y antihorario en el interior para genes de cadena ligera. Internamente se grafica el contenido de GC, no se observan ausencia de genes.

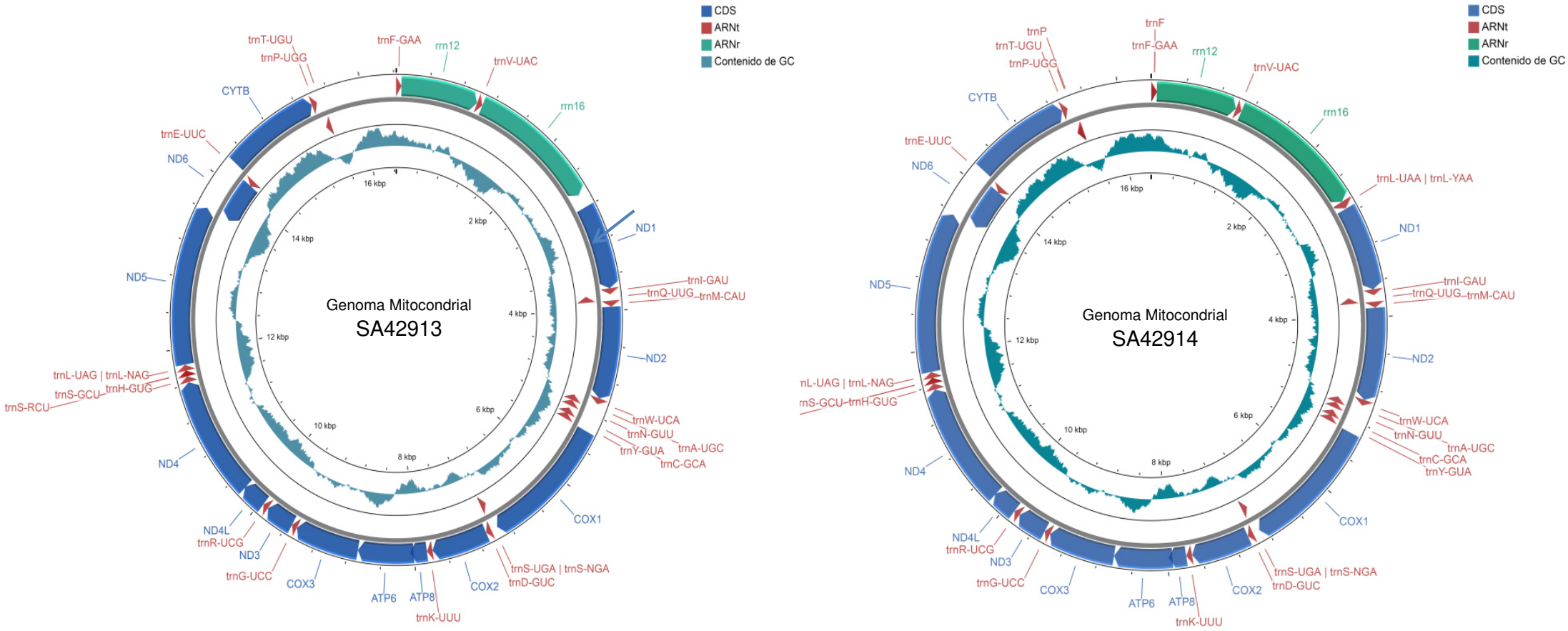


Figura 28. Anotación de genoma mitocondrial de códigos: SA42913 y SA42914. Se codifican por colores genes codificantes de proteínas (13 CDS), genes de ARN ribosomal (2 ARNr) y genes de ARN de transferencia (22/21 ARNt), se graficaron los genes en sentido horario en el exterior del circulo para genes de la cadena pesada y antihorario en el interior para genes de cadena ligera. Internamente se grafica el contenido de GC. Se observa para el genoma de código SA42913 ausencia del gen de ARNt Leucina (Flecha).

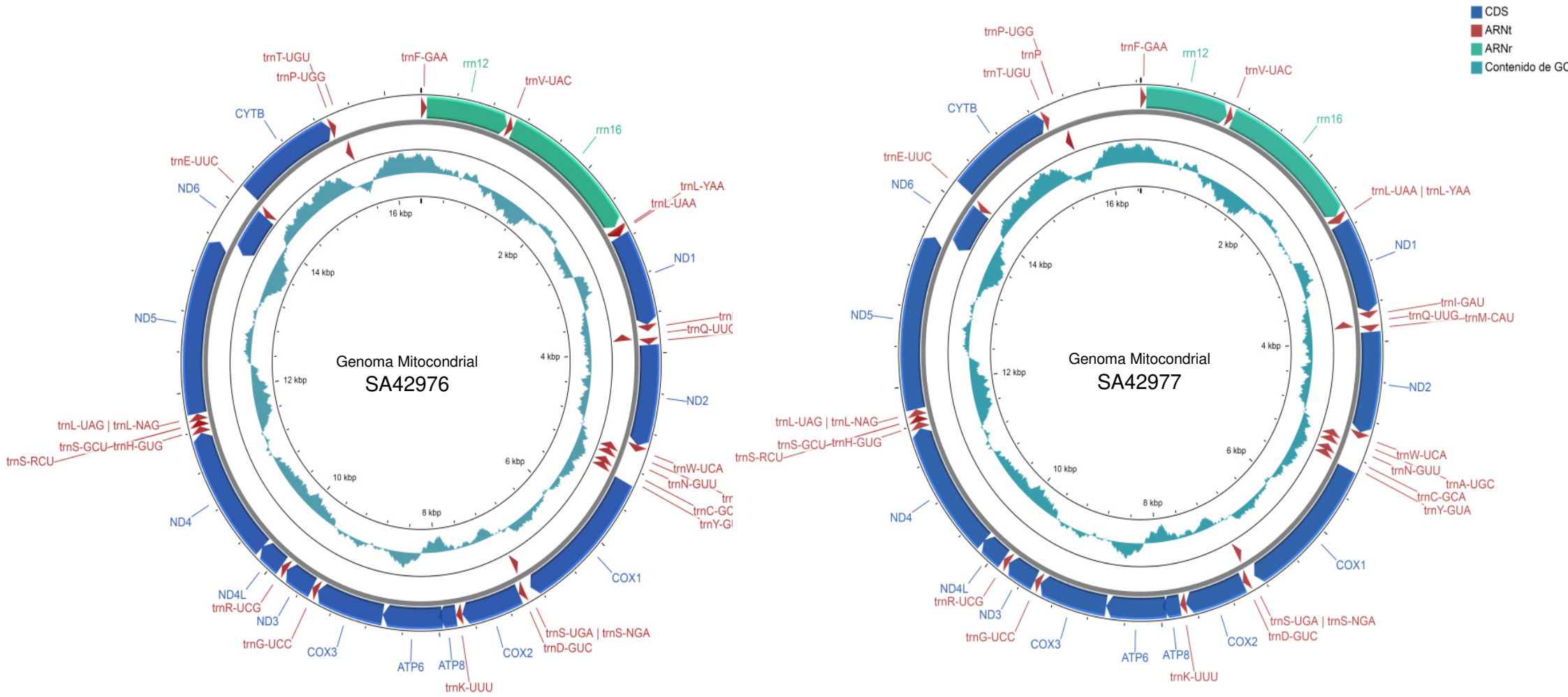


Figura 29. Anotación de genoma mitocondrial códigos: SA42976 y SA42977. Se codifican por colores genes codificantes de proteínas (13 CDS), genes de ARN ribosomal (2 ARNr) y genes de ARN de transferencia (22 ARNt), se graficaron los genes en sentido horario en el exterior del circulo para genes de la cadena pesada y antihorario en el interior para genes de cadena ligera, internamente se grafica el contenido de GC. No se observa ausencia de genes.

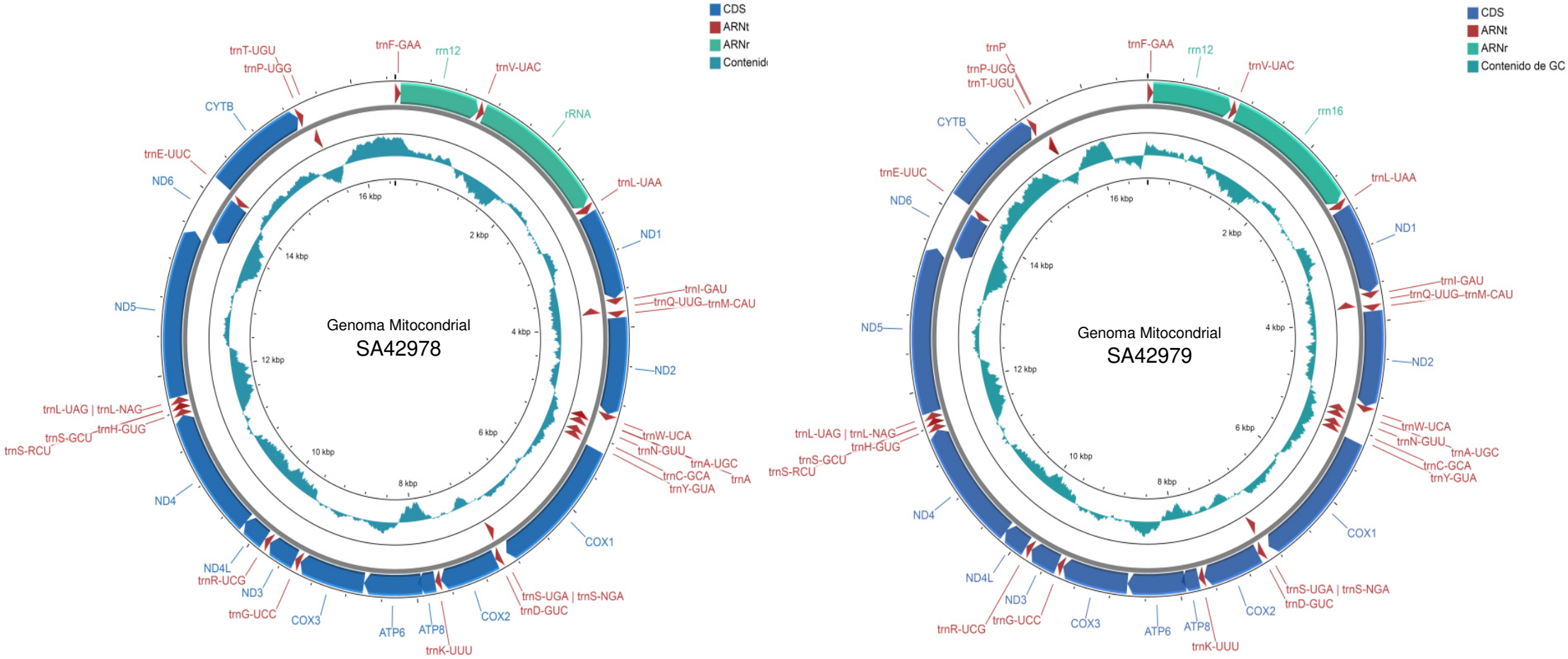


Figura 30. Anotación de genoma mitocondrial códigos: SA42978 y SA42979. Se codifican por colores genes codificantes de proteínas (13 CDS), genes de ARN ribosomal (2 ARNr) y genes de ARN de transferencia (22 ARNt), se graficaron los genes en sentido horario en el exterior del circulo para genes de la cadena pesada y antihorario en el interior para genes de cadena ligera, internamente se grafica el contenido de GC. No se observan ausencia de genes.

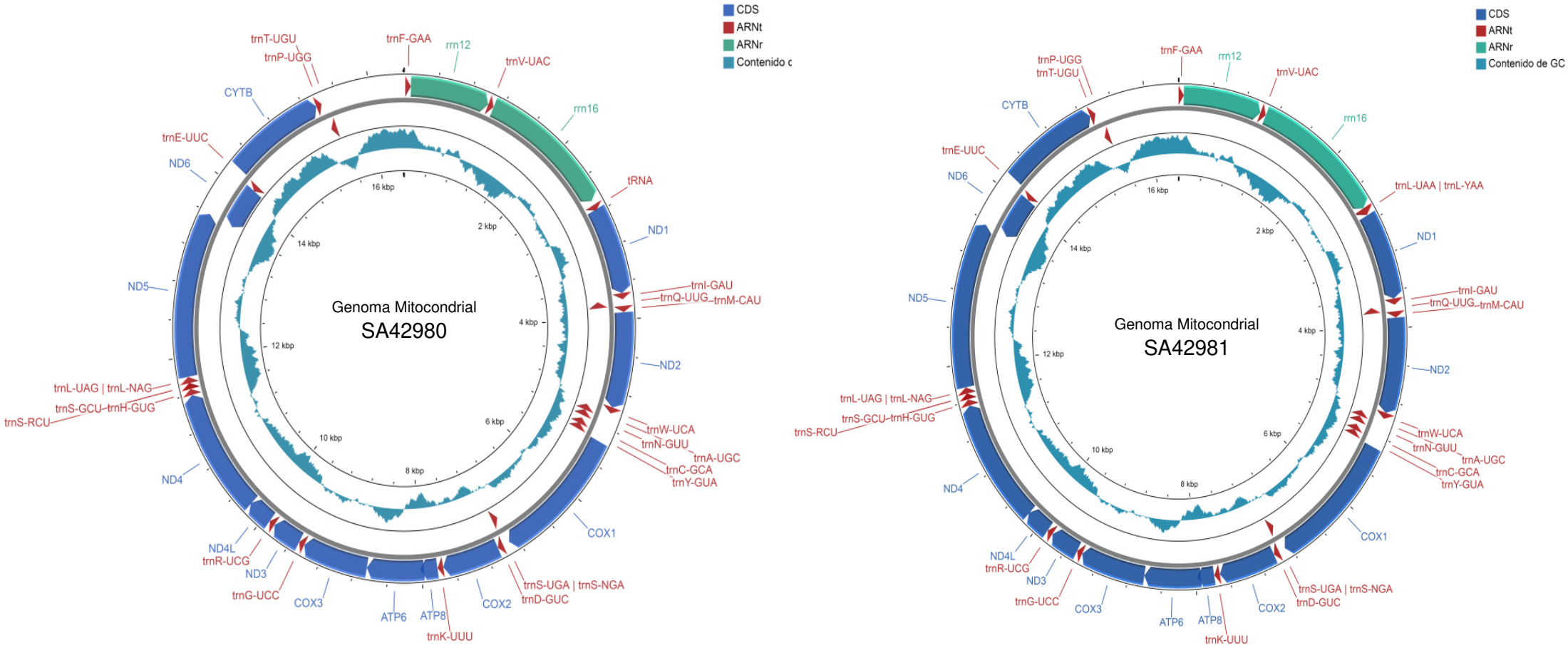


Figura 31. Anotación de genoma mitocondrial códigos: SA42980 y SA42981. Se codifican por colores genes codificantes de proteínas (13 CDS), genes de ARN ribosomal (2 ARNr) y genes de ARN de transferencia (22 ARNt), se graficaron los genes en sentido horario en el exterior del circulo para genes de la cadena pesada y antihorario en el interior para genes de cadena ligera. Internamente se grafica el contenido de GC. No se observa ausencia de genes.

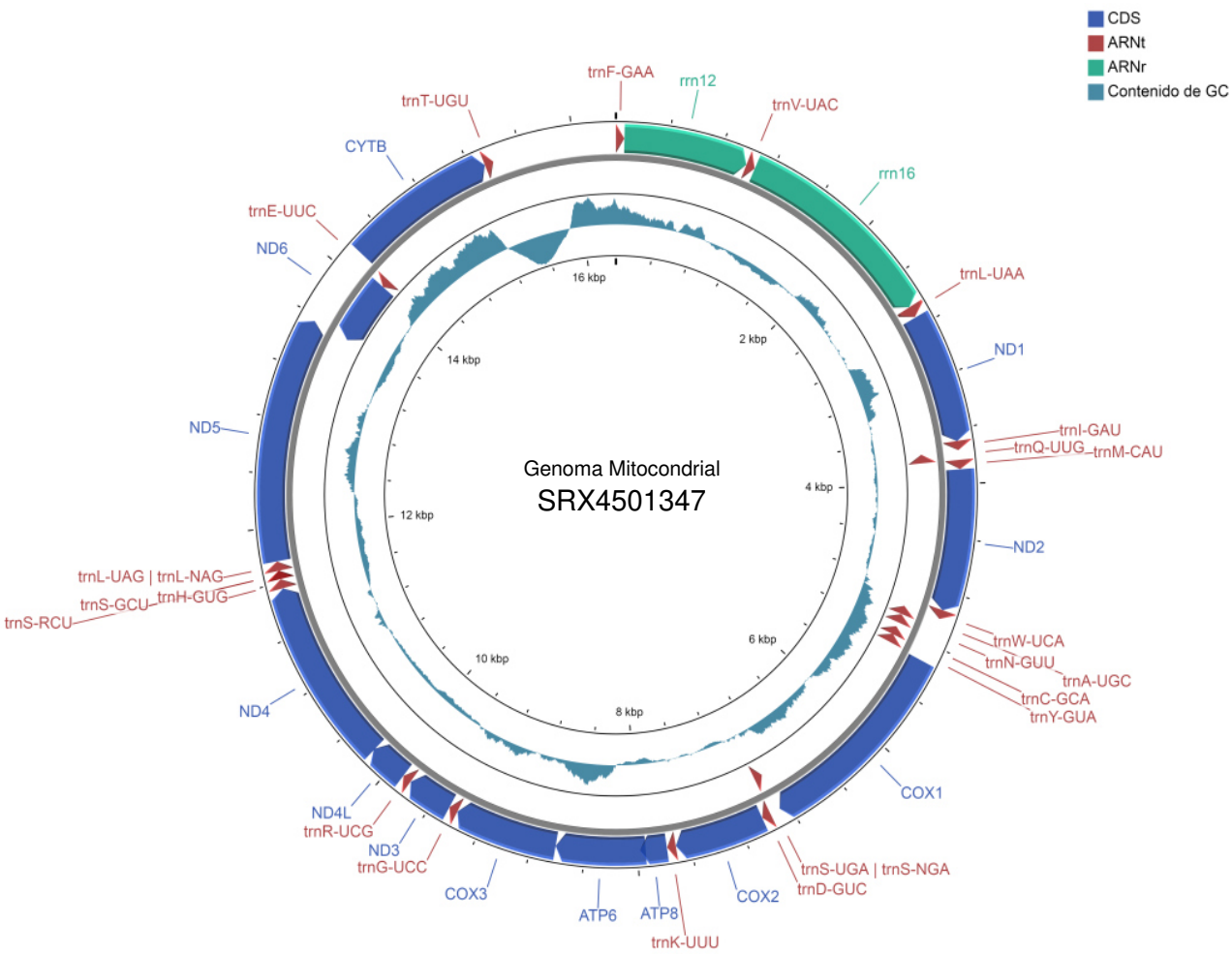


Figura 32. Anotación de genoma mitocondrial de la base de datos SRA. Se codifican por colores genes codificantes de proteínas (13 CDS), genes de ARN ribosomal (2 ARNr) y genes de ARN de transferencia (22 ARNt), no se observa ausencia de genes y se mantiene las mismas características de las figuras anteriores.

5.11 Distancias genéticas de mitogenomas.

Para los análisis de distancias se consideraron los ensamblajes por SPAdes, obviando las regiones de baja resolución de ensamblaje (región DLP - Figura 33A).

Los análisis de distancias genéticas muestran que a *C. tschudii* de Pachachaca como el individuo más distante genéticamente, en cambio las secuencias del genoma mitocondrial de *C. tschudii* de Pantanos de Villa son más cercanos a los *C. porcellus* “Criollos” y mejorados; estos últimos se separan en dos grupos de *C. porcellus*, constituido por los individuos “Criollos de EEA Chumbibamba” y mejorados (PERU). (Figura 33B).

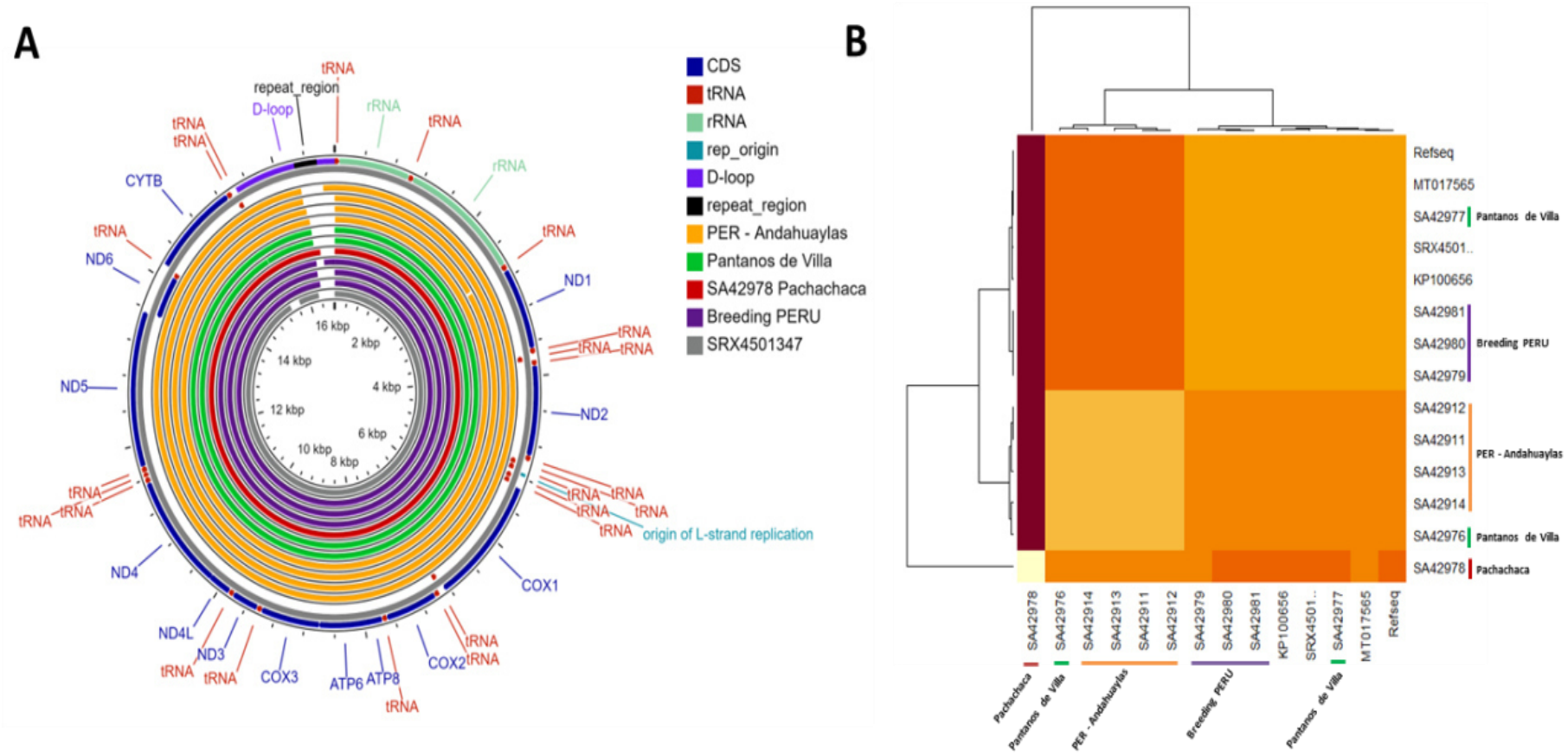


Figura 33. **A)** Comparación de ensamblajes de genomas mitocondriales con el genoma de referencia NC_000884.1. **B)** Heatmap de distancias genéticas (K80) de los genomas ensamblados.

5.12 Filogenómica y análisis demográficos - haplotipicos.

Los análisis filogenéticos revelan un clado monofilético que incluye a *C. porcellus* y *C. tschudii*. La muestra de *C. tschudii* proveniente de Pachachaca forma un clado basal al grupo de cuyes "criollos", mejorados y *C. tschudii* del refugio silvestre de Pantanos de Villa (Figura 34). Los análisis de máxima verosimilitud (Figura 34B) difiere en topología de los análisis bayesiano y de máxima parsimonia, sin embargo se puede apreciar el bajo soporte *bootstrap* que tienen las ramas en estos últimos análisis (Figura 34A). En las 3 filogenias se observa que el grupo que contiene cuyes "criollos" de EEA Chumbibamba - Andahuaylas se agrupa en un clado distinto a los cuyes mejorados "PERU" y el mitogenoma de *C. tschudii* reconstruido de datos públicos (NCBI: SRX4501347) está más emparentado a *C. tschudii* de Pantanos de villa que a los provenientes de Pachachaca. El análisis de *bayesian skyline plot* indica una disminución reciente del tamaño efectivo poblacional del linaje matrilineal, seguido de un reciente aumento (Figura 35).

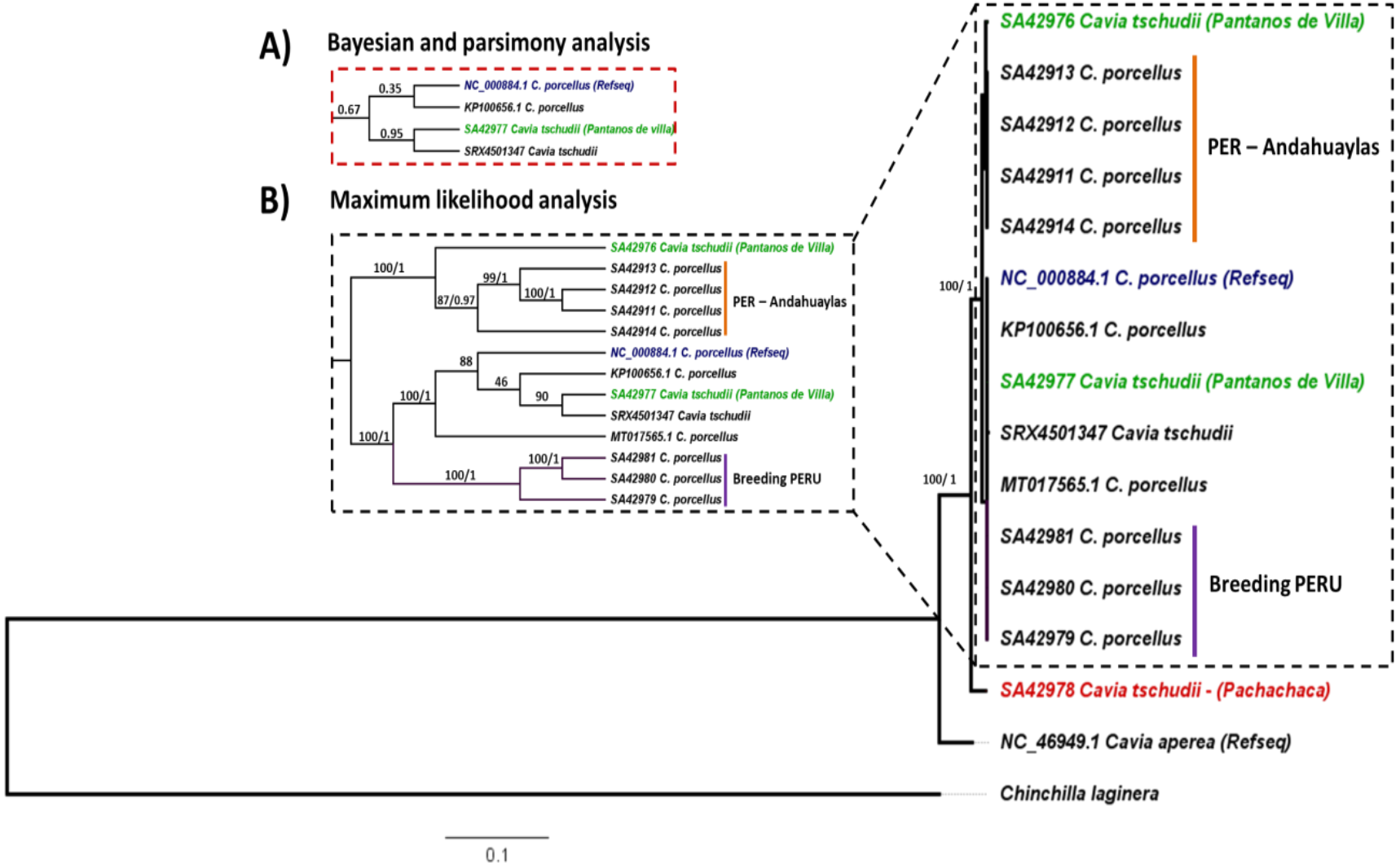


Figura 34. Reconstrucción filogenética a partir de mitogenomas ensamblados y disponibles en la base de datos NCBI. Las etiquetas de las ramas indican valores de *bootstrap* de 1000 replicaciones de análisis de máxima verosimilitud con RAxML-HPC v 8.2.11 y probabilidad posterior de análisis bayesianos con Beast v1.10.4 **A)** Análisis bayesiano y de máxima parsimonia con diferente topología **B)** Análisis de máxima verosimilitud de las secuencias de *C. porcellus* y *C. tschudii* mezclados.

