

UNIVERSIDAD MAYOR DE SAN ANDRÉS  
FACULTAD DE CIENCIAS FARMACÉUTICAS Y  
BIOQUÍMICAS,  
CARRERA DE BIOQUIMICA



ANÁLISIS METAGENÓMICO DE COMUNIDADES  
MICROBIANAS EN SUELOS DE CULTIVO DE *Chenopodium*  
*Quinoa* Y SU RELACIÓN CON LAS PROPIEDADES  
FISICOQUÍMICAS Y MANEJO DEL SUELO EN  
DIFERENTES REGIONES DEL ALTIPLANO BOLIVIANO.

“Tesis presentada para optar al grado de Licenciatura en Bioquímica”

POR: ARIEL DE LA FUENTE ARIAS

TUTOR: MARIA TERESA ARMINDA ALVAREZ ALIAGA PhD

LA PAZ - BOLIVIA

Abril, 2022

## Tabla de Contenido.

I. Planteamiento del Problema:	1
A. Justificación	1
B. Objetivos	2
1. Objetivo general	2
2. Objetivos específicos	2
II. Diseño Teórico	3
A. Marco Referencial:	3
1. Antecedentes generales sobre el problema de investigación	3
2. Descripción del ámbito de estudio	4
B. Marco Teórico	5
Diversidad biológica	5
Estudio de comunidades microbianas en ambientes naturales	5
Métodos independientes de cultivo	6
Fases fenológicas del ciclo de cultivo de la quinua	6
Requerimientos del cultivo orgánico de la quinua	7
Adaptabilidad, resistencia y simbiosis de la quinua	8
Estudio metagenómico en suelos de cultivo de Quinua	9
Metagenómica 16S y los proyectos globales	9
C. Marco Conceptual	12
III. Formulación de las Hipótesis de Investigación	14
IV. Operacionalización de las Variables en Estudio	14
V. Diseño Metodológico:	14
A. Población de estudio, Ámbito y Periodo de Investigación	14
B. Tipo de Investigación	15
C. Materiales, Métodos, Técnicas y Procedimientos a ser Empleados	15
1. Descripción de los sitios de estudio	15
a) Descripción del Altiplano Norte	17
b) Descripción del Altiplano Central	18
c) Descripción del Altiplano Sur	19
2. Obtención de las muestras	20
3. Extracción de ADN de muestras de suelos	23
a) Electroforesis horizontal	23
b) Cuantificación y estimación de la concentración de ADN	23
4. Amplificación y Secuenciación de las muestras de ADN	23
5. Pre-procesamiento bioinformático	24

a)	Validación del archivo de asignación .....	24
b)	Ensamblado de las lecturas .....	24
c)	Demultiplexado y filtrado de calidad .....	25
6.	Análisis bioinformático .....	25
a)	Agrupamiento en Unidades Taxonómicas Operativas (OTUs) .....	25
b)	Caracterización taxonómica .....	26
c)	Determinación de la diversidad Alfa y curvas de rarefacción .....	26
d)	Determinación de la diversidad Beta con Unifrac .....	26
7.	Análisis de Correlación Canónica del altiplano sur .....	27
D.	Procesamiento de los datos .....	28
1.	Recolección .....	28
2.	Elaboración .....	28
3.	Análisis .....	28
VI.	Resultados y discusión .....	28
A.	Extracción y cuantificación del ADN .....	28
B.	Características de la secuenciación Paired-End .....	31
C.	Grupos taxonómicos identificados por zona agroproductiva .....	40
D.	Grupos taxonómicos según la fase de crecimiento de la Quinoa .....	46
D.	Análisis de diversidad alfa ( $\alpha$ ) y beta ( $\beta$ ) .....	52
E.	Relaciones entre las comunidades microbianas y las características de los suelos .....	57
VII.	Conclusiones .....	60
VIII.	Bibliografía .....	62

## **Tabla de anexos.**

**Anexo 1.** Protocolo de extracción de ADN mediante kit comercial.

**Anexo 2.** Distribución de la calidad de las bases en las lecturas PairedEnd (forward y reverse) por cada muestra procesada.

**Anexo 3.** Abundancias de los géneros más abundantes en suelos de cultivo por zona agroproductiva.

**Anexo 4.** Géneros bacterianos potencialmente fitopatógenos presentes en los suelos de cada zona agroproductiva por fase de crecimiento de la quinua.

**Anexo 5.** Índices de diversidad alfa de las 30 muestras procesadas.

**Anexo 6.** Gráficas PCoA para los datos obtenidos por Unweighted Unifrac.

## **Tabla de Cuadros.**

**Cuadro 1.** Datos del agrupamiento y ensamblado.....44

**Cuadro 2.** Detalle de los conteos por muestra.....44

## **Indice de tablas.**

<b>Tabla 1.</b> Fases fenológicas del cultivo de quinua.....	6
<b>Tabla 2.</b> Requerimientos particulares del cultivo de Quinua.....	7
<b>Tabla 3.</b> Operacionalización de variables de estudio.....	14
<b>Tabla 4.</b> Características de las plantas de quinua obtenidas de los sitios de estudio.....	20
<b>Tabla 5.</b> Características principales de los puntos de muestro.....	21
<b>Tabla 6.</b> Codificación de las muestras obtenidas.....	22
<b>Tabla 7.</b> Datos agronómicos de los suelos según la zona agroproductiva por tipo de suelo y etapa de desarrollo.....	22
<b>Tabla 8.</b> Características fisicoquímicas de los puntos de muestreo en suelos puruma del altiplano sur.....	28
<b>Tabla 9.</b> Concentraciones e índices de pureza de las muestras de ADN.....	30
<b>Tabla 10.</b> Resultados de la secuenciación 16S para las lecturas Forward.....	40
<b>Tabla 11.</b> Porcentajes de abundancia en muestras puruma y cosecha de las 3 zonas agroproductivas del altiplano a nivel de reino.....	43
<b>Tabla 12.</b> Abundancia a nivel de reino en las muestras del altiplano norte.....	48
<b>Tabla 13.</b> Abundancia a nivel de reino en las muestras del altiplano central.....	49
<b>Tabla 14.</b> Abundancia a nivel de reino en las muestras del altiplano sur.....	50

## Tabla de Figuras.

<b>Figura 1.</b> Pipeline para metagenómica 16S. El pipeline es algo anticuado, pero aún vigente, la línea punteada indica un paso opcional.....	10
<b>Figura 2.</b> Mapa de división de las zonas agroproductivas en el altiplano boliviano.....	15
<b>Figura 3.</b> Precios alcanzados en bolivianos por quintal de quinua entre 2011 y 2015...	16
<b>Figura 4.</b> Imagen del archivo de mapeo (mapping file).....	25
<b>Figura 5.</b> Electroforesis de ADN total de las 30 muestras obtenidas. Se observa los pozos seguidas de las bandas de ADN durante la corrida.....	30
<b>Figura 6.</b> Distribución de la calidad de las lecturas de las bases en las secuencias para la muestra Ca.P.....	33
<b>Figura 7.</b> Se muestra el perfil de calidad de las bases para la muestra 3 (SaP.L).....	35
<b>Figura 8.</b> Calidad de la secuencia por "baldosa" (Per Tile Sequence Quality).....	36
<b>Figura 9.</b> Puntaje de calidad por secuencia. En el eje X se representa la media de la calidad de las secuencias.....	37
<b>Figura 10.</b> Contenido de la secuencia por base en sus posiciones.....	38
<b>Figura 11.</b> Contenido en guanina y citosina (GC) por secuencia.....	39
<b>Figura 12.</b> Contenido de Kmer.....	40
<b>Figura 13.</b> Clasificación taxonómica de los filos más abundantes en las muestras de estudio.....	44
<b>Figura 14.</b> Abundancias de filos en las muestras Puruma según la zona agroproductiva.....	45
<b>Figura 15.</b> Abundancias de los filos más abundantes en suelos de cultivo por zona agroproductiva.....	46
<b>Figura 16.</b> Géneros bacterianos potencialmente fitopatógenos presentes en los suelos de cada zona agroproductiva.....	48
<b>Figura 17.</b> Abundancias a nivel de filo de las tres zonas agroproductivos del altiplano boliviano por fase de desarrollo de la quinua.....	51
<b>Figura 18.</b> Abundancias a nivel de género de las tres zonas agroproductivos del altiplano boliviano por fase de desarrollo de la quinua.....	53
<b>Figura 19.</b> Curvas de rarefacción por muestras para el número de especies observadas (Observed_species).....	55
<b>Figura 20.</b> Índices de diversidad alfa por región altiplánica.....	55
<b>Figura 21.</b> Análisis PCoA por zona agroproductiva (Weighted Unifrac).....	57

<b>Figura 22.</b> Análisis PCoA por etapas de desarrollo de la quinua (Weighted Unifrac)...	58
<b>Figura 23.</b> Análisis PCoA por el tipo de suelo (Weighted Unifrac).....	59
<b>Figura 24.</b> Gráfica de Correlación Canónica de las muestras puruma de las tres zonas agroproductivas.....	60
<b>Figura 25.</b> Gráfica de Correlación Canónica de las muestras de suelos de cultivo de las tres zonas agroproductivas.....	61

#### **Tabla de abreviaturas.**

**DO:** Densidad óptica

**NGS:** Next Generation Sequencing

**PCR:** Reacción en cadena de la polimeraza

**PGPR:** Bacterias promotoras de crecimiento

**QIIME:** Quantitative Insights Into Microbial Ecology

**EMP:** Proyecto del microbio de la tierra

**RDP:** Proyecto de base de datos ribosómico

**HMP:** Proyecto del microbioma humano

**BMP:** Proyecto del microbioma brasileño

**ANAPQUI:** Asociación Nacional de Productores de Quinua

**SDS:** Dodecilsulfato sodico

**SBS:** Secuenciación por síntesis

**ACANA:** ACcurate ANchoring Alignment

**OTU:** Unidad taxonómica operativa

**PCoA:** Análisis de coordenadas principales

**CCA:** Análisis de correlación canónica

## **Resumen.**

Los microorganismos son entidades ubicuas y, dentro de estas, las comunidades bacterias conviven con toda clase de microorganismos y organismos como las plantas, efectuando principalmente una cadena de metabolismo compartido, permitiendo promover el desarrollo de estas así también como afectar o inhibir dicho desarrollo. La identificación de cada una de estas comunidades microbianas por los métodos microbiológicos tradicionales, sería dificultoso e imposible de llevarse a cabo. Es por eso que el presente trabajo reporta la diversidad bacteriana en suelos de tres zonas agroproductivas del altiplano boliviano donde se efectúa el cultivo de *Chenopodium quinua* Willd. empleando nuevas tecnologías de identificación microbiana como la secuenciación de nueva generación de la región V3 y V4 del gen 16S rRNA bacteriano. También se han estudiado algunas condiciones sobre el desarrollo de esta planta tomando en cuenta variables fisicoquímicas y algunas pertinentes con el manejo de los suelos, tratando de determinar su influencia en la diversidad de estas comunidades y su posible efecto en la producción de quinua. De esta forma se ha encontrado que la diversidad es mayor en la zona sur del altiplano seguida del centro y norte, donde la diversidad y riqueza de especies fue menor; coincidiendo en que la producción y las propiedades de las plantas son mejores en el sur y centro del altiplano. Proteobacteria, Actinobacteria, Acidobacteria, Planctomycetes, Bacteroidetes, Verrucomicrobia y Chloroflexi fueron los principales filos representados en todas las muestras, identificando también arqueas. También se ha identificado tanto promotores de crecimiento como fitopatogenos. Además, se vio que durante el desarrollo de la quinua la proporción de secuencias del gen 16S rRNA que no pudieron clasificarse a nivel de filo fue en aumento lo que indica posibles nuevos filos bacterianos presentes en estos suelos. También se observó que la riqueza de especies es mayor en suelos en descanso y son particulares al tipo de suelo y zona agroproductiva. Sin embargo, la diversidad cambia en los suelos de cultivo y pierde relación con las zonas del altiplano debido al manejo del suelo.

**Palabras clave:** Metagenómica, Diversidad microbiana, Secuenciación 16S, Cultivo de quinua

**Abstract.**

Microorganisms are ubiquitous entities and, within these, bacterial communities coexist with all kinds of microorganisms and organisms such as plants, mainly carrying out a chain of shared metabolism, allowing them to promote their development as well as affect or inhibit said development. The identification of each of these microbial communities by traditional microbiological methods would be difficult and impossible to carry out. That is why the present work reports the bacterial diversity in soils of three agro-productive zones of the Bolivian highlands where *Chenopodium quinoa* Willd is cultivated using new microbial identification technologies such as next-generation sequencing of the V3 and V4 region of the bacterial 16S rRNA gene. Some conditions on the development of this plant have also been studied, taking into account physicochemical variables and some ones pertinent to soil management, trying to determine their influence on the diversity of these communities and their possible effect on quinoa production. In this way, it has been found that diversity is greater in the southern area of the highlands followed by the center and north, where the diversity and richness of species was lower; agreeing that the production and properties of the plants are better in the south and center of the highlands. Proteobacteria, Actinobacteria, Acidobacteria, Planctomycetes, Bacteroidetes, Verrucomicrobia and Chloroflexi were the main phyla represented in all the samples and also identifying some archaea phylum. Both growth promoters and phytopathogens have also been identified. In addition, it was seen that during the development of quinoa, the proportion of 16S rRNA gene sequences that could not be classified at the phylum level increased, indicating possible new bacterial phyla present in these soils. It was also observed that the richness of species is greater in fallow soils and are particular to the type of soil and agro-productive zone. However, the diversity changes in the cultivated soils and loses relation with the high plateau areas due to soil management.

**Keywords:** Metagenomics, Microbial diversity, 16S sequencing, Quinoa cultivation

## **I. Planteamiento del Problema:**

### **A. Justificación.**

El altiplano boliviano es una extensa región ubicada a una altura promedio de 3800 msnm, una cuenca intramontañosa, cuyos sedimentos han sido mayoritariamente depositados en condiciones de endorreísmo y levantados por los mismos movimientos tectónicos que han dado origen a la cordillera Andina desde hace unos 15-25 millones de años (SempÈrÈ et al., 1991). Debido a sus características geográficas los climas extremos están presentes en toda su extensión, es por eso que la flora nativa es esteparia lo que permite sobrevivir y resistir estas condiciones. De la flora nativa hay especies que han sido aprovechadas, por ejemplo, para la actividad agrícola, como ciertas variedades de quinua (*Chenopodium quinoa* Willd.) que debido a sus características nutritivas, ha formado parte de la dieta en culturas ya antiguas (Montesinos, 2016).

En la actualidad estas actividades agrícolas se han visto afectadas principalmente por efecto del cambio climático, provocando pérdidas por las alteraciones producidas en las propiedades de las plantas y/o frutos como es el caso de la quinua, producidos en esta región para el consumo propio de los habitantes como para su comercialización y exportación. Este efecto en las propiedades de la quinua se debe a las características de los suelos, relacionadas no solo al cambio climático sino también a la actividad antropogénica. Las propiedades bióticas y abióticas (Cárdenas, Urquiza, Cárdenas, Fernández, & Vladimir, 2015) se ven alteradas, y estas alteraciones influyen en las comunidades microbianas que conforman parte de la simbiosis y mutualismo con las plantas. (Hiruma et al., 2016)

La identificación de cada una de estas comunidades microbianas por los métodos microbiológicos tradicionales, como el cultivo en placas Petri y otros, sería difícil y imposible de llevarse a cabo, debido a las condiciones que requieren estos microorganismos, es por eso que entre el 80 y 90% de estos son viables no cultivables (Hernández-León, Velázquez-Sepúlveda, Orozco-Mosqueda, & Santoyo, 2010).

En los últimos años han surgido nuevas tecnologías en el campo de la genómica, que permiten estudiar una amplia diversidad de microorganismos en una misma muestra (Rodríguez-Herrera et al., 2009). Estas tecnologías construyen las secuencias genéticas presentes en esas muestras por lo que el resultado en bruto es la obtención de una gran cantidad de información. Es por esta razón que es imprescindible el uso de programas informáticos (bases de datos) que facilitan el procesamiento de esa información y así efectuar la interpretación y el análisis de los resultados de ese procesamiento. De esta forma es posible contestar diferentes interrogantes, además, permite comprender las relaciones simbióticas, mostrando qué tan perjudiciales o beneficiosas podrían llegar a ser estas comunidades microbianas para el cultivo de la quinua.

Es así que algunas instituciones reconocidas en diferentes países han estado implementando estas nuevas tecnologías como la secuenciación y la bioinformática dirigido a diferentes ámbitos, como la biorremediación, estudios clínicos, etc. En este caso la aplicación en el estudio de suelos permitió comprender como los microorganismos perciben y reaccionan a los cambios en su medio ambiente e identificar los tipos de comunidades que habitan y así poder caracterizar algún potencial que favorezca por ejemplo el cultivo de la quinua como tal.

El desarrollo y mejora de estas tecnologías es constante, en el caso de la secuenciación de nueva generación NGS (por sus siglas en inglés), efectúa la secuenciación, rápida, masiva y a bajo costo, de regiones genómicas o genomas completos, haciendo posible estudios genómicos de organismos no modelo y estudios metagenómicos. Estas tecnologías permiten y prometen estudios ecológicos y evolutivos hacia asuntos relacionados con conservación y manejo de la diversidad biológica ante retos como es el cambio climático. (Escalante, Barbolla, Ramírez-Barahona, & Eguiarte, 2014)

Planteados estos elementos las cuestiones que nos llevará a efectuar el trabajo de investigación son: ¿Cuáles son las comunidades microbianas presentes en los suelos asociados al cultivo de la quinua? Y de estas ¿Cuál es su relación con las características fisicoquímicas del suelo y con el tipo de manejo de los suelos en el cultivo de quinua?