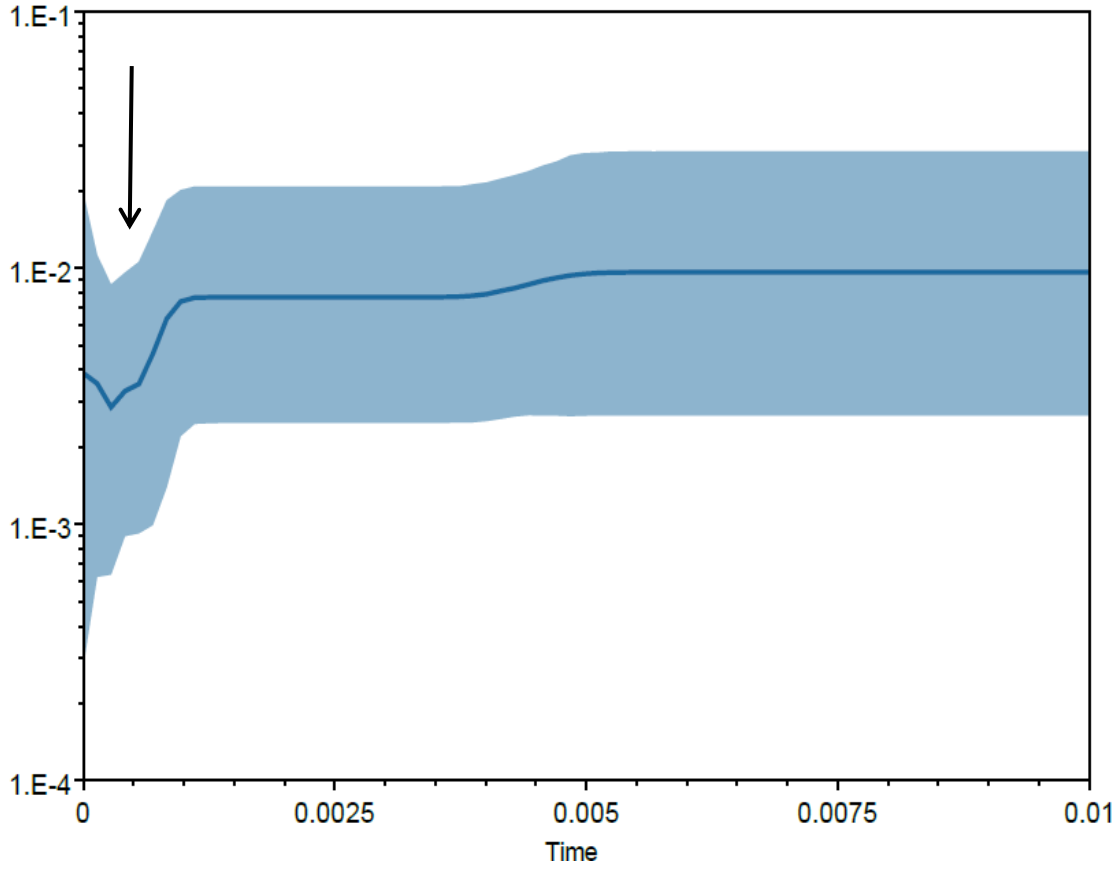


Figura 35. Análisis de coalescencia por *bayesian skyline* a partir de mitogenomas de *C. porcellus* y *C. tschudii*. Se indica la región de disminución del tamaño efectivo poblacional con una flecha.

Los análisis de haplotipos indican que los individuos de la raza "Perú" (diferenciados por 4 pasos mutacionales) están estrechamente relacionados agrupando incluso dos individuos en el haplotipo 9 (Hap_9). Los individuos de *C. tschudii* (Pantanos de villa), no parecen estar relacionados entre si directamente de modo que nos lleve a pensar que es una población aislada. Los C. porcellus "criollos", aunque presentan mayor cantidad de pasos mutacionales a diferencia de C. porcellus mejorados, se diferencian de los demás. Y algo que queda evidente es que el mitogenoma de *C. tschudii* proveniente de Pachachaca es el más distante de los haplotipos (Figura 36).

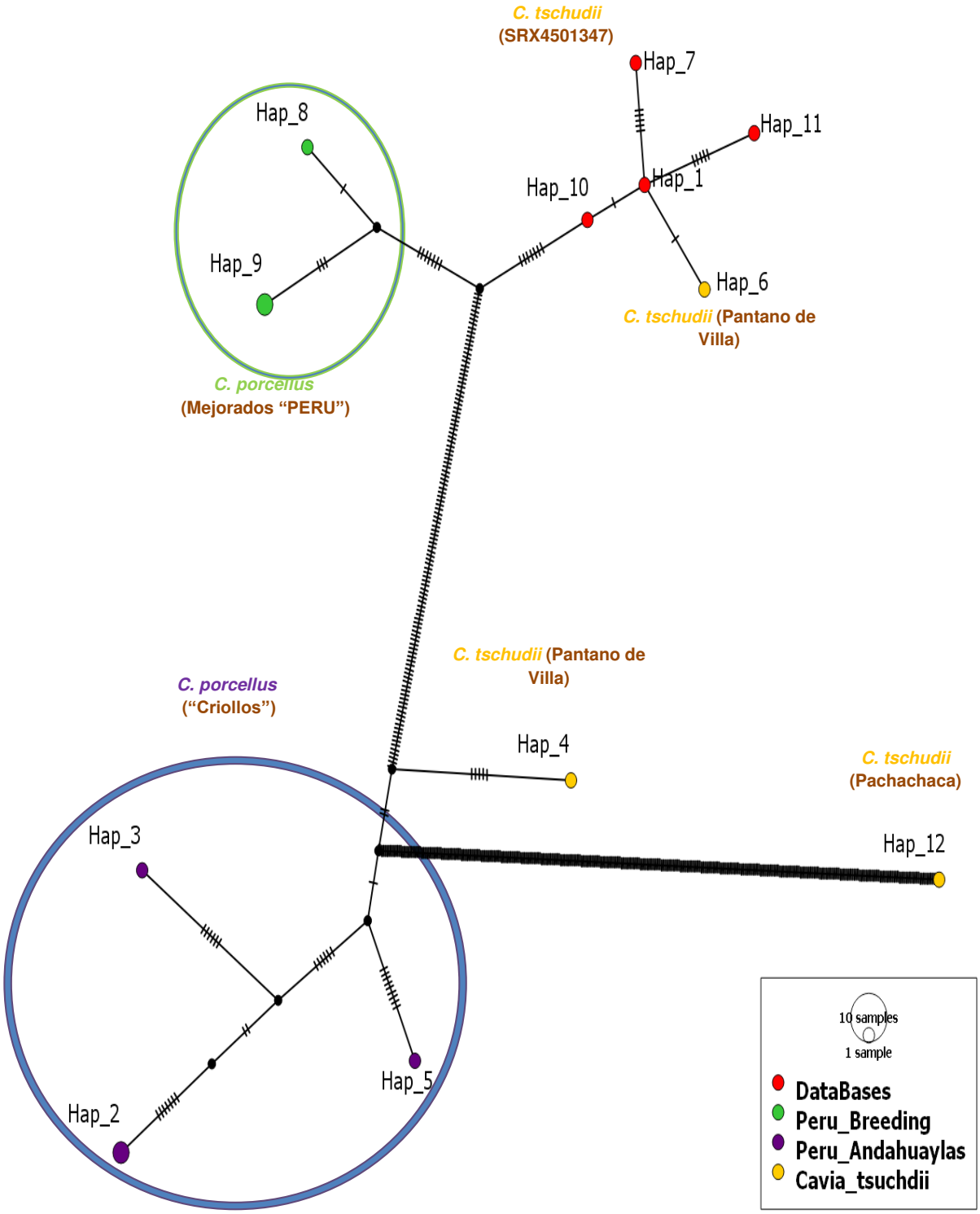


Figura 36. Median-joining network de haplotipos de *Cavia porcellus* y *C. tschudii* a partir de mitogenomas. Se aprecian agrupados los *C. porcellus* "criollos" y "mejorados", a diferencia de los *C. tschudii* y secuencias de la base de datos.

6. DISCUSIÓN

El genoma mitocondrial de mamíferos se caracteriza por su forma circular y tamaño (~16 kb), presentando una configuración y estructura genética conservada (Lavrov y Pett, 2016). La secuenciación del genoma mitocondrial es realizada actualmente a través de tecnologías NGS y debido al tamaño del genoma los trabajos son relativamente económicos, de muestreo significativo, y poco laborioso debido que se puede producir sin la necesidad de la separación del ADN nuclear (Smith, 2016; Marande y Burger, 2007). Debido a estas características no es sorpresa que las publicaciones de las secuencias mitocondriales y plastidiales sean numerosas y en aumento en las bases de datos públicas (Smith, 2016). Si realizamos una búsqueda en el NCBI con las palabras claves "*mitochondrial genome*" encontramos más de 100 mil resultados de más de 30 mil especies diferentes (Forni *et al.,* 2019). Esta gran cantidad de información disponible, además de su clara ortología convierte a estos marcadores como los más utilizados para estudios filogenéticos y sistemáticos (Ma *et al.,* 2012; Witt *et al.,* 2015; Smith, 2016; Sullivan *et al.,* 2017; Cabañas *et al.,* 2020), resolviendo incluso filogenias reportadas anteriormente con marcadores mitocondriales más pequeños (Williams *et al.,* 2014).

De los genomas mitocondriales reportados, los mitogenomas de mamíferos son los más numerosos, comprendiendo aproximadamente el 60% de las secuencias depositadas en el NCBI. Sin embargo, a pesar de que *C. porcellus* es usado como modelo biológico para diversas investigaciones en mamíferos (principalmente en humanos), no se cuentan con muchas secuencias de su mitogenoma. Cabe resaltar tambien que existen estudios que han secuenciado genomas mitocondriales antiguos (Lord *et al.,* 2020; Lord *et al.,* 2018), pero estos no se encuentran disponibles, por otra parte, dentro de las especies relacionadas a *C. porcellus* solo encontramos secuencias de *C. aperea* (Cao y Xia, 2016) y no se encuentran secuencias de mitogenomas de *C. tschudii*, a pesar que según estudios en esta especie, sean probablemente los parientes silvestres más cercanos a *C. porcellus* domésticos (Dunnum & Salazar-Bravo, 2010).

Es por ello que en este estudio se reporta por primera vez secuencias de mitogenomas de *C. tschudii*, además de secuencias de *C. porcellus* domésticos (“criollos”) y mejorados (Avilés *et al.,* 2015; Chauca F., 1997) difiriendo de las secuencias de *C. porcellus* de laboratorio reportados anteriormente.

Una justificación acorde al aporte de este trabajo, parte del hecho que los cuyes que se consumen en la actualidad en varios países, son de crianza doméstica o provienen de programas de mejoramiento. Una de las líneas mejoradas más demandadas son los denominados “peruanos”, que se ha exportado a diferentes países, y como consecuencia han generado un desplazamiento de la variabilidad genética de especies nativas (Burgos-Paz *et al.,* 2011; Solarte *et al.,* 2007).

En vista de una posible pérdida de diversidad y sumado a esto una inexistente estructuración (Avilés *et al.,* 2014), en este trabajo se reporta secuencias del genoma mitocondrial de *C. porcellus* mejorados que corresponden a machos productores del programa de mejoramiento de cuyes de la raza “Perú” de Instituto Nacional de Innovación Agraria (INIA), así como secuencias de cuyes domésticos “criollos” y silvestres, de modo que sirva de referencia para estudios posteriores de diversidad de esta especie.

6.1. Secuenciación y ensamblaje de genoma mitocondrial

El secuenciamiento de genoma completo (WGS) incluye además del genoma nuclear un alto número de secuencias extranucleares, ya sean mitogenomas (Jex *et al.,* 2010) o genomas de cloroplastos (Dierckxsens *et al.,* 2016). A partir de WGS de adecuada calidad se pueden reconstruir secuencias mitocondriales, que pueden ser altamente eficientes si la especie en cuestión presenta un genoma de referencia (Hahn *et al.,* 2013; Meng *et al.,* 2019). Y de no ser organismos modelos, otras metodologías también permiten rescatar secuencias mitocondriales de manera eficiente (Dierckxsens *et al.,* 2016; Al-Nakeeb *et al.* 2017).

Para este estudio se trabajó con muestras de diferentes tejidos (Tabla 2), que varía aparentemente la cantidad de secuencias mitocondriales recuperadas después de la secuenciación. Esto puede deberse principalmente al tipo de tejido (Fernández-Vizarra *et al.,* 2010) y la cantidad de ADN extraída (Figura 13) dado que la calidad de secuenciación es similar entre las muestras (entre 93 a 94% de Q30), con ligeras diferencias en las últimas posiciones entre lecturas *forward* y *reverse* (Figura 14). Los parámetros de calidad, comprueban que las secuencias en conjunto tienen una calidad muy buena que esta entre el rango de 30 a 40 *Phred score* (Figura 15 A) y una distribución homogénea de longitud de secuencias (Figura 15 B). Las variaciones del contenido de bases en las secuencias y la distribución del contenido de GC (Figura 16 A y B) tienen una ligera desviación, debido probablemente a la presencia de secuencias adaptadoras (**Anexo 5**). Por otra parte, el análisis de conteo de *k-mers* nos indica un único pico correspondiente al promedio de profundidad (Figura 17), otro correspondiente a los errores intrínsecos de secuenciamiento illumina y no se despliegan crestas de regiones repetitivas o heterocigocidad de alta frecuencia, que indiquen errores que hayan afectado la secuenciación (Marçais y Kingsford, 2011; Sohn y Nam, 2016). Sin embargo, algo que podemos apreciar es una diferencia en la frecuencia de *k-mers* de los picos de profundidad entre las muestras, siendo más frecuentes las muestras de *C. tschudii*, seguida por *C. porcellus* "criollos" de Andahuaylas (EEA Chumbibamba - Figura 17).

El filtrado de lecturas de baja calidad, adaptadores y lecturas de baja longitud, arrojan un alto porcentaje de recuperación (76,8%), con más del 50 % de lecturas emparejadas (Figura 18). Si bien las lecturas *reverse* tienen menor calidad en promedio (Kwon *et al.,* 2013), conseguimos poca cantidad de lecturas *reverse* impares rescatadas en comparación a secuencias *forward*, este sesgo se puede atribuir a la preparación de la librería o al protocolo de secuenciación (Pfeiffer *et al.,* 2018).

El porcentaje de alineamiento de las lecturas filtradas al genoma mitocondrial resultó en un 91,6%, muy similar entre las muestras, se obtuvo poco porcentaje de lecturas

impares (*singletons*) al igual que alineamientos a diferentes regiones del genoma (Figura 19). Como se indicó anteriormente se utilizó el genoma de referencia de *C. porcellus* (Cavpor v3.0) por lo que similares parámetros de alineamiento puede ser un indicativo de alta homología entre los genomas de C*. porcellus y C. tschudii*, como se reportó en algunos estudios citogenéticos (Walker *et al.* 2014). En cuanto a las lecturas mitocondriales recuperadas, apreciamos una variación en el número de lecturas entre genomas (Tabla 4), como se comentó anteriormente esto se puede atribuir al tipo de tejido y la cantidad de ADN extraído.

Las diferencias del contenido de mitocondrias entre tejidos es debido a la demanda energética, muy necesaria por ejemplo en músculos en comparación de otros tejidos (Fernández-Vizarra *et al.,* 2011). A pesar de esas diferencias, vemos que esto no limita considerablemente la secuenciación en los tejidos de cartílago y bulbo piloso, por lo que no los excluye para estudios de genómica, al contrario, los hace una alternativa viable para la colecta de muestras en animales silvestres, donde se debe evitar sacrificar al animal y se deben realizar muestreos menos invasivos (Henry *et al.,* 2011).

Las secuencias alineadas al genoma mitocondrial de referencia dentro de Cavpor v3 evidencian una profundidad variable, con regiones de alta profundidad, en detrimento de las regiones adyacentes (Figura 20 – 22), una de las posibles causas de esta sobrerrepresentación de secuencias puede deberse a la dinámica de replicación del genoma mitocondrial (Holt y Reyes, 2012); principalmente las regiones de origen de replicación, como se reportó en bacterias (Sirand-Pugnet *et al*., 2007). Por otra parte las regiones repetitivas pueden estar sobrerrepresentadas debido al número de copias presentes, tanto en la región en cuestión como en otros lugares del genoma (Ross *et al.,* 2013), en cambio las regiones que corresponden al gen NADH deshidrogenasa y fundamentalmente al gen de ARN 16S, no se puede atribuir una razón exacta de su profundidad, en la mitocondria el ARN16S no presenta múltiples copias a diferencia de las bacterias (Acinas *et al.,* 2004), aunque una explicación a su elevada profundidad puede ser a las similitudes con genes de ARN nucleares secuenciados paralelamente.

En el ensamblaje de *novo* realizado con SPAdes, la resolución del genoma decae en el gen ARNt-Leu y en la región DLP, en la Figura 23 la región repetitiva de los genoma mitocondriales está ausente, lo cual reduce el tamaño total del ensamblaje (Tabla 5) y la fracción de genoma cubierto (Tabla 7, 96% al ~99 %) en comparación del genoma de referencia (NC_000884.1: 16 801 pb). Dado que los ensamblajes obtuvieron 2 *contigs* como máximo (Tabla 5), que cubren más 50 % del genoma mitocondrial, los parámetros N50, NG50, L50 y LG50 reflejan el tamaño de los *contigs* más grandes (Tabla 7), las métricas de coherencia nos indican un alto porcentaje de alineamiento (mayor al 99 %), que en su mayoría es propiamente emparejada y con alto porcentaje de profundidad de cobertura. Las métricas de integridad nos indican la ausencia del gen ARNt-Leu en dos genomas ensamblados (Tabla 7), que se debe probablemente a la falta de resolución del ensamblaje en esa región.

El ensamblaje con NOVOplasty nos muestra que 9 de los 10 genomas ensamblados presentan 2 *contigs* y una mejor resolución de ensamblaje en la región repetitiva (Figura 24). Sin embargo, observamos que el tamaño total de ensamblaje es mayor al genoma de referencia (Tabla 6), como observamos en las métricas de contigüidad la fracción del genoma cubierto en 7 de los 10 genomas ensamblados supera el 100 %, que se debe a una expansión de la región repetitiva. Las métricas de coherencia muestran un alto porcentaje de alineamiento de lecturas al genoma ensamblado (Tabla 8), aunque están por debajo de las métricas del ensamblaje *de novo* por SPAdes. Una clara diferencia entre ambas metodologías lo encontramos en las métricas de integridad, donde se observa una duplicación del gen de ARNr 12S en el ensamblaje por NOVOplasty a diferencia de la ausencia de un ARNt en SPAdes, esta duplicación de genes se debe a la expansión de la región repetitiva dentro de la región DLP, que se encuentra adyacente al gen de ARNr 12S. Una explicación a estas diferencias se debe a que el programa NOVOplasty utiliza un algoritmo de ensamblaje similar a OLC basado en un alineamiento iterativo (Dierckxsens *et al.,* 2017), y al igual que otros métodos de ensamblaje (*de novo*) no puede determinar el tamaño de la expansión en regiones STR

por limitación de la librería (Tørresen *et al.,* 2019). Debemos recordar también que la resolución de las secuencias repetitivas en la secuenciación illumina es limitada por la longitud de bases secuenciadas y el tamaño de librería, de modo que pueden generarse ambigüedades en el genoma ensamblado (Treangen y Salzberg, 2011). A diferencia de NOVOplasty el ensamblaje de SPAdes es más flexible ya que se obtuvieron mejores resultados, variando los *k-mers* utilizados y reduciendo la profundidad de las secuencias al filtrar aleatoriamente lecturas con el programa Seqtk (Shen *et al.,* 2016), esto se realizó para no sobrepasar una óptima profundidad de ensamblaje (Desai *et al.,* 2013). Dadas las diferencias observadas entre los métodos de ensamblaje, se prefirió los ensamblajes obtenidos por SPAdes, por lo que los análisis realizados posteriormente se basaron en estas secuencias.

Por otro lado, las secuencias descargadas de la base de datos SRA presentan una mayor varianza de calidad, pero con buena calidad de secuenciamiento en promedio (Figura 25. A y B), la profundidad promedio de lecturas únicas del genoma es de aproximadamente 18X (Figura 25. C), el filtrado de secuencias al igual que los genomas anteriores obtuvo un porcentaje mayor de 90% (Figura 25. D). En comparación con los anteriores mitogenomas se recuperaron menos lecturas de genoma mitocondrial (Tabla 9). A pesar de presentar lecturas de mayor tamaño (2x250 pb), la región repetitiva no se logró resolver, al igual que parte del gen ARNt - Pro (Figura 26). Las métricas de contigüidad y coherencia nos indican un adecuado ensamblaje, con la ausencia del gen ARNt -Pro (Tabla 10).

6.2. Anotación de genomas y distancias genéticas entre *C. porcellus* y *tschudii*

En este estudio se reconstruyeron 11 mitogenomas entre "criollos", mejorados, silvestres y secuencias procedentes de la base de datos SRA-NCBI (Figura 33. A), encontrado en los ensamblajes una configuración de genes (Figura 27 - 32) similar al genoma de referencia de *C. porcellus* (D'Erchia *et al.,* 1996), a partir de una anotación de *novo* con GeSeq (Tillich *et al.,* 2017) y MITOS (Bernt *et al.,* 2013) se observan 13 genes que codifican para proteínas, 22 genes que codifican para ARNt, únicamente con la ausencia del gen ARNt – Leu en el genoma de SA42913, y dos genes que codifican para ARNr. No se observan variantes estructurales en los genomas ensamblados con SPAdes. A diferencia de otros animales los mitogenomas en mamíferos son consistentemente conservados, observándose estos cambios estructurales en el mitogenoma de animales no bilaterios (Lavrov y Pett, 2016) o plantas (Choi *et al.,* 2019), aunque no se puede descartar variaciones en el número de copias de la región repetitiva en *C. porcellus,* como se reportaron en algunas aves (Wang *et al.,* 2014). Los análisis de distancia genética indica que el mitogenoma de *C. tschudii* proveniente de Pachachaca es más distante, a diferencia de los otros dos *C. tschudii* provenientes del refugio de vida silvestre de Pantanos de Villa, que se agruparon finalmente tanto con las secuencias de *C. porcellus* "criollos" de Andahuaylas (EEA Chumbibamba) como con el bloque de mitogenomas de la bases de datos públicas y *C. porcellus* mejorados de la raza "PERU", donde se puede forzar una ligera distancia entre "criollos" y mejorados (Figura 33. B). La agrupación de animales de diferente procedencia (según el hábitat) va en contra de lo esperado entre animales domésticos y silvestres; como es sabido la relajación de la selección después de un proceso de domesticación, permite fijar variantes (deletéreas o de importancia fenotípica) en animales domésticos (Björnerfeldt *et al.,* 2006), en cambio la selección purificadora en animales silvestres juega un papel importante en su adaptación (Hughes y Hughes, 2007), estas diferentes fuerzas de selección parecen no estar actuando como se espera en los datos de distancia genética.

6.3. Análisis filogenéticos, demográficos y haplotipicos

Los análisis filogenéticos de máxima verosimilitud, máxima parsimonia y de inferencia bayesiana presentan una alta similitud en topología, con algunas diferencias entre los *C. porcellus* depositados en la base datos del NCBI. En general se obtiene un clado monofiletico entre *C. porcellus* y *C. tschudii,* observándose a *C. tschudii* de Pachachaca como clado basal y los *C. tschudii* de Pantanos de Villa agrupados entre *C. porcellus* domésticos y el grupo de la base de datos de NCBI y cuyes mejorados, similar a lo que se observa en las distancias genéticas (Figura 33. B). Asimismo, en un clado distinto se observa únicamente a *Cavia aperea,* con lo cual esta filogenia está en concordancia con los trabajos reportados anteriormente por Spotorno *et al.* en el 2006, que descarta a *C. aperea* como posible fuente de domesticación de *C. porcellus,* y lo emparenta a *C. tsuchdii*.

A diferencia del grupo polifilético reportado por Dunnum y Salazar-Bravo en el 2010, en este trabajo se reporta un grupo monofiletico entre *C. porcellus* y *C. tschudii*. Sin embargo, este resultado no es concluyente debido a la escasa información de mitogenomas secuenciados de *C. tschudii*. Los cuyes silvestres a los cuales atribuimos la fuente de domesticación de *C. porcellus* están presente en países como Chile (*Wild Lluta,* Spotorno *et al.,* 2006), Bolivia (*Cavia tschudii,* Dunnum y Salazar-Bravo, 2010), Argentina (*C. tschudii sodalis*) y Perú (*C. tschudii osgoodi, C. tschudii tschudii, C. tschudii arequipae* - Dunnum y Salazar-Bravo, 2010*, wild chiguata* – Spotorno *et al.,* 2006). Para este trabajo únicamente se reporta *C. tschudii* de dos locaciones en Perú, de manera que no podemos concretar el clado filogenético de *C. tschudii* por la ausencia de información, además no se debe descartar especies o subespecies que aún no están reportadas.

Por otro parte *C. porcellus* está separados en 2 grupos (Figura 34 B), por un lado se observan a los cuyes domésticos de Andahuaylas (EEA Chumbibamba) y conjuntamente se observan a los cuyes mejorados (EEA La Molina) y *C. porcellus*

depositados en la base de datos NCBI, lo que indica una diferencia entre los *C. porcellus* sometidos a selección artificial y animales domésticos (Avilés *et al.,* 2014), en concordancia con lo observado en otros animales sometidos a mejoramiento, ya que presentan por lo general una elevada endogamia y tienden a acumular alelos neutrales o deletéreas a diferencia de animales domésticos (Andersson y Georges, 2004). Algo que llama la atención en los resultados filogenéticos, es que las muestras de *C. tschudii* de Pantanos de Villa están más emparentados a *C. porcellus*, tanto que parecieran ser la misma especie. Como se mencionó anteriormente acerca de la relajación de la selección (presión de selección) a la cual están sometidos los animales silvestres, y no los animales domésticos, pareciera que encontramos una contradicción ya que lo normal sería una diferenciación entre silvestres y domésticos (Björnerfeldt *et al.,* 2006). Según los antecedentes en algunos estudios filogenéticos con especies domésticas y silvestres que comparten un mismo clado, se concluyó que los eventos de domesticación independiente o múltiples orígenes maternos puede ser la causa de estos resultados (Larson *et al.,* 2005; Liu *et al.,* 2006). Aunque sea probable una domesticación independiente de *C. porcellus* en algunas regiones de Colombia a partir de *Cavia aperea* (E. Lord *et al.,* 2020), se reafirma en este trabajo que la fuente de domesticación de los cuyes que son criados en la actualidad está emparentada a la especie *C. tschudii* (Figura 34). Además, los estudios en diferentes centros arqueológicos refuerzan el parentesco de *C. porcellus* con *C. tschudii* (E. Lord *et al.,* 2020, E. Lord *et al.,* 2018; Kimura *et al.,* 2016), lo cual nos indican que los animales domesticados en la antigüedad a partir de *C. tschudii* probablemente se hayan distribuido con mayor éxito o hayan gozado de mayor aceptación entre poblaciones precolombinas, resultado de migraciones humanas probablemente (Kimura *et al.,* 2016). Por lo que en el pasado los intercambios comerciales de cuyes pudieron haber desplazado animales nativos con diferente fuente de domesticación. Sin embargo, la información genética de animales domésticos criados en regiones rurales ya se en Perú o en otro país andino aún no se estudia a profundidad,

por lo que no podemos descartar la presencia de remantes de linajes mitocondriales de *C. tschudii* e incluso cuyes domesticados a partir de otra especie.

Otra posible explicación para la presencia de animales domésticos y silvestres dentro de un mismo clado es una posible introgresión de animales silvestres en poblaciones de animales domésticos (Giuffra *et al*., 2000). La crianza de cuyes "domésticos" es principalmente rural, y de consumo familiar (Avilés *et al.,* 2014) y se ha observado que se generan híbridos fértiles con animales silvestres, a pesar de presentar diferenciación cromosómica (Walker *et al.,* 2014), por ello la posibilidad de un contacto con animales silvestres no debería ser descartada, ya sea por la intervención de criadores de cuyes, al tratar de cruzar cuyes silvestres con cuyes domésticos, así como por la precariedad de los corrales de crianza, permitiendo el ingreso de cuyes silvestres a criaderos de cuyes. Sin embargo, debe considerarse también la posibilidad de una introgresión inversa, de poblaciones "criollos" a poblaciones silvestres. Para reforzar este último punto analizamos la historia demográfica de los cuyes por el análisis bayesiano de s*kyline*, que evidencia una disminución del tamaño efectivo poblacional reciente (Figura 35), respaldado por otros estudios esto se puede atribuir a un proceso de domesticación o a un cuello de botella en las poblaciones de cuyes domésticos. Si bien la domesticación de cuyes es una práctica llevada a cabo desde la antigüedad, esta no tiene un proceso tecnificado en el ámbito rural donde es más extendida la crianza, ni está orientado a la selección de una característica fenotípica en particular (Chauca F., 1997). Es por ello que a partir de los análisis demográficos en animales domésticos podemos formar una idea solamente del proceso de domesticación, con el cambio del tamaño efectivo poblacional a través del tiempo, si comparamos nuestro resultados con los estudios de otros animales domésticos, vemos que estos habitualmente son seleccionados a partir de un reducido número de animales, y por lo general se observa un reducido número de haplotipos fundadores, evidenciado en la disminución de su tamaño efectivo poblacional (Naderi *et al*., 2008; Achilli *et al.,* 2012; Colli *et al.,* 2015; Witt *et al.,* 2015; Sun *et al.,* 2020). De la misma forma la domesticación de *C. porcellus*

tuvo probablemente como origen un reducido número de individuos silvestres, no obstante, a diferencia de lo observado en otros animales domésticos, donde el desplazamiento de las poblaciones fuera de su centro de origen permite atribuirle a la selección principalmente la disminución del tamaño efectivo poblacional (Alberto *et al.,* 2018), la introgresión en cuyes puede obstaculizar una adecuada interpretación.

Muy poco se ha reportado sobre la crianza de cuyes en la antigüedad, incluso después de la colonización española, sin embargo sobre las poblaciones indígenas que están ligadas a la crianza de cuyes, se han reportado disminución poblacional considerable a causa de enfermedades y trabajos forzados (Riedel, 2005), también se documentó sobre la reubicación de poblaciones en centros poblados denominadas reducciones (Montero, 2018), que indefectiblemente debió reflejarse en el descuido o abandono de centros de crianza en la antigüedad; esto pudo mermar la cantidad de cuyes criados así como inducir el desplazamiento de animales domésticos a la vida silvestre incentivado la introgesión inversa, siendo este fenómeno una explicación para la disminución del tamaño efectivo poblacional. Otro hecho que refuerza esta hipótesis es que se han observado problemas en la caracterización filogenética de especies de *Cavia*, por ejemplo *Cavia cutleri* reportado por Bennett en 1836, que posteriormente fue modificado a *C. tschduii* por Fitzinger en 1867 (Weir, 1974), generó híbridos fértiles al mezclarlos con *C. porcellus*, a diferencia de lo reportado con otros *C. tschudii* donde la reproducción es más complicada (Walker *et al*., 2014), a partir de los resultados de este trabajo podemos observar que *C. tsuchdii* de Pachachaca está más distanciado que los cuyes domésticos, donde es probable que una hibridación con *C. porcellus* sea menos viable, caso contrario con *C. tschudii* de Pantanos de Villa (Figura 36). Asimismo, recientemente se reportaron que los cuyes de Pantanos de Villa presentan alelos deletéreos (Ramirez *et al.,* 2019), como se mencionó anteriormente estas características están presentes principalmente en animales sometidos a selección, por lo que puede reforzar una introgresión de *C. porcellus* a animales silvestres, incluso recientemente.

7. CONCLUSIONES

- Se reconstruyeron adecuadamente 10 genomas mitocondriales a partir de WGS, de cuyes “criollos”, mejorados, silvestres y 01 proveniente de la base de datos SRA-NCBI, obteniendo un mejor ensamblaje con el programa SPAdes.
- Se obtuvo un alto grado de conservación de genes entre los genomas mitocondriales ensamblados de *C. porcellus* y *C. tschudii.*
- Se encontraron diferencias en las distancias genéticas entre los *C. tschudii* secuenciados, observándose más distante el mitogenoma de *C. tschudii* de Pachachaca y más emparentados a *C. porcellus* y *C. tschudii* de Pantanos de Villa.
- Los análisis filogenéticos muestran a *C. tschudii* de Pachachaca como un grupo basal al grupo monofiletico de *C. porcellus* y *C. tschudii* de Pantanos de Villa.
- Se reafirma que la domesticación de *C. porcellus* fue a partir de *C. tschudii*, aunque con un posible caso de introgresión posterior de especies silvestres o viceversa, ya que se observa variada afinidad entre *C. porcellus* y algunos *C. tschudii.*

8. RECOMENDACIONES

- Es recomendable una adecuada extracción de ADN para el secuenciamiento (Cantidad y calidad) ya que influye en la calidad de secuenciamiento y el número de secuencias mitocondriales recuperadas.
- La secuenciación de la región repetitiva del genoma mitocondrial debe ser realizada con otra tecnología de secuenciación con lecturas de mayor longitud (como *Sanger* o PacBio), con la finalidad de circularizar los mitogenomas.
- Se sugiere desarrollar estudios de estructuración genética en poblaciones de *C. porcellus*, tomando en cuenta principalmente los ámbitos rurales, debido que pueden encontrase remanentes de diferentes linajes mitocondriales e incluso *C. porcellus* de diferentes fuentes de domesticación.
- Se sugiere realizar un muestreo más extenso en *C. tschudii* tanto en Perú con en otros países, debido a la heterogénea distancia genética a *C. porcellus* para esclarecer la filogenia entres estas especies.

9. REFERENCIAS BIBLIOGRÁFICAS

Achilli, Alessandro, Anna Olivieri, Pedro Soares, Hovirag Lancioni, Baharak Hooshiar Kashani, Ugo A. Perego, Solomon G. Nergadze, *et al*. «Mitochondrial genomes from modern horses reveal the major haplogroups that underwent domestication». *Proceedings of the National Academy of Sciences of the United States of America* 109, n.º 7 (14 de febrero de 2012): 2449-54. https://doi.org/10.1073/pnas.1111637109.

Acinas, Silvia G., Luisa A. Marcelino, Vanja Klepac-Ceraj, y Martin F. Polz. «Divergence and Redundancy of 16S rRNA Sequences in Genomes with Multiple rrn Operons». *Journal of Bacteriology* 186, n.º 9 (mayo de 2004): 2629-35. https://doi.org/10.1128/JB.186.9.2629-2635.2004.

Ackerman, M. J. «The Visible Human Project: A Resource for Education». *Academic Medicine* 74, n.º 6 (junio de 1999): 667–70.

Adkins, Ronald M., Eric L. Gelke, Diane Rowe, y Rodney L. Honeycutt. «Molecular Phylogeny and Divergence Time Estimates for Major Rodent Groups: Evidence from Multiple Genes». *Molecular Biology and Evolution* 18, n.º 5 (1 de mayo de 2001): 777-91. https://doi.org/10.1093/oxfordjournals.molbev.a003860.

Alberto, Florian J., Frédéric Boyer, Pablo Orozco-terWengel, Ian Streeter, Bertrand Servin, Pierre de Villemereuil, Badr Benjelloun, et al. «Convergent genomic signatures of domestication in sheep and goats». *Nature Communications* 9 (6 de marzo de 2018). https://doi.org/10.1038/s41467-018-03206-y.

Alberts, Bruce, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, y Peter Walter. *Biología Molecular de la Célula*. 4a ed. Garland Science, 2002.

Al-Hasan, Yazan M., Gerard A. Pinkas, y Loren P. Thompson. «Prenatal Hypoxia Reduces Mitochondrial Protein Levels and Cytochrome c Oxidase Activity in Offspring Guinea Pig Hearts». *Reproductive Sciences (Thousand Oaks, Calif.)* 21, n.º 7 (julio de 2014): 883-91. https://doi.org/10.1177/1933719113518981.

Alic, Andy S., David Ruzafa, Joaquin Dopazo, y Ignacio Blanquer. «Objective Review of de Novo Stand-Alone Error Correction Methods for NGS Data». *WIREs Computational Molecular Science* 6, n.º 2 (2016): 111-46. https://doi.org/10.1002/wcms.1239.

Allen, Naomi, Cathie Sudlow, Paul Downey, Tim Peakman, John Danesh, Paul Elliott, John Gallacher, et al. «UK Biobank: Current status and what it means for epidemiology». *Health Policy and Technology* 1, n.º 3 (1 de septiembre de 2012): 123-26. https://doi.org/10.1016/j.hlpt.2012.07.003.

Al-Nakeeb, Kosai, Thomas Nordahl Petersen, y Thomas Sicheritz-Pontén. «Norgal: extraction and de novo assembly of mitochondrial DNA from whole-genome sequencing data». *BMC Bioinformatics* 18, n.º 1 (21 de noviembre de 2017): 510. https://doi.org/10.1186/s12859-017-1927-y.

Anderson, S., A. T. Bankier, B. G. Barrell, M. H. de Bruijn, A. R. Coulson, J. Drouin, I. C. Eperon, et al. «Sequence and Organization of the Human Mitochondrial Genome». *Nature* 290, n.º 5806 (9 de abril de 1981): 457-65. https://doi.org/10.1038/290457a0.

Andersson, Leif, y Michel Georges. «Domestic-Animal Genomics: Deciphering the Genetics of Complex Traits». *Nature Reviews. Genetics* 5, n.º 3 (marzo de 2004): 202-12. https://doi.org/10.1038/nrg1294.

Andrews, S. «FastQC A Quality Control tool for High Throughput Sequence Data», 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Arampatzis, Asterios, Ioanna Papagiouvanni, Doxakis Anestakis, y Magda Tsolaki. «A Classification and Comparative Study of European Biobanks: an Analysis of Biobanking Activity and its Contribution to Scientific Progress». *Archives of Medicine* 8, n.º 3 (9 de junio de 2016). http://www.archivesofmedicine.com/abstract/a-classification-and-comparative-study-of-european-biobanks-an-analysis-of-biobanking-activity-andits-contribution-to-scientific-p-9575.html.

Avilés, D. F., A. M. Martínez, V. Landi, y J. V. Delgado. «El Cuy (*Cavia Porcellus*): Un Recurso Andino de Interés Agroalimentario The Guinea Pig (*Cavia Porcellus*): An Andean Resource of Interest as an Agricultural Food Source». *Animal Genetic Resources/Resources Génétiques Animales/Recursos Genéticos Animales* 55 (diciembre de 2014): 87-91. https://doi.org/10.1017/S2078633614000368.

Aviles, Diana, Vincenzo Landi, Juan Vicente Delgado, José Luis Vega-Pla, y Amparo Martinez. «Isolation and Characterisation of a Dinucleotide Microsatellite Set for a Parentage and Biodiversity Study in Domestic Guinea Pig (*Cavia Porcellus*)». *Italian Journal of Animal Science* 14, n.º 4 (1 de enero de 2015): 3960. https://doi.org/10.4081/ijas.2015.3960.

Aviles Esquivel, Diana Fernanda Avilés. «"Caracterización genética del cuy doméstico en America"», s. f., 143.

Ayad, Lorraine A. K., y Solon P. Pissis. «MARS: improving multiple circular sequence alignment using refined sequences». *BMC Genomics* 18 (14 de enero de 2017). https://doi.org/10.1186/s12864-016-3477-5.

Ayres, Daniel L., Aaron Darling, Derrick J. Zwickl, Peter Beerli, Mark T. Holder, Paul O. Lewis, John P. Huelsenbeck, et al. «BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics». *Systematic Biology* 61, n.º 1 (1 de enero de 2012): 170-73. https://doi.org/10.1093/sysbio/syr100.

Bankevich, Anton, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, et al. «SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing». *Journal of Computational Biology* 19, n.º 5 (mayo de 2012): 455-77. https://doi.org/10.1089/cmb.2012.0021.

Barilari. «Relaciones filogenéticas de Cávidos sudamericanos basados en la región control del DNA mitcondrial», s. f., 59.

Bartolák-Suki, Erzsébet, Jasmin Imsirovic, Yuichiro Nishibori, Ramaswamy Krishnan, y Béla Suki. «Regulation of Mitochondrial Structure and Dynamics by the Cytoskeleton and Mechanical Factors». *International Journal of Molecular Sciences* 18, n.º 8 (21 de agosto de 2017). https://doi.org/10.3390/ijms18081812.

Batzoglou, Serafim, David B. Jaffe, Ken Stanley, Jonathan Butler, Sante Gnerre, Evan Mauceli, Bonnie Berger, Jill P. Mesirov, y Eric S. Lander. «ARACHNE: A Whole-Genome Shotgun Assembler». *Genome Research* 12, n.º 1 (enero de 2002): 177-89. https://doi.org/10.1101/gr.208902.

Bentley, David R. «Whole-Genome Re-Sequencing». *Current Opinion in Genetics & Development* 16, n.º 6 (diciembre de 2006): 545-52. https://doi.org/10.1016/j.gde.2006.10.009.

Bentley, David R., Shankar Balasubramanian, Harold P. Swerdlow, Geoffrey P. Smith, John Milton, Clive G. Brown, Kevin P. Hall, et al. «Accurate Whole Human Genome Sequencing using Reversible Terminator Chemistry». *Nature* 456, n.º 7218 (6 de noviembre de 2008): 53-59. https://doi.org/10.1038/nature07517.

Bernt, Matthias, Alexander Donath, Frank Jühling, Fabian Externbrink, Catherine Florentz, Guido Fritzsch, Joern Pütz, Martin Middendorf, y Peter F. Stadler. «MITOS: Improved de

Novo Metazoan Mitochondrial Genome Annotation». *Molecular Phylogenetics and Evolution* 69, n.º 2 (noviembre de 2013): 313-19. https://doi.org/10.1016/j.ympev.2012.08.023.

Bevers, Roel P. J., Maria Litovchenko, Adamandia Kapopoulou, Virginie S. Braman, Matthew R. Robinson, Johan Auwerx, Brian Hollis, y Bart Deplancke. «Extensive Mitochondrial Population Structure and Haplotype-Specific Phenotypic Variation in the Drosophila Genetic Reference Panel». *BioRxiv*, 9 de noviembre de 2018, 466771. https://doi.org/10.1101/466771.

Bignell, G. R., A. R. Miller, y I. H. Evans. «Isolation of Mitochondrial DNA». *Methods in Molecular Biology (Clifton, N.J.)* 53 (1996): 109-16. https://doi.org/10.1385/0-89603-319-8:109.

Birsoy, Kıvanç, Tim Wang, Walter Chen, Elizaveta Freinkman, Monther Abu-Remaileh, y David M. Sabatini. «An essential role of the mitochondrial electron transport chain in cell proliferation is to enable aspartate synthesis». *Cell* 162, n.º 3 (30 de julio de 2015): 540-51. https://doi.org/10.1016/j.cell.2015.07.016.

Björnerfeldt, Susanne, Matthew T. Webster, y Carles Vilà. «Relaxation of Selective Constraint on Dog Mitochondrial DNA Following Domestication». *Genome Research* 16, n.º 8 (agosto de 2006): 990-94. https://doi.org/10.1101/gr.5117706.

Blachly-Dyson, Elizabeth, y Michael Forte. «VDAC Channels». *IUBMB Life* 52, n.º 3-5 (2001): 113-18. https://doi.org/10.1080/15216540152845902.

Boettcher, P. J., M. Tixier-Boichard, M. A. Toro, H. Simianer, H. Eding, G. Gandini, S. Joost, D. Garcia, L. Colli, y P. Ajmone-Marsan. «Objectives, Criteria and Methods for Using Molecular Genetic Data in Priority Setting for Conservation of Animal Genetic Resources». *Animal Genetics* 41, n.º s1 (2010): 64-77. https://doi.org/10.1111/j.1365-2052.2010.02050.x.

Boixés Saña, D., y M. Badia Mallorquí. «[Peripheral specimen collecting centers. Preanalytical control]». *Revista De Enfermeria (Barcelona, Spain)* 21, n.º 235 (marzo de 1998): 13-16.

Bolger, Anthony M., Marc Lohse, y Bjoern Usadel. «Trimmomatic: a flexible trimmer for Illumina sequence data». *Bioinformatics* 30, n.º 15 (1 de agosto de 2014): 2114-20. https://doi.org/10.1093/bioinformatics/btu170.

Bradbury, Peter J., Zhiwu Zhang, Dallas E. Kroon, Terry M. Casstevens, Yogesh Ramdoss, y Edward S. Buckler. «TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples». *Bioinformatics (Oxford, England)* 23, n.º 19 (1 de octubre de 2007): 2633-35. https://doi.org/10.1093/bioinformatics/btm308.

Brandon, Marty C., Eduardo Ruiz-Pesini, Dan Mishmar, Vincent Procaccio, Marie T. Lott, Kevin Cuong Nguyen, Syawal Spolim, Upen Patil, Pierre Baldi, y Douglas C. Wallace. «MITOMASTER – A Bioinformatics Tool For the Analysis of Mitochondrial DNA Sequences». *Human mutation* 30, n.º 1 (enero de 2009): 1-6. https://doi.org/10.1002/humu.20801.

Breton, Sophie, Liliana Milani, Fabrizio Ghiselli, Davide Guerra, Donald T. Stewart, y Marco Passamonti. «A Resourceful Genome: Updating the Functional Repertoire and Evolutionary Role of Animal Mitochondrial DNAs». *Trends in Genetics: TIG* 30, n.º 12 (diciembre de 2014): 555-64. https://doi.org/10.1016/j.tig.2014.09.002.

Burgos-Paz, William, Mario Cerón-Muñoz, y Carlos Solarte-Portilla. «Genetic diversity and population structure of the Guinea pig (*Cavia porcellus*, Rodentia, caviidae) in Colombia». *Genetics and Molecular Biology* 34, n.º 4 (2011): 711-18. https://doi.org/10.1590/S1415-47572011005000057.

Butler, Jonathan, Iain MacCallum, Michael Kleber, Ilya A. Shlyakhter, Matthew K. Belmonte, Eric S. Lander, Chad Nusbaum, y David B. Jaffe. «ALLPATHS: De novo assembly of whole-genome shotgun microreads». *Genome Research* 18, n.º 5 (mayo de 2008): 810-20. https://doi.org/10.1101/gr.7337908.

Bycroft, Clare, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T. Elliott, Kevin Sharp, Allan Motyer, et al. «The UK Biobank Resource with Deep Phenotyping and Genomic Data». *Nature* 562, n.º 7726 (octubre de 2018): 203. https://doi.org/10.1038/s41586-018-0579-z.

Cabañas, Noel, Arturo Becerra, David Romero, Tzipe Govezensky, Jesús Javier Espinosa-Aguirre, y Rafael Camacho-Carranza. «Repetitive DNA profile of the amphibian mitogenome». *BMC Bioinformatics* 21 (19 de mayo de 2020). https://doi.org/10.1186/s12859-020-3532-8.

Camino M., Javier, y Víctor Hidalgo L. «EVALUACIÓN DE DOS GENOTIPOS DE CUYES (*Cavia porcellus*) ALIMENTADOS CON CONCENTRADO Y EXCLUSIÓN DE FORRAJE VERDE.» *Revista de Investigaciones Veterinarias del Perú* 25, n.º 2 (3 de julio de 2014). https://doi.org/10.15381/rivep.v25i2.8490.

Cao, Wei, y Yan Xia. «The Complete Mitochondrial Genome of the *Cavia Aperea*». *Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis* 27, n.º 6 (2016): 4429-30. https://doi.org/10.3109/19401736.2015.1089557.

Cao, Y., J. Adachi, T. Yano, y M. Hasegawa. «Phylogenetic place of guinea pigs: no support of the rodent-polyphyly hypothesis from maximum-likelihood analyses of multiple protein sequences.» *Molecular Biology and Evolution* 11, n.º 4 (1 de julio de 1994): 593-604. https://doi.org/10.1093/oxfordjournals.molbev.a040139.

CATCHEN, JULIAN, PAUL A. HOHENLOHE, SUSAN BASSHAM, ANGEL AMORES, y WILLIAM A. CRESKO. «Stacks: an analysis tool set for population genomics». *Molecular ecology* 22, n.º 11 (junio de 2013): 3124-40. https://doi.org/10.1111/mec.12354.

Catchen, Julian M., Angel Amores, Paul Hohenlohe, William Cresko, y John H. Postlethwait. «Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences». *G3: Genes, Genomes, Genetics* 1, n.º 3 (1 de agosto de 2011): 171-82. https://doi.org/10.1534/g3.111.000240.

Cavalier-Smith, T. «Archamoebae: The Ancestral Eukaryotes?» *Biosystems*, Papers presented at the 8th Biennial Conference of the International Society for Evolutionary Protistology, 25, n.º 1 (1 de enero de 1991): 25-38. https://doi.org/10.1016/0303-2647(91)90010-I.

Chan, Cheong Xin, y Mark A Ragan. «Next-generation phylogenomics». *Biology Direct* 8 (22 de enero de 2013): 3. https://doi.org/10.1186/1745-6150-8-3.

Chan, David C. «Mitochondria: Dynamic Organelles in Disease, Aging, and Development». *Cell* 125, n.º 7 (30 de junio de 2006): 1241-52. https://doi.org/10.1016/j.cell.2006.06.010.

Chauca F., Lilia. «Producción de cuyes (*Cavia porcellus*) -», 1997. http://www.fao.org/3/w6562s/w6562s00.htm.

Chen, Shifu, Tanxiao Huang, Yanqing Zhou, Yue Han, Mingyan Xu, y Jia Gu. «AfterQC: automatic filtering, trimming, error removing and quality control for fastq data». *BMC Bioinformatics* 18, n.º Suppl 3 (14 de marzo de 2017). https://doi.org/10.1186/s12859-017-1469-3.

Chen, T. J., R. G. Boles, y L. J. Wong. «Detection of Mitochondrial DNA Mutations by Temporal Temperature Gradient Gel Electrophoresis». *Clinical Chemistry* 45, n.º 8 Pt 1 (agosto de 1999): 1162-67.

Cho, Sang Yun, Eun Jung Hong, Jung Min Nam, Bogkee Han, Chaeshin Chu, y Ok Park. «Opening of the National Biobank of Korea as the Infrastructure of Future Biomedical Science in Korea». *Osong Public Health and Research Perspectives* 3, n.º 3 (septiembre de 2012): 177-84. https://doi.org/10.1016/j.phrp.2012.07.004.

Choi, In-Su, Erika N. Schwarz, Tracey A. Ruhlman, Mohammad A. Khiyami, Jamal S. M. Sabir, Nahid H. Hajarah, Mernan J. Sabir, Samar O. Rabah, y Robert K. Jansen. «Fluctuations in Fabaceae mitochondrial genome size and content are both ancient and recent». *BMC Plant Biology* 19, n.º 1 (25 de octubre de 2019): 448. https://doi.org/10.1186/s12870-019-2064-8.

Cock, Peter J. A., Christopher J. Fields, Naohisa Goto, Michael L. Heuer, y Peter M. Rice. «The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants». *Nucleic Acids Research* 38, n.º 6 (abril de 2010): 1767-71. https://doi.org/10.1093/nar/gkp1137.

Cohen, Julie S., y Barbara B. Biesecker. «Quality of Life in Rare Genetic Conditions: A Systematic Review of the Literature». *American journal of medical genetics. Part A* 152A, n.º 5 (mayo de 2010): 1136-56. https://doi.org/10.1002/ajmg.a.33380.

Colareda, Germán A., María I. Ragone, y Alicia E. Consolini. «Sex Differences in the Mechano-Energetic Effects of Genistein on Stunned Rat and Guinea Pig Hearts». *Clinical and Experimental Pharmacology & Physiology* 43, n.º 1 (enero de 2016): 102-15. https://doi.org/10.1111/1440-1681.12500.

Cole, Logan W. «The Evolution of Per-cell Organelle Number». *Frontiers in Cell and Developmental Biology* 4 (18 de agosto de 2016). https://doi.org/10.3389/fcell.2016.00085.

Colli, Licia, Hovirag Lancioni, Irene Cardinali, Anna Olivieri, Marco Rosario Capodiferro, Marco Pellecchia, Marcin Rzepus, et al. «Whole mitochondrial genomes unveil the impact of domestication on goat matrilineal variability». *BMC Genomics* 16 (29 de diciembre de 2015). https://doi.org/10.1186/s12864-015-2342-2.

Compeau, Phillip E. C., Pavel A. Pevzner, y Glenn Tesler. «Why are de Bruijn graphs useful for genome assembly?» *Nature biotechnology* 29, n.º 11 (8 de noviembre de 2011): 987-91. https://doi.org/10.1038/nbt.2023.

Craig, David W., John V. Pearson, Szabolcs Szelinger, Aswin Sekar, Redma Margot, Jason J. Corneveaux, Traci L. Pawlowski, et al. «IDENTIFICATION OF GENETIC VARIANTS USING BARCODED MULTIPLEXED SEQUENCING». *Nature methods* 5, n.º 10 (octubre de 2008): 887-93. https://doi.org/10.1038/nmeth.1251.

Cui, Xin-Gang, Cheng-Yao Liu, Bo Wei, Wen-Jian Zhao, y Wen-Feng Zhang. «Whole Mitochondrial Genome Sequence for an Osteoarthritis Model of Guinea Pig (Caviidae; *Cavia*)». *Mitochondrial DNA. Part A, DNA Mapping, Sequencing, and Analysis* 27, n.º 6 (2016): 3915-16. https://doi.org/10.3109/19401736.2014.987263.

D. R. Notter. «importance of genetic diversity in livestock populations of the future1 | Journal of Animal Science | Oxford Academic». Accedido 15 de julio de 2019. https://academic.oup.com/jas/article-abstract/77/1/61/4625321.

Danecek, Petr, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert E. Handsaker, et al. «The variant call format and VCFtools». *Bioinformatics* 27, n.º 15 (1 de agosto de 2011): 2156-58. https://doi.org/10.1093/bioinformatics/btr330.

Darling, Aaron C.E., Bob Mau, Frederick R. Blattner, y Nicole T. Perna. «Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements». *Genome Research* 14, n.º 7 (julio de 2004): 1394-1403. https://doi.org/10.1101/gr.2289704.

D'Erchia, Anna Maria, Carmela Gissi, Graziano Pesole, Cecilia Saccone, y Ulfur Arnason. «The Guinea-Pig Is Not a Rodent». *Nature* 381, n.º 6583 (junio de 1996): 597-600. https://doi.org/10.1038/381597a0.

Desai, Aarti, Veer Singh Marwah, Akshay Yadav, Vineet Jha, Kishor Dhaygude, Ujwala Bangar, Vivek Kulkarni, y Abhay Jere. «Identification of Optimum Sequencing Depth Especially for De Novo Genome Assembly of Small Genomes Using Next Generation Sequencing Data». *PLOS ONE* 8, n.º 4 (12 de abril de 2013): e60204. https://doi.org/10.1371/journal.pone.0060204.

Días Rivadeneira, Marcela Alexandra Díaz. «ESTUDIO FILOGENÉTICO DE TRES LINEAS DE CUYES (*Cavia porcellus* L.), PERÚ, ANDINA E INTI EN LA HACIENDA "EL PRADO"», s. f., 106.

Dierckxsens, Nicolas, Patrick Mardulyn, y Guillaume Smits. «NOVOPlasty: de novo assembly of organelle genomes from whole genome data». *Nucleic Acids Research* 45, n.º 4 (28 de febrero de 2017): e18. https://doi.org/10.1093/nar/gkw955.

Dillard, E. U., R. Vaccaro, J. Lozano, y O. W. Robison. «Phenotypic and Genetic Parameters for Growth in Guinea Pigs». *Journal of Animal Science* 34, n.º 2 (1 de febrero de 1972): 193-95. https://doi.org/10.2527/jas1972.342193x.

Dominguez Del Angel, Victoria, Erik Hjerde, Lieven Sterck, Salvadors Capella-Gutierrez, Cederic Notredame, Olga Vinnere Pettersson, Joelle Amselem, et al. «Ten steps to get started in Genome Assembly and Annotation». *F1000Research* 7 (5 de febrero de 2018). https://doi.org/10.12688/f1000research.13598.1.

Drummond, A. J., A. Rambaut, B. Shapiro, y O. G. Pybus. «Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences». *Molecular Biology and Evolution* 22, n.º 5 (1 de mayo de 2005): 1185-92. https://doi.org/10.1093/molbev/msi103.

Duan, Mengqin, Jing Tu, y Zuhong Lu. «Recent Advances in Detecting Mitochondrial DNA Heteroplasmic Variations». *Molecules : A Journal of Synthetic Chemistry and Natural Product Chemistry* 23, n.º 2 (3 de febrero de 2018). https://doi.org/10.3390/molecules23020323.

Dunnum, Jonathan L., y Jorge Salazar-Bravo. «Molecular Systematics, Taxonomy and Biogeography of the Genus *Cavia* (Rodentia: Caviidae)». *Journal of Zoological Systematics and Evolutionary Research* 48, n.º 4 (2010): 376-88. https://doi.org/10.1111/j.1439-0469.2009.00561.x.