## **B. Objetivos**

### **1. Objetivo general.**

Establecer la relación entre las comunidades microbianas de suelos de cultivo de quinua (*Chenopodium quinua* Wild) con las propiedades fisicoquímicas que presenta el suelo y el tipo de manejo efectuado a los suelos de cultivo empleando el análisis metagenómico.

### **2. Objetivos específicos.**

- a) Describir el análisis metagenómico del gen 16S rRNA de comunidades microbianas procedentes de muestras de suelos utilizados en el cultivo de quinua.
- b) Determinar las diferencias entre las abundancias y riqueza de las comunidades microbianas presentes en suelos de cultivo de quinua de acuerdo a su distribución en las zonas agroproductivas como también a la fase de desarrollo de la quinua.
- c) Analizar la correlación multivariante entre las abundancias bacterianas versus variables fisicoquímicas correspondientes a los suelos y a las zonas agroproductivas del cultivo de quinua.

## II. Diseño Teórico

### A. Marco Referencial:

#### 1. Antecedentes generales sobre el problema de investigación.

La quinua es un alimento muy importante por sus propiedades nutricionales, un alimento ancestral de los andes sudamericanos. Su cultivo se ha extendido a otros continentes gracias a su variabilidad genética, que permite su cultivo en diversas zonas agroecológicas y en diferentes condiciones climatológicas. Esto permitió aumentar el nivel de producción así como el nivel de exportación y dado el gran interés en este alimento en los últimos años su valor en mercado se a quintuplicado desde 1980 (Juan Risi, 2015).

En Bolivia, su expansión se dio principalmente en el altiplano sur, en los departamentos de Oruro y Potosí que circundan los salares de Uyuni y Coipasa, llegando a ser una expansión desmedida, poniendo en riesgo recursos naturales, además empleando prácticas de cultivo intensivas que causan un serio deterioro del suelo (Juan Risi, 2015).

El suelo de la región andina se caracteriza por las diferentes condiciones de acidez, alcalinidad, presencia de sales, texturas diversas, pedregosidad, contenido de materia orgánica variable. La región Norte se caracteriza por un mayor contenido en rocas, además posee una mayor precipitación fluvial; la región Central es semi-árida con suelos pedregosos a arenosos ligeramente salinos o alcalinos; y la región Sur es más árida y presenta suelos arenosos (Juan Risi, 2015). Además, los suelos de la región sur tienen una baja capacidad de retención de agua, sumado a las bajas precipitaciones (entre 250 y 550 mm/año) (Gomez-Montano et al., 2013), hacen que se de mayor importancia el estudio de esta planta en conjunto a las características y propiedades del suelo.

Se han realizado diversos estudios concernientes a las características geográficas que presenta el altiplano boliviano y se han identificado factores que han alterado y desgastado la calidad de los suelos, como el uso de plaguicidas, la actividad agrícola y más importante el cambio climático. (Ruiz et al., 2013).

Este efecto en la calidad del suelo tiene un impacto negativo en el cultivo de los diferentes productos agrícolas como la quinua, además de soportar los climas extremos su desarrollo se ve afectado por las alteraciones en las propiedades de los suelo, llegando a producir pérdidas en el sector. (Montesinos & Montesinos, 2017)

Otros estudios del suelo demuestran el efecto que tiene el cambio climático en las comunidades microbianas que forman parte de la simbiosis con el crecimiento de la quinua. Además, estas comunidades no solo son afectadas por el cambio climático sino por el manejo del suelo como el uso de

agroquímicos, el monocultivo, el tipo de suelo, etc. (Gomez-Montano et al., 2013)

El reciente interés en mantener la calidad del suelo ha sido estimulado por un conocimiento renovado de la importancia de la condición del suelo para la sostenibilidad de los sistemas de producción agrícola y la calidad del medio ambiente.

Determinar qué tipos de comunidades microbianas están presentes en estos suelos permitiría comprender la relación y efecto que podrían tener con el cultivo de la quinua, y qué condiciones serían favorables para esta actividad.

Un estudio de suelos en el altiplano norte y central que empleó la pirosecuenciación demostró que la dinámica de comunidades bacterianas y fúngicas se ve alterada por el tipo de manejo de los suelos y la remoción de plantas nativas, lo que permitió determinar que los periodos de barbecho influye negativamente a las comunidades microbianas. Esto permitió considerar programas de restauración y recomendaciones para los agricultores en el manejo de suelos. (Gomez-Montano et al., 2013)

La aplicación de programas bioinformáticos en nuestro medio es aun reducida, teniendo en cuenta que la información generada por las tecnologías de secuenciación de nueva generación (NGS) nos dan la capacidad de gestionar y analizar la relación de las comunidades microbianas con los diferentes ecosistemas, tanto de ríos como del sistema digestivo humano entre muchos otros. (Zamora, Malaver, & Ramos, 2012)

## **2. Descripción del ámbito de estudio.**

Las bacterias presentes en el suelo cumplen en conjunto con otros tipos de microorganismos funciones muy importantes que permiten la nutrición y desarrollo de las plantas, como la transformación de compuestos orgánicos e inorgánicos, la humificación, etc. Es por esta razón su interés en el cultivo de la quinua lo que permitirá determinar su relación con las propiedades de los suelos como con el manejo de los mismos. Además de poder determinar si existen cambios temporales y espaciales de estas en los suelos de diferentes localidades del altiplano boliviano.

Dada su importancia, poder identificar estas comunidades empleando técnicas tradicionales resulta complicada e imposible de llevar a cabo, es por esta razón que el surgimiento de nuevas tecnologías en el campo de la secuenciación permiten identificar esas comunidades microbianas, su predominancia, a través de las secuencias nucleotídicas, generando conocimiento de la diversidad microbiana presente en los suelos donde se realiza el cultivo de la quinua.

La quinua (*Chenopodium quinua* Willd.), es un pseudocereal nativo de las regiones andinas de Sudamérica (Jancurová, 2009). La quinua es cultivada en

rotación con otras especies agrícolas en el norte del altiplano, estas plantas de quinua llegan a una altura de 1.06 m, mientras en el sur el cultivo no esta en rotación con otras especies, y su tamaño alcanza los 1.38 m. Este estudio permitirá determinar si existen diferencias entre las comunidades microbianas de las 3 regiones del altiplano y si tiene relación o no con los monocultivos.

## **B. Marco Teorico.**

### **Diversidad biológica.**

Se define a la biodiversidad o diversidad biológica como "la variabilidad de organismos vivos de cualquier fuente, incluidos, entre otras cosas, los ecosistemas terrestres y marinos y otros ecosistemas acuáticos y los complejos ecológicos de los que forman parte. Para generar conocimiento sobre la diversidad biológica se requiere considerar los diferentes niveles jerárquicos de organización de la vida (genes, especies, poblaciones, comunidades y ecosistemas), junto con su composición, estructura y funcionalidad, lo cual puede abordarse a partir de tres preguntas: ¿qué elementos la componen?, ¿cómo están organizados? y ¿cómo interactúan? (Noss, 1990)

### **Estudio de comunidades microbianas en ambientes naturales**

El desarrollo de técnicas para la extracción directa del ADN o ARN de comunidades a partir de muestras ambientales y su análisis ha hecho posible un mayor conocimiento de la diversidad microbiana. Las técnicas para la extracción de ADN usualmente son una combinación de métodos físicos y químicos para romper las células tratando de minimizar el daño causado a los ácidos nucleicos, sin embargo, estas técnicas no logran romper la pared celular de todas las células en la comunidad y la pureza de los ácidos nucleicos pueden presentar problemas para su posterior amplificación por medio de la reacción en cadena de la polimerasa (PCR). La PCR permite la detección de fragmentos de ADN que se encuentran en poca cantidad en muestras ambientales, no obstante, esta técnica presenta sesgos relacionados con la diferencia en el acceso del iniciador (primer) a la cadena molde y la eficiencia en la extensión. (Polz & Cavanaugh, 1998)

El objetivo clave en la metagenómica comparativa es poder identificar grupos microbianos que son responsables de otorgar características específicas a cierto ecosistema. Además, existen algoritmos que proporcionan una visión de las interacciones intermicrobianas, permiten la agrupación de los metagenomas basándose en patrones intermicrobianos probables y permiten a los usuarios realizar un análisis comparativo a través de microbiomas. (Wikipedia, 2019) La metagenómica como una herramienta adicional en los sitios de contaminación, puede mejorar las estrategias para

monitorear el impacto de estos en los ecosistemas y para limpiar los ambientes contaminados. (Techtmann & Hazen, 2016)

Las tecnologías de NGS disminuyen los costes de secuenciación de las bibliotecas de ADN, permitiendo su accesibilidad a usuarios o instituciones que requieran aplicar a estudios de diferente escala. Estas tecnologías recientes nos permiten secuenciar el ADN y el ARN mucho más rápido en comparación con la secuenciación de Sanger, y como tal han revolucionado el estudio de la genómica y la biología molecular (Rodríguez-Santiago & Armengol, 2012).

### **Métodos independientes de cultivo.**

Los métodos moleculares, desde los años 90, iniciaron nuevas formas de comprensión dentro la microbiología y la ecología microbiana. Estos métodos llamados independientes de cultivo, son capaces de detectar e identificar microorganismos directamente en una muestra sin cultivarlos ni aislarlos, ya que analizan su ADN o ARN.

Estos métodos emplean la extracción directa de los ácidos nucleicos de las células, que luego son analizados mediante métodos que permiten resaltar las características de una o varias comunidades microbianas.

### **Fases fenológicas del ciclo de cultivo de la quinua.**

Las duraciones de las fases del cultivo están dadas por los factores ambientales y condiciones edáficas, por ejemplo, estas fases se alargan más en un suelo franco arcilloso que en un suelo franco arenoso, esto por la fácil retención de humedad del primero.

**Tabla 1.** Fases fenológicas del cultivo de quinua.

Fase	Descripción	Tiempo (después de la siembra)
Pre-emergencia	Comprende la germinación, enraizamiento y comienzo de la emergencia.	Una semana
Emergencia	La plántula emerge del suelo y extiende las hojas cotiledonales.	Dos semanas
Dos hojas verdaderas	Dos hojas verdaderas extendidas de forma lanceolada y en la yema apical el siguiente par de hojas.	10 a 15 días
4 y 6 hojas verdaderas	Tres pares de hojas	25 a 45 días
Ramificación	Ocho hojas verdaderas extendidas con presencia de hojas axilares hasta el tercer nudo, las hojas cotiledonales se caen y dejan cicatrices en el tallo	45 a 50 días

Fase	Descripción	Tiempo (después de la siembra)
Inicio y desarrollo del panojamiento	La inflorescencia va emergiendo del ápice y aglomeración de hojas.	55 a 75 días
Inicio de floración y floración	La flor hermafrodita apical se abre mostrando los estambres Separados, anteras protegidas por el perigonio de un color verde limón. Esta fase es muy sensible a las heladas y granizadas	75 a 80 días
Grano lechoso y pastoso	Los frutos en los glomérulos contienen una sustancia lechosa en esta fase y no debe existir un déficit hídrico. Y luego presenta una consistencia pastosa	100 a 160 días
Madurez fisiológica	Cuando el grano presenta resistencia a la presión o penetración, contenido de humedad del grano varía de 14 a 16%, constituye el periodo de llenado del grano, la presencia de lluvia es perjudicial, empieza a secarse y es el momento de realizar la cosecha.	160 a 180 días

Fuente: SEPHU “CULTIVO DE LA QUÍNOA ORGÁNICA”

### Requerimientos del cultivo orgánico de la quínoa.

La quínoa presenta una amplia variedad y diversidad, alcanzada por las diferentes características geográficas de la región en la que se originó y desarrolló, es por eso que esta planta tiene algunos requerimientos particulares como se indica en la Tabla 2.

**Tabla 2.** Requerimientos particulares del cultivo de Quínoa.

Requerimiento	Descripción
Suelos	Se adapta a diferentes tipos de suelo, pero es preferible los franco-arenosos a los franco-arcillosos, llanos o con pendientes moderadas, con profundidad media.
pH	Es preferibles suelos con pH de 6 a 8.5 y una conductividad eléctrica de 12 mhos/cm
Agua	Son óptimas las precipitaciones de 300 a 500 mm y máximo de 600 a 800 mm, aunque posee características que la hacen soportar la falta de humedad.
Temperatura	La óptima comprende los 8 a 15 °C, pero puede soportar hasta -4 °C y heladas hasta 15 o 20 días.
Sequias	Soporta hasta 60 días (excepto en la etapa de germinación, florescencia y madurez)
Humedad	El exceso de humedad es dañino.
Radiación solar	Es favorables sectores de alta iluminación solar, requiriendo un promedio de 12 horas de luz.
Altitud	Hasta los 4000 msnm

### **Adaptabilidad, resistencia y simbiosis de la quinua.**

La quinua ha sido, durante varios años, objeto de diferentes tipos de estudios e investigaciones, todas relacionadas principalmente a sus características nutritivas, diversidad o variedad, condiciones edáficas, a su importancia económica y en los últimos años referidas al cambio climático y su efecto en esta, tanto por su rendimiento y calidad como por las consecuencias económicas. Así, el uso de nuevas variedades de quinua de alto rendimiento, los agricultores podrán mejorar sus ingresos.

Montesinos, (2016) realizó una evaluación de la adaptación de la quinua al cambio climático en el Perú. Dado que el cultivo de esta planta ha enfrentado los nuevos factores abióticos y bióticos producidos por el cambio en las precipitaciones, sequías, heladas, etc. La evaluación se realizó durante 4 etapas de crecimiento de la planta, que son el crecimiento (a los 30 días), floración (a los 60 días), fructificación (a los 120 días) y madurez (a los 150 días). Se evaluó la altura de la planta y la altura de la panoja. Se seleccionaron 5 variedades de semillas y una nativa, las cuales fueron tratadas y cultivadas en 2 regiones. En general las semillas seleccionadas tuvieron una mayor producción en comparación con la semilla nativa en aproximadamente 22% y 32%. Los resultados de análisis estadísticos indican que la densidad de cultivo ha incrementado la producción y por lo tanto favorece los ingresos de los agricultores.

Sivila & Hervé, (1999) realizaron un estudio de los suelos de cultivo y su relación con la microbiota y los factores que intervienen en su presencia y actividad. Este estudio se realizó en las comunidades de Huaraco, Pumani y Patarani (provincia Aroma) a una altura promedio de 3800 msnm. Su estudio reflejó una amplia diferencia en la microbiota presente en el suelo entre el cultivo de quinua y el cultivo de papa, cebada y avena, resaltando a la quinua por poseer una mayor resistencia a micorrizas y hongos principalmente.

Entre estos estudios, también están aquellos que estudiaron las variedades de quinua que tienen una mayor resistencia a infecciones, especialmente al Mildiu, un conjunto de enfermedades causadas por un hongo (Gabriel, Pereira, Angulo, Magne, & Gandarillas, 2012). Estos señalan que las variedades de quinua que presentaban un mayor diámetro tanto de tallo como de panoja (procedente de valles), tenían una mayor resistencia al mildiu, sin embargo, el diámetro y grosor de los granos era menor en comparación a variedades del altiplano, una característica de importancia económica.

También se estudian bacterias endófitas o bacterias promotoras de crecimiento (PGPR), por su rol en el metabolismo de compuestos requeridos por las plantas, como la solubilización de fósforo, fijación de nitrógeno y otros, siendo los del género *Paenibacillus* las promotoras de crecimiento más promisorias (Angulo, Mollinedo, Claros, & Ortuño, 2013). Las bacterias

endófitas permiten la síntesis de ciertas sustancias reguladoras de crecimiento, como giberelinas, citocininas y auxinas, haciendo que las plantas sean más vigorosas, productivas y tolerantes a condiciones climáticas adversas. En un estudio se obtuvo 164 aislados bacterianos endófitos procedentes de 3 cultivares (Oruro) de los cuales 72 fueron solubilizadoras de fosfato, 59 producen la fitohormona del ácido indol acético y 31 fueron fijadoras de nitrógeno.

### **Estudio metagenómico en suelos de cultivo de Quinoa.**

Con el auge de las nuevas tecnologías biotecnológicas, el estudio de microorganismos presentes en diversos tipos de ambientes está en actual desarrollo y mejora. Tal es el caso de la metagenómica, que ha permitido identificar microorganismos que no hubieran sido posibles aislarlos por los métodos microbiológicos tradicionales (Hernández-León et al., 2010).

Un estudio reciente de la diversidad microbiana se llevó a cabo en cultivos de quinoa, los objetivos de éste fueron: Evaluar los efectos del periodo de barbecho y la presencia de tola en las características fisicoquímicas del suelo, así como en la diversidad microbiana de bacterias y hongos. Este estudio tuvo lugar en dos comunidades del departamento de La Paz (Umala y Ancoraimes). Los resultados refutaron la hipótesis inicial que periodos largos de barbecho aumentan la diversidad microbiana y que por ende enriquecen el suelo.

En este estudio se observó que los periodos de barbecho largos tuvieron efectos diferentes entre las 2 comunidades, en Umala la diversidad de comunidades fúngicas disminuyó en comparación con Ancoraimes y en comunidades bacterias sucedió lo contrario, un género muy frecuente en Umala fue *Paenibacillus*, pero estuvo ausente en Ancoraimes, este género es importante en la fijación de nitrógeno. Pero también Umala presentó géneros que incluyen patógenos de plantas y antagonistas de patógenos de plantas. También se han identificado patógenos de hongos, algunos aumentaron mientras que otros disminuyeron con periodos largos de barbecho. Adicionalmente se han hecho relaciones entre las comunidades bacterianas y la presencia de plantas silvestres (como la tola) en las que se observó que estas tienen un efecto en la composición de diversidad microbiana en periodos largos de barbecho, por ejemplo, la presencia de tola disminuyó la frecuencia de hongos patógenos como también de bacterias útiles a las plantas.