Earl, Dent A., y Bridgett M. vonHoldt. «STRUCTURE HARVESTER: A Website and Program for Visualizing STRUCTURE Output and Implementing the Evanno Method». *Conservation Genetics Resources* 4, n.º 2 (1 de junio de 2012): 359-61. https://doi.org/10.1007/s12686-011-9548-7.

Earl, Dent, Keith Bradnam, John St. John, Aaron Darling, Dawei Lin, Joseph Fass, Hung On Ken Yu, et al. «Assemblathon 1: A competitive assessment of de novo short read assembly methods». *Genome Research* 21, n.º 12 (diciembre de 2011): 2224-41. https://doi.org/10.1101/gr.126599.111.

Ekblom, Robert, y Jochen B W Wolf. «A field guide to whole-genome sequencing, assembly and annotation». *Evolutionary Applications* 7, n.º 9 (noviembre de 2014): 1026-42. https://doi.org/10.1111/eva.12178.

El, Carla G van, y Martina C Cornel. «Genetic testing and common disorders in a public health framework». *European Journal of Human Genetics* 19, n.º 4 (abril de 2011): 377-81. https://doi.org/10.1038/ejhg.2010.176.

Elliott, Paul, y Tim C. Peakman. «The UK Biobank sample handling and storage protocol for the collection, processing and archiving of human blood and urine». *International Journal of Epidemiology* 37, n.º 2 (1 de abril de 2008): 234-44. https://doi.org/10.1093/ije/dym276.

El-Metwally, Sara, Taher Hamza, Magdi Zakaria, y Mohamed Helmy. «Next-Generation Sequence Assembly: Four Stages of Data Processing and Computational Challenges». *PLoS Computational Biology* 9, n.º 12 (12 de diciembre de 2013). https://doi.org/10.1371/journal.pcbi.1003345.

Elshire, Robert J., Jeffrey C. Glaubitz, Qi Sun, Jesse A. Poland, Ken Kawamoto, Edward S. Buckler, y Sharon E. Mitchell. «A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species». *PLoS ONE* 6, n.º 5 (4 de mayo de 2011). https://doi.org/10.1371/journal.pone.0019379.

Emelyanov, Victor V. «Mitochondrial Connection to the Origin of the Eukaryotic Cell». *European Journal of Biochemistry* 270, n.º 8 (2003): 1599-1618. https://doi.org/10.1046/j.1432-1033.2003.03499.x.

Ewels, Philip, Måns Magnusson, Sverker Lundin, y Max Käller. «MultiQC: Summarize Analysis Results for Multiple Tools and Samples in a Single Report». *Bioinformatics* 32, n.º 19 (1 de octubre de 2016): 3047-48. https://doi.org/10.1093/bioinformatics/btw354.

Ewing, B., L. Hillier, M. C. Wendl, y P. Green. «Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment». *Genome Research* 8, n.º 3 (marzo de 1998): 175-85. https://doi.org/10.1101/gr.8.3.175.

Excoffier, Laurent, y Heidi E. L. Lischer. «Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows». *Molecular Ecology Resources* 10, n.º 3 (2010): 564-67. https://doi.org/10.1111/j.1755-0998.2010.02847.x.

F, Ronquist, Teslenko M, van der Mark P, Ayres Dl, Darling A, Höhna S, Larget B, Liu L, Suchard Ma, y Huelsenbeck Jp. «MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice across a Large Model Space.» *Systematic Biology* 61, n.º 3 (22 de febrero de 2012): 539-42. https://doi.org/10.1093/sysbio/sys029.

Fabre, Pierre-Henri, Lionel Hautier, Dimitar Dimitrov, y Emmanuel J P Douzery. «A glimpse on the pattern of rodent diversification: a phylogenetic approach». *BMC Evolutionary Biology* 12 (14 de junio de 2012): 88. https://doi.org/10.1186/1471-2148-12-88.

Fajardo, Violeta, Isabel González, María Rojas, Teresa García, y Rosario Martín. «A review of current PCR-based methodologies for the authentication of meats from game animal species». *Trends in Food Science & Technology* 21, n.º 8 (1 de agosto de 2010): 408-21. https://doi.org/10.1016/j.tifs.2010.06.002.

Falkenberg, Maria. «Mitochondrial DNA replication in mammalian cells: overview of the pathway». *Essays in Biochemistry* 62, n.º 3 (7 de junio de 2018): 287-96. https://doi.org/10.1042/EBC20170100.

Fernández-Vizarra, Erika, José A. Enríquez, Acisclo Pérez-Martos, Julio Montoya, y Patricio Fernández-Silva. «Tissue-Specific Differences in Mitochondrial Activity and Biogenesis». *Mitochondrion* 11, n.º 1 (enero de 2011): 207-13. https://doi.org/10.1016/j.mito.2010.09.011.

Fitzpatrick, Scott M., y William F. Keegan. «Human Impacts and Adaptations in the Caribbean Islands: An Historical Ecology Approach». *Earth and Environmental Science Transactions of The Royal Society of Edinburgh* 98, n.º 1 (marzo de 2007): 29-45. https://doi.org/10.1017/S1755691007000096.

Flicek, Paul, y Ewan Birney. «Sense from Sequence Reads: Methods for Alignment and Assembly». *Nature Methods* 6, n.º 11 (noviembre de 2009): S6-12. https://doi.org/10.1038/nmeth.1376.

Florez, Jose C. «Genetics and Biobanks Converge to Resolve a Vexing Knowledge Gap in Diabetes». *The Lancet Diabetes & Endocrinology* 6, n.º 2 (1 de febrero de 2018): 87-89. https://doi.org/10.1016/S2213-8587(17)30399-6.

Fonseca, Rute R da, Warren E Johnson, Stephen J O'Brien, Maria João Ramos, y Agostinho Antunes. «The adaptive evolution of the mammalian mitochondrial genome». *BMC Genomics* 9 (4 de marzo de 2008): 119. https://doi.org/10.1186/1471-2164-9-119.

Fonseca, Rute R. da, Warren E. Johnson, Stephen J. O'Brien, Maria João Ramos, y Agostinho Antunes. «The adaptive evolution of the mammalian mitochondrial genome». *BMC Genomics* 9, n.º 1 (4 de marzo de 2008): 119. https://doi.org/10.1186/1471-2164-9-119.

Forni, Giobbe, Guglielmo Puccio, Thomas Bourguignon, Theodore Evans, Barbara Mantovani, Omar Rota-Stabelli, y Andrea Luchetti. «Complete mitochondrial genomes from transcriptomes: assessing pros and cons of data mining for assembling new mitogenomes». *Scientific Reports* 9 (15 de octubre de 2019). https://doi.org/10.1038/s41598-019-51313-7.

Foster, D. Brian, Ting Liu, Kai Kammers, Robert O'Meally, Ni Yang, Kyriakos N. Papanicolaou, C. Conover Talbot, Robert N. Cole, y Brian O'Rourke. «Integrated Omic Analysis of a Guinea Pig Model of Heart Failure and Sudden Cardiac Death». *Journal of Proteome Research* 15, n.º 9 (02 de 2016): 3009-28. https://doi.org/10.1021/acs.jproteome.6b00149.

Freitas, Patrícia Domingues de, Fernando Luis Mendez, Karla Chávez-Congrains, Pedro Manoel Galetti, Luiz Lehmann Coutinho, Alcides Pissinatti, y Carlos Daniel Bustamante. «Next-Generation Sequencing of the Complete Mitochondrial Genome of the Endangered Species Black Lion Tamarin Leontopithecus Chrysopygus (Primates) and Mitogenomic Phylogeny Focusing on the Callitrichidae Family». *G3: Genes, Genomes, Genetics* 8, n.º 6 (1 de junio de 2018): 1985-91. https://doi.org/10.1534/g3.118.200153.

Friedman, Jonathan R., y Jodi Nunnari. «Mitochondrial form and function». *Nature* 505, n.º 7483 (16 de enero de 2014): 335-43. https://doi.org/10.1038/nature12985.

Fu, Limin, Beifang Niu, Zhengwei Zhu, Sitao Wu, y Weizhong Li. «CD-HIT: accelerated for clustering the next-generation sequencing data». *Bioinformatics* 28, n.º 23 (1 de diciembre de 2012): 3150-52. https://doi.org/10.1093/bioinformatics/bts565.

Gabella, Giorgio. «Mitochondria in smooth muscle cells of viscera». *Journal of Smooth Muscle Research* 54 (2018): 51-70. https://doi.org/10.1540/jsmr.54.51.

Garcia, Monique, Jenny Downs, Alyce Russell, y Wei Wang. «Impact of biobanks on research outcomes in rare diseases: a systematic review». *Orphanet Journal of Rare Diseases* 13, n.º 1 (12 de noviembre de 2018): 202. https://doi.org/10.1186/s13023-018-0942-z.

Garrido-Cardenas, Jose Antonio, Federico Garcia-Maroto, Jose Antonio Alvarez-Bermejo, y Francisco Manzano-Agugliaro. «DNA Sequencing Sensors: An Overview». *Sensors (Basel, Switzerland)* 17, n.º 3 (14 de marzo de 2017). https://doi.org/10.3390/s17030588.

Ghodsi, Mohammadreza, Christopher M Hill, Irina Astrovskaya, Henry Lin, Dan D Sommer, Sergey Koren, y Mihai Pop. «De novo likelihood-based measures for comparing genome assemblies». *BMC Research Notes* 6 (22 de agosto de 2013): 334. https://doi.org/10.1186/1756-0500-6-334.

Giovas, Christina M., Michelle J. LeFebvre, y Scott M. Fitzpatrick. «New Records for Prehistoric Introduction of Neotropical Mammals to the West Indies: Evidence from

Carriacou, Lesser Antilles». *Journal of Biogeography* 39, n.º 3 (2012): 476-87. https://doi.org/10.1111/j.1365-2699.2011.02630.x.

Giuffra, E, J M Kijas, V Amarger, O Carlborg, J T Jeon, y L Andersson. «The origin of the domestic pig: independent domestication and subsequent introgression.» *Genetics* 154, n.º 4 (abril de 2000): 1785-91.

Glaubitz, Jeffrey C., Terry M. Casstevens, Fei Lu, James Harriman, Robert J. Elshire, Qi Sun, y Edward S. Buckler. «TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline». *PLoS ONE* 9, n.º 2 (28 de febrero de 2014). https://doi.org/10.1371/journal.pone.0090346.

Gnerre, Sante, Iain MacCallum, Dariusz Przybylski, Filipe J. Ribeiro, Joshua N. Burton, Bruce J. Walker, Ted Sharpe, et al. «High-quality draft assemblies of mammalian genomes from massively parallel sequence data». *Proceedings of the National Academy of Sciences of the United States of America* 108, n.º 4 (25 de enero de 2011): 1513-18. https://doi.org/10.1073/pnas.1017351108.

Godard, Béatrice, Jörg Schmidtke, Jean-Jacques Cassiman, y Ségolène Aymé. «Data Storage and DNA Banking for Biomedical Research: Informed Consent, Confidentiality, Quality Issues, Ownership, Return of Benefits. A Professional Perspective». *European Journal of Human Genetics* 11, n.º S2 (13 de enero de 2004): S88-122. https://doi.org/10.1038/sj.ejhg.5201114.

Goig, Galo A., Silvia Blanco, Alberto L. Garcia-Basteiro, y Iñaki Comas. «Contaminant DNA in bacterial sequencing experiments is a major source of false genetic variability». *BMC Biology* 18, n.º 1 (2 de marzo de 2020): 24. https://doi.org/10.1186/s12915-020-0748-z.

Gossmann, Toni I., Achchuthan Shanmugasundram, Stefan Börno, Ludovic Duvaux, Christophe Lemaire, Heiner Kuhl, Sven Klages, et al. «Ice-Age Climate Adaptations Trap the Alpine Marmot in a State of Low Genetic Diversity». *Current Biology* 29, n.º 10 (20 de mayo de 2019): 1712-1720.e7. https://doi.org/10.1016/j.cub.2019.04.020.

Grant, Jason R., y Paul Stothard. «The CGView Server: a comparative genomics tool for circular genomes». *Nucleic Acids Research* 36, n.º Web Server issue (1 de julio de 2008): W181-84. https://doi.org/10.1093/nar/gkn179.

Gray, Michael W. «Mitochondrial Evolution». *Cold Spring Harbor Perspectives in Biology* 4, n.º 9 (septiembre de 2012). https://doi.org/10.1101/cshperspect.a011403.

Green, Douglas R. «Apoptotic Pathways: The Roads to Ruin». *Cell* 94, n.º 6 (18 de septiembre de 1998): 695-98. https://doi.org/10.1016/S0092-8674(00)81728-6.

Greiner, Stephan, Pascal Lehwark, y Ralph Bock. «OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes». *Nucleic Acids Research* 47, n.º W1 (2 de julio de 2019): W59-64. https://doi.org/10.1093/nar/gkz238.

Grey, Aubrey D. N. J. de. «Mitochondrial Mutations in Mammalian Aging: An over-Hasty about-Turn?» *Rejuvenation Research* 7, n.º 3 (2004): 171-74. https://doi.org/10.1089/rej.2004.7.171.

Guindon, Stéphane, Jean-François Dufayard, Vincent Lefort, Maria Anisimova, Wim Hordijk, y Olivier Gascuel. «New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0». *Systematic Biology* 59, n.º 3 (1 de mayo de 2010): 307-21. https://doi.org/10.1093/sysbio/syq010.

Gurevich, Alexey, Vladislav Saveliev, Nikolay Vyahhi, y Glenn Tesler. «QUAST: Quality Assessment Tool for Genome Assemblies». *Bioinformatics* 29, n.º 8 (15 de abril de 2013): 1072-75. https://doi.org/10.1093/bioinformatics/btt086.

Hahn, Christoph, Lutz Bachmann, y Bastien Chevreux. «Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach». *Nucleic Acids Research* 41, n.º 13 (1 de julio de 2013): e129-e129. https://doi.org/10.1093/nar/gkt371.

Harati, Mitra D., Ryan R. Williams, Masoud Movassaghi, Amin Hojat, Gregory M. Lucey, y William H. Yong. «An Introduction to Starting a Biobank». *Methods in Molecular Biology (Clifton, N.J.)* 1897 (2019): 7-16. https://doi.org/10.1007/978-1-4939-8935-5_2.

Hardison, Ross C. «Comparative Genomics». *PLoS Biology* 1, n.º 2 (noviembre de 2003). https://doi.org/10.1371/journal.pbio.0000058.

Havlak, Paul, Rui Chen, K. James Durbin, Amy Egan, Yanru Ren, Xing-Zhi Song, George M. Weinstock, y Richard A. Gibbs. «The Atlas Genome Assembly System». *Genome Research* 14, n.º 4 (abril de 2004): 721-32. https://doi.org/10.1101/gr.2264004.

Heather, James M., y Benjamin Chain. «The sequence of sequencers: The history of sequencing DNA». *Genomics* 107, n.º 1 (enero de 2016): 1-8. https://doi.org/10.1016/j.ygeno.2015.11.003.

Hel, Olga L. van der, Rob B. van der Luijt, H. B. Bueno de Mesquita, Paul A. H. van Noord, Barbara Slothouber, Mark Roest, Yvonne T. van der Schouw, Diederick E. Grobbee, Peter L. Pearson, y Petra H. M. Peeters. «Quality and Quantity of DNA Isolated from Frozen Urine in Population-Based Research». *Analytical Biochemistry* 304, n.º 2 (15 de mayo de 2002): 206-11. https://doi.org/10.1006/abio.2002.5634.

Henriquez, Fiona L., Thomas A. Richards, Fiona Roberts, Rima McLeod, y Craig W. Roberts. «The Unusual Mitochondrial Compartment of Cryptosporidium Parvum». *Trends in Parasitology* 21, n.º 2 (1 de febrero de 2005): 68-74. https://doi.org/10.1016/j.pt.2004.11.010.

Henry, Philippe, Alison Henry, y Michael A. Russello. «A Noninvasive Hair Sampling Technique to Obtain High Quality DNA from Elusive Small Mammals». *Journal of Visualized Experiments : JoVE*, n.º 49 (13 de marzo de 2011). https://doi.org/10.3791/2791.

Herrmann, Johannes M, y Walter Neupert. «Protein Transport into Mitochondria». *Current Opinion in Microbiology* 3, n.º 2 (1 de abril de 2000): 210-14. https://doi.org/10.1016/S1369-5274(00)00077-1.

Holland, Nina T., Martyn T. Smith, Brenda Eskenazi, y Maria Bastaki. «Biological Sample Collection and Processing for Molecular Epidemiological Studies». *Mutation Research* 543, n.º 3 (junio de 2003): 217-34.

Holt, Ian J., y Aurelio Reyes. «Human Mitochondrial DNA Replication». *Cold Spring Harbor Perspectives in Biology* 4, n.º 12 (diciembre de 2012). https://doi.org/10.1101/cshperspect.a012971.

Hood, Leroy, y Lee Rowen. «The Human Genome Project: big science transforms biology and medicine». *Genome Medicine* 5, n.º 9 (13 de septiembre de 2013): 79. https://doi.org/10.1186/gm483.

Hoogenboom, Bart W., Kitaru Suda, Andreas Engel, y Dimitrios Fotiadis. «The Supramolecular Assemblies of Voltage-Dependent Anion Channels in the Native Membrane». *Journal of Molecular Biology* 370, n.º 2 (6 de julio de 2007): 246-55. https://doi.org/10.1016/j.jmb.2007.04.073.

Huang, Xiaoqiu, y Anup Madan. «CAP3: A DNA Sequence Assembly Program». *Genome Research* 9, n.º 9 (septiembre de 1999): 868-77.

Hubisz, Melissa J., Daniel Falush, Matthew Stephens, y Jonathan K. Pritchard. «Inferring weak population structure with the assistance of sample group information». *Molecular*

ecology resources 9, n.º 5 (septiembre de 2009): 1322-32. https://doi.org/10.1111/j.1755-0998.2009.02591.x.

Hughes, Austin L. «Accumulation of Slightly Deleterious Mutations in the Mitochondrial Genome: a Hallmark of Animal Domestication». *Gene* 515, n.º 1 (15 de febrero de 2013): 28-33. https://doi.org/10.1016/j.gene.2012.11.064.

Hughes, Austin L., y Mary Ann K. Hughes. «Coding Sequence Polymorphism in Avian Mitochondrial Genomes Reflects Population Histories». *Molecular Ecology* 16, n.º 7 (abril de 2007): 1369-76. https://doi.org/10.1111/j.1365-294X.2007.03242.x.

Hunt, Martin, Taisei Kikuchi, Mandy Sanders, Chris Newbold, Matthew Berriman, y Thomas D Otto. «REAPR: a universal tool for genome assembly evaluation». *Genome Biology* 14, n.º 5 (2013): R47. https://doi.org/10.1186/gb-2013-14-5-r47.

Hwang, Kyu-Baek, In-Hee Lee, Honglan Li, Dhong-Geon Won, Carles Hernandez-Ferrer, Jose Alberto Negron, y Sek Won Kong. «Comparative analysis of whole-genome sequencing pipelines to minimize false negative findings». *Scientific Reports* 9 (1 de marzo de 2019). https://doi.org/10.1038/s41598-019-39108-2.

Iqbal, Zamin, Mario Caccamo, Isaac Turner, Paul Flicek, y Gil McVean. «De novo assembly and genotyping of variants using colored de Bruijn graphs». *Nature genetics* 44, n.º 2 (8 de enero de 2012): 226-32. https://doi.org/10.1038/ng.1028.

Ivannikov, Maxim V., y Gregory T. Macleod. «Mitochondrial Free Ca2+ Levels and Their Effects on Energy Metabolism in Drosophila Motor Nerve Terminals». *Biophysical Journal* 104, n.º 11 (4 de junio de 2013): 2353-61. https://doi.org/10.1016/j.bpj.2013.03.064.

Ivanova, Natalia V., Elizabeth L. Clare, y Alex V. Borisenko. «DNA Barcoding in Mammals». *Methods in Molecular Biology (Clifton, N.J.)* 858 (2012): 153-82. https://doi.org/10.1007/978-1-61779-591-6_8.

Jewell, Scott D., Mythily Srinivasan, Linda M. McCart, Nita Williams, William H. Grizzle, Virginia LiVolsi, Greg MacLennan, y Daniel D. Sedmak. «Analysis of the Molecular Quality of Human Tissues: An Experience from the Cooperative Human Tissue Network». *American Journal of Clinical Pathology* 118, n.º 5 (noviembre de 2002): 733-41. https://doi.org/10.1309/VPQL-RT21-X7YH-XDXK.

Jex, Aaron R., D. Timothy J. Littlewood, y Robin B. Gasser. «Toward Next-Generation Sequencing of Mitochondrial Genomes--Focus on Parasitic Worms of Animals and Biotechnological Implications». *Biotechnology Advances* 28, n.º 1 (febrero de 2010): 151-59. https://doi.org/10.1016/j.biotechadv.2009.11.002.

Johnsen, Jill M., Deborah A. Nickerson, y Alex P. Reiner. «Massively Parallel Sequencing: The New Frontier of Hematologic Genomics». *Blood* 122, n.º 19 (7 de noviembre de 2013): 3268-75. https://doi.org/10.1182/blood-2013-07-460287.

Johri, Parul, Georgi K. Marinov, Thomas G. Doak, y Michael Lynch. «Population Genetics of Paramecium Mitochondrial Genomes: Recombination, Mutation Spectrum, and Efficacy of Selection». *Genome Biology and Evolution* 11, n.º 5 (1 de mayo de 2019): 1398-1416. https://doi.org/10.1093/gbe/evz081.

Jombart, Thibaut. «Adegenet: A R Package for the Multivariate Analysis of Genetic Markers». *Bioinformatics* 24, n.º 11 (1 de junio de 2008): 1403-5. https://doi.org/10.1093/bioinformatics/btn129.

Karger, Barry L., y Andras Guttman. «DNA Sequencing by Capillary Electrophoresis». *Electrophoresis* 30, n.º Suppl 1 (junio de 2009): S196-202. https://doi.org/10.1002/elps.200900218.

Kelt, Douglas A., Enrique P. Lessa, James L. Patton, y Jorge Salazar-Bravo. *The Quintessential Naturalist: Honoring the Life and Legacy of Oliver P. Pearson*. University of California Press, 2007.

Kimura, Birgitta K., Michelle J. LeFebvre, Susan D. deFrance, Hilary I. Knodel, Michelle S. Turner, Natalie S. Fitzsimmons, Scott M. Fitzpatrick, y Connie J. Mulligan. «Origin of pre-Columbian guinea pigs from Caribbean archeological sites revealed through genetic analysis». *Journal of Archaeological Science: Reports* 5 (1 de febrero de 2016): 442-52. https://doi.org/10.1016/j.jasrep.2015.12.012.

Kishikawa, Toshihiro, Yukihide Momozawa, Takeshi Ozeki, Taisei Mushiroda, Hidenori Inohara, Yoichiro Kamatani, Michiaki Kubo, y Yukinori Okada. «Empirical evaluation of variant calling accuracy using ultra-deep whole-genome sequencing data». *Scientific Reports* 9 (11 de febrero de 2019). https://doi.org/10.1038/s41598-018-38346-0.

Klont, R. E., L. Brocks, y G. Eikelenboom. «Muscle fibre type and meat quality». *Meat Science*, Meat Consumption and Culture 44th International Congress of Meat Science and Technology, 49 (1 de enero de 1998): S219-29. https://doi.org/10.1016/S0309-1740(98)90050-X.

Knaus, Brian J., y Niklaus J. Grünwald. «Vcfr: A Package to Manipulate and Visualize Variant Call Format Data in R». *Molecular Ecology Resources* 17, n.º 1 (enero de 2017): 44-53. https://doi.org/10.1111/1755-0998.12549.

Koren, Sergey, y Adam M Phillippy. «One Chromosome, One Contig: Complete Microbial Genomes from Long-Read Sequencing and Assembly». *Current Opinion in Microbiology*, Host–microbe interactions: bacteria • Genomics, 23 (1 de febrero de 2015): 110-20. https://doi.org/10.1016/j.mib.2014.11.014.

Kubota, Yasue, Hikaru Hashitani, Hiroyasu Fukuta, Hiroki Kubota, Kenjiro Kohri, y Hikaru Suzuki. «Role of Mitochondria in the Generation of Spontaneous Activity in Detrusor Smooth Muscles of the Guinea Pig Bladder». *The Journal of Urology* 170, n.º 2 Pt 1 (agosto de 2003): 628-33. https://doi.org/10.1097/01.ju.0000069428.46133.d5.

Kukat, Christian, Christian A. Wurm, Henrik Spåhr, Maria Falkenberg, Nils-Göran Larsson, y Stefan Jakobs. «Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA». *Proceedings of the National Academy of Sciences of the United States of America* 108, n.º 33 (16 de agosto de 2011): 13534-39. https://doi.org/10.1073/pnas.1109263108.

Kumar, Sudhir, Glen Stecher, Michael Li, Christina Knyaz, y Koichiro Tamura. «MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms». *Molecular Biology and Evolution* 35, n.º 6 (junio de 2018): 1547-49. https://doi.org/10.1093/molbev/msy096.

Kuznetsov, Andrey V., Sabzali Javadov, Rita Guzun, Michael Grimm, y Valdur Saks. «Cytoskeleton and regulation of mitochondrial function: the role of beta-tubulin II». *Frontiers in Physiology* 4 (22 de abril de 2013). https://doi.org/10.3389/fphys.2013.00082.

Kwon, Sunyoung, Seunghyun Park, Byunghan Lee, y Sungroh Yoon. «In-Depth Analysis of Interrelation between Quality Scores and Real Errors in Illumina Reads». *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference* 2013 (2013): 635-38. https://doi.org/10.1109/EMBC.2013.6609580.

Kyle, R. «New species for meat production». *The Journal of Agricultural Science* 123, n.º 1 (agosto de 1994): 1-8. https://doi.org/10.1017/S0021859600067708.

Ladoukakis, Emmanuel D., y Eleftherios Zouros. «Evolution and inheritance of animal mitochondrial DNA: rules and exceptions». *Journal of Biological Research* 24 (31 de enero de 2017). https://doi.org/10.1186/s40709-017-0060-4.

Laetsch, Dominik R., y Mark L. Blaxter. «BlobTools: Interrogation of Genome Assemblies». *F1000Research* 6 (31 de julio de 2017): 1287. https://doi.org/10.12688/f1000research.12232.1.

Langmead, Ben, y Steven L Salzberg. «Fast gapped-read alignment with Bowtie 2». *Nature methods* 9, n.º 4 (4 de marzo de 2012): 357-59. https://doi.org/10.1038/nmeth.1923.

Larizza, Alessandra, Graziano Pesole, Aurelio Reyes, Elisabetta Sbisà, y Cecilia Saccone. «Lineage Specificity of the Evolutionary Dynamics of the MtDNA D-Loop Region in Rodents». *Journal of Molecular Evolution* 54, n.º 2 (1 de febrero de 2002): 145-55. https://doi.org/10.1007/s00239-001-0063-4.

Larson, Greger, Keith Dobney, Umberto Albarella, Meiying Fang, Elizabeth Matisoo-Smith, Judith Robins, Stewart Lowden, et al. «Worldwide Phylogeography of Wild Boar Reveals Multiple Centers of Pig Domestication». *Science (New York, N.Y.)* 307, n.º 5715 (11 de marzo de 2005): 1618-21. https://doi.org/10.1126/science.1106927.

Lavrov, Dennis V., y Walker Pett. «Animal Mitochondrial DNA as We Do Not Know It: mt-Genome Organization and Evolution in Nonbilaterian Lineages». *Genome Biology and Evolution* 8, n.º 9 (24 de agosto de 2016): 2896-2913. https://doi.org/10.1093/gbe/evw195.

Leggett, Richard M., Ricardo H. Ramirez-Gonzalez, Bernardo J. Clavijo, Darren Waite, y Robert P. Davey. «Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics». *Frontiers in Genetics* 4 (17 de diciembre de 2013). https://doi.org/10.3389/fgene.2013.00288.

Legros, Frédéric, Florence Malka, Paule Frachon, Anne Lombès, y Manuel Rojo. «Organization and Dynamics of Human Mitochondrial DNA». *Journal of Cell Science* 117, n.º 13 (1 de junio de 2004): 2653-62. https://doi.org/10.1242/jcs.01134.

Leigh, Jessica W., y David Bryant. «Popart: Full-Feature Software for Haplotype Network Construction». *Methods in Ecology and Evolution* 6, n.º 9 (2015): 1110-16. https://doi.org/10.1111/2041-210X.12410.

Li, Heng. «A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data». *Bioinformatics* 27, n.º 21 (1 de noviembre de 2011): 2987-93. https://doi.org/10.1093/bioinformatics/btr509.

Li, Heng, y Richard Durbin. «Fast and accurate short read alignment with Burrows–Wheeler transform». *Bioinformatics* 25, n.º 14 (15 de julio de 2009): 1754-60. https://doi.org/10.1093/bioinformatics/btp324.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, y Richard Durbin. «The Sequence Alignment/Map format and SAMtools». *Bioinformatics* 25, n.º 16 (15 de agosto de 2009): 2078-79. https://doi.org/10.1093/bioinformatics/btp352.

Li, Heng, Robert E. Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, y Richard Durbin. «The Sequence Alignment/Map Format and SAMtools». *Bioinformatics*, 2009. https://doi.org/10.1093/bioinformatics/btp352.

Li, Ruiqiang, Hongmei Zhu, Jue Ruan, Wubin Qian, Xiaodong Fang, Zhongbin Shi, Yingrui Li, et al. «De novo assembly of human genomes with massively parallel short read sequencing». *Genome Research* 20, n.º 2 (febrero de 2010): 265-72. https://doi.org/10.1101/gr.097261.109.

Li, Zhentang, Yi Wang, y Fei Wang. «A study on fast calling variants from next-generation sequencing data using decision tree». *BMC Bioinformatics* 19 (19 de abril de 2018). https://doi.org/10.1186/s12859-018-2147-9.

Li, Zhenyu, Yanxiang Chen, Desheng Mu, Jianying Yuan, Yujian Shi, Hao Zhang, Jun Gan, et al. «Comparison of the Two Major Classes of Assembly Algorithms: Overlap-Layout-Consensus and de-Bruijn-Graph». *Briefings in Functional Genomics* 11, n.º 1 (enero de 2012): 25-37. https://doi.org/10.1093/bfgp/elr035.

Lichardusova, Lucia, Zuzana Tatarkova, Andrea Calkovska, Daniela Mokra, Ivan Engler, Peter Racay, Jan Lehotsky, y Peter Kaplan. «Proteomic Analysis of Mitochondrial Proteins in the Guinea Pig Heart Following Long-Term Normobaric Hyperoxia». *Molecular and Cellular Biochemistry* 434, n.º 1-2 (octubre de 2017): 61-73. https://doi.org/10.1007/s11010-017-3037-6.

Lightowlers, R. N., P. F. Chinnery, D. M. Turnbull, y N. Howell. «Mammalian Mitochondrial Genetics: Heredity, Heteroplasmy and Disease». *Trends in Genetics: TIG* 13, n.º 11 (noviembre de 1997): 450-55. https://doi.org/10.1016/s0168-9525(97)01266-3.

Lischer, Heidi E. L., y Kentaro K. Shimizu. «Reference-guided de novo assembly approach improves genome reconstruction for related species». *BMC Bioinformatics* 18 (10 de noviembre de 2017). https://doi.org/10.1186/s12859-017-1911-6.

Liu, Angen, y Kai Pollard. «Biobanking for Personalized Medicine». *Advances in Experimental Medicine and Biology* 864 (2015): 55-68. https://doi.org/10.1007/978-3-319-20579-3_5.

Liu, Yi-Ping, Gui-Sheng Wu, Yong-Gang Yao, Yong-Wang Miao, Gordon Luikart, Mumtaz Baig, Albano Beja-Pereira, Zhao-Li Ding, Malliya Gounder Palanichamy, y Ya-Ping Zhang. «Multiple Maternal Origins of Chickens: Out of the Asian Jungles». *Molecular Phylogenetics and Evolution* 38, n.º 1 (enero de 2006): 12-19. https://doi.org/10.1016/j.ympev.2005.09.014.

Logan, David C. «Mitochondrial fusion, division and positioning in plants». *Biochemical Society Transactions* 38, n.º 3 (1 de junio de 2010): 789-95. https://doi.org/10.1042/BST0380789.

Lord, E., C. Collins, S. deFrance, M. J. LeFebvre, F. Pigière, P. Eeckhout, C. Erauw, et al. «Ancient DNA of Guinea Pigs (*Cavia* Spp.) Indicates a Probable New Center of Domestication and Pathways of Global Distribution». *Scientific Reports* 10, n.º 1 (1 de junio de 2020): 8901. https://doi.org/10.1038/s41598-020-65784-6.

Lord, Edana, Catherine Collins, Susan deFrance, Michelle J. LeFebvre, y Elizabeth Matisoo-Smith. «Complete mitogenomes of ancient Caribbean Guinea pigs (*Cavia porcellus*)». *Journal of Archaeological Science: Reports* 17 (1 de febrero de 2018): 678-88. https://doi.org/10.1016/j.jasrep.2017.12.004.

Lowe, Todd M., y Patricia P. Chan. «tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes». *Nucleic Acids Research* 44, n.º Web Server issue (8 de julio de 2016): W54-57. https://doi.org/10.1093/nar/gkw413.

Luo, Ruibang, Binghang Liu, Yinlong Xie, Zhenyu Li, Weihua Huang, Jianying Yuan, Guangzhu He, et al. «SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler». *GigaScience* 1 (27 de diciembre de 2012): 18. https://doi.org/10.1186/2047-217X-1-18.

Ma, Chuan, Pengcheng Yang, Feng Jiang, Marie-Pierre Chapuis, Yasen Shali, Gregory A. Sword, y Le Kang. «Mitochondrial Genomes Reveal the Global Phylogeography and Dispersal Routes of the Migratory Locust». *Molecular Ecology* 21, n.º 17 (septiembre de 2012): 4344-58. https://doi.org/10.1111/j.1365-294X.2012.05684.x.

Manchanda, Nancy, John L. Portwood, Margaret R. Woodhouse, Arun S. Seetharam, Carolyn J. Lawrence-Dill, Carson M. Andorf, y Matthew B. Hufford. «GenomeQC: a quality assessment tool for genome assemblies and gene structure annotations». *BMC Genomics* 21 (2 de marzo de 2020). https://doi.org/10.1186/s12864-020-6568-2.

Manekar, Swati C, y Shailesh R Sathe. «A benchmark study of k-mer counting methods for high-throughput sequencing». *GigaScience* 7, n.º 12 (22 de octubre de 2018). https://doi.org/10.1093/gigascience/giy125.

Mannella, Carmen A. «Structure and Dynamics of the Mitochondrial Inner Membrane Cristae». *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, Mitochondrial Dynamics in Cell Life and Death, 1763, n.º 5 (1 de mayo de 2006): 542-48. https://doi.org/10.1016/j.bbamcr.2006.04.006.

Marande, William, y Gertraud Burger. «Mitochondrial DNA as a Genomic Jigsaw Puzzle». *Science (New York, N.Y.)* 318, n.º 5849 (19 de octubre de 2007): 415. https://doi.org/10.1126/science.1148033.

Marçais, Guillaume, Arthur L. Delcher, Adam M. Phillippy, Rachel Coston, Steven L. Salzberg, y Aleksey Zimin. «MUMmer4: A fast and versatile genome alignment system». *PLoS Computational Biology* 14, n.º 1 (26 de enero de 2018). https://doi.org/10.1371/journal.pcbi.1005944.

Marçais, Guillaume, y Carl Kingsford. «A fast, lock-free approach for efficient parallel counting of occurrences of *k-mers*». *Bioinformatics* 27, n.º 6 (15 de marzo de 2011): 764-70. https://doi.org/10.1093/bioinformatics/btr011.

Matijn, Joran, Julian Vosseberg, Lionel Guy, Pierre Offre, y Thijs J. G. Ettema. «Origen mitocondrial profundo fuera de las alfaproteobacterias muestreadas | Naturaleza», 21 de octubre de 2019. https://www.nature.com/articles/s41586-018-0059-5.

McBride, Heidi M., Margaret Neuspiel, y Sylwia Wasiak. «Mitochondria: More than Just a Powerhouse». *Current Biology: CB* 16, n.º 14 (25 de julio de 2006): R551-560. https://doi.org/10.1016/j.cub.2006.06.054.

McCombie, W. Richard, John D. McPherson, y Elaine R. Mardis. «Next-Generation Sequencing Technologies». *Cold Spring Harbor Perspectives in Medicine*, 26 de noviembre de 2018. https://doi.org/10.1101/cshperspect.a036798.

McKenna, Aaron, Matthew Hanna, Eric Banks, Andrey Sivachenko, Kristian Cibulskis, Andrew Kernytsky, Kiran Garimella, et al. «The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data». *Genome Research* 20, n.º 9 (septiembre de 2010): 1297-1303. https://doi.org/10.1101/gr.107524.110.

McMillin, Jeanie B, y William Dowhan. «Cardiolipin and Apoptosis». *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, Lipids in apoptosis, 1585, n.º 2 (30 de diciembre de 2002): 97-107. https://doi.org/10.1016/S1388-1981(02)00329-3.

Meng, Guanliang, Yiyuan Li, Chentao Yang, y Shanlin Liu. «MitoZ: A Toolkit for Animal Mitochondrial Genome Assembly, Annotation and Visualization». *Nucleic Acids Research* 47, n.º 11 (20 de junio de 2019): e63-e63. https://doi.org/10.1093/nar/gkz173.

Meza R, Elmer, Yiyson Rojas Y, Jorge Raymondi C, y Cesar Olaguivel F. «Estimación de heterosis individual en cuyes (*Cavia porcellus*) F1 cruces de los genotipos Peru, Inti y Andina». *Revista de Investigaciones Veterinarias del Perú* 29, n.º 2 (abril de 2018): 495-506. https://doi.org/10.15381/rivep.v29i2.14488.

Miller, Jason R., Sergey Koren, y Granger Sutton. «Assembly Algorithms for Next-Generation Sequencing Data». *Genomics* 95, n.º 6 (junio de 2010): 315-27. https://doi.org/10.1016/j.ygeno.2010.03.001.

Montero, Raquel Gil. «Reseña de Reducciones. La concentración forzada de las poblaciones indígenas en el Virreinato del Perú, de Akira Saito y Claudia Rosas Lauro (eds.). Lima : Colección Estudios Andinos, National Museum of Ethnology, Pontificia Universidad Católica del Perú, 2017. 680 págs.» *Corpus. Archivos virtuales de la alteridad americana*, n.º Vol. 8, No 1 (17 de junio de 2018). https://doi.org/10.4000/corpusarchivos.2013.

Morales M., Augusto, Fernando Carcelén C., Miguel Ara G., Teresa Arbaiza F., y Lilia Chauca F. «Evaluación de dos niveles de energía en el comportamiento productivo de cuyes (*cavia porcellus*) de la raza Perú». *Revista de Investigaciones Veterinarias del Perú* 22, n.º 3 (septiembre de 2011): 177-82.

Mouchaty, S. K., F. Catzeflis, A. Janke, y U. Arnason. «Molecular Evidence of an African Phiomorpha-South American Caviomorpha Clade and Support for Hystricognathi Based on the Complete Mitochondrial Genome of the Cane Rat (Thryonomys Swinderianus)». *Molecular Phylogenetics and Evolution* 18, n.º 1 (enero de 2001): 127-35. https://doi.org/10.1006/mpev.2000.0870.

Müller, Miklós, y William Martin. «The genome of Rickettsia prowazekii and some thoughts on the origin of mitochondria and hydrogenosomes - Müller - 1999 - BioEssays - Wiley Online Library», 21 de octubre de 2019. https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291521-1878%28199905%2921%3A5%3C377%3A%3AAID-BIES4%3E3.0.CO%3B2-W.

Mullikin, James C. «The evolution of comparative genomics». *Molecular Genetics & Genomic Medicine* 2, n.º 5 (septiembre de 2014): 363-68. https://doi.org/10.1002/mgg3.112.

Muruve, Daniel A., Michelle C. Mann, Kevin Chapman, Josee F. Wong, Pietro Ravani, Stacey A. Page, y Hallgrimur Benediktsson. «The biobank for the molecular classification of kidney disease: research translation and precision medicine in nephrology». *BMC Nephrology* 18, n.º 1 (26 de julio de 2017): 252. https://doi.org/10.1186/s12882-017-0669-4.

Muti, Paola. «[Biological specimen banks: epidemiological tools and scientific enterprises]». *Epidemiologia E Prevenzione* 27, n.º 2 (abril de 2003): 121-25.

Myers, E. W., G. G. Sutton, A. L. Delcher, I. M. Dew, D. P. Fasulo, M. J. Flanigan, S. A. Kravitz, et al. «A Whole-Genome Assembly of Drosophila». *Science (New York, N.Y.)* 287, n.º 5461 (24 de marzo de 2000): 2196-2204. https://doi.org/10.1126/science.287.5461.2196.

Nabholz, Benoit, Jean-François Mauffrey, Eric Bazin, Nicolas Galtier, y Sylvain Glemin. «Determination of Mitochondrial Genetic Diversity in Mammals». *Genetics* 178, n.º 1 (1 de enero de 2008): 351-61. https://doi.org/10.1534/genetics.107.073346.

Naderi, Saeid, Hamid-Reza Rezaei, François Pompanon, Michael G. B. Blum, Riccardo Negrini, Hamid-Reza Naghash, Özge Balkız, et al. «The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals». *Proceedings of the National Academy of Sciences of the United States of America* 105, n.º 46 (18 de noviembre de 2008): 17659-64. https://doi.org/10.1073/pnas.0804782105.

Nakamura, Eri, Yoshihiko Kito, Hikaru Hashitani, y Hikaru Suzuki. «Metabolic Component of the Temperature-Sensitivity of Slow Waves Recorded from Gastric Muscle of the Guinea-Pig». *Journal of Smooth Muscle Research = Nihon Heikatsukin Gakkai Kikanshi* 42, n.º 1 (febrero de 2006): 33-48.

Nederhand, R. J., S. Droog, C. Kluft, M. L. Simoons, M. P. M. de Maat, y Investigators of the EUROPA trial. «Logistics and Quality Control for DNA Sampling in Large Multicenter

Studies». *Journal of Thrombosis and Haemostasis: JTH* 1, n.º 5 (mayo de 2003): 987-91.

Ngoula, Ferdinand, Maryvonne Guemdjo Tekam, Augustave Kenfack, Cyrille D'Alex Tadondjou Tchingo, Sandrine Nouboudem, Herman Ngoumtsop, Borice Tsafack, et al. «Effects of heat stress on some reproductive parameters of male cavie (*Cavia porcellus*) and mitigation strategies using guava (Psidium guajava) leaves essential oil». *Journal of Thermal Biology* 64 (1 de febrero de 2017): 67-72. https://doi.org/10.1016/j.jtherbio.2017.01.001.

Ning, Kang, y Hon Wai Leong. «Towards a better solution to the shortest common supersequence problem: the deposition and reduction algorithm». *BMC Bioinformatics* 7, n.º Suppl 4 (12 de diciembre de 2006): S12. https://doi.org/10.1186/1471-2105-7-S4-S12.

Okonechnikov, Konstantin, Olga Golosova, y Mikhail Fursov. «Unipro UGENE: A Unified Bioinformatics Toolkit». *Bioinformatics* 28, n.º 8 (15 de abril de 2012): 1166-67. https://doi.org/10.1093/bioinformatics/bts091.

Opazo, Juan C. «A Molecular Timescale for Caviomorph Rodents (Mammalia, Hystricognathi)». *Molecular Phylogenetics and Evolution* 37, n.º 3 (1 de diciembre de 2005): 932-37. https://doi.org/10.1016/j.ympev.2005.05.002.

Padilla-Carlin, Danielle J., David N. McMurray, y Anthony J. Hickey. «The Guinea Pig as a Model of Infectious Diseases». *Comparative Medicine* 58, n.º 4 (agosto de 2008): 324-40.

Pagliarini, David J., y Jared Rutter. «Hallmarks of a new era in mitochondrial biochemistry». *Genes & Development* 27, n.º 24 (15 de diciembre de 2013): 2615-27. https://doi.org/10.1101/gad.229724.113.

Paradis, Emmanuel, y Klaus Schliep. «Ape 5.0: An Environment for Modern Phylogenetics and Evolutionary Analyses in R». *Bioinformatics* 35, n.º 3 (1 de febrero de 2019): 526-28. https://doi.org/10.1093/bioinformatics/bty633.

Patil, Shankargouda, Barnali Majumdar, Kamran Habib Awan, Gargi S. Sarode, Sachin C. Sarode, Amol R. Gadbail, y Shailesh Gondivkar. «Cancer Oriented Biobanks: A Comprehensive Review». *Oncology Reviews* 12, n.º 1 (30 de enero de 2018). https://doi.org/10.4081/oncol.2018.357.

Patowary, Ashok, Ryan Nesbitt, Marilyn Archer, Raphael Bernier, y Zoran Brkanac. «Next Generation Sequencing Mitochondrial DNA Analysis in Autism Spectrum Disorder». *Autism Research* 10, n.º 8 (agosto de 2017): 1338-43. https://doi.org/10.1002/aur.1792.

Peakman, Tim C., y Paul Elliott. «The UK Biobank Sample Handling and Storage Validation Studies». *International Journal of Epidemiology* 37 Suppl 1 (abril de 2008): i2-6. https://doi.org/10.1093/ije/dyn019.

Pearce, Sarah F., Pedro Rebelo-Guiomar, Aaron R. D'Souza, Christopher A. Powell, Lindsey Van Haute, y Michal Minczuk. «Regulation of Mammalian Mitochondrial Gene Expression: Recent Advances». *Trends in Biochemical Sciences* 42, n.º 8 (1 de agosto de 2017): 625-39. https://doi.org/10.1016/j.tibs.2017.02.003.

Peralta, Susana, Xiao Wang, y Carlos T. Moraes. «Mitochondrial transcription: Lessons from mouse models». *Biochimica et biophysica acta* 1819, n.º 9-10 (septiembre de 2012): 961-69. https://doi.org/10.1016/j.bbagrm.2011.11.001.

Peterson, Brant K., Jesse N. Weber, Emily H. Kay, Heidi S. Fisher, y Hopi E. Hoekstra. «Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species». *PLOS ONE* 7, n.º 5 (31 de mayo de 2012): e37135. https://doi.org/10.1371/journal.pone.0037135.

Pevzner, Pavel A., Haixu Tang, y Michael S. Waterman. «An Eulerian path approach to DNA fragment assembly». *Proceedings of the National Academy of Sciences of the United States of America* 98, n.º 17 (14 de agosto de 2001): 9748-53. https://doi.org/10.1073/pnas.171285098.

Pfeiffer, Franziska, Carsten Gröber, Michael Blank, Kristian Händler, Marc Beyer, Joachim L. Schultze, y Günter Mayer. «Systematic Evaluation of Error Rates and Causes in Short Samples in Next-Generation Sequencing». *Scientific Reports* 8, n.º 1 (19 de julio de 2018): 1-14. https://doi.org/10.1038/s41598-018-29325-6.

Piras, Daniela, Maria Grazia Doro, Giuseppina Casu, Paola Maria Melis, Simona Vaccargiu, Ignazio Piras, Debora Parracciani, et al. «Haplotype Affinities Resolve a Major Component of Goat (Capra hircus) MtDNA D-Loop Diversity and Reveal Specific Features of the Sardinian Stock». *PLoS ONE* 7, n.º 2 (17 de febrero de 2012). https://doi.org/10.1371/journal.pone.0030785.

Poland, Jesse A., Patrick J. Brown, Mark E. Sorrells, y Jean-Luc Jannink. «Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by-Sequencing Approach». *PLoS ONE* 7, n.º 2 (28 de febrero de 2012). https://doi.org/10.1371/journal.pone.0032253.

Portilla, Carlos Solarte, Heiber Cárdenas Henao, Carolina Rosero Galindo, y William Burgos Paz. «Caracterización molecular de tres líneas de *Cavia porcellus* mediante la aplicación de AFLP». *Revista Colombiana de Ciencias Pecuarias* 20, n.º 1 (2007): 49-58.

Quijandria, B., L. Chauca de Zaldivar, y O. W. Robison. «Selection in Guinea Pigs: I. Estimation of Phenotypic and Genetic Parameters for Litter Size and Body Weight». *Journal of Animal Science* 56, n.º 4 (abril de 1983): 814-19.

Quinlan, Aaron R., y Ira M. Hall. «BEDTools: a flexible suite of utilities for comparing genomic features». *Bioinformatics* 26, n.º 6 (15 de marzo de 2010): 841-42. https://doi.org/10.1093/bioinformatics/btq033.

Rajmil, Luis, Lilisbeth Perestelo-Pérez, y Michael Herdman. «Quality of Life and Rare Diseases». *Advances in Experimental Medicine and Biology* 686 (2010): 251-72. https://doi.org/10.1007/978-90-481-9485-8_15.

Rambaut, A. «GitHub - rambaut/figtree: Automatically exported from code.google.com/p/figtree», 2012. https://github.com/rambaut/figtree.

Rambaut, Andrew, Alexei J. Drummond, Dong Xie, Guy Baele, y Marc A. Suchard. «Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7». *Systematic Biology* 67, n.º 5 (1 de septiembre de 2018): 901-4. https://doi.org/10.1093/sysbio/syy032.

Ramirez, Dámaso W., Manuel Quispe-López, Diego Marcelo-Carranza, y Víctor Pacheco. «Primer reporte de albinismo para el cuy silvestre *Cavia tschudii* (Mammalia: Rodentia)». *Revista Peruana de Biología* 26, n.º 4 (15 de diciembre de 2019): 521-24. https://doi.org/10.15381/rpb.v26i4.17218.

Riedel, Stefan. «Edward Jenner and the history of smallpox and vaccination». Proceedings (Baylor University. Medical Center) 18, n.º 1 (enero de 2005): 21-25.

Rodríguez L, Humberto, Mabel Palomino T, Víctor Hidalgo L, y Gustavo Gutiérrez R. «Efectos de factores fijos y al azar sobre el peso al nacimiento y al destete en cuyes de la costa central del Perú». *Revista de Investigaciones Veterinarias del Perú* 24, n.º 1 (2013): 16-24.

Ross, Michael G, Carsten Russ, Maura Costello, Andrew Hollinger, Niall J Lennon, Ryan Hegarty, Chad Nusbaum, y David B Jaffe. «Characterizing and measuring bias in sequence data». *Genome Biology* 14, n.º 5 (2013): R51. https://doi.org/10.1186/gb-2013-14-5-r51.

Rozas, Julio, Albert Ferrer-Mata, Juan Carlos Sánchez-DelBarrio, Sara Guirao-Rico, Pablo Librado, Sebastián E. Ramos-Onsins, y Alejandro Sánchez-Gracia. «DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets». *Molecular Biology and Evolution* 34, n.º 12 (1 de diciembre de 2017): 3299-3302. https://doi.org/10.1093/molbev/msx248.

Ruffalo, Matthew, Mehmet Koyutürk, Soumya Ray, y Thomas LaFramboise. «Accurate estimation of short read mapping quality for next-generation genome sequencing». *Bioinformatics* 28, n.º 18 (15 de septiembre de 2012): i349-55. https://doi.org/10.1093/bioinformatics/bts408.

Ruffalo, Matthew, Thomas LaFramboise, y Mehmet Koyutürk. «Comparative Analysis of Algorithms for Next-Generation Sequencing Read Alignment». *Bioinformatics (Oxford, England)* 27, n.º 20 (15 de octubre de 2011): 2790-96. https://doi.org/10.1093/bioinformatics/btr477.

Rustagi, Navin, Anbo Zhou, W. Scott Watkins, Erika Gedvilaite, Shuoguo Wang, Naveen Ramesh, Donna Muzny, et al. «Extremely low-coverage whole genome sequencing in South Asians captures population genomics information». *BMC Genomics* 18 (22 de mayo de 2017). https://doi.org/10.1186/s12864-017-3767-6.

Sachidanandam, Ravi, David Weissman, Steven C. Schmidt, Jerzy M. Kakol, Lincoln D. Stein, Gabor Marth, Steve Sherry, et al. «A Map of Human Genome Sequence Variation Containing 1.42 Million Single Nucleotide Polymorphisms». *Nature* 409, n.º 6822 (febrero de 2001): 928-33. https://doi.org/10.1038/35057149.

Salzberg, Steven L., Adam M. Phillippy, Aleksey Zimin, Daniela Puiu, Tanja Magoc, Sergey Koren, Todd J. Treangen, et al. «GAGE: A critical evaluation of genome assemblies and assembly algorithms». *Genome Research* 22, n.º 3 (marzo de 2012): 557-67. https://doi.org/10.1101/gr.131383.111.

Sánchez-Macías, Davinia, Lida Barba-Maggi, Antonio Morales-delaNuez, y Julio Palmay-Paredes. «Guinea Pig for Meat Production: A Systematic Review of Factors Affecting the Production, Carcass and Meat Quality». *Meat Science* 143 (septiembre de 2018): 165-76. https://doi.org/10.1016/j.meatsci.2018.05.004.

Sanger, F., S. Nicklen, y A. R. Coulson. «DNA sequencing with chain-terminating inhibitors». *Proceedings of the National Academy of Sciences of the United States of America* 74, n.º 12 (diciembre de 1977): 5463-67.

Sato, Naoki. «Revisiting the Theoretical Basis of the Endosymbiotic Origin of Plastids in the Original Context of Lynn Margulis on the Origin of Mitosing, Eukaryotic Cells». *Journal of Theoretical Biology* 434 (07 de 2017): 104-13. https://doi.org/10.1016/j.jtbi.2017.08.028.

Schattner, Peter, Angela N. Brooks, y Todd M. Lowe. «The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs». *Nucleic Acids Research* 33, n.º Web Server issue (1 de julio de 2005): W686-89. https://doi.org/10.1093/nar/gki366.

Schliep, Klaus Peter. «phangorn: phylogenetic analysis in R». *Bioinformatics* 27, n.º 4 (15 de febrero de 2011): 592-93. https://doi.org/10.1093/bioinformatics/btq706.

Schmieder, Robert, y Robert Edwards. «Quality control and preprocessing of metagenomic datasets». *Bioinformatics* 27, n.º 6 (15 de marzo de 2011): 863-64. https://doi.org/10.1093/bioinformatics/btr026.

Schwartz, Scott, W. James Kent, Arian Smit, Zheng Zhang, Robert Baertsch, Ross C. Hardison, David Haussler, y Webb Miller. «Human–Mouse Alignments with BLASTZ». *Genome Research* 13, n.º 1 (1 de enero de 2003): 103-7. https://doi.org/10.1101/gr.809403.

Sengupta, Srikumar, Jennifer M. Bolin, Victor Ruotti, Bao Kim Nguyen, James A. Thomson, Angela L. Elwell, y Ron Stewart. «Single Read and Paired End mRNA-Seq Illumina Libraries from 10 Nanograms Total RNA». *Journal of Visualized Experiments : JoVE*, n.º 56 (27 de octubre de 2011). https://doi.org/10.3791/3340.

Shen, Wei, Shuai Le, Yan Li, y Fuquan Hu. «SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation». *PLoS ONE* 11, n.º 10 (5 de octubre de 2016). https://doi.org/10.1371/journal.pone.0163962.

Simão, Felipe A., Robert M. Waterhouse, Panagiotis Ioannidis, Evgenia V. Kriventseva, y Evgeny M. Zdobnov. «BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs». *Bioinformatics* 31, n.º 19 (1 de octubre de 2015): 3210-12. https://doi.org/10.1093/bioinformatics/btv351.

Simpson, Jared T., y Richard Durbin. «Efficient de novo assembly of large genomes using compressed data structures». *Genome Research* 22, n.º 3 (marzo de 2012): 549-56. https://doi.org/10.1101/gr.126953.111.

Simpson, Jared T., Kim Wong, Shaun D. Jackman, Jacqueline E. Schein, Steven J.M. Jones, y İnanç Birol. «ABySS: A parallel assembler for short read sequence data». *Genome Research* 19, n.º 6 (junio de 2009): 1117-23. https://doi.org/10.1101/gr.089532.108.

Sirand-Pugnet, Pascal, Carole Lartigue, Marc Marenda, Daniel Jacob, Aurélien Barré, Valérie Barbe, Chantal Schenowitz, et al. «Being Pathogenic, Plastic, and Sexual while Living with a Nearly Minimal Bacterial Genome». *PLoS Genetics* 3, n.º 5 (mayo de 2007). https://doi.org/10.1371/journal.pgen.0030075.

Slatko, Barton E., Andrew F. Gardner, y Frederick M. Ausubel. «Overview of Next Generation Sequencing Technologies». *Current protocols in molecular biology* 122, n.º 1 (abril de 2018): e59. https://doi.org/10.1002/cpmb.59.

Smith, David Roy. «RNA-Seq Data: A Goldmine for Organelle Research». *Briefings in Functional Genomics* 12, n.º 5 (septiembre de 2013): 454-56. https://doi.org/10.1093/bfgp/els066.

Smith, David Roy, y Patrick J. Keeling. «Mitochondrial and Plastid Genome Architecture: Reoccurring Themes, but Significant Differences at the Extremes». *Proceedings of the National Academy of Sciences of the United States of America* 112, n.º 33 (18 de agosto de 2015): 10177. https://doi.org/10.1073/pnas.1422049112.