### **Metagenómica 16S y los proyectos globales.**

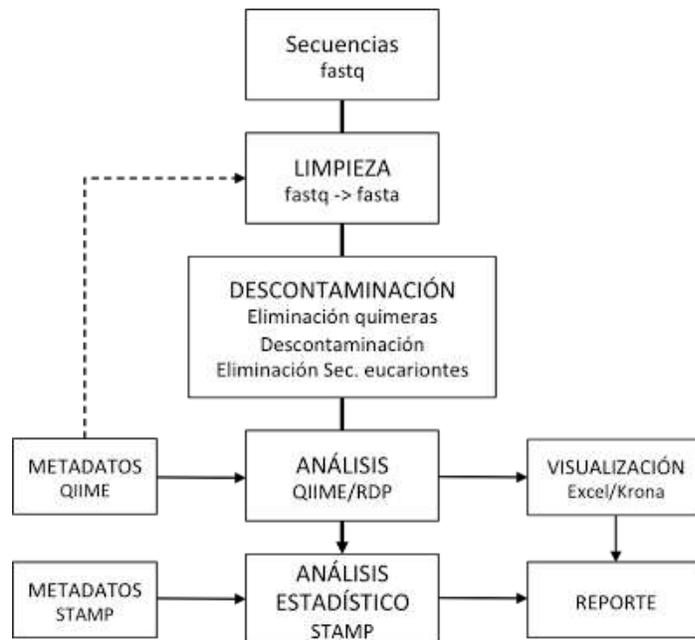
La metagenómica 16S forma parte de la metagenómica dirigida, es decir se da enfoque a amplicones, en este caso del gen ARNr 16S presente en procariontes, arqueas, entre otros (Ramírez Muñoz & González Pazmiño,

2015). Este tipo de estudio, parte de la bioinformática, permite realizar una clasificación a nivel taxonómico ya que este gen presenta regiones conservadas y variables. Además, permite identificar ‘QUIENES ESTÁN PRESENTES’ en una sola muestra e indicar ‘QUÉ RELACIÓN TIENEN’ con factores o variables de su entorno, algo que por métodos tradicionales es prácticamente imposible (Hernández-León et al., 2010).

El estudio metagenómico hace uso de NGS y de herramientas informáticas, evitando técnicas o análisis microbiológicos y/o bioquímicos tradicionales (Mocali & Benedetti, 2010), para llegar a determinar a microorganismos presentes y la proporción en que se encuentran en esa muestra.

Estas herramientas informáticas que funcionan en ‘cadena’ de procesos son conocidas como ‘Pipelines’, es decir una combinación de paquetes o programas que se arman de acuerdo a un estándar o requerimiento, que procesan la información generada por NGS. Actualmente existen diferentes Pipelines disponibles, tanto online o para su uso autónomo en una computadora personal.

Uno de los Pipelines más empleados es Qiime (Perspectivas cuantitativas en ecología microbiana), la característica de este software, que ostenta un flujo de trabajo generalizado (Figura 1), es que permite al usuario programar un Pipeline para un análisis particular. Siendo un software abierto, permite interactuar con otros programas o bases de datos de terceros, que amplía las posibilidades en el procesamiento y análisis de datos.



**Figura 1.** Pipeline para metagenómica 16S. El pipeline es algo anticuado, pero aún vigente, la línea punteada indica un paso opcional. (Bruno Gómez Gil y Raúl Llera).

En general existe un pipeline frecuente en el procesamiento de datos de secuenciación 16S, varios de estos programas van agregando nuevas funciones lo que permite mejorar y ampliar los alcances en sus aplicaciones, integrando prácticamente un pipeline completo, es decir desde el preprocesamiento de las lecturas crudas hasta obtener los resultados estadísticos.

En paralelo, también se han ido creando organizaciones o comunidades de científicos que han publicado estándares para el análisis de secuencias 16S y 18S, todo enfocado a NGS, como el Proyecto del Microbioma de la Tierra creado en 2010, EMP (por sus siglas en inglés) <http://press.igsb.anl.gov/earthmicrobiome>, que como indica “es un intento sistemático de caracterizar la diversidad microbiana taxonómica y funcional a nivel mundial en beneficio del planeta y la humanidad”, su visión: construir el mapa microbiano de la Tierra, este sitio ofrece servicios de secuenciación (16S o 18S), procesamiento y análisis, software (plugin para Qiime) y brinda protocolos y estándares (desde la extracción del ADN hasta el preprocesamiento de secuencias) para poder adicionar estudios individuales o particulares dentro este proyecto (Gilbert, Jansson, & Knight, 2014).

Otros proyectos o instituciones enfocados a la metagenómica son: Proyecto de Base de Datos Ribosómicos (RDP), ampliamente usado, proporciona una Base de Datos de secuencias rARN 16S de Bacterias y Arqueas muy controladas, alineadas y de calidad, y recientemente secuencias de ARNr 28S Fúngicos, también proporciona un conjunto de herramientas de análisis (Cole et al., 2014).

El Proyecto del microbioma humano (HMP), cuya misión es la caracterización integral del microbioma humano y el análisis de su rol en la salud humana y las enfermedades. Provee flujos de trabajo, protocolos estandarizados y herramientas para diferentes tecnologías de secuenciación y plataformas (Huttenhower et al., 2012).

El Proyecto de Microbioma Brasileño (BMP) apunta a reunir un Consorcio/ Base de datos metagenómico brasileño sea de diferentes hábitats, entornos o humanos. Que coordina y estandariza los diversos proyectos que se realizan en Brasil, junto con proyectos futuros, mediante recursos como protocolos y herramientas (16S, 18S e ITS). También cuenta con un Sistema Operativo basado en Linux (Pylro et al., 2014).

Y entre otros como SILVAngs, MG-RAST, MG7 cuentan con sus pipelines estandarizados siendo el mayor enfoque la secuenciación 16S, 18S e ITS, y a excepción de MG7, una institución privada, la mayoría es de acceso libre requiriendo solo un registro de usuario, permitiendo su uso o modificación de acuerdo a sus requerimientos o criterios.

## **C. Marco Conceptual.**

### **CAMBIO CLIMATICO**

Se trata de una variación en el estado del sistema climático que perdura durante periodos de tiempo suficientemente largos para suponer que se ha alcanzado un nuevo equilibrio. Puede afectar tanto a los valores medios del clima como a su variabilidad y extremos. Son graduales o abruptos debido a causas diversas, como las relacionadas con los cambios en los parámetros orbitales, la deriva continental o periodos de vulcanismo intenso. El cambio climático actual es de origen antropogénico y se relaciona principalmente con la intensificación del efecto invernadero debido a las emisiones humanas procedentes de la quema de combustibles fósiles.

### **IMPACTO AMBIENTAL**

Modificaciones al medio ambiente ocasionadas por actividades antropogénicas o por fenómenos naturales.

### **INDICADOR DE CALIDAD DE LA TIERRA**

Los ICT son datos estadísticos que informan sobre la condición y la calidad del recurso tierra y también sobre las relaciones causa-efecto que pueden dar lugar a cambios en su calidad y las respuestas de la sociedad a esos cambios.

Capacidad de campo: contenido de agua o humedad que es capaz de retener el suelo luego de la saturación o de haber sido mojado abundantemente y dejado drenar libremente.

### **CALIDAD DEL SUELO**

La calidad del suelo se basa en la premisa de que su manejo puede deteriorar, estabilizar o mejorar las funciones del ecosistema suelo.

### **DIVERSIDAD BIOLOGICA**

La variedad de organismos vivos considerando todos los niveles de organización, incluyendo los genes, especies y a los niveles taxonómicos más altos. La diversidad biológica incluye también la variedad de hábitat y ecosistemas, así como los procesos naturales que ocurren dentro de los mismos.

### **DIVERSIDAD ALFA**

La diversidad alfa se considera como la riqueza biológica de un determinado hábitat o bien el número de especies en una localidad.

### **DIVERSIDAD BETA**

A la diversidad beta también se le considera como el recambio de especies de una localidad a otra o bien la riqueza de especies promedio de las localidades de cierta región.

## SECUENCIACIÓN DEL ADN

La secuenciación del ADN es un conjunto de métodos y técnicas bioquímicas cuya finalidad es la determinación del orden de los nucleótidos (A, C, G y T) en un oligonucleótido de ADN. La secuencia de ADN constituye la información genética heredable que forman la base de los sistemas de desarrollo de los seres vivos. Así pues, determinar la secuencia de ADN es útil en el estudio de la investigación básica de los procesos biológicos fundamentales, así como en campos aplicados.

## METAGENÓMICA

La metagenómica se define como el estudio del material genético, el cual es recuperado directamente de muestras ambientales. Este amplio campo también puede ser conocido como: genómica ambiental, ecogenómica o genómica de la comunidad.

## SECUENCIACIÓN DE NUEVA GENERACIÓN (NGS)

La secuenciación de nueva generación (NGS, por sus siglas en inglés), también conocida como secuenciación de alto rendimiento, es el término general utilizado para describir una serie de diferentes tecnologías de secuenciación modernas, incluyendo: Secuenciación de Illumina (Solexa), Secuenciación Roche 454, Ion torrent: secuenciación Proton/PGM y secuenciación SOLiD.

## SECUENCIACION PAIRED-END

La secuenciación de ‘extremo pareado’ permite a los usuarios secuenciar ambos extremos de un fragmento y generar datos de secuencia alineables de alta calidad. La secuenciación de ‘extremo pareado’ facilita la detección de reordenamientos genómicos y elementos de secuencia repetitiva, así como fusiones genéticas y transcripciones novedosas.

## SECUENCIACION POR SINTESIS

La secuenciación por síntesis SBS, es una más de las tecnologías de secuenciación masiva de nueva generación que utilizan la amplificación clonal in vitro. Los fragmentos a secuenciar son preparados con adaptadores en los extremos, para permitir que sus cadenas desnaturalizadas puedan unirse a una superficie sólida y así poder ser enriquecidas por medio de una amplificación tipo puente, generando clusters de cadenas del mismo fragmento, con el fin de amplificar la señal generada por los nucleótidos fluorescentes incorporados a la cadena en crecimiento y así puedan ser detectados por una cámara, esto durante un proceso cíclico que permite alargar las lecturas dependiendo de la plataforma seleccionada.

### III. Formulación de las Hipótesis de Investigación.

Se ha encontrado una diferencia significativa en la diversidad microbiana entre las diferentes regiones del altiplano debido a sus características físicoquímicas, lo que incide en las propiedades de la quinua.

Las variaciones de las comunidades y estructura microbiana son influenciadas tanto por el cambio climático como por las características del cultivo.

### IV. Operacionalización de las Variables en Estudio.

**Tabla 3.** Operacionalización de variables de estudio.

Variable	Indicador	Unidad de medida
Índice o tasa de biodiversidad	Grupos identificados Taxa • Género • Familia • Orden • Especie	Porcentaje
	Dominancia OTUs	Índice de dominancia Número
Propiedades físicoquímicas	pH	Unidad
	Conductividad eléctrica	uS
	Potasio	kg/ha
	Fosforo	ppm
	Aluminio	ppm
	Humus	rango
	Hierro férrico	kg/ha
Nitrato nitrogenado	kg/ha	
Manejo de suelos	Region	Altiplano Norte Altiplano Centro Altiplano Sur
	Tipo de suelo	Puruma Pre-siembra Fase lechosa Pre-cosecha

### V. Diseño Metodológico:

#### A. Población de estudio, Ámbito y Periodo de Investigación.

La población de estudio implicó 30 muestras provenientes de suelos de cultivo de quinua y purumas, de las 3 zonas agroproductivas del altiplano boliviano en tres fases durante un año agrario. Estas se obtuvieron de seis localidades, en el norte: Sullcataca baja (Municipio de Laja) y Estación Experimental de Choquenaira (UMSA); en la zona central: Cañaviri (Municipio de Umala) y Estación Experimental de Patacamaya (UMSA); en el Sur: Irpani y Saitoco (Municipio Salinas de Garcí Mendoza). El periodo de la investigación se da en los meses de Marzo-Abril (fase lechosa), Junio (pre cosecha) y Septiembre y Octubre (pre siembra) de 2016, cubriendo una gestión agraria.

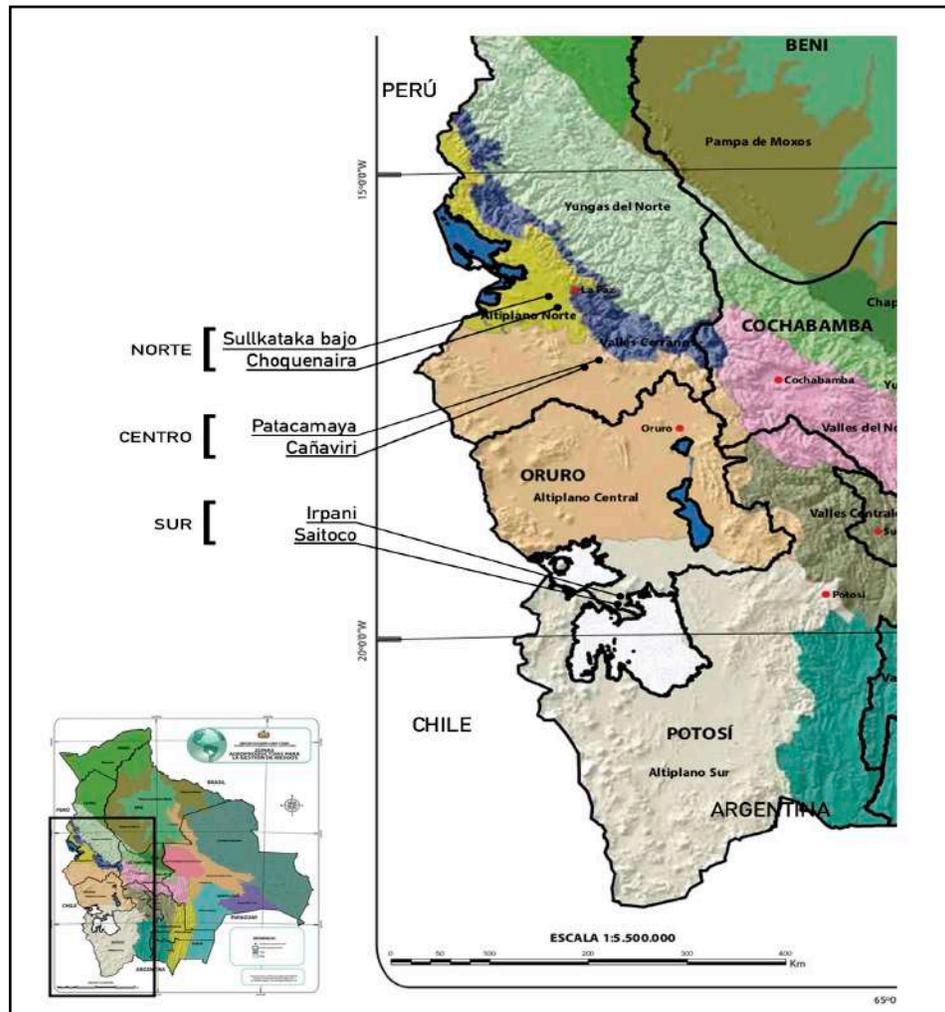
## B. Tipo de Investigación.

Esta investigación es de tipo descriptiva, cuali-cuantitativa, experimental *in silico* e hipotética-deductiva longitudinal.

## C. Materiales, Métodos, Técnicas y Procedimientos a ser Empleados.

### 1. Descripción de los sitios de estudio.

El Altiplano boliviano es una planicie que se halla entre las cordilleras occidental y oriental. Se encuentra por encima de los 3000 msnm, con temperaturas promedio entre -6 a 16 °C y precipitaciones que fluctúan entre los 200 a 900 mm/año. Esta región se sub divide en tres Zonas Agroproductivas: Altiplano Norte, Altiplano Central y Altiplano Sur (Figura 2). Estas zonas tienen variación según los diferentes pisos ecológicos, altitud y fisiografía que determinan las características en la producción agrícola y pecuaria (Aragón Oraquine, Arias Irusta, Chacolla Arias, & Tito Villca, 2014).