Sobrero, Raúl, Oscar Inostroza-Michael, Cristián E. Hernández, y Luis A. Ebensperger. «Phylogeny Modulates the Effects of Ecological Conditions on Group Living across Hystricognath Rodents». *Animal Behaviour* 94 (1 de agosto de 2014): 27-34. https://doi.org/10.1016/j.anbehav.2014.05.008.

Sohn, Jang-Il, y Jin-Wu Nam. «The Present and Future of de Novo Whole-Genome Assembly». *Briefings in Bioinformatics* 19, n.º 1 (01 de 2018): 23-40. https://doi.org/10.1093/bib/bbw096.

Solarte, Carlos, Heiber Cárdenas Henao, Carol Rosero Galindo, y William Burgos Paz. «Caracterización Molecular de Tres Líneas de *Cavia Porcellus* Mediante La Aplicación de AFLP». *Revista Colombiana de Ciencias Pecuarias* 20, n.º 1 (2007): 49-58.

Song, Hong, Bhanu P. Telugu, y Loren P. Thompson. «Sexual Dimorphism of Mitochondrial Function in the Hypoxic Guinea Pig Placenta». *Biology of Reproduction* 100, n.º 1 (1 de enero de 2019): 208-16. https://doi.org/10.1093/biolre/ioy167.

Spotorno, A. E., J. C. Marín, G. Manríquez, J. P. Valladares, E. Rico, y C. Rivas. «Ancient and Modern Steps during the Domestication of Guinea Pigs (*Cavia Porcellus* L.)». *Journal of Zoology* 270, n.º 1 (2006): 57-62. https://doi.org/10.1111/j.1469-7998.2006.00117.x.

Spotorno, A. E., J. C. Marin, J. P. Valladares, E. Rico, y C. Rivas. «Ancient and modern steps during the domestication of guinea pigs *(Cavia porcellus* L.)», 2006. https://zslpublications.onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-7998.2006.00117.x.

Spotorno, Ángel E., John P. Valladares, Juan C. Marín, y Horacio Zeballos. «Diversidad molecular entre cuyes domésticos (Cavia porcellus) y su relación filogenética cercana con la especie silvestre andina Cavia tschudii». *Revista chilena de historia natural* 77, n.º 2 (junio de 2004): 243-50. https://doi.org/10.4067/S0716-078X2004000200004.

Stamatakis, Alexandros. «RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies». *Bioinformatics* 30, n.º 9 (1 de mayo de 2014): 1312-13. https://doi.org/10.1093/bioinformatics/btu033.

Stein, Catherine M., Lindsay Sausville, Christian Wejse, Rafal S. Sobota, Nicola M. Zetola, Philip C. Hill, W. Henry Boom, William K. Scott, Giorgio Sirugo, y Scott M. Williams. «Genomics of human pulmonary tuberculosis: from genes to pathways». *Current genetic medicine reports* 5, n.º 4 (diciembre de 2017): 149-66. https://doi.org/10.1007/s40142-017-0130-9.

Suchard, Marc A, Philippe Lemey, Guy Baele, Daniel L Ayres, Alexei J Drummond, y Andrew Rambaut. «Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10». *Virus Evolution* 4, n.º 1 (8 de junio de 2018). https://doi.org/10.1093/ve/vey016.

Sullivan, Kevin A. M., Roy N. Platt, Robert D. Bradley, y David A. Ray. «Whole mitochondrial genomes provide increased resolution and indicate paraphyly in deer mice». *BMC Zoology* 2, n.º 1 (18 de agosto de 2017): 11. https://doi.org/10.1186/s40850-017-0020-3.

Sun, Ting, Guangyun Huang, Junli Sun, Zihao Wang, Shaohua Teng, Yanhong Cao, Quratulain Hanif, Ningbo Chen, Chuzhao Lei, y Yuying Liao. «Mitogenome Diversity and Maternal Origins of Guangxi Buffalo Breeds». *Animals* 10, n.º 4 (abril de 2020): 547. https://doi.org/10.3390/ani10040547.

Taanman, Jan-Willem. «The mitochondrial genome: structure, transcription, translation and replication». *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1410, n.º 2 (9 de febrero de 1999): 103-23. https://doi.org/10.1016/S0005-2728(98)00161-3.

Tamura, Koichiro, Daniel Peterson, Nicholas Peterson, Glen Stecher, Masatoshi Nei, y Sudhir Kumar. «MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods». *Molecular Biology and Evolution* 28, n.º 10 (octubre de 2011): 2731-39. https://doi.org/10.1093/molbev/msr121.

Taylor, Steven W., Eoin Fahy, Bing Zhang, Gary M. Glenn, Dale E. Warnock, Sandra Wiley, Anne N. Murphy, et al. «Characterization of the Human Heart Mitochondrial Proteome». *Nature Biotechnology* 21, n.º 3 (marzo de 2003): 281-86. https://doi.org/10.1038/nbt793.

Thirstrup, J. P., L. A. Bach, V. Loeschcke, y C. Pertoldi. «Population Viability Analysis on Domestic Horse Breeds (Equus Caballus)». *Journal of Animal Science* 87, n.º 11 (1 de noviembre de 2009): 3525-35. https://doi.org/10.2527/jas.2008-1760.

Thompson, J D, D G Higgins, y T J Gibson. «CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.» *Nucleic Acids Research* 22, n.º 22 (11 de noviembre de 1994): 4673-80.

Tillich, Michael, Pascal Lehwark, Tommaso Pellizzer, Elena S. Ulbricht-Jones, Axel Fischer, Ralph Bock, y Stephan Greiner. «GeSeq - Versatile and Accurate Annotation of Organelle Genomes». *Nucleic Acids Research* 45, n.º W1 (03 de 2017): W6-11. https://doi.org/10.1093/nar/gkx391.

Tørresen, Ole K, Bastiaan Star, Pablo Mier, Miguel A Andrade-Navarro, Alex Bateman, Patryk Jarnot, Aleksandra Gruca, et al. «Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases». *Nucleic Acids Research* 47, n.º 21 (2 de diciembre de 2019): 10994-6. https://doi.org/10.1093/nar/gkz841.

Treangen, Todd J., y Steven L. Salzberg. «Repetitive DNA and next-generation sequencing: computational challenges and solutions». *Nature Reviews. Genetics* 13, n.º 1 (29 de noviembre de 2011): 36-46. https://doi.org/10.1038/nrg3117.

Trivedi, Urmi H., Timothée Cézard, Stephen Bridgett, Anna Montazam, Jenna Nichols, Mark Blaxter, y Karim Gharbi. «Quality control of next-generation sequencing data without a reference». *Frontiers in Genetics* 5 (6 de mayo de 2014). https://doi.org/10.3389/fgene.2014.00111.

Upham, Nathan S., y Bruce D. Patterson. «Diversification and Biogeography of the Neotropical Caviomorph Lineage Octodontoidea (Rodentia: Hystricognathi)». *Molecular Phylogenetics and Evolution* 63, n.º 2 (1 de mayo de 2012): 417-29. https://doi.org/10.1016/j.ympev.2012.01.020.

Wagner, J.E. y Manning, P.J. 1976. The biology of the guinea pig págs. 79-98. Londres, Academic Press.

Walker, Laura I., Miguel A. Soto, y Ángel E. Spotorno. «Similarities and differences among the chromosomes of the wild guinea pig *Cavia tschudii* and the domestic guinea pig *Cavia porcellus* (Rodentia, Caviidae)». *Comparative Cytogenetics* 8, n.º 2 (14 de julio de 2014): 153-67. https://doi.org/10.3897/CompCytogen.v8i2.7509.

Wang, Xiaoyang, Nian Liu, Hongli Zhang, Xiao-Jun Yang, Yuan Huang, y Fumin Lei. «Extreme Variation in Patterns of Tandem Repeats in Mitochondrial Control Region of Yellow-Browed Tits (Sylviparus Modestus , Paridae)». *Scientific Reports* 5, n.º 1 (19 de agosto de 2015): 13227. https://doi.org/10.1038/srep13227.

Wang, Xuelin, Feng Cheng, Dekai Rohlsen, Changwei Bi, Chunyan Wang, Yiqing Xu, Suyun Wei, Qiaolin Ye, Tongming Yin, y Ning Ye. «Organellar Genome Assembly Methods and Comparative Analysis of Horticultural Plants». *Horticulture Research* 5, n.º 1 (10 de enero de 2018): 1-13. https://doi.org/10.1038/s41438-017-0002-1.

Wei, An-Chi, Ting Liu, Sonia Cortassa, Raimond L. Winslow, y Brian O'Rourke. «Mitochondrial Ca2+ Influx and Efflux Rates in Guinea Pig Cardiac Mitochondria: Low and High Affinity Effects of Cyclosporine A». *Biochimica Et Biophysica Acta* 1813, n.º 7 (julio de 2011): 1373-81. https://doi.org/10.1016/j.bbamcr.2011.02.012.

Weir BJ. (1974) Notes on the origin of the domestic guinea pig. Symposia of the Zoological Society of London 34: 437–446 [Google Scholar]

Weyrich, Alexandra, Tino Schüllermann, Felix Heeger, Marie Jeschek, Camila J. Mazzoni, Wei Chen, Kathrin Schumann, y Joerns Fickel. «Whole Genome Sequencing and Methylome Analysis of the Wild Guinea Pig». *BMC Genomics* 15 (28 de noviembre de 2014): 1036. https://doi.org/10.1186/1471-2164-15-1036.

Weyrich, Alexandra, Tino Schüllermann, Felix Heeger, Marie Jeschek, Camila J Mazzoni, Wei Chen, Kathrin Schumann, y Joerns Fickel. «Whole genome sequencing and methylome analysis of the wild guinea pig». *BMC Genomics* 15, n.º 1 (28 de noviembre de 2014). https://doi.org/10.1186/1471-2164-15-1036.

Williams, S. T., P. G. Foster, y D. T. J. Littlewood. «The Complete Mitochondrial Genome of a Turbinid Vetigastropod from MiSeq Illumina Sequencing of Genomic DNA and Steps towards a Resolved Gastropod Phylogeny». *Gene* 533, n.º 1 (1 de enero de 2014): 38-47. https://doi.org/10.1016/j.gene.2013.10.005.

Wing E. (1986) Domestication of Andean mammals. In: Vuilleumier F, Monasterio M. (Eds) High altitude tropical biogeography. Oxford University Press, New York, 246–264

Witt, Kelsey E., Kathleen Judd, Andrew Kitchen, Colin Grier, Timothy A. Kohler, Scott G. Ortman, Brian M. Kemp, y Ripan S. Malhi. «DNA Analysis of Ancient Dogs of the Americas: Identifying Possible Founding Haplotypes and Reconstructing Population Histories». *Journal of Human Evolution* 79 (febrero de 2015): 105-18. https://doi.org/10.1016/j.jhevol.2014.10.012.

Witte, John S. «Genome-Wide Association Studies and Beyond». *Annual review of public health* 31 (2010): 9-20. https://doi.org/10.1146/annurev.publhealth.012809.103723.

Wood, N. J., y S. H. Phua. «Variation in the Control Region Sequence of the Sheep Mitochondrial Genome». *Animal Genetics* 27, n.º 1 (1996): 25-33. https://doi.org/10.1111/j.1365-2052.1996.tb01173.x.

Wyman, Stacia K., Robert K. Jansen, y Jeffrey L. Boore. «Automatic Annotation of Organellar Genomes with DOGMA». *Bioinformatics* 20, n.º 17 (22 de noviembre de 2004): 3252-55. https://doi.org/10.1093/bioinformatics/bth352.

Yamada A, Graciela, Víctor Bazán R, y Nadia Fuentes N. «Parámetros productivos de cuyes G en la costa central del Perú». *Revista de Investigaciones Veterinarias del Perú* 29, n.º 3 (julio de 2018): 877-81. https://doi.org/10.15381/rivep.v29i3.14748.

Ye, Chengxi, Christopher M. Hill, Shigang Wu, Jue Ruan, y Zhanshan (Sam) Ma. «DBG2OLC: Efficient Assembly of Large Genomes Using Long Erroneous Reads of the Third Generation Sequencing Technologies». *Scientific Reports* 6 (30 de agosto de 2016): 31900. https://doi.org/10.1038/srep31900.

Yu, Guangchuang, Tommy Tsan-Yuk Lam, Huachen Zhu, y Yi Guan. «Two Methods for Mapping and Visualizing Associated Data on Phylogeny Using Ggtree». *Molecular Biology and Evolution* 35, n.º 12 (diciembre de 2018): 3041-43. https://doi.org/10.1093/molbev/msy194.

Zabetian, Cyrus P., Ignacio F. Mata, y Latin American Research Consortium on the Genetics of PD (LARGE-PD). «LARGE-PD: Examining the Genetics of Parkinson's Disease in Latin America». *Movement Disorders: Official Journal of the Movement Disorder Society* 32, n.º 9 (2017): 1330-31. https://doi.org/10.1002/mds.27081.

Zerbino, Daniel R., y Ewan Birney. «Velvet: Algorithms for de novo short read assembly using de Bruijn graphs». *Genome Research* 18, n.º 5 (mayo de 2008): 821-29. https://doi.org/10.1101/gr.074492.107.

Zimin, Aleksey V., Guillaume Marçais, Daniela Puiu, Michael Roberts, Steven L. Salzberg, y James A. Yorke. «The MaSuRCA genome assembler». *Bioinformatics* 29, n.º 21 (1 de noviembre de 2013): 2669-77. https://doi.org/10.1093/bioinformatics/btt476.

10. ANEXOS

Anexo 1. Autorización de colecta de material genetico de cuyes

PERÚ | Ministerio de Agricultura y Riego

"Decenio de la Igualdad de oportunidades para mujeres y hombres"
"Año de la Lucha contra la Corrupción y la Impunidad"

MEMORANDUM Nº 815 -2019-MINAGRI-INIA-DGIA/SDRIA

A	:	**Dr. JUAN CARLOS GUERRERO ABAD** Director General Dirección de Recursos Genéticos y Biotecnología
Asunto	:	Informa sobre solicitud de autorización de colecta de material genético de cuyes
Referencia	:	Informe Nº 073-2019-MINAGRI-INIA/DGIA-SDRIA-ARAPOV
Fecha	:	La Molina, 02 OCT. 2019

Me dirijo a usted, con relación al documento de la referencia, concerniente a la solicitud de autorización de colecta de material genético de cuyes en el marco del Proyecto Nº 091_PI "*Utilización de herramientas moleculares para la caracterización genética de las razas de cuy Perú, Inti y Andina generada por el INIA*", financiado por el PNIA y liderado por el Ing. Eudosio Amancio Veli Rivera.

Al respecto, se adjunta el Informe de la referencia, que recomienda entre otros, que en el caso de hacer uso del recurso genético del cuy peruano, para los objetivos diferentes a los mencionados en su plan de investigación y que tengan un fin comercial, el solicitante, deberá obligatoriamente solicitar el acceso a los recursos genéticos mediante un contrato, de acuerdo a los requisitos establecidos en el link: **http//www-inia.gob.pe/requisitos-acceso-rrgg/.**

Atentamente

INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA
DIRECCIÓN DE GESTIÓN DE LA INNOVACIÓN AGRARIA
ING. JESÚS F. CALDAS CUEVA, MSc.
DIRECTOR GENERAL

JCC/JAD/amqg

CUT: 16274-2017

Av. La Molina 1981, La Molina
T: (051) 240 2100 anexo (indicar)
www.inia.gob.pe
www.minagri.gob.pe

EL PERÚ PRIMERO

DIRECCION DE RECURSOS GENETICOS Y BIOTECNOLOGIA

Pase a: SDB

Acción: Conocimiento y fines

La Molina 02/10/19

DIRECCIÓN DE RECURSOS GENÉTICOS Y BIOTECNOLOGÍA JUAN CARLOS GUERRERO ABAD PHD

INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA
SUB DIRECCIÓN DE BIOTECNOLOGÍA
DBRG

Lima 07 de Octubre, 2019

Pase a: Blga. Fredesvinda Carrillo

Acción: Conocimiento, Trámite respectivo

INIA

15.10.2019.

Blga. Claudia Falta

Conocimiento y Trámite Respectivo.

PERÚ | Ministerio de Agricultura y Riego

MINISTERIO DE AGRICULTURA Y RIEGO - INIA
Dirección de Gestión de la Innovación Agraria - DGIA
inia 06
INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA
Dirección de Gestión de Innovación Agraria - DGI
Sub Dirección de Regulación de la Innovación Agraria
30 SEP 2019
INGRESADO
Hora: Firma:

"Decenio de la Igualdad de Oportunidades para mujeres y hombres"
"Año de la Lucha contra la Corrupción y la Impunidad"

INFORME Nº 073-2019-MINAGRI-INIA/DGIA-SDRIA-ARAPOV

PARA	:	**Blgo. JORGE ALCÁNTARA DELGADO** Director Subdirección de Regulación de la Innovación Agraria
ASUNTO	:	Solicitud de Autorización de colecta de material genético de cuyes
REFERENCIA	:	a) OFICIO N° 170-2016-INIA-DRGB/DG b) Informe Técnico N° 008-2017-INIA-DGIA/SDRIA/ARAPOV-RARG. c) MEMORANDO N° 804-2017-MINAGRI-INIA-SG/OAJ
FECHA	:	La Molina, 30 de setiembre de 2019

Tengo el agrado de dirigirme a usted, a fin de informarle lo siguiente:

I. ANTECEDENTES

1.1 Mediante el Oficio de la referencia a) de fecha 17 de mayo del 2016, dirigida al Ing. Juan Álvaro Loayza Valdivia Director General de la Dirección de Gestión Agraria (SDRIA) del INIA, la Dra. Rosa Sánchez Díaz Directora General de la Dirección de Recursos Genéticos y Biotecnología (DRGB) del INIA, solicita la autorización de colecta de material genético de cuyes, adjuntando la solicitud s/n y sus anexos, de fecha 12 de mayo de 2017, en el marco del Proyecto N° 091_PI denominado: ***"Utilización de herramientas moleculares para la Caracterización genética de las razas de cuy Perú, Inti y Andina generadas por el INIA"***, financiado por el PNIA y liderado por el Ing. Eudosio Amancio Veli Rivera del Instituto Nacional de Innovación Agraria (INIA).

1.2 Mediante Memorándum N° 877-2016-INIA-DRGB/SDB de fecha 27 de diciembre de 2016, dirigido al Ing. Juan Loayza Valdivia, la Dra. Rosa Sánchez Díaz, hace llegar el levantamiento de observaciones a la solicitud en cuestión, elaborado por el Ing. Eudosio Veli Rivera, especialista de la Subdirección de Biotecnología de la DRGB, adjuntando el Oficio N° 0066-2016-INIA-DRGB-SDB/EVR, en el cual entre otros indica que el proyecto en mención tiene como objetivo la identificación filogenética de los cuyes peruanos, con carácter de investigación no persiguiendo otro fin y adjunta la solicitud de registro como institución especializada en caracterización genética de plantas cultivadas, animales y microorganismos relacionados a estos.

SUBDIRECCIÓN DE REGULACIÓN DE LA INNOVACIÓN AGRARIA
Ing. Roger Becerra Gallardo
V°B°
Reg. Asesoría Jurídica INIA - DGIA
SDRIA

1.3 Mediante correo electrónico de fecha 12 de julio de 2017, el Ing. Eudosio Veli Rivera, especialista de la Subdirección de Biotecnología de la DRGB, hace de conocimiento que viene conduciendo dos proyectos financiado por PNIA, con

fines de determinar la variabilidad genética de 02 especies nativas, siendo una de ellas los cuyes, e indica que esta especie está sufriendo cruzamiento con razas comerciales; lo cual está ocasionando erosión genética de estos recursos nativos y manifiesta que los cuyes son recursos genéticos estratégicos para los pobladores de bajos recursos, puesto que contribuyen a la seguridad alimentaria.

1.4 Mediante el documento de la referencia b), el suscrito emite informe técnico, sobre la solicitud materia del presente, para lo cual se realizó un análisis sobre el marco legal relacionado al aprovechamiento de los recursos naturales (Ley N° 26821), a la Conservación y aprovechamiento sostenible de la diversidad biológica (Ley N° 26839 y su reglamento) y al Acceso a los Recursos Genéticos (D.S. N° 003-2009-MINAM).

1.5 Mediante el memorando indicado en la referencia c), emitido por la Oficina de Asesoría Jurídica del INIA, entre otros indica que dicha oficina no tiene claro cómo se realizará la investigación para determinar la estructura genética y diversidad genética sin usar el recurso genético del cuy nativo, considerando que las muestras colectadas (pelo) van a ser objeto de extracción de ADN en el laboratorio de biología molecular y genómica del INIA.

1.6 Al respecto, es necesario informar, que en reiteradas oportunidades, se le comunico en forma verbal y mediante correo electrónico, al Ing. Eudosio Veli, la necesidad de reunirnos y conversar sobre sus solicitudes, teniendo en cuenta que el INIA en su condición de Autoridad de Acceso a RRGG no cuenta a la fecha con ningún procedimiento administrativo para regular la colecta de recursos biológicos que contienen recursos genéticos con fines de investigación para especies cultivadas o domésticas continentales, por lo tanto la colecta para estos caso no tenía restricciones legales, por lo que no es necesario contar con autorización para la colecta de tales recursos biológicos.

1.7 Mediante correo electrónico de fecha 27 setiembre de 2019, dirigido al suscrito, la Blga. Claudia Yalta Macedo, genetista y biotecnóloga de la Dirección de Recursos y Biotecnología (DRGB) del INIA, por encargo de la Dra. Dina Gutiérrez Reynoso, Directora de la Subdirección de Biotecnología de la DRGB, solicita se le informe sobre el estado de los documentos relacionados al Memorándum N° 877-2016-INIA-DRGB/SDB, referido a la solicitud para el permiso de colecta de material biológico de animales domésticos del proyecto 091 (cuyes).

II. ANÁLISIS TÉCNICO - LEGAL

2.1 De acuerdo a las conclusiones indicadas en el Informe Técnico N° 008-2017-INIA-DGIA/SDRIA/ARAPOV-RARG, entre otras se determinó lo siguiente:

- En general, la presente solicitud se encuentra dentro de los alcances del marco jurídico contemplados en la Ley N° 26821 sobre Aprovechamiento de los Recursos Naturales y la Ley N° 26839 referida a la Conservación y

Aprovechamiento Sostenible de la Diversidad Biológica, los cuales tienen por objetivo ***promover y regular el aprovechamiento sostenible de los recursos naturales, renovables y no renovables, así como la CONSERVACIÓN de estos recursos,*** en tanto constituyen **PATRIMONIO DE LA NACIÓN** promoviendo la investigación científica y tecnológica sobre la diversidad (incluyendo la biotecnología), para lo cual podrán otorgarse **PERMISOS PARA INVESTIGACIÓN.**

- De acuerdo a los objetivos, métodos y técnicas detallados en el plan de investigación, el área técnica de Acceso a Recursos Genéticos de especies cultivadas y domésticas continentales incluyendo a los animales domésticos de origen peruano, da opinión técnica favorable y considera pertinente otorgar la autorización a la Dra. Rosa Sánchez Díaz y sus colaboradores, para realizar investigación científica en el marco del Proyecto N° 091_PI denominado: ***"Utilización de herramientas moleculares para la Caracterización genética de las razas de cuy Perú, Inti y Andina generadas por el INIA"***, financiado por el PNIA, liderado por el Ing. Eudosio Amancio Veli Rivera del Instituto Nacional de Innovación Agraria (INIA).

2.2 Si bien es cierto que a través del Informe Técnico mencionado líneas arriba, se dio opinión favorable y se consideró pertinente otorgar la autorización, bajo la modalidad de **PERMISOS PARA INVESTIGACIÓN,** en función a lo contemplado en el marco jurídico contemplado en otras leyes diferentes a las de acceso a recursos genéticos, mencionadas en el numeral anterior; el cual incluye la colecta de pelo del cuy nativo; es necesario indicar que en general para la colecta de cualquier recurso biológico que contenga algún recurso genético de interés, de especies cultivadas o domésticas continentales (ejemplar completo o parte de ello), no existe ninguna restricción técnica – legal; más aún a la fecha el INIA no cuenta con una directiva o procedimiento específico para autorizar el uso de los recursos biológicos de las especies cultivadas o domésticas continentales, mediante colectas para fines de conservación, estudios taxonómicos, filogenéticos u otro que no implique el uso de un recurso genético.

2.3 De acuerdo a los objetivos y a la justificación de la colecta mencionados en el Plan de Investigación de la presente solicitud, está referida a realizar estudios de caracterización molecular en poblaciones domésticas de cuyes nativos, con fines de conservación e investigación para determinar la estructura genética, diversidad genética y estudios sobre relaciones filogenéticas, haciendo uso de la biotecnología moderna; y de esta manera conocer la variabilidad genética, identificar poblaciones vulnerables a la pérdida de diversidad como producto de aislamiento genético (el subrayado es nuestro), mejoramiento genético, detectar cuellos de botella genético y conocer los efectos de selección de los cuyes nativos. Por lo que según estos objetivos van a hacer uso, necesariamente del recurso genético contenido en el pelo del cuy peruano.

DIRECCIÓN DE REGULACIÓN DE LA INNOVACIÓN Ing. Roger Becerra Gallardo VºBº

2.4 Al respecto es necesario informar que en el mes de febrero del 2019, a través de diferentes reuniones, las autoridades competentes para el acceso a los recursos genéticos del país (INIA, SERFOR, PRODUCE) y el ente rector (MINAM, elaboraron el documento denominado "Criterios de Identificación de un Proyecto o Actividad de Acceso a Recursos Genéticos", como una herramienta de trabajo para determinar cuándo un proyecto o actividad presentada en el marco de una solicitud de acceso, requiere ser atendida y autorizada mediante la celebración de un Contrato de Acceso ya sea para fines de investigación con fines no comerciales o comerciales. Habiéndose determinado entre otros que una solicitud, no requiere de contrato de acceso, cuando su ámbito de aplicación comprende actividades de investigación científica básica en diversidad biológica con fines no comerciales y que involucren actividades de identificación, delimitación y clasificación de especies relacionado con estudios moleculares con fines sistemáticos, filogenéticos, filogeográficos, biogeográficos, evolutivos; así como de genética de la conservación y ecología molecular sin fines comerciales. Dado que el uso de secuencias de ADN para análisis filogenéticos o diversidad genética no es uso de un recurso genético, puesto que estos no tienen un valor real o potencial debido a que son fragmentos de secuencias no codificantes altamente conservadas o aleatorias, en función al marcador empleado. Por ejemplo: Secuencia del ARN ribosomal 16S (ARNr 16S), secuencias de ADN empleadas para Barcoding (para fin taxonómico), caracterización molecular por ISSR (Inter-Secuencias Simples Repetitivas); siendo éste el caso de la presente solicitud, por lo tanto en mérito a los criterios indicados, no se requiere de un contrato de acceso.

III. CONCLUSIONES

3.1 La presente solicitud, contempla actividades de colecta del recurso biológico que contiene el recurso genético del cuy nativo (pelo) y tiene por objetivo principal realizar estudios de caracterización molecular en poblaciones domésticas de cuyes nativos, con fines de conservación e investigación para determinar la estructura genética, diversidad genética y estudios sobre relaciones filogenéticas, haciendo uso de la biotecnología moderna; y de esta manera conocer la variabilidad genética, identificar poblaciones vulnerables a la pérdida de diversidad como producto de aislamiento genético (el subrayado es nuestro), mejoramiento genético, detectar cuellos de botella genético y conocer los efectos de selección de los cuyes nativos.

3.2 El INIA a la fecha no cuenta con una directiva o procedimiento específico para autorizar el uso de los recursos biológicos de las especies cultivadas o domésticas continentales, mediante colectas para fines de conservación, estudios taxonómicos, filogenéticos u otro que no implique el uso de un recurso genético.

3.3 En general para actividades de colecta de cualquier recurso biológico que contenga algún recurso genético de interés, de especies cultivadas o domésticas continentales (ejemplar completo o parte de ello), no existe ninguna restricción técnico – legal para su realización, bastando solamente que la persona natural o

jurídica que va a realizar la colecta, cuente con el consentimiento fundamentado previo (PIC por sus siglas en inglés) y las condiciones mutuamente acordadas (MAT por sus siglas en inglés) con el poseedor o proveedor del recurso biológico que contenga al recurso genético. Por lo tanto la presente solicitud referida a la colecta de pelo de cuy de origen peruano, no necesita de ninguna autorización para llevarla a cabo la colecta materia del presente, en el marco del proyecto indicado.

3.4 La investigación a realizarse en el marco del proyecto ***"Utilización de herramientas moleculares para la Caracterización genética de las razas de cuy Perú, Inti y Andina generadas por el INIA"***, contempla realizar estudios de caracterización molecular en poblaciones domésticas de cuyes nativos, con fines de conservación e investigación para determinar la estructura genética, diversidad genética y estudios sobre relaciones filogenéticas, haciendo uso de la biotecnología moderna; y de esta manera conocer la variabilidad genética, identificar poblaciones vulnerables a la pérdida de diversidad como producto de aislamiento genético, lo cual está contemplado en el documento "Criterios de Identificación de un Proyecto o Actividad de Acceso a Recursos Genéticos", en el cual se determinó entre otros que, una solicitud, no requiere de contrato de acceso, cuando su ámbito de aplicación comprende actividades de investigación científica básica en diversidad biológica con fines no comerciales y que involucren actividades de identificación, delimitación y clasificación de especies relacionado con estudios moleculares con fines sistemáticos, filogenéticos, filogeográficos, biogeográficos, evolutivos; así como de genética de la conservación y ecología molecular sin fines comerciales; tal como se indica en el numeral 2.4 del presente informe; siendo este el caso del proyecto a ejecutar en el marco de la solicitud presentada.

IV. RECOMENDACIONES

4.1 Se recomienda al solicitante, que en caso de hacer uso del recurso genético del cuy peruano, para los objetivos diferentes a los mencionados en su plan de investigación y que tengan un fin comercial, deberá obligatoriamente solicitar el acceso a los Recursos Genéticos mediante un contrato; para lo cual deberá tener en cuenta los requisitos establecidos por el ARAPOV en el siguiente link de la página web del INIA: http://www.inia.gob.pe/requisitos-acceso-rrgg/

4.2 Remitir copia del presente informe, a la Dra. Dina Gutiérrez Reynoso, Directora de la Subdirección de Biotecnología de la Dirección de Recursos Genéticos y Biotecnología, para su conocimiento y fines pertinentes.

Sin otro particular, saludo a usted.

Atentamente,

INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA
DIRECCIÓN DE GESTIÓN DE LA INNOVACIÓN AGRARIA
Subdirección de Regulación de la Innovación Agraria
Ing. ROGER BECERRA GALLARDO
Coordinador
Área de Regulación en Acceso a los Recursos Genéticos y Protección de Obtentores Vegetales

CUT N° 16274-2017-INIA

Av. La Molina 1981, La Molina
T: (051) 240 2100 anexo 278
www.inia.gob.pe
www.minagri.gob.pe

INIA - DGIA - SDRIA / ARAPOV
Regulación del Acceso a Recursos Genéticos 1

PERÚ Ministerio de Agricultura y Riego | Instituto Nacional de Innovación Agraria

Dirección de Recursos Genéticos y Biotecnología

Dirección de Gestión de la Innovación Agraria - DGIA
Sub Dirección de Regulación de la Innovación Agraria
"Decenio de las Personas con Discapacidad en el Perú"
"Año de la consolidación del Mar de Grau"

Folio Nº 03

INGRESADO
Hora: 3:27 p.m. Firma:

17 MAYO 2016
INGRESADO
Hora: 11.55 Firma:

La Molina, 17 MAYO 2016

OFICIO Nº 170 -2016-INIA-DRGB/DG

Señor Ing.
JUAN ALVARO LOAYZA VALDIVIA
Director General de la
Dirección de Gestión de la Innovación Agraria

Asunto: **Solicito Autorización de Colecta de material genético de cuyes**

Referencia: **Oficio N° 0004-2016-INIA-PNIA-PROYECTO N° 091**

Es grato dirigirme a usted para expresarle un cordial saludo y manifestarle que en marco del Proyecto N° 091_PI denominado: **Utilización de herramientas moleculares para la caracterización genética de las razas de cuy Perú, Inti y Andina generadas por el INIA**, financiado por el PNIA, se ha programado la colecta de material genético (folículo pilosos) de cuyes criollos y de las razas Perú, Inti y Andina.

Al respecto, agradeceré se emita la autorización de colecta de material genético, para tal fin se adjunta:

Anexo 1A: Modelo de Solicitud,
Anexo 2A: Contenido del Plan de Investigación
Anexo 3A: Carta de Presentación de Participante del Proyecto
Anexo 4A: Datos de los Participantes
Anexo 6A: Carta de Compromiso
Currículo vitae.

Agradeciendo su atención, reitero mi saludo.

Cordialmente,

INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA
ROSA A. SÁNCHEZ DÍAZ PhD.
DIRECTORA GENERAL
DIRECCIÓN DE RECURSOS GENÉTICOS Y BIOTECNOLOGÍA

INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA
DIRECCIÓN DE GESTIÓN DE LA INNOVACIÓN AGRARIA
PASE a: SDRIA
PARA: Conocimiento y fines.
La Molina

Cc: SDB - 2209.91
RASO/can.

CUT Nº 65266 DRGB

www.inia.gob.pe
Av. La Molina 1981
La Molina. Lima 12. Perú
T: (511) 349-2600 Anexos 313, 315
E: drgb@inia.gob.pe



MINISTERIO DE AGRICULTURA Y RIEGO - INIA
Dirección de Gestión de la Innovación Agraria - DGIA
Folio Nº 02

PERÚ | Ministerio de Agricultura y Riego

inia
Instituto Nacional de Innovación Agraria

"Año del buen servicio al ciudadano"

INFORME TÉCNICO N° 008-2017-INIA-DGIA/SDRIA/ARAPOV-RARG

PARA	:	**BLGO. JORGE ALCÁNTARA DELGADO** Director Subdirección de Regulación de la Innovación Agraria - SDRIA
ASUNTO	:	Opinión técnica sobre solicitud de Autorización para realizar colecta de recursos genéticos de especies animales domésticos con fines de investigación científica. Cuyes.
REFERENCIA	:	1) OFICIO N° 170-2016-INIA-DRGB/DG 2) Solicitud s/n de fecha 12.05.2016
FECHA	:	La Molina, 21 de Julio del 2017

Tengo el agrado de dirigirme a Ud. a fin de informarle lo siguiente:

1.0 ANTECEDENTES

1.1 Mediante Oficio N°170-2016-INIA-DRGB/DG, de fecha 17 de mayo del 2016, dirigida al Ing. Juan Álvaro Loayza Valdivia Director General de la Dirección de Gestión Agraria (SDRIA) del INIA, la Dra. Rosa Sánchez Díaz Directora General de la Dirección de Recursos Genéticos y Biotecnología (DRGB) del INIA, solicita la autorización de colecta de material genético de cuyes, adjuntando la solicitud indicada en el documento 2) de la referencia y sus anexos, en el marco del Proyecto N° 091_PI denominado: ***"Utilización de herramientas moleculares para la Caracterización genética de las razas de cuy Perú, Inti y Andina generadas por el INIA"***, financiado por el PNIA y liderado por el Ing. Eudosio Amancio Veli Rivera del Instituto Nacional de Innovación Agraria (INIA).

1.2 Al mencionado oficio se adjunta los siguientes documentos:

- Anexo 1A: Solicitud.
- Anexo 2A: Plan de Investigación en idioma español.
- Anexo 3A: Carta de presentación de la representante del Proyecto
- Anexo 4A: Datos de los Participantes.
- Anexo 6A: Carta de compromiso
- Curriculum vitae

1.3 Con fecha 27 de mayo del 2016, la Unidad Técnica del componente de Acceso a los Recursos Genéticos, da por admitida la presente solicitud al haber cumplido con los requisitos mínimos establecidos en el Artículo 26 de la Decisión 391, dándose apertura al expediente N° 0006-2016-INIA/RARG/SDRIA.

SUBDIRECCIÓN DE REGULACIÓN DE LA INNOVACIÓN AGRARIA - SDRIA - Ing. Roger Becerra Gallardo V°B° Responsable RARG INIA - DGIA

1.4 Mediante Memorándum N° 474-2016-INIA-DGIA/SDRIA de fecha 24 de Agosto de 2016, el Ing. Juan Loayza Valdivia, solicita a la Dra. Rosa Sánchez Díaz subsanar las observaciones y requerimientos, referidos a precisar el número y nombres de las regiones a realizar el estudio; informar si las áreas de estudio consideradas en el Plan de

Investigación involucra Áreas de Reserva Natural y/o Zonas de Amortiguamiento de Áreas Naturales Protegidas; presentar la autorización del propietario, poseedor o administrador del predio donde se encuentra el recurso biológico que contenga el recurso genético del cuy; remitir carta de presentación de la Ing. Lilia Chauca Francia de Zaldivar y de la Blga. Claudia Esther Yalta Macedo y la solicitud sobre Registro como Institución Especializada en caracterización genética de animales domésticos.

1.5 Mediante Memorándum N° 877-2016-INIA-DRGB/SDB de fecha 27 de diciembre de 2016, dirigido al Ing. Juan Loayza Valdivia, la Dra. Rosa Sánchez Díaz, hace llegar el levantamiento de observaciones a la solicitud en cuestión, elaborado por el Ing. Eudosio Veli Rivera, especialista de la Subdirección de Biotecnología de la DRGB, adjuntando el Oficio N° 0066-2016-INIA-DRGB-SDB/EVR, en el cual detalla las respuestas a las observaciones realizadas, en relación al número y nombres de las regiones a realizar el estudio, indicando que se ha modificado el Plan de Investigación y que las colectas se van a realizar en los siguientes departamentos: en la zona norte: Cajamarca, Lambayeque, La Libertad; en la zona sur: Cusco, Huancavelica, Apurímac, Puno, Ayacucho y en la zona centro: Ancash, Huánuco y Junín; asimismo, indica que las áreas mencionadas no incluyen Áreas de Reserva Natural y/o Zonas de Amortiguamiento de Áreas Naturales Protegidas; las autorizaciones de los propietarios del recurso biológico, serán solicitados durante la colecta; adjunta las cartas de presentación de la Ing. Lilia Chauca Francia de Zaldívar y de la Blga. Claudia Esther Yalta Macedo; e indica también, que el proyecto en mención tiene como objetivo la identificación filogenética de los cuyes peruanos, con carácter de investigación no persiguiendo otro fin y adjunta la solicitud de registro como institución especializada en caracterización genética de plantas cultivadas, animales y microorganismos relacionados a estos.

1.6 Mediante correo electrónico de fecha 12 de julio de 2017, el Ing. Eudosio Veli Rivera, especialista de la Subdirección de Biotecnología de la DRGB, hace de conocimiento que viene conduciendo dos proyectos financiado por PNIA, con fines de determinar la variabilidad genética de 02 especies nativas, siendo una de ellas los cuyes, e indica que esta especie está sufriendo cruzamiento con razas comerciales; lo cual está ocasionando erosión genética de estos recursos nativos y manifiesta que los cuyes son recursos genéticos estratégicos para los pobladores de bajos recursos, puesto que contribuyen a la seguridad alimentaria.

2.0 CONSIDERACIONES LEGALES:

2.1 El Convenio de Diversidad Biológica (CDB)[1], reconoce los derechos soberanos de los Estados sobre sus recursos naturales y, en consecuencia la facultad de los gobiernos nacionales para regular dicho acceso, el mismo que se encuentra sometido a la legislación nacional.

SUBDIRECCIÓN DE REGULACIÓN DE LA INNOVACIÓN AGRARIA - SDRIA
Ing. Roger Becerra Gallardo
V°B°

2.2 El artículo 66° de la Constitución Política del Perú, promulgada el 29 de diciembre de 1993, establece que los recursos naturales, renovables y no renovables, son patrimonio de la Nación. El Estado es soberano en su aprovechamiento; asimismo, en su artículo 68°

[1] Aprobado en la Cumbre de Río en junio de 1992, ratificado por el Perú a través de la Resolución Legislativa N° 26181. Tiene como objetivos la conservación de la biodiversidad, la utilización sostenible de sus componentes y la participación justa y equitativa en los beneficios que se deriven de la utilización de los recursos genéticos.

MINISTERIO DE AGRICULTURA Y RIEGO - INIA
Dirección de Gestión de la Innovación Agraria - DGIA
Folio Nº 01

Jefatura

Oficina de Asesoría Jurídica

MEMORANDO N°804-2017-MINAGRI-INIA-SG/OAJ

PARA : **ING. JESÚS FRANCISCO CALDAS CUEVA**
Director General de la Dirección de la Gestión de la Innovación Agraria

ASUNTO: Acceso a recursos genéticos de cuyes para investigación.

REF. : a) Memorando N° 517-2017-INIA/DGIA/SDRIA.
b) Informe Técnico N° 008-2017-INIA-DGIA/SDRIA/ARAPOV-RARG.

FECHA : La Molina, 07 NOV 2017

Tengo a bien dirigirme a usted, en relación al documento de la referencia a), a través del cual su Despacho solicitó opinión legal sobre la solicitud de acceso a recursos genéticos de cuyes presentada por la Directora General de la Dirección de Recursos Genéticos y Biotecnología (DRGB) para el proyecto N° 091_PI "Utilización de herramientas moleculares para la caracterización genética de las razas de cuy Perú, Inti y Andina generadas por el INIA", financiado por el Programa Nacional de Innovación Agraria – PNIA (en adelante la SOLICITUD).

Funciones de los órganos del INIA

De manera preliminar, se debe indicar que esta Oficina tiene como una de sus funciones emitir opinión sobre la interpretación y aplicación de las normas que le son derivadas por los órganos del INIA; por lo que en el presente caso, en base a la opinión técnica contenida en el Informe N° 008-2017-INIA-DGIA/SDRIA/ARAPOV-RARG del 21 de julio de 2017, elaborado por la Dirección a su cargo, se procederá a opinar sobre las normas que resultan de aplicación a la SOLICITUD.

Sin perjuicio de ello, se debe indicar que la Dirección de Gestión de la Innovación Agraria (DGIA) es responsable de elaborar y proponer normas en relación al acceso a los recursos genéticos en las especies cultivadas o domésticas continentales, así como dirigir la implementación de las disposiciones técnicas correspondientes.

Análisis de la SOLICITUD

En el Informe Técnico N° 008-2017-INIA/DGIA-SDRIA-ARAPOV-RBG, se estableció en el numeral 5.2 (Conclusiones) que la SOLICITUD consiste en realizar la colecta del recurso biológico que contiene el recurso genético del cuy nativo, para fines de conservación e investigación para determinar la estructura genética, diversidad genética, etc[1]. Asimismo, en el numeral 4.8 del citado Informe, la DGIA señaló que las muestras colectadas van a ser objeto de extracción de ADN en el laboratorio de biología molecular y genómica del INIA.

[1] El resaltado y subrayado es nuestro.

PERÚ | Ministerio de Agricultura y Riego | Instituto Nacional de Innovación Agraria

Oficina de Asesoría Jurídica

"Año del Buen Servicio al Ciudadano"

En relación a lo anterior, la DGIA en el numeral 6.1 del Informe Técnico N° 008-2017-INIA/DGIA-SDRIA-ARAPOV-RBG, declaró procedente la SOLICITUD referida a la colecta e investigación del recurso biológico (pelo) que contiene el recurso genético del cuy nativo[2], con el objeto de realizar estudios de caracterización molecular, con fines de conservación e investigación. Asimismo, en el numeral 6.4 de dicho Informe, la DGIA opinó: *"(...) Si como consecuencia de la presente investigación científica, se logra superar los riesgos de que el cuy nativo sufra erosión genética y se cuente con datos de pasaporte, de sus características genéticas y moleculares debidamente identificadas, sobre la estructura genética, diversidad genética y relaciones filogenéticas; y se garanticen que están debidamente conservadas en el banco de germoplasma de ADN del INIA; el solicitante u otro tercero que quiera realizar estudios haciendo uso del recurso genético del cuy nativo con fines de investigación aplicada, industrial, comercial, tendrán que solicitar el Acceso a los Recursos Genéticos (...)".*

De lo expuesto, se debe indicar que esta Oficina no tiene claro, cómo se realizará la investigación para determinar la estructura genética y diversidad genética sin usar el recurso genético del cuy nativo, considerando que las muestras colectadas (pelo) van a ser objeto de extracción de ADN en el laboratorio de biología molecular y genómica del INIA.

Por lo expuesto, esta Oficina solicita que se determine de manera precisa sí en el marco del proyecto de investigación se va acceder a recursos genéticos, teniendo en consideración que por acceso a recursos genéticos, se entiende a las actividades de obtención y utilización con fines de investigación. De ser afirmativa la respuesta, es decir que el solicitante acceda al recurso genético para fines de investigación, sírvase precisar si resulta aplicable las disposiciones sobre las Limitaciones al Acceso establecidas en el artículo 27° del Reglamento de Acceso a Recursos Genéticos.

Atentamente,

ABOG. CÉSAR ARTURO REVOREDO CASTAÑEDA
Director General
Oficina de Asesoría Jurídica

Instituto Nacional de Innovación Agraria - INIA - Oficina de Asesoría Jurídica

INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA
DIRECCIÓN DE GESTIÓN DE LA INNOVACIÓN AGRARIA - DGIA
PASE a: SDRIA.
PARA: Analizar documento de OAJ.
La Molina 8/XI/17

CRC/lr

CUT 16274-17

INSTITUTO NACIONAL DE INNOVACIÓN AGRARIA
DGIA - SDRIA
PASE a: ARAPOV
PARA: Atender solicitud de OAJ
Proyecto respuesta
La Molina:
FIRMA

[2] El resaltado es nuestro.

Av. La Molina 1981 - La Molina - Lima
T: (511) 240-2100
www.inia.gob.pe
www.minagri.gob.pe

Trabajando para todos los peruanos

RESOLUCIÓN DEL JEFE DEL REFUGIO DE VIDA SILVESTRE LOS PANTANOS DE VILLA N°005-2019-SERNANP-JEF

Chorrillos, 12 de abril de 2019

VISTO:

La solicitud de fecha 07.04.2019 presentada por Dámaso Wilfredo Ramírez Huaroto, docente de la Universidad Científica del Sur, responsable de realizar la investigación titulada: "Estado poblacional y conservación del cuy silvestre (*Cavia tshudii*) en el humedal Ramsar Los Pantanos de Villa (Lima-Perú)", por el periodo de dos años.

CONSIDERANDO:

Que, según lo previsto en los incisos g) e i) del artículo 2° de la Ley N° 26834, Ley de Áreas Naturales Protegidas, unos de sus principales objetivos de protección es servir de sustento y proporcionar medios y oportunidades para el desarrollo de la investigación científica;

Que, en concordancia con ello, en el artículo 29° de la precitada Ley, se establece que el Estado reconoce la importancia de las Áreas Naturales Protegidas para el desarrollo de la investigación científica básica y aplicada, siempre que no afecte los objetivos de conservación, se respete la zonificación y las condiciones establecidas en el Plan Maestro;

Que, la actualización del Plan Director de las Áreas Naturales Protegidas, aprobada por Decreto Supremo Nº016-2009-MINAM, refiere que la investigación científica constituye una herramienta básica para la generación de información que permita mejorar el conocimiento sobre la diversidad biológica, así como para el manejo de recursos naturales y la gestión de riesgos y amenazas;

Que, mediante la Resolución Presidencial N°250-2013-SERNANP, publicado el 26 de diciembre del 2013, se aprobó el Certificado de Procedencia de los recursos naturales renovables forestales, flora y/o fauna silvestre provenientes de las Áreas Naturales Protegidas de administración nacional;

Que, mediante Decreto Supremo N°010-2015-MINAM, publicado el 23 de setiembre de 2015, se declara de interés nacional el desarrollo de investigaciones al interior de las Áreas Naturales Protegidas de administración nacional, determinándose su gratuidad, así como los procedimientos de aprobación automática y evaluación previa para su otorgamiento;

Que, en el artículo 4° del mencionado Decreto Supremo, se prevé cinco supuestos en los que la autorización de investigación requiere de evaluación previa: a) ingreso a ámbitos de acceso restringido, b) la colecta o extracción de muestras biológicas, c) se prevea la alteración del entorno o instalación de infraestructura en el caso de áreas naturales protegidas de administración nacional, d) el uso de equipo o infraestructura perteneciente a las ANP de administración nacional, e) investigación en predios privados;

Que, mediante Resolución Presidencial Nº287-2015-SERNANP, publicada el 20 de enero de 2016, se aprueban las Disposiciones Complementarias al Reglamento de la Ley de Áreas Naturales Protegidas en materia de investigación, las mismas que establecen las normas y

lineamientos que regulan las investigaciones realizadas al interior de las Áreas Naturales Protegidas de administración nacional;

Que, en el artículo 23° de las precitadas Disposiciones Complementarias se establecen los criterios de evaluación del Plan de Investigación;

Que, mediante Resolución Ministerial N°35-2017 del 03 de febrero del 2017, modifica, entre otros, el Procedimiento N°4 del Texto Único de Procedimientos Administrativos-TUPA del SERNANP, aprobado por Decreto Supremo N°002-2012-MINAM y modificado por Resolución Ministerial N°152-2016-MINAM y Resolución Ministerial N°315-2016-MINAM;

Que, a través del documento del visto, el Sr. Dámaso Wilfredo Ramírez Huaroto, responsable de realizar la investigación científica: "Estado poblacional y conservación del cuy silvestre (*Cavia tshudii*) en el humedal Ramsar Los Pantanos de Villa (Lima-Perú), por el periodo de dos años;

Que, mediante Informe Técnico N°015-2019-SERNANP-RVSLPV-MDS de fecha 11 de Abril del 2019, se evalúa la solicitud del Sr. Dámaso Wilfredo Ramírez Huaroto concluyendo que el expediente cumple con los requisitos establecidos en el artículo 18° de las Disposiciones Complementarias al Reglamento de la Ley de Áreas Naturales Protegidas en materia de investigación, y que el Plan de Investigación se encuentra conforme a los criterios establecidos en el artículo 23° de las Disposiciones Complementarias en mención;

En uso de las atribuciones conferidas por el numeral 2.1 del artículo 2° del Decreto Supremo N° 010-2015-MINAM, el artículo 14° de las Disposiciones Complementarias al Reglamento de la Ley de Áreas Naturales Protegidas en materia de investigación, aprobadas por Resolución Presidencial Nº 287-2015-SERNANP, y el artículo 27° del Reglamento de Organización y Funciones del SERNANP, aprobado mediante Decreto Supremo Nº 006-2008-MINAM.

SE RESUELVE:

Artículo 1º.- Autorizar la investigación científica: "Estado poblacional y conservación del cuy silvestre (*Cavia tshudii*) en el humedal Ramsar Los Pantanos de Villa (Lima-Perú), a favor del Sr. Dámaso Wilfredo Ramírez Huaroto, por el período de dos años, contados a partir de la fecha de emisión de la presente Resolución;

Artículo 2°. - Autorizar el ingreso al Área Natural Protegida Refugio de Vida Silvestre Los Pantanos de Villa, a las siguientes personas;

Apellidos y Nombres	DNI	País	Cargo	Institución
Dámaso Wilfedo Ramírez Huaroto	41807437	Perú	Docente universitario/ coordinador del área de biología Investigador	Universidad Científica del Sur
			colaborador del Museo de Historia Natural	Un. Nacional Mayor de San Marcos
Pacheco Torres Víctor Raúl	06651155	Perú	Jefe del Departamento de Mastozoología	Un. Nacional Mayor de San Marcos

			del Museo de Historia Natural).	
Aponte Ubillús Héctor Alonso	42815994	Perú	Docente universitario/ Fac. Biología Marina	Universidad Científica del Sur
			Investigador asociado del Museo de Historia Natural	Un. Nacional Mayor de San Marcos
Quispe López Arturo Manuel	47066758	Perú	Estudiante de Biología	Uni. Nac. Agraria La Molina
Marcelo Carranza Diego Miguel	74162446	Perú	Estudiante de Biología- Investig. Colaborador departamento de Mastozoología	Un. Nacional Mayor de San Marcos

Artículo 3°. – Autorizar la captura temporal de los ejemplares vivos para el marcaje y correspondiente liberación en campo.

Excepcionalmente, se permitirá la colecta definitiva de ejemplares muertos hallados en el ámbito de estudio, durante la investigación, previa verificación realizada por el personal del ANP.

La presente autorización no otorga derechos sobre los recursos genéticos o productos derivados de las muestras colectadas.

Artículo 4º.- Los integrantes del equipo de investigación son responsables de conocer y cumplir las disposiciones contenidas en la Ley Nº26834, Ley de Áreas Naturales Protegidas, y su Reglamento, aprobado mediante Decreto Supremo Nº038-2001-AG, modificado por Decreto Supremo N°010-2015-MINAM, así como en la Resolución Presidencial N°287-2015-SERNANP. Asimismo, los investigadores deberán cumplir con las normas que la Jefatura y su personal dispongan durante el desarrollo de la investigación;

Artículo 5º.- El Sr. Dámaso Wilfredo Ramírez Huaroto, docente de la Universidad Científica del Sur, responsable de la investigación se compromete a:

a. Presentar copia de la presente autorización al personal del ANP que lo solicite.
b. Tramitar con carácter obligatorio los permisos adicionales que se requieran para la investigación.
c. Comunicar al SERNANP cualquier descubrimiento nuevo para la ciencia, debiendo entregar una copia del depósito del holotipo del nuevo taxa en una institución científica nacional.
d. Realizar una presentación a la JANP con los resultados de la investigación.
e. Gestionar los permisos de exportación ante la autoridad competente, cuando se requiera enviar al extranjero parte de las muestras colectadas.
f. Entregar una vez publicado los resultados de la investigación, una copia digital de la publicación al SERNANP y autorizar su registro en la biblioteca digital del SERNANP.
g. Entregar a la jefatura del ANP un informe, en el caso de investigaciones que generan información prioritaria para la gestión del ANP y que justificó el apoyo del SERNANP.

h. No generar altercados con la población local que puedan desencadenar un conflicto social.
i. Entregar a la jefatura del ANP una copia digital del informe final (incluye versión digital) conteniendo el material fotográfico y la lista de ejemplares colectados.

El incumplimiento injustificado de estos compromisos producirá el ingreso del investigador en la lista de investigadores inhabilitados para próximas autorizaciones emitidas por el SERNANP.

Artículo 6°.- La autorización a la que se refiere el Artículo 1° caducará automáticamente al vencer el plazo concedido, por el incumplimiento injustificado de los compromisos adquiridos o por cualquier daño al patrimonio natural, sin perjuicio de las responsabilidades administrativas, civiles o penales que pudieran originarse.

Artículo 7°.- Todas las muestras colectadas serán reportadas a la Jefatura del Refugio de Vida Silvestre Los Pantanos de Villa inmediatamente después de la culminación de las labores de campo en el Puesto de Control o sede administrativa más cercana, para la emisión del Certificado de Procedencia correspondiente.

Artículo 8°. - El SERNANP se abstiene de toda responsabilidad por los accidentes o daños que puedan sufrir los integrantes del investigador durante el desarrollo del proyecto de investigación, con el conocimiento al responsable de la investigación sobre el contexto social del Refugio de Vida Silvestre Los Pantanos de Villa.

Artículo 9°. - Regístrese la presente Resolución en el Módulo de Seguimiento a las autorizaciones de investigación del SERNANP, en el archivo de autorizaciones del Refugio de Vida Silvestre Los Pantanos de Villa y publíquese en la página web del SERNANP (www.sernanp.gob.pe).

Regístrese y comuníquese,

ING. OMAR UBILLUS TOLENTINO
JEFE RVSLPV
Servicio Nacional de Áreas Naturales
Protegidas por el Estado - SERNANP

Anexo 2. Protocolo *A* (**Mammalian Tissue)** *PureLink™ Genomic Mini Kit*

1. Adicionar 200 ul PureLink™ Genomic/Lysis buffer y mezclar bien con vortex.
2. Adicionar 200 ul de etanol abosluto al lisado. mezclar bien por 5 segundos.
3. Proceder a **Binding DNA**. Armar un PureLink Spin Column y colocarlo en un Collection Tube.