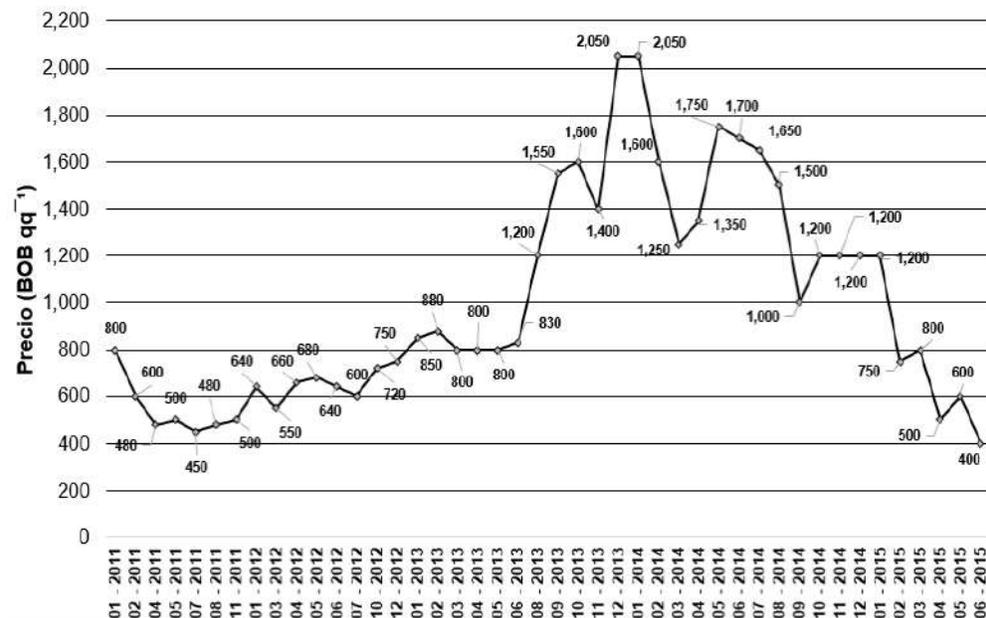


**Figura 2.** Mapa de división de las zonas agroproductivas en el altiplano boliviano.

En cuanto a la producción de quinua la superficie cultivada en Bolivia representa solo el 2% del total y aproximadamente el 80% de los productores de quinua se sitúan en el altiplano boliviano, la mayoría de los productores de Potosí y Oruro comercializan su producto a través de la Asociación Nacional de Productores de Quinua (ANAPQUI, 2017). La feria de Challapata se constituye en una de las ferias más importantes para los productores del altiplano sur, a esta arriban generalmente productores de las comunidades de Huari, Salinas de Garci Mendoza, Uyuni, Quillaca, Coipasa y Pampa Aullagas. En el altiplano central, la feria de Patacamaya donde llegan productores de las comunidades de Umala, Llanga, San Martín, Kiluhiri, Pomposilo y Huaylanicon. La feria del Desaguadero se caracteriza por su actividad de intercambio bilateral en la que participan los países de Bolivia y Perú, la quinua recolectada proviene básicamente de la feria de Challapata (Sarco S., 2011).

Los poblados que son parte del Municipio de Salinas de Garci Mendoza del altiplano sur son las que tienen una mayor producción, en la gestión 2009 se tuvo un rendimiento de 17.565,72 Kg/ha, en contraste a otras regiones del altiplano, en el 2009 entre Municipio de Umala y Patacamaya el rendimiento fue 3.270 Kg/ha. Las variedades más producidas en estas regiones son la quinua Real blanca y Toledo.

En los últimos años el precio de la quinua tuvo variaciones muy perceptibles, llegando a costar desde Bs. 400 hasta Bs. 2050 por quintal (Figura 3), este último se debe a la declaración en 2013 del año internacional de la quinua.



**Figura 3.** Precios alcanzados en bolivianos por quintal de quinua entre 2011 y 2015. Fuente: (Miranda, 2015)

#### **a) Descripción del Altiplano Norte.**

El Altiplano norte tiene una extensión de 13600 km<sup>2</sup>, se encuentra circunscrito en los márgenes del lago Titicaca y comprende las provincias Manco Kapac, Omasuyos, Ingavi, Los Andes y Pacajes. Esta zona se caracteriza por zonas fisiográficas de la cordillera Occidental, Altiplano y Cordillera Oriental. Posee un clima templado frío (Koppen, 1936). Según la clasificación de la UNESCO, se presentan dos tipos de climas: clima semiárido de verano templado e invierno fresco, clima subhúmedo seco de verano cálido e invierno cálido. Por sus factores climáticos y por su altitud, recibe una mayor cantidad de energía solar que una superficie similar ubicada a nivel del mar.

Gran parte de la cuenca del Altiplano de La Paz sufre de una gran deficiencia de agua, situación que afecta gravemente a la atención de las diferentes demandas existentes. Ocasionalmente se presentan sequías fuertes, cuya ocurrencia varía de 1 cada 2 años (Aragón Oraquine et al., 2014).

De forma genérica, en esta región los suelos son superficiales a moderadamente profundos con texturas franco-arcillo-arenosos y abundantes fragmentos de piedras, expuestos a un proceso erosivo de tipo laminar y en cárcavas, con pendientes moderadamente escarpadas (15-30%). Los suelos que pertenecen a la cadena montañosa de la Cordillera Oriental son superficiales, bien drenados, con mucha pedregosidad superficial y afloramientos rocosos, además con erosión severa.

El manejo del suelo es principalmente para fines agropecuarios, es decir se emplean los suelos para el pastoreo de ganado y cultivos agrícolas (Laja, 2005). En esta zona se siembra usualmente leguminosas, tubérculos y quinua en una extensión menor a una hectárea, por lo que es una costumbre efectuar rotación de los suelos, generalmente entre quinua y papa. Los suelos están en descanso por el tiempo de dos años, y los empleados se encuentran en las planicies y en menor frecuencia las serranías y laderas.

Las variedades de quinua que se siembran para su comercialización son con mayor frecuencia la real blanca y criollo, este grano es tanto para autoconsumo o vendido a mercados nacionales y locales. El rendimiento obtenido antes (20 años atrás) en promedio era de 7 quintales por hectárea, hoy en día se obtiene en promedio 5 quintales por hectárea.

La mayoría de los agricultores destinan su producción a la comercialización sólo cuando la producción agrícola es buena. Los mayores ingresos lo obtienen de la venta del ganado, también existe en

pequeña escala el intercambio o trueque entre diferentes productos presentes en las ferias locales.

#### **b) Descripción del Altiplano Central.**

El Altiplano central ocupa 91079 km<sup>2</sup>, está ubicado en la parte sur del departamento de La Paz. Comprende las provincias Pacajes, Aroma, Gualberto Villarroel e Inquisivi, el departamento de Oruro, norte de Potosí y la provincia Bolívar del departamento de Cochabamba. Esta zona también es conocida como puna seca.

En general la, fisiografía es plana a levemente ondulada, siendo el río Desaguadero el principal afluente de agua. La producción agropecuaria es limitada debido a los factores climáticos y edafológicos.

La sequía es uno de los riesgos de mayor afectación en la zona del Altiplano central, debido a que su ocurrencia es de 1 cada 2 años de sequía en la parte del Desaguadero y 4 cada 5 años en la parte sur en las que se presentaron sequías meteorológicas e hidrológicas (Aragón Oraquine et al., 2014). La topografía es ondulada, con una precipitación anual de 450 mm, clima frío, con heladas en los meses de invierno y fuertes granizadas.

Los suelos varían desde arcillas fuertemente salinas, arcillas no salinas y arenas, hasta bastante pedregosas. Los recursos hídricos provienen de los ríos Desaguadero, Umala, Calteu y Kheto. La vegetación es de pastura seca con arbustos.

La principal actividad es la pecuaria, con la cría de ganado vacuno, ovino y camélido, práctica que genera ingresos económicos significativos para la población. La agricultura varía con cultivos de papa, oca, haba, quinua, alfalfa y cebada en grano, esta última destinada a la alimentación del ganado (Patacamaya, 2010; Sica\_Sica, 2010).

En esta zona, por lo general se siembra quinua en una extensión mayor a una hectárea. Es una costumbre efectuar rotación de los suelos, un año es cultivado en planicie y al siguiente en serranía o ladera.

Las variedades de quinua más sembradas actualmente son: Real Blanca, seguido de jacha grano, La mayoría de los productores llevan su producción para comercializar cuando la producción agrícola es buena. Los mayores ingresos obtienen de la venta de los ganados, también existe en pequeña escala el intercambio denominado trueque entre diferentes productos presentes en las ferias.

### **c) Descripción del Altiplano Sur.**

El Altiplano sur con 73983 km<sup>2</sup>, tiene características de puna con escasa cobertura vegetal, bajas temperaturas y escasa precipitación (Monasterio, Andressen, & Terceros, 2007).

Los paisajes de esta zona se caracterizan por la presencia de conos volcánicos, mesetas de lava, acumulaciones de ceniza e ignimbritas. Esta región encierra en su interior a varios salares, como el de Uyuni, Chiguana, Empexa, con una suave inclinación desde el suroeste hacia el Salar de Uyuni. En la parte plana, el Altiplano potosino tiene una altura de 3.650 a 3.700 msnm, y en este se encuentran llanuras de arenas, arcillas, limos y gravas modeladas por el escurrimiento hídrico superficial y la acción de fuertes vientos.

Debido a las condiciones ecológicas limitadas, la producción agropecuaria es de menor proporción con relación a las zonas norte y central del altiplano; sin embargo, se destaca la producción de quinua y haba como los productos principales, siendo la actividad ganadera de menor proporción con relación a camélidos y ovinos.

La principal amenaza para la zona son las heladas y sequías, cuya frecuencia es de 1 cada 2 años y de 4 a 5 años respectivamente (Aragón Oraquine et al., 2014).

La principal actividad de los pobladores es la crianza de llamas, las condiciones de los suelos con presencia de una variedad de pastos nativos favorecen esta actividad. La actividad agrícola es enfocada principalmente al cultivo de la quinua, principal fuente de producción económica. Las Comunidades que se sitúan son colindantes a los salares, se dedican a la prestación de servicios manuales en el acopio y transporte de sal (Salinas\_Garci\_Mendoza, 2002). Los productores hacen uso de maquinaria en comparación a las zonas norte y centro, donde realizan manualmente o con azadón y yuntas.

Las variedades más sembradas de quinua son la variedad Real Blanca, Pandela, Toledo, Pisankalla, Negra un 23%, jacha grano. Se hace uso de fertilizantes y en menor proporción biofertilizantes y superfosfato. El abono más utilizado por los productores y disponible en el altiplano es el estiércol de camélido y muy poco se utiliza el estiércol de ovino.

La comercialización de la Quinua en la situación actual se realiza principalmente por venta directa en ferias o a minoristas (Feria de Challapata), 42% los productores entregan su producción a la institución COPROQUIRC que lo comercializan internacionalmente, y un 4% es para el consumo del productor. El rendimiento promedio actualmente es

de 12 quintales por hectárea, antes se podía obtener un promedio de 17 quintales por hectárea.

Las variedades de quinua sembradas en los sitios de estudio fueron Jach'a grano y Real blanca. En la Tabla 4 se muestran las características de las plantas cosechadas en cada sitio. De estas resaltan las de Cañaviri (altiplano centro), Saitoco e Irpani (altiplano sur) por tener dimensiones (longitud de la planta, diámetro de la panoja, diámetro de grano, etc.) superiores a las de las comunidades Sullcataca Baja y Patacamaya. El mejor rendimiento obtenido de grano fue de la comunidad de Saitoco, seguido de Cañaviri.

**Tabla 4.** Características de las plantas de quinua obtenidas de los sitios de estudio (Datos no obtenidos para Choquenaira).

Descripción De La Planta	ALTIPLANO				
	NORTE	CENTRO		SUR	
	Sullcataca Baja	Cañaviri	Patacamaya	Saitoco	Irpani
<b>Altura (m)</b>	1,06	1,205	1	1,44	1,325
<b>Diámetro de panoja (cm)</b>	4,29	6,055	0,32	6,16	9,935
<b>Longitud de panoja (cm)</b>	26,3	41,4	2,9	52,65	67
<b>Diámetro de tallo (cm)</b>	1,44	2,005	1,05	1,74	1,375
<b>Longitud de raíz (cm)</b>	21,5	38,25	0,21	43,5	47,75
<b>Diámetro de grano (mm)</b>	1,9	1,75	0,92	4,225	2,5
<b>Peso total de grano (g)</b>	6	125,185	25,92	154,545	113

## 2. Obtención de las muestras.

Para el presente trabajo las muestras de suelo fueron obtenidas a partir de muestras madre recolectadas dentro el proyecto “Estudio del efecto del cambio climático sobre la biodiversidad microbiana de suelos donde se realizan cultivos de quinua (*Chenopodium quinoa* Willd.) procedentes de distintos ecosistemas del altiplano” (COSUDE 2016), almacenadas en el Laboratorio de la Facultad de Agronomía (UMSA) en la ciudad de Viacha.

La obtención de las muestras madre se efectuó cubriendo una gestión agraria de quinua en las tres regiones de estudio, solicitando los permisos correspondientes de las autoridades de las diferentes comunidades. Se emplearon dos técnicas de muestreo, la primera se trata de las calicatas

realizadas en los suelos con cultivo y en los suelos puruma, que consiste en efectuar excavaciones de un metro cuadrado de superficie por 1.5 m de profundidad, se identifican los estratos u horizontes del suelo comenzando desde la superficie con la letra A en adelante, de estos estratos se tomó el primer horizonte para el presente estudio, llamado también la fase orgánica. La segunda técnica de zigzag fue empleada en los suelos con cultivo que consistía en coleccionar 30 sub-muestras que fueron tomadas en forma de zigzag a lo largo del campo a una profundidad de 20 cm. Este volumen de sub-muestras se homogeneizó y se procedió al cuarteo de la muestra hasta obtener 2 kilogramos de muestra de suelo, esta última fue trasladada al laboratorio.

Estas muestras madre fueron tomadas en 6 localidades (2 localidades por cada zona agroproductiva) en tres tiempos: Fase lechosa (marzo-abril), Pre-cosecha (junio) y Pre-siembra (septiembre-octubre). Adicionalmente se tomaron muestras control colectadas de suelos “Puruma” o suelos en descanso por más de 12 años, en la Tabla 5 se muestran las características de cada punto de muestreo.

**Tabla 5.** Características principales de los puntos de muestreo.

Localidad	Muestra	Localización (UTM)	Altura (msnm)	Textura del suelo	Capacidad de campo
<i>Altiplano norte</i>					
Sullcataca baja	Puruma	19K, 563450X,8171387Y	3751	Franco	2,08 %
	Cultivo	19K, 563446X,8171349Y	3736	Franco-arcilloso	41,6 %
Choquenaira	Puruma	19K, 567854X,8146810Y	3851	Franco-arcilloso	31,98 %
	Cultivo	19K, 576125X,8154219Y	3871	Franco-arcilloso	26,3 %
<i>Altiplano central</i>					
Cañaviri	Puruma	19K,603945X,8086311Y	3842	Franco-arenoso	10,52 %
	Cultivo	19K, 639461X,8086296Y	3833	Franco-arenoso	9,85 %
Patacamaya	Puruma	19K, 628492X,8091175Y	3611	Franco-arenoso	10,24 %
	Cultivo	19K, 612034X,8090975Y	3588	Franco-arcillo-arenoso	30,9 %
<i>Altiplano sur</i>					
Irpani	Puruma	19K, 638624X,7815836Y	3676	Arenoso-franco	22,46 %
	Cultivo	19K, 638498X,7815767Y	3671	Franco-arenoso	24,06 %
Saitoco	Puruma	19K, 636802X,7812420Y	3728	Franco-arenoso	46,31 %
	Cultivo	19K, 636174X,7812964Y	3646	Arenoso-franco	11,17 %

De las muestras madre se tomaron 0,25g en tubos Eppendorf a 3 repeticiones y fueron conservadas en frío (4 °C) durante su transporte al laboratorio donde fueron almacenadas a -20 °C hasta la extracción de ADN. En la Tabla 6 se muestran los Códigos asignados a las muestras indicando las fases en que las muestras madre fueron tomadas.

**Tabla 6.** Codificación de las muestras obtenidas (P: puruma, S: pre-siembra, L: fase lechosa, C: pre-cosecha).

Localidad	Tiempo			
	Puruma	Pre-siembra	Fase lechosa	Pre-cosecha
<i>Altiplano norte</i>				
Sullcataca baja	Su.P	Su.S	Su.L	Su.C
Choquenaira	Ch.P	Ch.S	Ch.L	Ch.C
<i>Altiplano central</i>				
Cañaviri serrania	Ca.P	CaS.S	CaS.L	CaS.C
Cañaviri planicie		CaP.S	CaP.L	CaP.C
Patacamaya	Pa.P	Pa.S		Pa.C
<i>Altiplano sur</i>				
Irpani serrania	Ir.P	IrS.S	IrS.L	IrS.C
Irpani planicie		IrP.S		IrP.C
Saitoco serrania	Sa.P	SaS.S		SaS.C
Saitoco planicie		SaP.S	SaP.L	SaP.C

En las localidades de Cañaviri, Irpani y Saitoco se obtuvieron además muestras de los cultivos localizados en las serranías o laderas. Por problemas técnicos no se pudo obtener las muestras en fase lechosa de las localidades de Patacamaya, planicie de Irpani y serranía de Saitoco.

Adicionalmente, se solicitó los datos fisicoquímicos y agronómicos del análisis de las muestras madre en los sitios de estudio (Tabla 7).

**Tabla 7.** Datos agronómicos de los suelos según la zona agroproductiva por tipo de suelo y etapa de desarrollo.

Zona	Suelo	Humedad gravimétrica (%)	Capacidad de campo (%)	Etapa	Humedad gravimétrica (%)	Capacidad de campo (%)
Norte	Puruma	38,04 ±15,1	37,5±11,56	Pre-Siembra	10,7 ±1,72	9,6 ±4,07
	Cultivo	10,1 ±1,1	25,03 ±3,1	Fase Lechosa	9,3 ±3,67	25,51 ±9,32
				Pre-Cosecha	10,3 ±3,5	39,9 ±3,9

Centro	Puruma	3,7 ±2,14	16,24 ± 14,9	Pre-Siembra	5,2 ±0,18	4,4 ±0,19
	Cultivo	4 ± 0,45	17,3 ±5,9	Fase Lechosa	4,77 ±1,03	19,78 ±11,4
				Pre-Cosecha	1,9 ±0,83	27,7 ±2,6
Sur	Puruma	7,8 ±6,03	30,9 ±4,1	Pre-Siembra	0,35 ±0,36	5,4 ±6,64
	Cultivo	1,9 ±0,65	18,03 ±2,3	Fase Lechosa	3,62 ±1,66	21,37 ±8,7
				Pre-Cosecha	1,82 ±1,05	27,31 ±4,86

### 3. Extracción de ADN de muestras de suelos.

Para la extracción del material genético se empleo el Kit de extracción de ADN de suelos Powersoil® (Mobio, Carlsbad, CA, EE. UU) según las indicaciones del fabricante (Anexo 1).

#### a) Electroforesis horizontal.

Se verificó la calidad de la extracción de ADN mediante electroforesis en gel de agarosa al 1 % con tampón amortiguador TAE 1X (Tris-Acetato-EDTA), posteriormente se visualizó el ADN extraído con bromuro de etidio (EtBr) 0,5 ug/mL utilizando un transiluminador de luz UV (Sharp et al. 1973).

#### b) Cuantificación y estimación de la concentración de ADN.

El ADN fue cuantificado mediante espectrofotometría, se prepararon diluciones 1/100 de cada muestra y se empleo como blanco agua bidestilada, se anotaron las lecturas a 230, 260 y 280 nm (Nair, Vincent, & Bhat, 2014). Las absorbancias fueron determinadas mediante el uso de espectrofotómetro UV/vis Cary 50 (Varian, Inc. 190-1100 nm).

La concentración total se obtuvo mediante la siguiente fórmula:

$$[C]_{ADN} = DO * 50_{\mu g/mL} * FD$$

Donde, DO es la densidad óptica a 260nm y FD es el factor de dilución. Para determinar la pureza del ADN extraído se estableció mediante la relación de absorbancias A260/280 (Huberman, 1995), y en conjunto con la relación A260/230 se determinó si existe alguna contaminación proteica, compuestos aromáticos, carbohidratos, fenólica o de ARN (J. F & Russell, 2001; PROGRAMA CONTROL DE CALIDAD DE MUESTRAS, 2010).

### 4. Amplificación y Secuenciación de las muestras de ADN.

La amplificación del gen ARN ribosomal 16S de las muestras de ADN extraídas se efectuó empleando los primers o cebadores 341F (CCTACGGGAGGCAGCAG) y 805R (GACTACHVGGGTATCTAATCC), que cubren las regiones V3 y V4, de aproximadamente 460pb (Illumina; Klindworth et al., 2012).

La secuenciación se realizó en el sistema Illumina MiSeq (Omega Bioservices, Georgia USA) mediante la técnica Paired-End con lecturas de 2x300pb. El sistema MiSeq utiliza la secuenciación por síntesis (SBS), y las lecturas son almacenadas en archivos FASTQ comprimidos.

## 5. Pre-procesamiento bioinformático.

Para comprobar la calidad de las bases de las lecturas se empleó FastQC v0.11.5 (Andrews et al., 2012), que creó reportes de la calidad y características de las lecturas obtenidas en los archivos FASTQ.

Para el procesamiento y análisis de las lecturas se empleó la versión Unix de la suite Qiime (Quantitative Insights Into Microbial Ecology) MacQiime 1.9.1 (WernerLab), una instalación precompilada de Qiime con toda sus dependencias. También se empleó la versión nativa en el sistema Linux (J. Gregory Caporaso et al., 2010). Este software, de código abierto, posee pipelines y herramientas para el procesamiento y análisis de datos de secuenciación de alto rendimiento.

### a) Validación del archivo de asignación.

En primer lugar se construyó el archivo de asignación (`mapping_file`) en formato de texto en columnas tabuladas como se muestra en la Figura 4, que contiene información básica relevante a las muestras. Éste se validó mediante el comando `validate_mapping_file.py`, que crea un archivo HTML con un reporte detallado.

#SampleID	BarcodeSequence	LinkerPrimerSequence	ReversePrimer	pH	Fase	Region	Description
Ca.P	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.3	Puruma	Centro	franco_arenoso	
Ir.P	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	7.8	Puruma	Sur	arenoso_franco	
SaP.L	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.8	Fase_lechosa	Sur	arenoso_franco	
IrS.L	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	5.4	Fase_lechosa	Sur	franco_arenoso	
CaP.L	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.0	Fase_lechosa	Centro	franco_arenoso	
Sa.P	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	9.1	Puruma	Sur	franco_arenoso	
Su.P	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.9	Puruma	Norte	franco	
Su.L	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.8	Fase_lechosa	Norte	franco_arculloso	
Ch.L	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.2	Fase_lechosa	Norte	franco_arculloso	
Pa.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.2	Pre_cosecha	Centro	franco_arcullo_arenoso	
CaP.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.0	Pre_cosecha	Centro	franco_arenoso	
Sa.S.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	8.0	Pre_cosecha	Centro	franco_arenoso	
Su.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.0	Pre_cosecha	Norte	franco_arculloso	
Ch.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.2	Pre_cosecha	Norte	franco_arculloso	
SaP.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	8.0	Pre_cosecha	Sur	arenoso_franco	
IrS.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.2	Pre_cosecha	Sur	arenoso_franco	
SaS.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.2	Pre_cosecha	Sur	arenoso_franco	
IrP.C	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	8.0	Pre_cosecha	Sur	arenoso_franco	
SaS.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.0	Pre_siebra	Sur	arenoso_franco	
IrS.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.4	Pre_siebra	Sur	arenoso_franco	
SaP.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.6	Pre_siebra	Sur	arenoso_franco	
IrP.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	8.0	Pre_siebra	Sur	arenoso_franco	
CaS.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.8	Pre_siebra	Centro	franco_arenoso	
Su.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.2	Pre_siebra	Norte	franco_arculloso	
CaP.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	7.0	Pre_siebra	Centro	franco_arenoso	
Pa.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.8	Pre_siebra	Centro	franco_arcullo_arenoso	
Ch.S	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.8	Pre_siebra	Norte	franco_arculloso	
Pa.P	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.5	Puruma	Centro	franco_arenoso	
Ch.P	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	8.3	Puruma	Norte	franco_arculloso	
CaS.L	CCTACGGGNGGCWGCAG	GACTACHVGGGTATCTAATCC	6.0	Fase_lechosa	Centro	franco_arenoso	

Figura 4. Imagen del archivo de mapeo (mapping file).

### b) Ensamblado de las lecturas.

Las lecturas Paired-End, fueron ensambladas en contigs mediante el método `fastq-join` (Aronesty, 2013) con un porcentaje de similitud mayor o igual al 80%, este algoritmo es más eficiente y admite algunas comparaciones y

ajustes automáticos que otras herramientas, por ejemplo toma una secuencia A 5'-3' y alinea automáticamente la secuencia B 3'-5' sobre la secuencia A. Esta basado en un algoritmo que aplica el Alineamiento Anclado Preciso (ACANA: ACcurate ANchoring Alignment) (Huang, Umbach, & Li, 2005).

Básicamente lo que realiza es superponer las bases de dos secuencias y si éstas coinciden se toma la base con mayor puntaje de calidad, incrementando hasta en 3 puntos; y si las bases no coinciden se comprueba si un puntaje de calidad es mayor a 3, el puntaje resultante es la diferencia entre los dos puntajes (calidad reducida debido a la falta de coincidencia) o caso contrario el puntaje será 3.

### **c) Demultiplexado y filtrado de calidad.**

Las lecturas ensambladas fueron demultiplexadas y filtradas empleando el script *split\_libraries\_fastq.py* que genera un solo archivo con las secuencias de todas las muestras en formato FASTA, necesario para posteriores análisis (Bokulich et al., 2012). Los siguientes parámetros fueron aplicados: el puntaje mínimo de calidad debe ser mayor o igual a 20, el tipo de código de barras es 'not-barcoded', el número máximo de bases consecutivas de baja calidad a 3.

Para descartar secuencias quiméricas que posiblemente se hayan generado durante la amplificación por PCR primero se identificaron mediante el comando *identify\_chimeric\_seqs.py* empleando USEARCH v6.1 (Edgar, 2010; Edgar, Haas, Clemente, Quince, & Knight, 2011), posteriormente las quimeras identificadas se utilizaron para filtrar las secuencias del archivo FASTA. Por último, para ver el resultado de la filtración se realizó un conteo de secuencias finales con el script *count\_seqs.py*.

## **6. Análisis bioinformático.**

### **a) Agrupamiento en Unidades Taxonómicas Operativas (OTUs).**

Las secuencias filtradas fueron agrupadas en OTUs con el protocolo Open-Reference (*pick\_open\_reference\_otus.py*), en este proceso, las secuencias son alineadas frente a una base de datos de referencia, GreenGenes v13.8 (DeSantis et al., 2006; McDonald et al., 2011), aquellas secuencias que no se encuentren en la base de datos son agrupadas mediante el protocolo 'De Novo' con el algoritmo UCLUST (Edgar, 2010), que agrupa tomando secuencias centroides o representantes de grupo con los que se va comparando al 97% de similitud entre las secuencias restantes, las secuencias que no coincidan van formando otros grupos con el cual se siguen comparando e ir construyendo los OTUs con todas las secuencias restantes. Aquellas secuencias, que no formaron parte de un OTU o un OTU que posea menos de 3 miembros fueron descartados (J. G. Caporaso et al., 2009). También se obtendrá el árbol filogenético en formato TREE (Price, Dehal, &

Arkin, 2010) aplicando el remuestreo estadístico (bootstrapping) de los datos dentro de FastTree2 como método de máxima verosimilitud. Las tablas OTU fueron construidas en formato BIOM (Biological Observation Matrix) y por ultimo las secuencias representativas de cada OTU junto con las nuevas secuencias de referencia en el archivo rep\_set.fna.

Un resumen estadístico de las tablas OTU fueron obtenidas con el comando *biom summarize-table*.

#### **b) Caracterización taxonómica.**

Las tablas OTUs fueron resumidas por niveles taxonómicos con los comandos *summarize\_taxa\_trough\_plots.py* y *summarize\_taxa.py*, ambos crearon tablas en archivos de texto con columnas tabuladas para cada nivel taxonómico: filo, clase, orden, familia, género y especie, estas tablas fueron importadas en Office Excel (Microsoft Inc.).

Después de realizar el compendio de las comunidades los resultados obtenidos (archivos TXT) fueron ordenados en Excel, de estos datos se calculó los porcentajes de abundancia a nivel de reino por localidad y tipo de muestra. También se generó una gráfica en barras de las proporciones a nivel de filo mayores o iguales al 0.4% en cada muestra. Además, se incluye una gráfica de las proporciones a nivel de género de cada muestra con abundancias mayores al 0.5%.

#### **c) Determinación de la diversidad Alfa y curvas de rarefacción.**

Se empleó el comando *alpha\_diversity.py* para obtener los índices de diversidad, incluyendo PD\_whole\_tree para la diversidad filogenética (Faith & Baker, 2007), la riqueza estimada y observada como el numero de observaciones (Observed\_Otus) y el estimador Chao1 (Gotelli & Colwell, 2011), además del índice de Shannon (Magurran, 1988) y el índice de heterogeneidad de Simpson (Peet, 1974; Simpson, 1949). Posteriormente se obtuvieron las respectivas curvas de rarefacción mediante el comando *alpha\_rarefaction.py*. Finalmente se efectuó el calculo de la significancia y la normalización de cada estimador de diversidad alfa mediante el comando *compare\_alpha\_diversity.py* y *add\_alpha\_to\_mapping\_file.py*.

#### **d) Determinación de la diversidad Beta con Unifrac.**

Se calculó la diversidad beta para identificar las diferencias entre las comunidades bacterianas de los puntos del altiplano, por zona agroproductiva, por etapas del desarrollo de la quinua y por tipo de suelo. Para esto se utilizó las métricas de UniFrac, que son medidas filogenéticas utilizadas ampliamente como ser: Weighted Unifrac (Cuantitativo), que permitió revelar las diferencias de la comunidad que se deben a cambios en la abundancia relativa de taxones, y Unweighted Unifrac (Cualitativo), que es más informativo sobre la presencia o ausencia de individuos en la muestra

(C. Lozupone & Knight, 2005). Mediante *beta\_diversity\_through\_plots.py* se construyó las matrices de distancias y se realizó el Análisis de Coordenadas Principales (PCoA).

## 7. Análisis de Correlación Canónica del altiplano sur.

Para este análisis se obtuvieron los datos de las mediciones fisicoquímicas de la Base de Datos del proyecto COSUDE (Tabla 8). El análisis se realizó con las muestras puruma y cultivo de las tres zonas agroproductivas.

**Tabla 8.** Características fisicoquímicas de los puntos de muestreo en suelos puruma del altiplano sur.

Analito	Altiplano Norte		Altiplano Centro		Altiplano Sur	
	Puruma	Cultivo	Puruma	Cultivo	Puruma	Cultivo
pH	7,60	6,31	6,40	6,68	8,45	6,79
Nitrógeno De Nitrato (Kg/ha)	25,00	25,42	17,50	30,16	10,00	28,75
Potasio (Kg/ha)	320,00	321,67	292,50	313,33	325,00	298,33
Fosforo (Kg/ha)	162,50	118,75	106,25	75,24	137,50	128,13
Humus (%)	3,25	2,08	1,50	0,75	2,50	1,21
Magnesio (ppm)	11,25	100,33	28,75	90,08	132,50	91,63
Calcio (ppm)	1925,00	1716,67	1200,00	1476,19	1475,00	1933,33
Sulfatos (ppm)	287,50	277,08	650,00	389,68	275,00	670,83
Aluminio (ppm)	7,50	11,88	5,00	11,35	7,50	11,25
Cloro (ppm)	275,00	229,17	175,00	154,76	175,00	255,00
Hierro Férrico (Kg/ha)	5,00	26,25	16,25	24,68	18,75	26,88
Nitrógeno De Nitrito (ppm)	5,50	6,50	1,00	2,03	1,00	2,42
Nitrógeno De Amonio (ppm)	6,25	10,00	13,75	6,03	5,00	15,00

Con la información obtenida se determinó la correlación existente entre las características y propiedades de los suelos con las comunidades microbianas identificadas. Se empleó CANOCO v.4.5 para realizar el Análisis de Correlación Canónica (CCA), en este caso observar las variaciones de las comunidades microbianas con las condiciones dadas de los suelos respecto a la zona de cultivo de quinua, y observar la respuesta de estas comunidades a los cambios del suelo debido al cultivo de quinua. Además, permitió encontrar cierta consistencia o previsibilidad de estas variaciones.

## **D. Procesamiento de los datos.**

### **1. Recolección.**

Los datos iniciales de la secuenciación fueron obtenidos en archivos de texto en formato FASTQ, la información y metadatos de cada muestra fueron almacenados en archivos de Texto en columnas tabuladas y la información detallada como los datos fisicoquímicos fueron entregados en hojas de calculo en Microsoft Excel (2016).

### **2. Elaboración.**

La elaboración de los datos se efectuó en Tablas en formato TXT tabulados, tablas BIOM, hojas de calculo en Numbers (5.0 2018) y para el análisis CCA se transformaron las tablas en archivos DTA utilizando CANOCO v.4.5. Las gráficas de abundancia se elaboraron en Numbers, las gráficas de calidad y curvas de rarefacción fueron creadas en archivos HTML.

### **3. Análisis.**

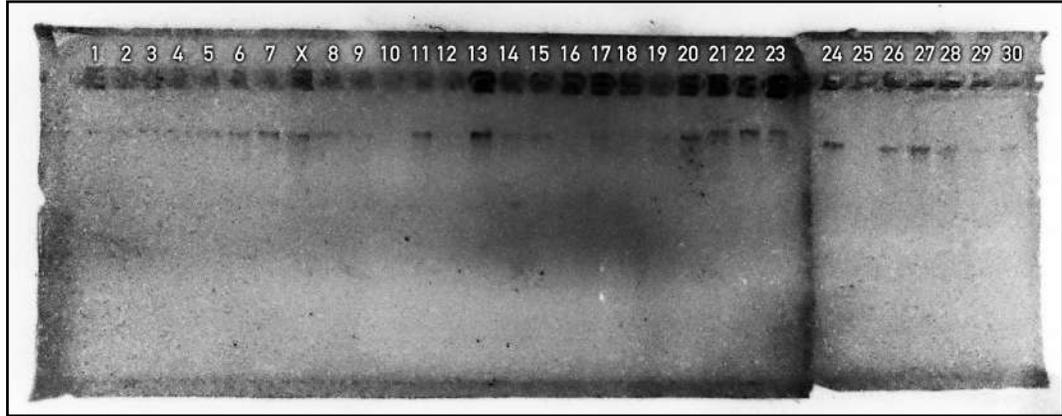
El análisis se realizará mediante la construcción de gráficas de proporciones al 100% de las abundancias de las comunidades microbianas por muestra, por fase de cultivo y por zona agroproductiva dentro de Numbers (5.0 2018) resaltando los grupos taxonómicos más representativos, también se obtuvieron los índices de diversidad alfa de cada muestra así como las gráficas de rarefacción en Qiime. Además se efectuó un análisis de la diversidad beta entre las muestras que representan las fases fonológicas del cultivo de quinua, representados con los gráficos de Análisis de Coordenadas Principales (PCoA) empleando Emperor (Vázquez-Baeza, Pirrung, Gonzalez, & Knight, 2013). Ambos análisis, diversidad alfa y beta, se efectuaron dentro de Qiime. Por último se realizó el análisis de correlación multivariable entre las propiedades fisicoquímicas de los suelos y las comunidades microbianas más representativas mediante gráficas de Correlación Canónica (CCA) que fueron elaboradas en CANODraw v4.14.

## **VI. Resultados y discusión.**

### **A. Extracción y cuantificación del ADN**

Después de la extracción de ADN se utilizaron 10  $\mu$ L para la electroforesis y 10  $\mu$ l para la cuantificación por espectrofotometría. La electroforesis mostró bandas en la mayoría de las muestras (Figura 5), de estos algunos pozos mostraron bandas tenues y otros no presentaron bandas. Varios factores están presentes en las causas de bandas erráticas, como el manejo de los reactivos, manejo de la muestra y el proceso de sembrado; ya que los resultados obtenidos en el calculo de la concentración de ADN por espectrofotometria indicaron la presencia ADN (Tabla 9) en todas las muestras. Además, se observó los rastros dispersos en las bandas que indica una contaminación leve, estos pueden ser por carbohidratos o

residuos fenólicos o algún residuos provenientes de las soluciones empleadas en el Kit de extracción.