Binding DNA	1. Adicionar el lisado aprox. 640 ul preparada con el PureLink™ Genomic Lysis /Binding buffer y etanol al Purelink Spin Column.
	2. Centrifugar la columna a 10, 000 x g por 1 minuto a temperatura ambiente.
	3. Descartar el Collection Tube y colocar en un clean PureLink™ Collection Tube del kit.
	4. Poceder a **Washing DNA.**
Washing DNA	5. Adicionar 500 ul **Wash Buffer 1** preparada con etanol a la columna
	6. Centrifugar la columna a temperatura ambiente a 10,000 g por 1 minuto.
	7. Descartar el tubo de colección y colocar el spin column en un tubo de colección limpio.
	8. Adicionar 500 ul **Wash Buffer 2** preparada con etanol a la columna
	9. Centrifugar la columna a máxima velocidad (en gravedades o rcf) por 3 minutos a temperatura ambiente. Descartar el tubo de colección.
	10. Proceder a **Eluting DNA**
Eluting DNA	11. Colocar el spin column en un tubo de microcentrifuga de 1.5 ml.
	12. Adicionar 100 ul de PureLinkTM Genomic Elution Buffer a la columna
	13. Incubar a temperatura ambiente por 1 minuto.
	14. Centrifugar la columna a una máxima velocidad por 1 minuto a temperatura ambiente. (El tubo contiene DNA genómico purificado)
	15. Para recuperar mayor DNA , realizar una segunda elución usando el mismo volumen que el primero en otro tubo de microcentrifuga 1.5 ml.
	16. Centrifugar la columna a una máxima velocidad por 1 minuto a temperatura ambiente
	17. Remover y descartar la columna.
	18. Almacenar el DNA purificado a -20°C

Anexo 3. Estadísticas de las lecturas filtradas por Trimmomatic

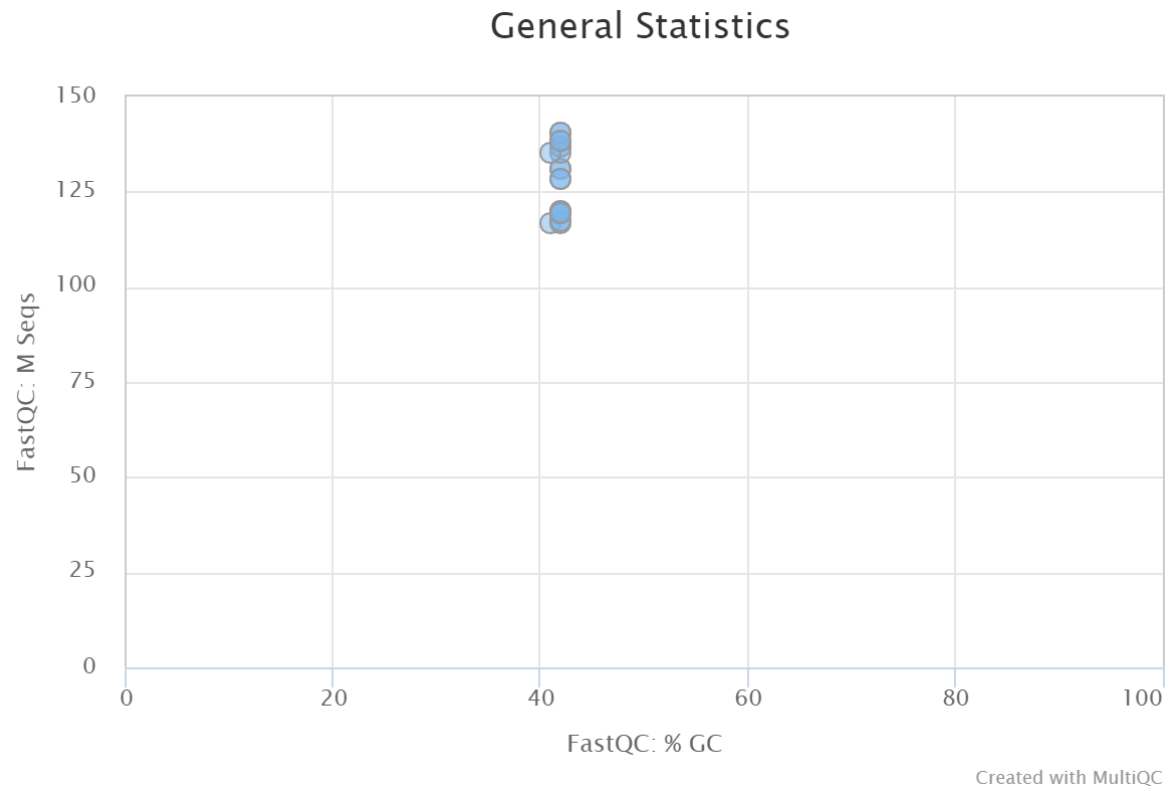
ID de muestra	Total de lecturas (input)	Filtrado				
		Total	Pares	*Forward*	*Reverse*	Removido
SA42911	234680628	183483991 (78.18%)	69545149 (59,27%)	44036251 (37,53%)	357442 (0,30%)	3401472 (2,90%)
SA42912	261574232	203460084 (77.78%)	76248489 (58,30%)	50576026 (38,67%)	387080 (0,30%)	3575521 (2,73%)
SA42913	280736706	217589421 (77.51%)	81087754 (57,77%)	55004432 (39,19%)	409481 (0,29%)	3866686 (2,75%)
SA42914	277282488	212677649 (76.70%)	77749612 (56,08%)	56768235 (40,95%)	410190 (0,30%)	3713207 (2,68%)
SA42976	238325494	176363557 (74%)	62344628 (52,32%)	51324617 (43,07%)	349684 (0,29%)	5143818 (4,32%)
SA42977	273231448	204425160 (74.81%)	72896837 (53,36%)	58230741 (42,62%)	400745 (0,29%)	5087401 (3,72%)
SA42978	232986904	180033491 (77.27%)	66819622 (57,36%)	46034423 (39,52%)	359824 (0,31%)	3279583 (2,82%)
SA42979	270304094	208333691 (77.07%)	77729073 (57,51%)	52493902 (38,84%)	381643 (0,28%)	4547429 (3,36%)
SA42980	239931906	184039930 (76.71%)	67990977 (56,68%)	47697853 (39,76%)	360123 (0,30%)	3917000 (3,27%)
SA42981	256244138	200754847 (78.35%)	77184750 (60,24%)	45992886 (35,90%)	392461 (0,31%)	4551972 (3,55%)

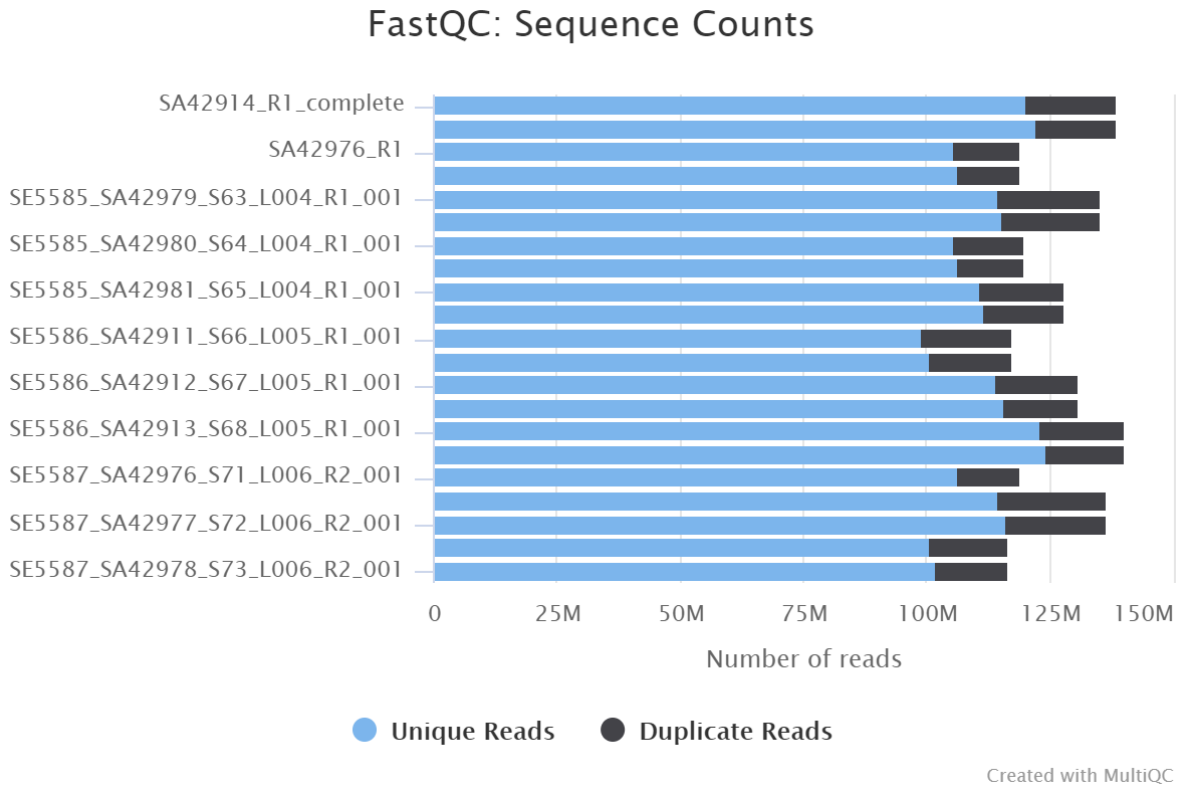
Anexo 4. Estadísticas de alineamiento al genoma de referencia con Bowtie2.

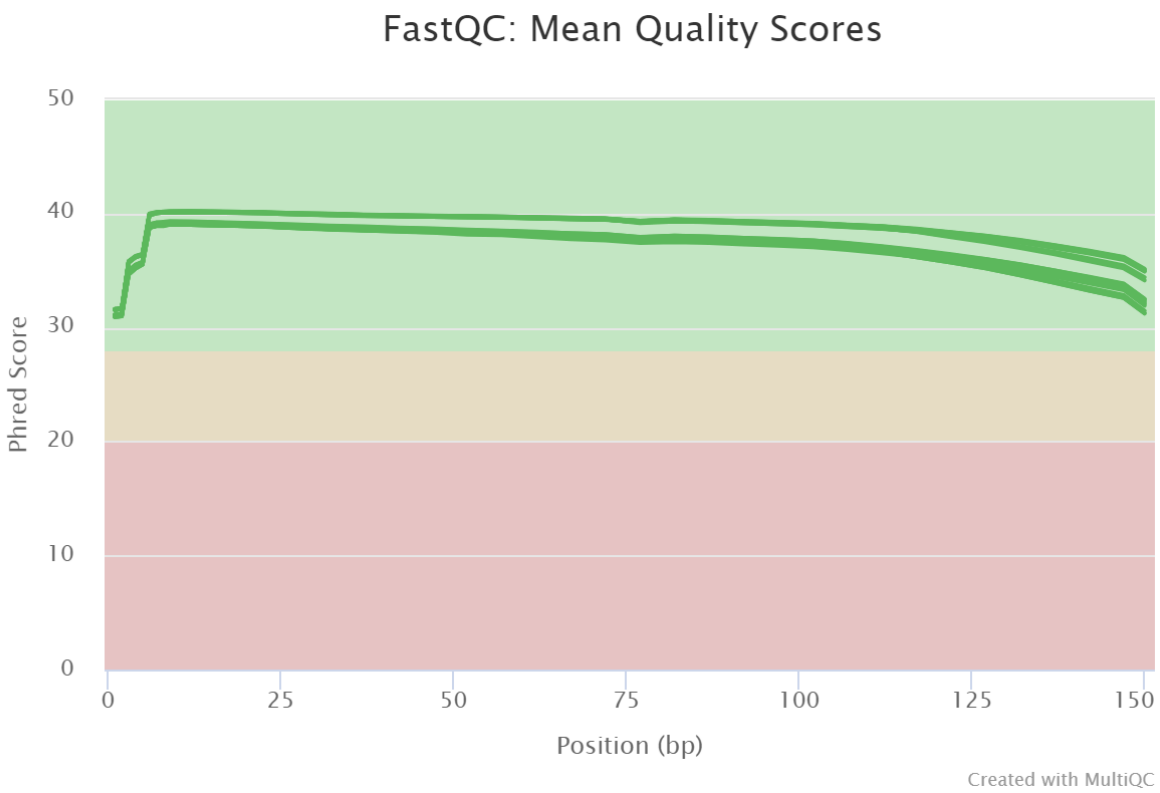
ID de muestra	Total de lecturas (input)	Alineamiento al genoma de referencia de *Cavia porcellus* (Cavpor v3.0)				
		Mapeados	Pares Total	Mapeados propiamente pareados (*)	Singletons (*)	Mapeados a otro cromosoma
SA42911	183483991	165321633 (90.10%)	139090298	116232776 (83.57%)	2017300 (1.45%)	1537040
SA42912	203460084	188240372 (92.52%)	152496978	130864848 (85.81%)	2009144 (1.32%)	1757974
SA42913	217589421	203805382(93.67%)	162175508	140894710 (86.88%)	2022564 (1.25%)	1848414
SA42914	212677649	193804591 (91.13%)	155499224	130803792 (84.12%)	2073025 (1.33%)	1831876
SA42976	176363557	163787034 (92.87%)	124689256	107184272 (85.96%)	1765486 (1,42%)	1291138
SA42977	204425160	184210221 (90.11%)	145793674	120353538 (82.55%)	2417650 (1.66%)	1852026
SA42978	180033491	164099707 (91.15%)	133639244	112697228 (84.33%)	1890455 (1.41%)	1639270
SA42979	208333691	187476377 (89.99%)	155458146	129090146 (83.04%)	2410551 (1.55%)	2091182
SA42980	184039930	170520570 (92.65%)	135981954	116846214 (85.93%)	1767636 (1.30%)	1460966
SA42981	200754847	185283326 (92.29%)	154369500	132828894 (86.05%)	2173371 (1.41%)	1769570

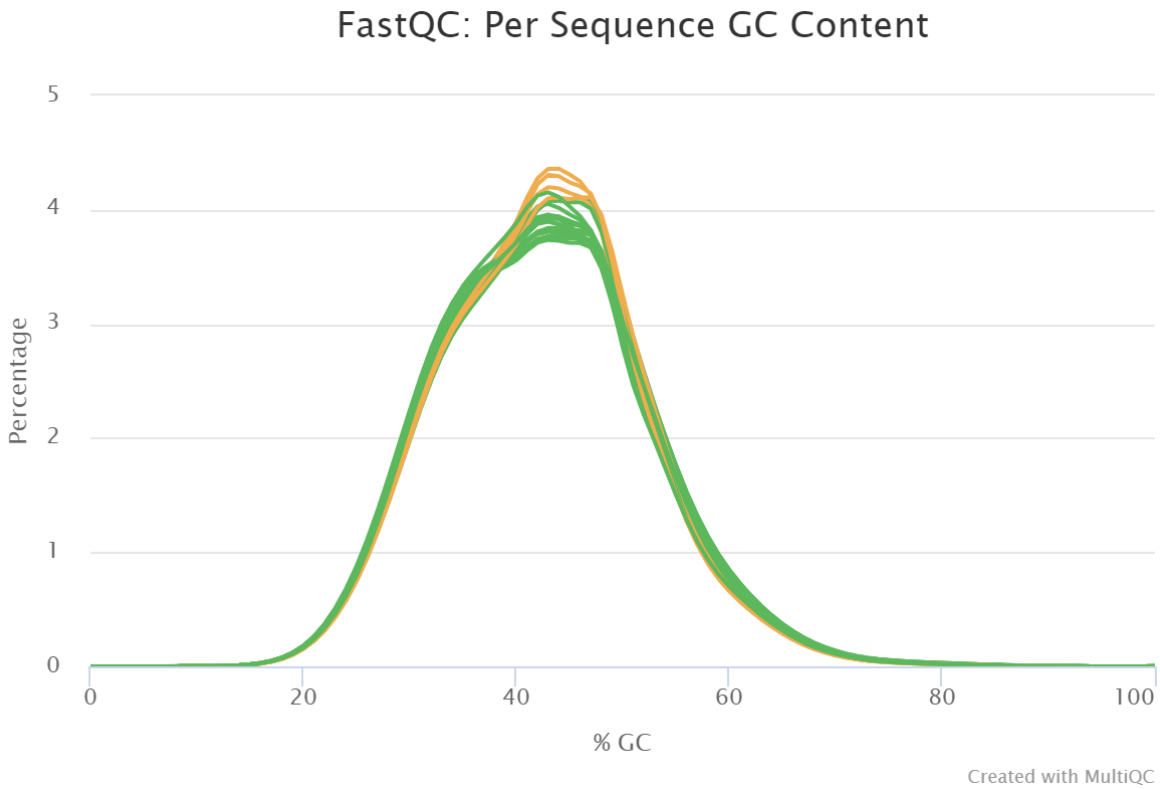
** Porcentaje calculado a partir del total secuencias pares*

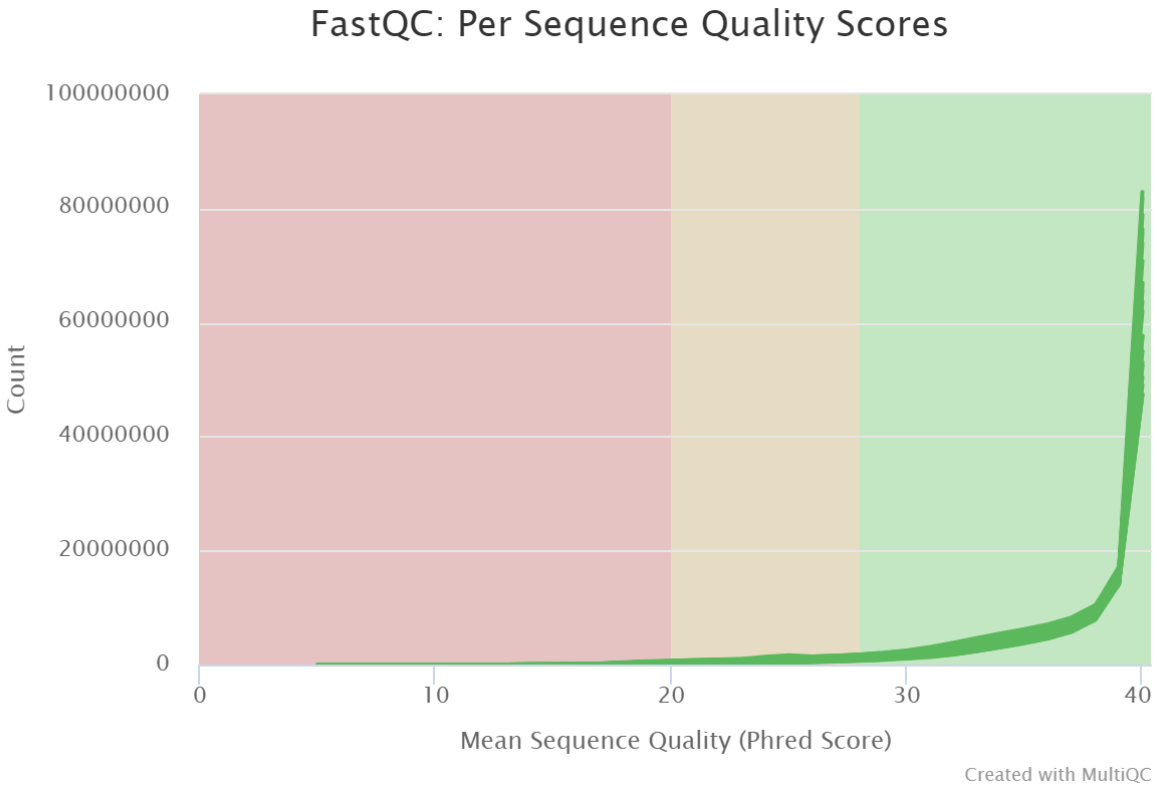
Anexo 5 - Reportes FastQC de genomas

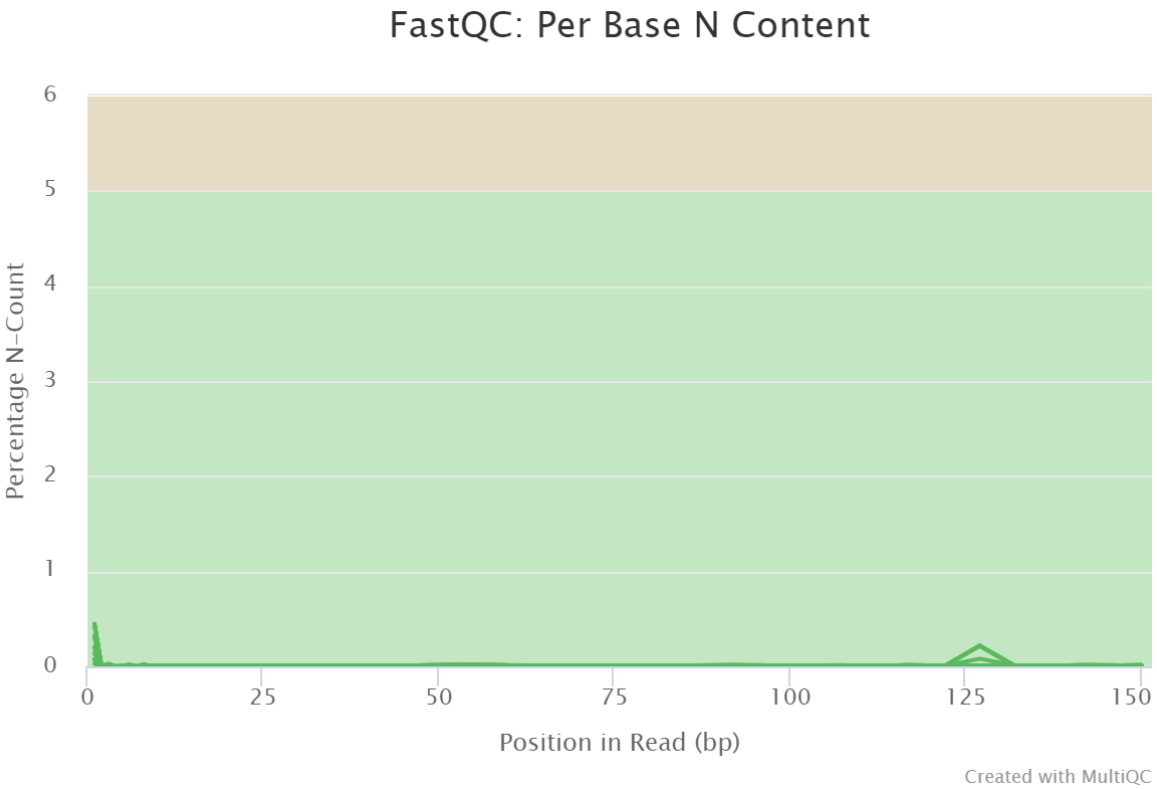


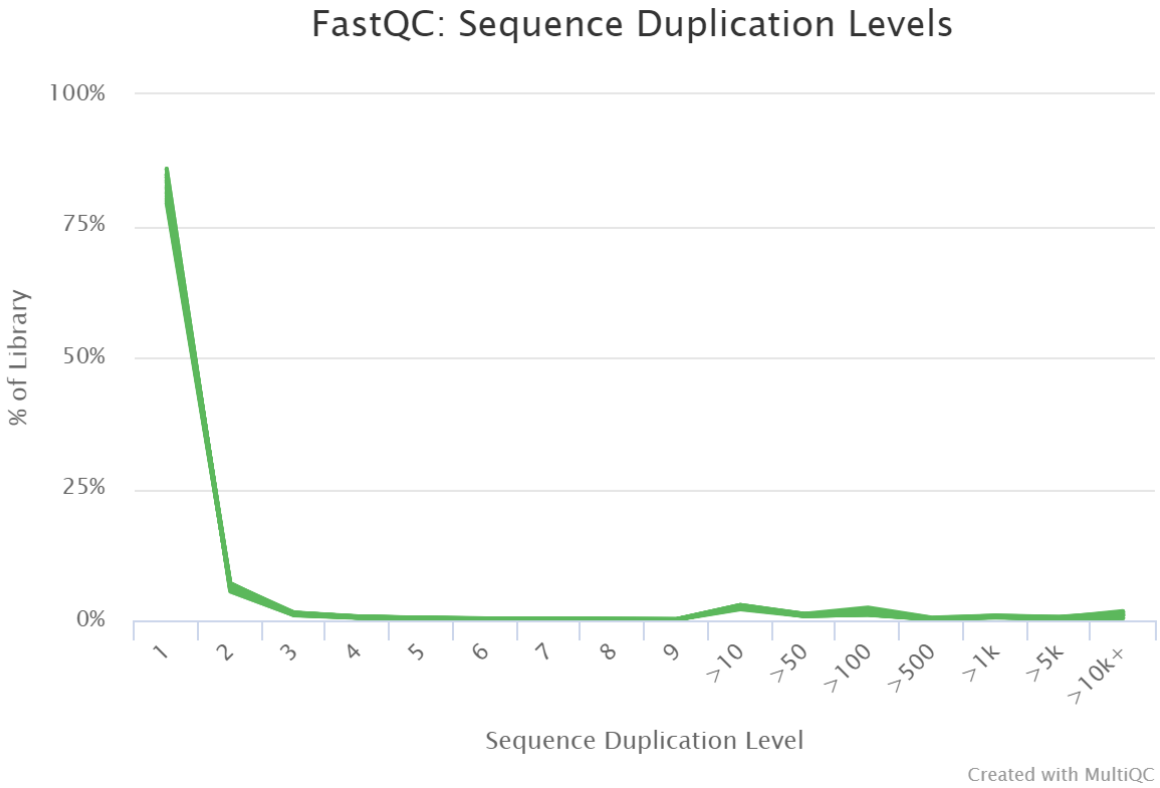


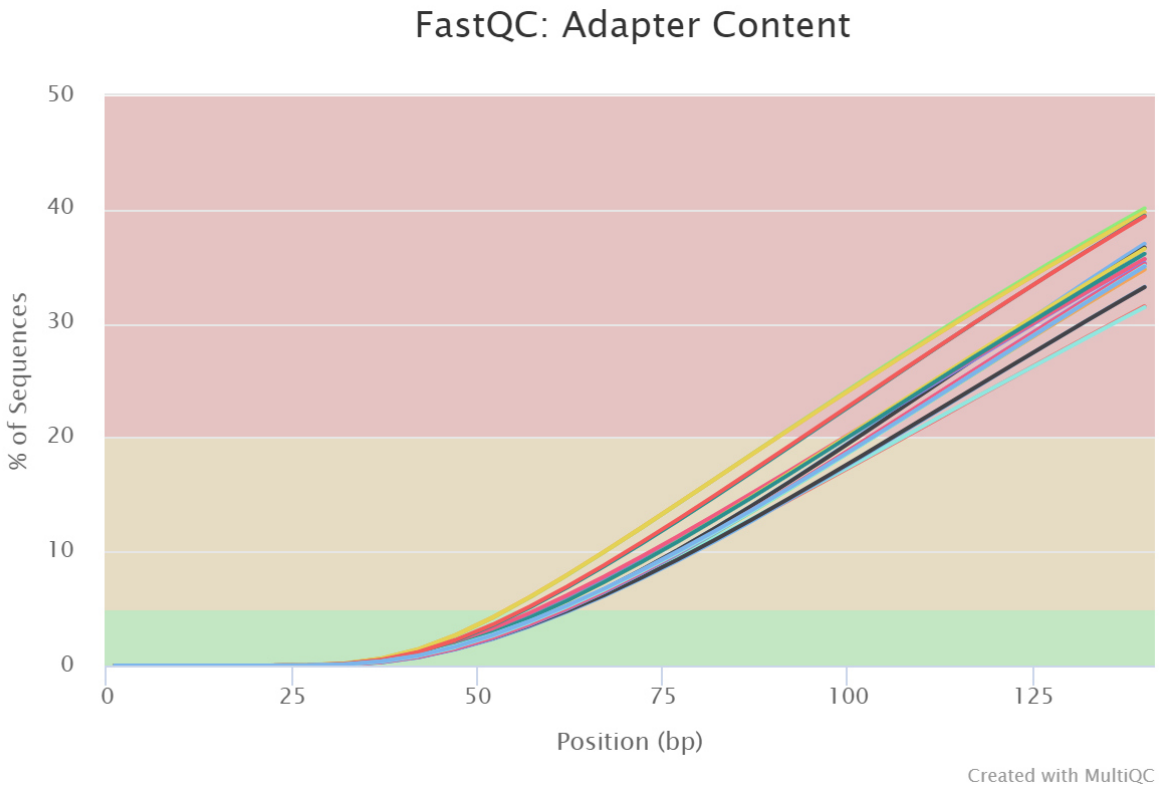

FastQC: Mean Quality Scores
50
40
30
20
10
0
Phred Score
0
25
50
75
100
125
150
Position (bp)
Created with MultiQC


FastQC: Per Sequence GC Content
5
4
3
2
1
0
Percentage
0
20
40
60
80
100
% GC
Created with MultiQC


FastQC: Per Sequence Quality Scores
100000000
80000000
60000000
40000000
20000000
0
Count
0
10
20
30
40
Mean Sequence Quality (Phred Score)
Created with MultiQC


FastQC: Per Base N Content
6
5
4
3
2
1
0
Percentage N-Count
0
25
50
75
100
125
150
Position in Read (bp)
Created with MultiQC


FastQC: Sequence Duplication Levels
100%
75%
50%
25%
0%
% of Library
1
2
3
4
5
6
7
8
9
>10
>50
>100
>500
>1k
>5k
>10k+
Sequence Duplication Level
Created with MultiQC


FastQC: Adapter Content
50
40
30
20
10
0
% of Sequences
0
25
50
75
100
125
Position (bp)
Created with MultiQC