**Figura 5.** Electroforesis de ADN total de las 30 muestras obtenidas. Se observa los pozos seguidas de las bandas de ADN durante la corrida. El carril X es una muestra repetida.

Se calculó el rendimiento y la pureza del ADN con las mediciones de la DO a 260nm y las relaciones de A260/A280 respectivamente (Glasel, 1995; Nair et al., 2014), que se observa en la Tabla 9, las concentraciones estuvieron entre los 2 y los 36.5  $\mu\text{g/mL}$ , la media fue de 14.02  $\mu\text{g/mL}$  ( $\pm 9.56$ ), existen variaciones con los resultados obtenidos en la electroforesis, estas variaciones se deben a dos motivos, el primero al error de muestreo de la alicuota o mala resuspensión del ADN, y el segundo a que la determinación de la concentración mediante espectrofotometría tiene sesgos, debido principalmente a que el equipo presentó una variabilidad en las mediciones por el recorrido del equipo.

**Tabla 9.** Concentraciones e índices de pureza de las muestras de ADN. Se indica la concentración: índice de pureza (IP), y la relación A260/A230 (indicador de contaminación fenólica y/o proteica).

Código	Concentración $\mu\text{g/mL}$	IP A260/280	A260/230
Ca.P	15,5	1.24	0.479
Ir.P	4,5	0.65	0.686
SaP.L	24,5	1.54	0.629
IrS.L	9	2.25	0.257
CaP.L	8,5	4.25	0.378
Sa.P	2,5	1.48	0.430
Su.P	13,5	1.68	0.508
Su.L	27,5	1.41	0.608
Ch.L	13	1.87	0.583
Pa.C	13,5	1.38	0.603
CaP.C	8,5	1.67	0.476

<b>Código</b>	<b>Concentración µg/mL</b>	<b>IP A260/280</b>	<b>A260/230</b>
CaS.C	4	4.00	0.131
Su.C	17,5	1.43	0.430
Ch.C	33,5	1.31	0.587
SaP.C	9	1.40	0.500
IrS.C	11,5	2.88	0.354
SaS.C	12	2.67	0.316
IrP.C	13,5	1.31	0.579
SaS.S	13	1.30	0.522
IrS.C	15,5	1.39	0.548
SaP.S	4	2.67	0.267
IrP.S	2	1.14	0.452
CaS.S	23,5	0.82	0.673
Su.S	15	1.80	0.714
CaP.S	36,5	1.94	3.667
Pa.S	8	16.00	0.340
Ch.S	11,5	1.85	0.500
Pa.P	3,5	1.16	0.649
Ch.P	10	1.32	0.621
CaS.L	36,5	1.30	0.775

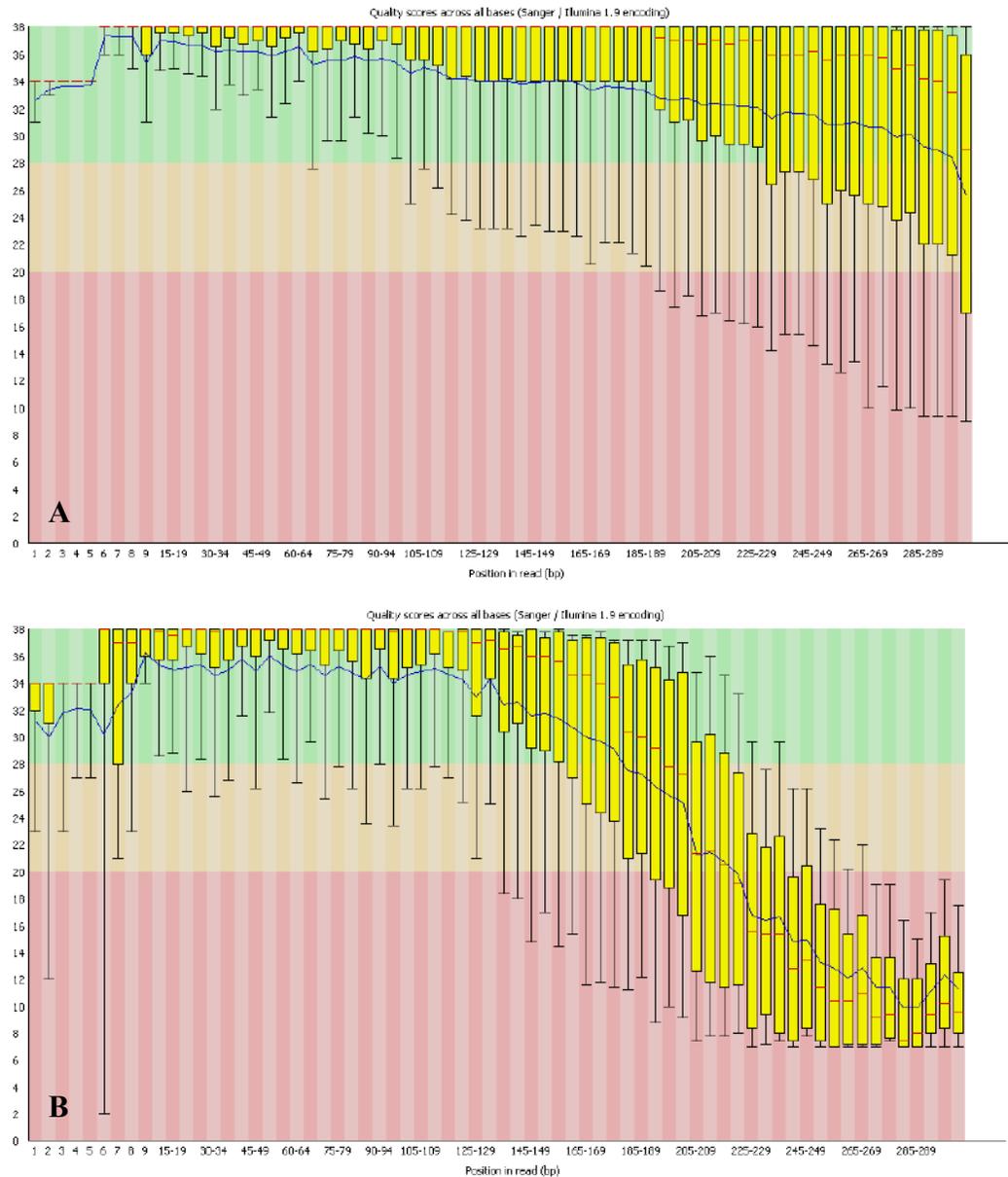
De las 30 muestras, seis presentaron concentraciones bajas entre 2 y 4.5 µg/mL siendo la concentración óptima requerida para el servicio de secuenciación de 5 µg/mL. Las muestras 6 y 22, las de menores concentraciones, muestran bandas notorias en el gel de la corrida electroforética, dada las variaciones del equipo de espectrofotometría, estos valores de rendimiento y pureza son tomados como una verificación opcional de la extracción de ADN, teniendo en cuenta que el uso del kit MoBio de extracción de ADN de suelos, con un amplio uso, proporciona buenos rendimientos en estos tipos de muestras.

Otras observaciones en este paso indican: un IP de 16 para la muestra 26, este valor esta sobreestimado, la absorbancia obtenida a 260nm para esta muestra fue de 0.0016 y a 280nm fue de 0.0001, rango en el que las proteínas absorben la luz, en particular los aminoácidos aromáticos (J. F & Russell, 2001), ya que las mediciones se hicieron por triplicado y se promediaron los resultados no significan errores sino los sesgos por el manejo de los materiales y muestra; en el caso de la muestra 25 se observa una relación A260/A230 de 3.67, esto indica que existe una mayor proporción de ácidos nucleicos que otros componentes que absorben a 230nm, como residuos de fenol o carbohidratos (Jenway), componentes presentes en las demás muestras observadas.

## **B. Características de la secuenciación Paired-End.**

Las lecturas RAW obtenidas de la secuenciación fueron guardadas en dos archivos comprimidos (GNU ZIP) por muestra, uno con las lecturas Forward con el sufijo R1\_001.fastq.gz y el otro con las lecturas Reverse con el sufijo R2\_001.fastq.gz en formato FASTQ, éstas se descargaron a travez del servidor BaseSpace (<https://basespace.illumina.com>) de Illumina Inc.

Los resultados de los reportes de calidad obtenidos con FastQC v0.11.5 muestran, en general, que las lecturas Forward tuvieron mejores puntajes de calidad durante la secuenciación en comparación con las lecturas Reverse (Figura 6). Al comienzo de la secuenciación las bases de cada secuencia tienen una buena calidad de lectura, y a medida que la secuenciación transcurre la calidad baja, mostrando una mayor dispersión respecto a la media representada con la línea azul, algunas bases tienen buena calidad y otras muy mala, por esta razón el cuartil es tan grande (Andrews et al., 2012). Este decaimiento de la calidad durante el proceso de secuenciación es común en este tipo de tecnología, y los errores son numerosos, en general los grupos de secuencias (clusters) no crecen al mismo tiempo y lentamente se va perdiendo la sincronía de las señales en la placa, también esta relacionado con la densidad de los grupos que interfieren la 'llamada' de base, además de que los elementos o reactivos van desnaturalizándose durante la corrida (Illumina; Inc.).

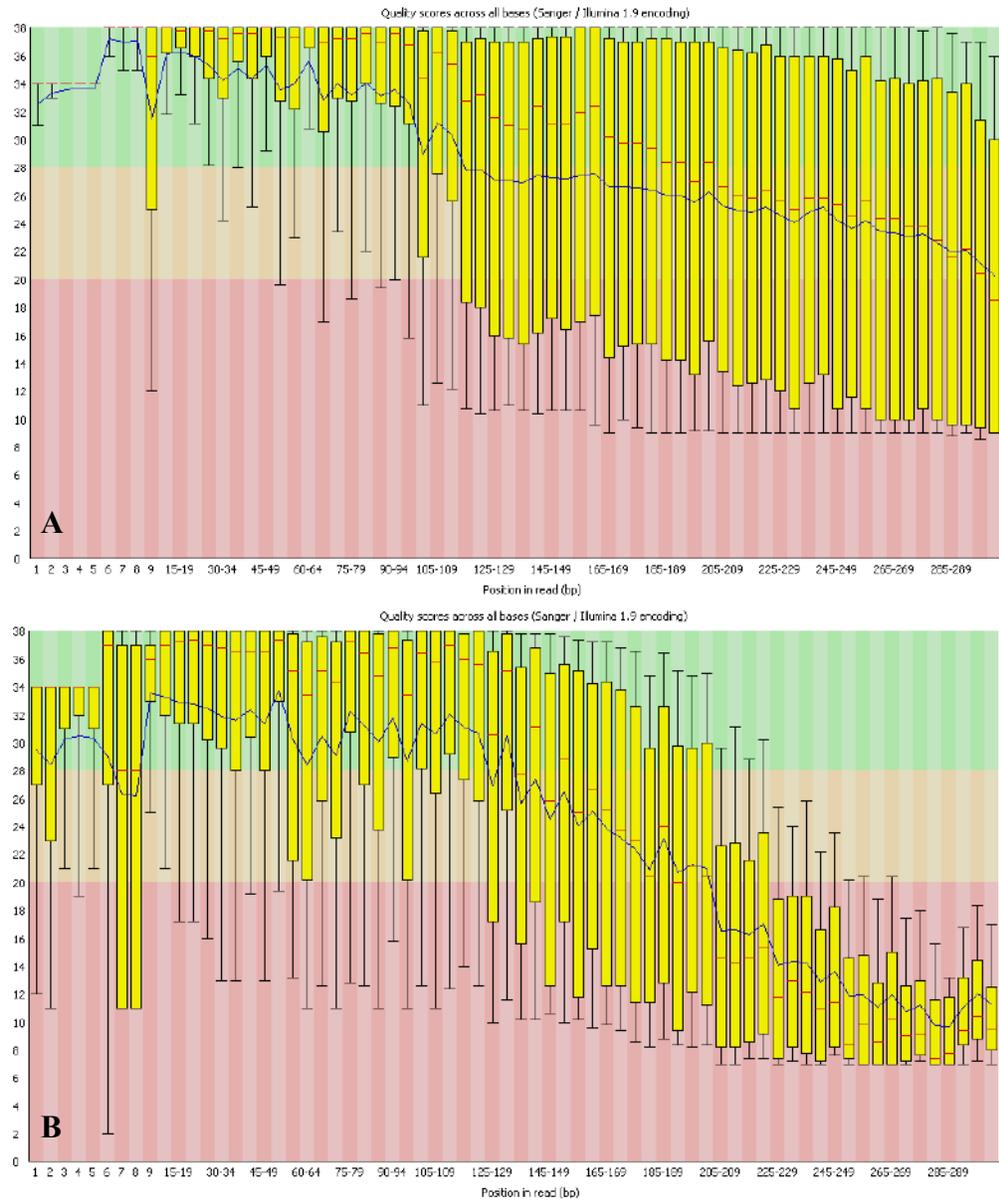


**Figura 6.** Distribución de la calidad de las lecturas de las bases en las secuencias para la muestra Ca.P. Se representa en amarillo los cuartiles, en azul la media, zona verde: buena calidad (28-34), zona naranja: calidad intermedia (20-28), zona roja: mala calidad (0-20). **(A)** Lecturas Forward, **(B)** Lecturas Reverse.

Luego de la secuenciación el sistema crea los archivos de las lecturas por muestra alcanzando 2.41 GB de tamaño. La técnica Paired-End es muy recomendada, porque proporcionan una alineación superior a través de las regiones del ADN que contienen secuencias repetitivas, y producen contigs más largos rellorando espacios en la secuencia consenso (Illumina). También detecta

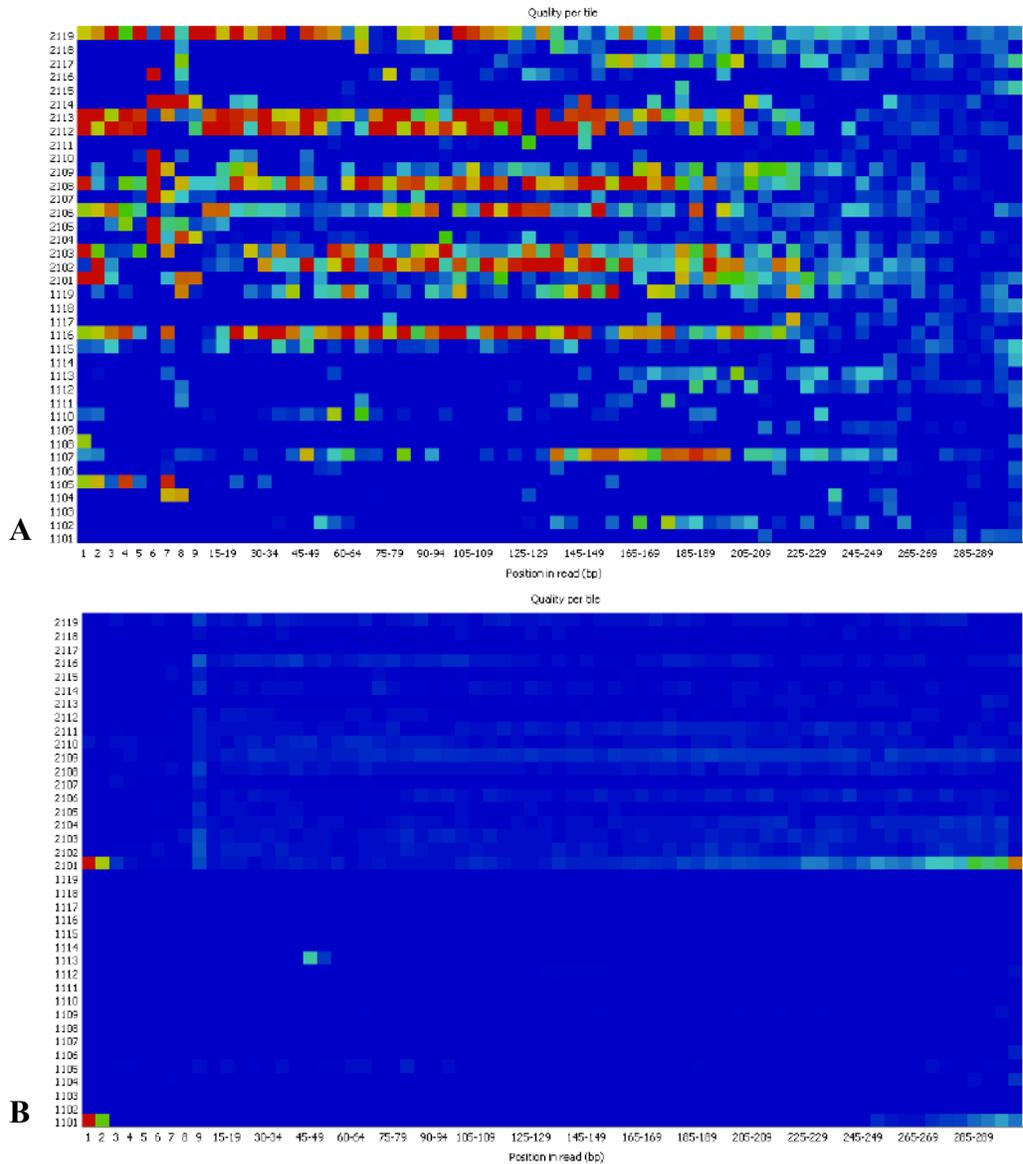
reordenamientos tales como inserciones, eliminaciones e inversiones entre otras aplicaciones dentro la genómica.

La distribución de la calidad tuvo el mismo comportamiento en todas las muestras (Anexo 2), con excepción de la muestra 3 (SaP.L: Saitoco Fase lechosa) que presenta una distribución muy errática (Figura 7), en esta muestra la secuenciación tuvo un mal desempeño, según los datos de rendimiento y pureza de ADN en la Tabla 8, posee valores aceptables que superan la cantidad optima requerida para llevar a cabo la secuenciación, descartando que el problema provenga de la muestra, sino más bien del proceso de secuenciación. Como se mencionó anteriormente, existen muchos factores; pero esto sucede cuando existe una sobresaturación de los grupos o clusters en las celdas de flujo, lo que produce que el dispositivo reconozca menos grupos (Illumina; Schirmer, D'Amore, Ijaz, Hall, & Quince, 2016).



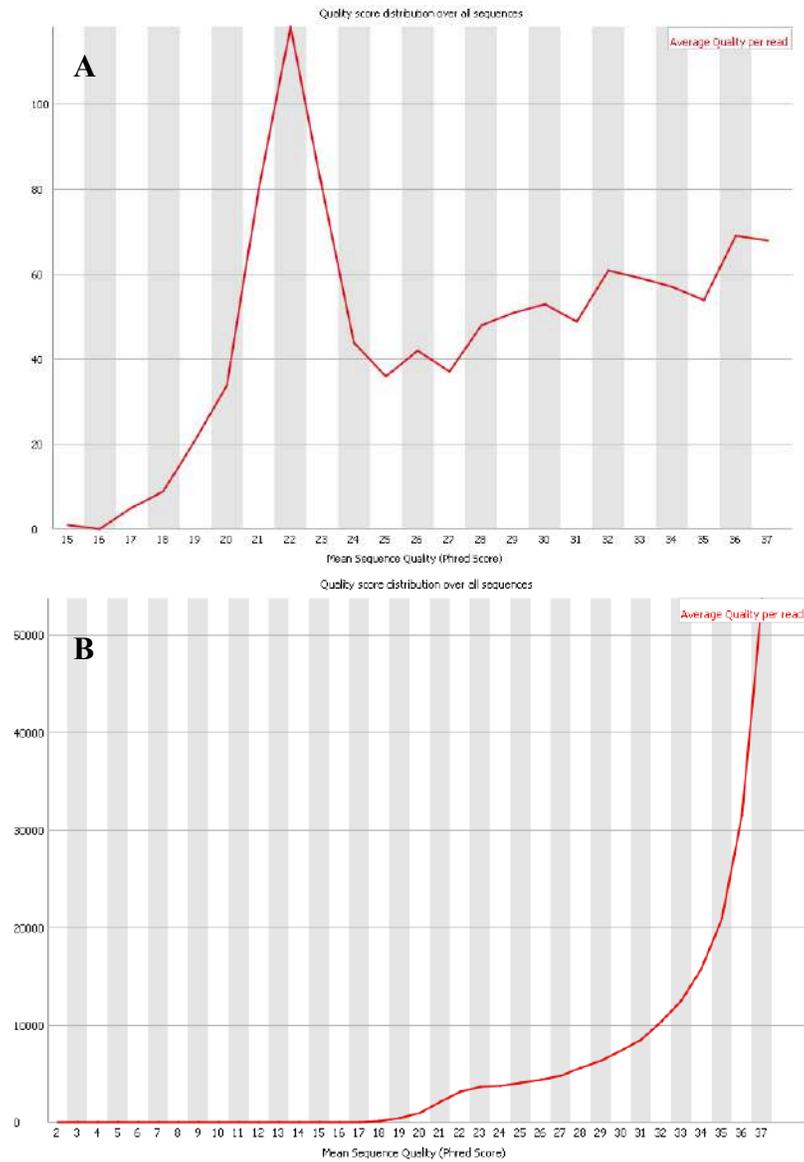
**Figura 7.** Se muestra el perfil de calidad de las bases para la muestra 3 (SaPL), la variación de las calidades se da desde el inicio de la corrida. **(A)** Lecturas Forward, **(B)** Lecturas Reverse.

Para comprender los factores o sesgos de calidad inherentes de la secuenciación, FastQC proporciona gráficas adicionales para entender las posibles causas. Uno de esos factores se muestra en la Figura 8A que representa si la celda de flujo fue filtrada de burbujas o interferentes, como los reactivos. Si se efectuó un buen filtrado será completamente azul, las marcas de color (diferente de azul) indican que se tiene mala calidad, ya que las burbujas quedan en la celda de flujo y produce interferencia. Como referencia se muestra también la gráfica para la muestra 1 (Ca.P) que presenta mejor calidad (Figura 8B).



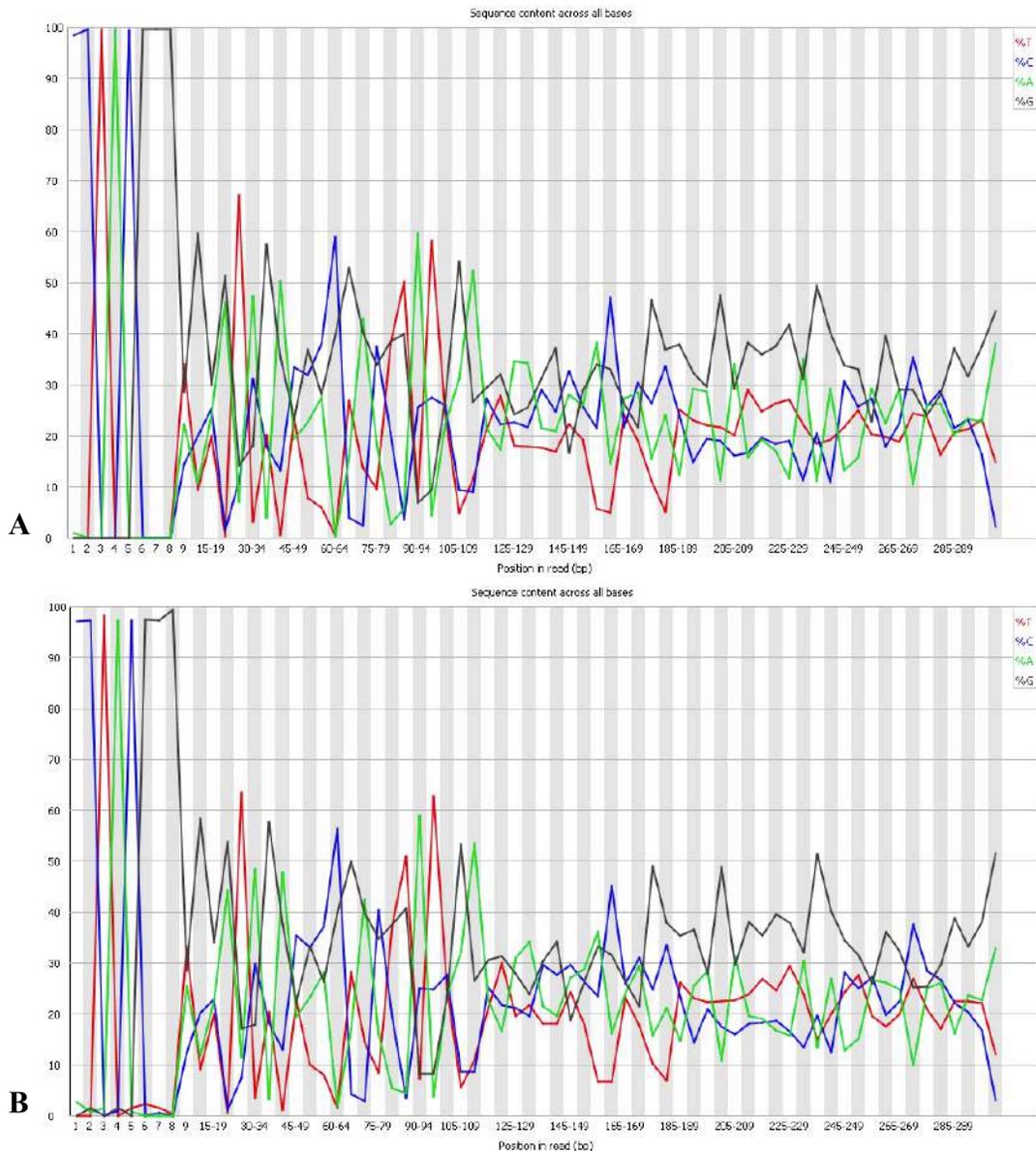
**Figura 8.** Calidad de la secuencia por "baldosa" (Per Tile Sequence Quality).  
**(A)** Representa la muestra 3:SaP.L, que tuvo la corrida más errática. **(B)**  
 Representa la muestra 1:Ca.P, que en conjunto con el resto de las muestras  
 tuvieron mejores puntajes de calidad.

La siguiente gráfica muestra los puntajes de calidad por secuencia, muestra el número de lecturas distribuido a través de la media de los puntajes de calidad de las secuencias. Con esta gráfica se consideró el número aproximado de lecturas que se tuvieron que quitar, en la Figura 9A se observa la distribución de la muestra 3:SaP.L la cual tuvo pocas lecturas y que aproximadamente 120 lecturas tienen un puntaje de calidad media de 22. En el resto de las muestras, las lecturas se agrupan hacia los mejores puntajes de calidad (Figura 9B),



**Figura 9.** Puntaje de calidad por secuencia. En el eje X se representa la media de la calidad de las secuencias. Mientras que en el eje Y, se representa el numero se secuencias o lecturas al que corresponde esa media. **(A)** Muestra 3:SaP.L, se observa que la mayor cantidad de secuencias tienen un puntaje de calidad de 22. **(B)** Muestra 1:Ca.P, más de 50000 lecturas tienen una media de 37 en puntaje de calidad.

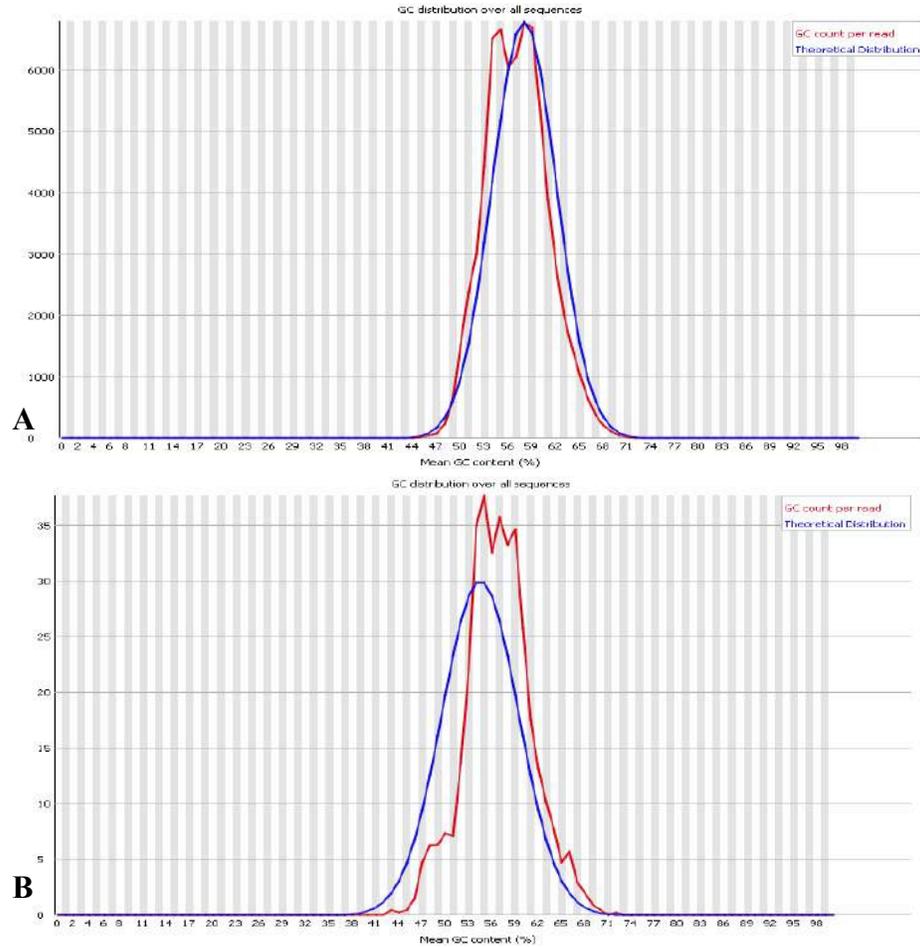
El siguiente apartado nos muestra el contenido de secuencias a través de cada base o la proporción de cada una de las bases que hay en las secuencias, en esta gráfica no debería existir variación de las líneas entre A (Adenina) vs T (Tiamina) y G (Guanina) vs C (Citosina) (Figura 10), en este caso, para la muestra 1 y 3, se observa diferencias entre unas bases y otras cuando la cantidad de A debería ser igual a la de T y lo mismo para G con C.



**Figura 10.** Contenido de la secuencia por base en sus posiciones, la diferencia entre los proporciones de las bases deberían tener un comportamiento paralelo. La gráfica superior representa la muestra 1:Ca.P (A) y la inferior a la muestra 3:SaP.L (B).

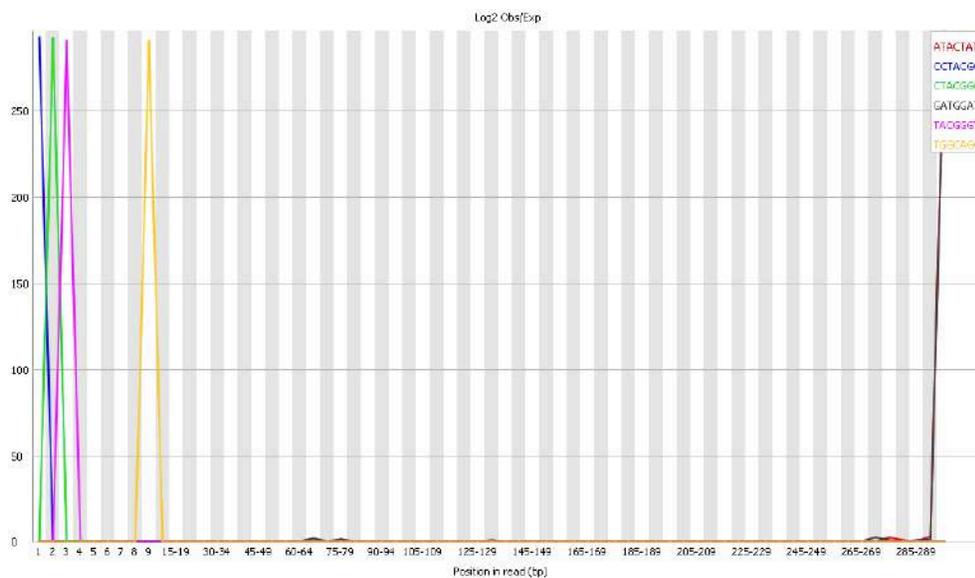
Luego tenemos el contenido de G y C en las secuencias, en el que se mide la proporción GC en toda la secuencia y se compara con una distribución normal teórica en función a la media y desviación estándar de las lecturas, en la Figura 11A, perteneciente a la muestra 1:Ca.P, se observa que la distribución de los datos (línea roja) tiende a solapar con la distribución teórica (línea azul), pero muestra dos picos, el primero que sobresale indica que hay secuencias con menor contenido GC. Si la variación de estas dos líneas es peor, como sucede con la muestra 3:SaP.L (Figura 11B), nos indica que existió contaminación,

principalmente de los adaptadores, también indican contaminación con otro ADN o la incorrecta llamada de base durante la secuenciación, es decir se coloca una G o C donde no debería ir. Este apartado nos confirma que la secuenciación, para esta última muestra, no se llevó correctamente.



**Figura 11.** Contenido en guanina y citosina (GC) por secuencia. La línea roja representa la distribución GC de toda la secuencia y la línea azul representa la distribución teórica del contenido GC. Distribución para la muestra 1:Ca.P (A) y la muestra 3:SaP.L (B).

Por último el apartado Contenido de Kmer, que indica los subfragmentos de ADN de un tamaño determinado que se posicionan en la secuencia, estos no deberían aparecer en la región de interés o target, como se observa en la Figura 12, están presentes en la región donde se encuentran los adaptadores o cebadores, indicándonos que las secuencias todavía cuentan con estos elementos.



**Figura 12.** Contenido de Kmer. Este apartado mide el número de cada k-mer de tamaño 7 en cada posición de la librería y usa un análisis binomial para buscar desviaciones significativas de una cobertura en todas las posiciones.

Dentro de FastQC para cada uno de los apartados analizados se muestra un indicador de aprobación si no tiene errores o ha cumplido con los requisitos teóricos, un indicador de advertencia para que el usuario tenga en cuenta al momento de la filtración y un indicador de error cuando posee muchas inexactitudes o variaciones.

Posteriormente se realizó el filtrado de calidad, corte de adaptadores y primers de las lecturas Paired-End ya ensambladas. En la Tabla 10 se muestran estos resultados de los pasos durante el pre-procesamiento. En promedio se alcanzó un 74.82% de secuencias filtradas, que pasaron el control de calidad para el procesamiento y análisis final. Se observó que la muestra puruma de Patacamaya (Pa.P) tuvo el mejor rendimiento en la secuenciación, con 82.2% de secuencias que pasaron el filtro, teniendo muy pocas lecturas cortas.

**Tabla 10.** Resultados de la secuenciación 16S para las lecturas Forward.

ID	N° de lecturas	Después del ensamblaje	Lecturas muy cortas	Conteo de caracteres N	Longitud media	Después del filtrado	Secuencias limpias
Ca.P	200456	125018	9876	21478	449,00	93664	74,92 %
Ca.P.C	188391	121198	8318	22077	445,00	90803	74,92 %
Ca.P.L	230360	154938	11224	27665	445,00	116049	74,90 %
Ca.P.S	202383	129008	9582	22120	454,00	97306	75,43 %
Ca.S.C	188282	122004	8292	21841	459,00	91871	75,30 %
Ca.S.L	198251	130369	9837	22947	445,00	97585	74,85 %
Ca.S.S	213668	143085	10525	26205	460,00	106355	74,33 %
Ch.C	202528	128915	8680	22963	447,00	97272	75,45 %

ID	N° de lecturas	Después del ensamblaje	Lecturas muy cortas	Conteo de caracteres N	Longitud media	Después del filtrado	Secuencias limpias
Ch.L	105757	69028	4973	12455	454,00	51600	74,75 %
Ch.P	180866	121078	8651	21345	445,00	91082	75,23 %
Ch.S	203614	138148	9546	24807	448,00	103795	75,13 %
Ir.P	242096	157273	10735	28685	445,00	117853	74,94 %
IrP.C	144342	94604	7187	16834	445,00	70583	74,61 %
IrP.S	211704	141799	10168	26246	454,00	105385	74,32 %
IrS.C	140494	94371	6868	17489	447,00	70014	74,19 %
IrS.L	215295	127115	10384	22425	445,00	94306	74,19 %
IrS.S	186725	123048	8751	21678	454,00	92619	75,27 %
Pa.C	203509	122897	8409	20918	447,00	93570	76,14 %
Pa.P	196513	127181	95	22542	445,00	104544	82,20 %
Pa.S	214087	141284	9842	25452	449,00	105990	75,02 %
Sa.P	122819	76512	5278	13355	449,00	57879	75,65 %
SaP.C	116824	73976	5603	12763	451,00	55610	75,17 %
SaP.L	1076	271	59	37	445,00	175	64,58 %
SaP.S	208316	144616	10207	26909	459,00	107500	74,33 %
SaS.C	149109	100393	6780	18622	445,00	74991	74,70 %
SaS.S	123776	82769	5894	14962	448,00	61913	74,80 %
Su.C	199002	124581	9201	21976	458,00	93404	74,97 %
Su.L	143793	93376	6697	16852	457,00	69827	74,78 %
Su.P	146990	100273	6820	18495	458,00	74958	74,75 %
Su.S	191237	130542	9116	23794	446,00	97632	74,79 %

Debido a que se obtuvo muy poca cantidad de secuencias (175 secuencias) en la muestra de SaP.L de Saitoco (planicie en fase lechosa), será descartada para los análisis de diversidad, ya que no será compatible con el mínimo de secuencias en común para todas las muestras.

Finalmente, a partir de estas secuencias ensambladas se armó la biblioteca de secuencias con todas las muestras que fue empleada para todos los análisis posteriores (*split\_libraries\_fastq.py*),

### C. Grupos taxonómicos identificados por zona agroproductiva.

De todas las lecturas obtenidas, más del 99% fueron asignadas taxonomicamente empleando la base de datos de Green Genes. Esta clasificación taxonómica permitió obtener los siete niveles taxonómicos, sin embargo, a nivel de especie no se obtuvo una asignación en la mayoría de los OTUs, debido a que Green Genes no cuenta con la profundidad de identificación. Esta profundidad o alcance depende de las actualizaciones a la base de datos.

Se empleó el método Open Reference (*pick\_open\_reference\_otus.py*) para realizar la clasificación, este método es apropiado para muestras ambientales ya que existe una porción considerable de organismos que aún no han sido

identificados y estas deben conformar parte de los análisis posteriores. Lo contrario sucede en el caso la microbiota humana, que se tiene identificada prácticamente la totalidad de microorganismos y en este caso es adecuado emplear el método Closed Reference.

### Cuadro 1. Datos del agrupamiento y ensamblado

<b>Numero de muestras</b>	30
<b>Numero de observaciones</b>	12417
<b>Conteo Total</b>	1437277
Resumen del conteo por muestra:	
<b>Min</b>	111,0
<b>Max</b>	70071,0
<b>Mediana</b>	50487,500
<b>Promedio</b>	47909,233
<b>Desviación estándar</b>	13833,126

De las 1437277 lecturas filtradas (Cuadro 1) se obtuvieron 12417 OTUs, la desviación es amplia debido a la muestra de Saitoco (SaP.L). En el Cuadro 2 se detallan las cantidades de secuencias obtenidas en cada muestra.

### Cuadro 2. Detalle de los conteos por muestra

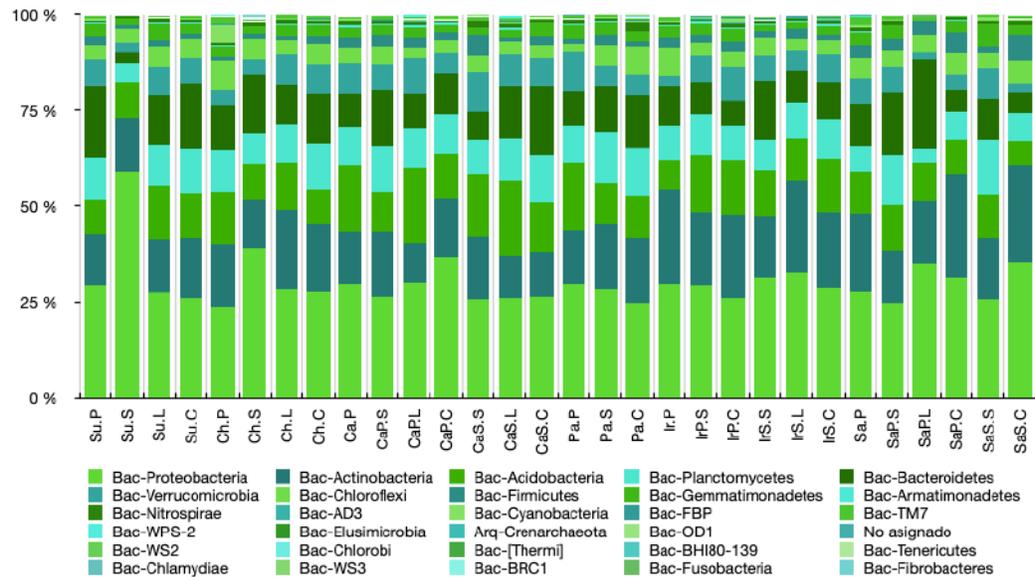
<b>SaP.L</b>	111,0	<b>CaS.L</b>	51189,0
<b>Sa.P</b>	27683,0	<b>Su.C</b>	52773,0
<b>Ch.L</b>	30695,0	<b>CaP.C</b>	52839,0
<b>SaP.C</b>	31335,0	<b>CaS.C</b>	52927,0
<b>SaS.S</b>	38106,0	<b>Ir.P</b>	54145,0
<b>IrS.C</b>	38173,0	<b>CaP.S</b>	55326,0
<b>IrP.C</b>	39186,0	<b>IrP.S</b>	55389,0
<b>Su.L</b>	42173,0	<b>Ch.C</b>	58325,0
<b>Pa.C</b>	42241,0	<b>Pa.P</b>	58326,0
<b>SaS.C</b>	42399,0	<b>Ch.S</b>	59755,0
<b>Ca.P</b>	44306,0	<b>SaP.S</b>	61329,0
<b>IrS.S</b>	44369,0	<b>CaP.L</b>	63372,0
<b>Su.P</b>	45220,0	<b>CaS.S</b>	64422,0
<b>Ch.P</b>	46082,0	<b>Pa.S</b>	65224,0
<b>IrS.L</b>	49786,0	<b>Su.S</b>	70071,0

A nivel de Reino en el altiplano norte, las bacterias alcanzaron 99.94% en comparación con arqueas (0.026%) y los no identificados alcanzaron 0,035%. En el altiplano central se alcanzó al 99.85% para bacterias y 0.036% para arqueas, los OTUs no identificados llegaron al 0.109%; y en el altiplano sur se alcanzó 99.77% para bacterias y 0.043% para arqueas, dejando el 0.186% para OTUs no asignados. Se observa que los OTUs no asignados y la presencia de arqueas tiende a aumentar hacia el sur del altiplano. En las Tablas 11 a 13 se muestra en detalle, a nivel de reino, los OTUs identificados por cada zona de muestreo.

**Tabla 11.** Porcentajes de abundancia en muestras puruma y cosecha de las 3 zonas agroproductivas del altiplano a nivel de reino.

Reino	Norte		Centro		Sur	
	Puruma	Cosecha	Puruma	Cosecha	Puruma	Cosecha
Arquea	0,026 %	0,016 %	0,036 %	0,029 %	0,043 %	0,029 %
Bacteria	99,940 %	99,951 %	99,855 %	99,948 %	99,771 %	99,912 %
No asignado	0,034 %	0,033 %	0,109 %	0,023 %	0,186 %	0,059 %

A nivel de filo se ha alcanzado a identificar 47 taxones bacterianos en todas las muestras, de las cuales las más abundantes fueron: Proteobacteria, Actinobacteria, Acidobacteria, Planctomycetes, Bacteroidetes, Verrucomicrobia, Chloroflexi, entre otros (Figura 13). Entre las arqueas, Crenarchaeota estuvo entre las más abundantes.

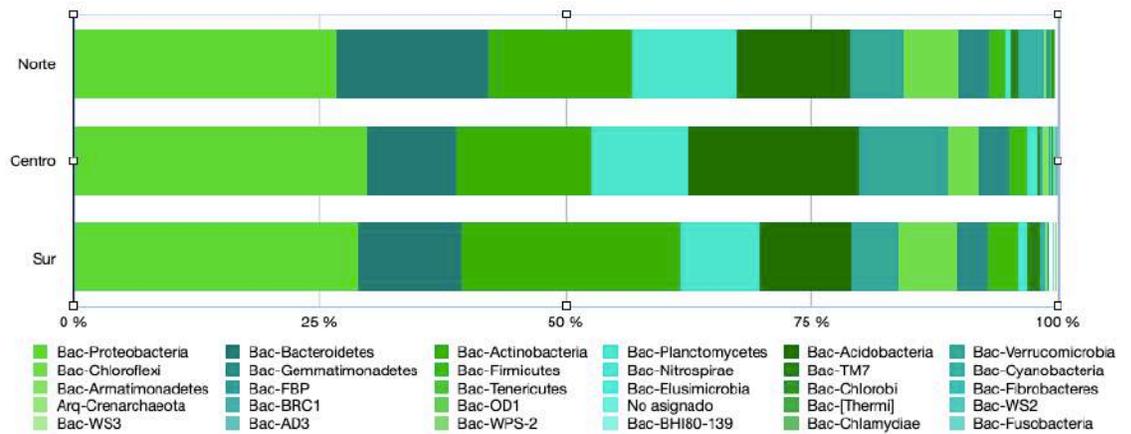


**Figura 13.** Clasificación taxonómica de los filos más abundantes en las muestras de estudio.

Por zona agroproductiva, en el norte los filos más abundantes fueron: Proteobacteria, Bacteroidetes, Actinobacteria, Acidobacteria, Planctomycetes, Verrumicrobia entre otros. En la zona central Proteobacteria, Acidobacteria, Actinobacteria, Planctomycetes, Verrumicrobia Bacteroidetes fueron los más abundantes, y en el Sur Proteobacteria, Actinobacteria, Bacteroidetes, Acidobacteria, Planctomycetes, Chloroflexi (Figura 14).

Estos filos conforman la microbiota presente en el altiplano ya mencionado por Gomez-Montano et al. (2013) en el que se emplearon la pirosecuenciación, e identificaron a Proteobacteria, Actinobacteria, Firmicutes, Acidobacteria, Verrucomicrobia y Bacteroidetes como los filos mas abundantes en las comunidades de Umala y Ancoraimes (Altiplano norte).

Solo algunos de los filos identificados variaron por zona agroproductiva; por ejemplo, para proteobacterias, los suelos en descanso no presentaron variación significativa según la zona, pero sí se observo que en los suelos de cultivo tuvieron mayor proporción en la zona norte en comparación con las otras zonas.



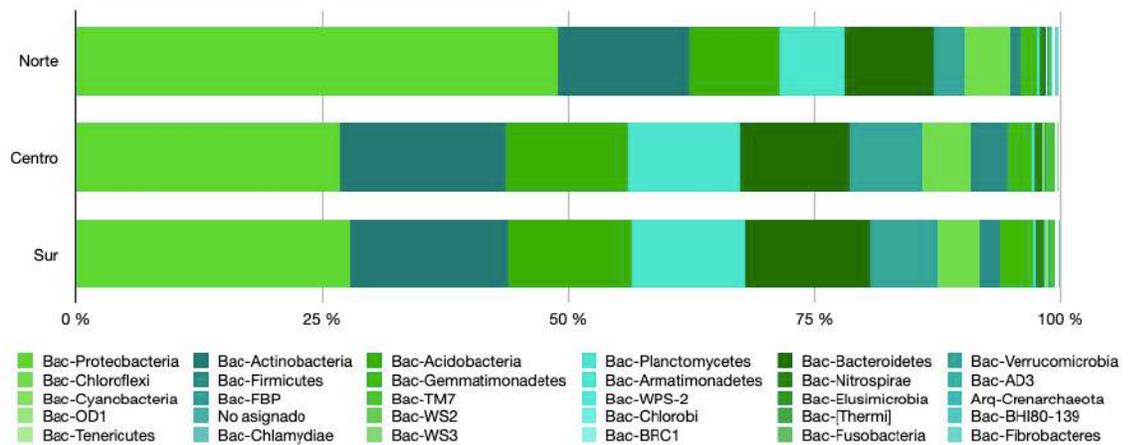
**Figura 14.** Abundancias de filos en las muestras Puruma según la zona agroproductiva.

En el norte, refiriéndose a suelos en descanso, se observó Bacteroidetes, Cyanobacteria, FBP, Chlorobi y WS3 como los filos más abundantes en comparación con las otras zonas. Las abreviaciones fueron los nombres iniciales o alternativos hasta la ultima versión disponible de la base de datos de GreenGenes. FBP se refiere a la Abditibacteria, una bacteria muy extendida en ambientes extremos, desde ecosistemas polares, desiertos hasta aguas residuales y sitios mineros contaminados (Tahon, Tytgat, Lebbe, Carlier, & Willems, 2018), pertenece actualmente al filo Armatimonadetes (Parks et al., 2018). WS3 es Latescibacteria un filo candidato de bacterias recientemente propuesto, están presentes en diversos tipos de algas (Farang, 2015).

En la zona central los filos más abundantes fueron: Acidobacteria, Verrumicrobia, AD3, Elusimicrobia y BHI80-139. AD3 es un filo que esta

relacionado con plantas, este se desarrolla dentro de las raíces, pero permanece en el suelo (Billings et al., 2018). BHI80-139 es el filo recientemente propuesto como Aerophobetes, los análisis genómicos revelan una amplia distribución en los sedimentos de aguas profundas (Wang et al., 2016). En el sur el filo Actinobacteria tuvo mayor presencia, así como Firmicutes y Fibrobacteres.

Estas abundancias observadas fueron muy diferentes en los suelos de cultivo (Figura 15). En el norte, además de Proteobacterias ya mencionado, también abundaron Fibrobacteres. Por otro lado se observó una reducida presencia de Planctomycetes, Verrucomicrobia y Cyanobacterias en comparación con las otras zonas, este último contrario a los suelos Puruma, donde Cyanobacteria es más abundante. En el centro Firmicutes y WS3 tuvieron mayor abundancia sin ninguna otra observación. En el sur la única observación que se tiene es la presencia de BHI80-139, filo que no se identificó en las otras zonas.



**Figura 15.** Abundancias de los filos más abundantes en suelos de cultivo por zona agroproductiva.

Estas variaciones parecen tener influencia por el manejo de suelo, puruma o cultivo, así también las propiedades edáficas de los suelos pueden tener influencia sobre estos. Esto indica que en los suelos en descanso las bacterias tuvieron un desarrollo muy diferente en comparación con los suelos donde se realiza el cultivo habitual.

En cuanto al género, muchos OTUs no fueron identificados hasta este nivel taxonómico, debido al alcance de GreenGenes. Entre estos, un género de la clase WD2101 del filo Planctomycetes fue el más abundante en el norte y centro (Anexo 3). Esta clase es Phycisphaerae (WD2101), son cocos Gram negativos, presentan flagelos y son aerobias o facultativas, se los ha encontrado asociado a diferentes tipos de algas (Yoon, Jang, & Kasai, 2013). En suelos de cultivo este género tuvo mayor abundancia, principalmente en el centro y sur.

Kaistobacter fue el más abundante en dos muestras, una de la zona norte (6.6%) y otra de la zona central (5.1%), en el resto de las muestras se observó una

abundancia promedio del 2%. Este género habita principalmente suelos orgánicos (rectificados) y tiene una alta tasa de crecimiento (Schwartz et al., 2014). En suelos de cultivo fue más abundante en todas las muestras llegando hasta un 7.7%.

El género *Flavobacterium* tuvo 0.1% en promedio en todas las muestras, pero tuvo una mayor abundancia en la muestra de Sullcataca con 3.7%. Se conoce que habita suelos y cuerpos de agua dulce y en una variedad de ambientes, algunos son patógenos de peces (BERNARDET et al., 1996). En cuanto a los suelos de cultivo este genero tuvo entre el 0.2 a 1.3% en todas las muestras.

Un genero del filo Verrucomicrobia descrito como DA101 fue muy abundante en el altiplano central (entre el 5 y 8%) y en una muestra del altiplano sur (5%). Dentro este genero, a pesar de haber sido descrito en 1998 mantiene ese nombre y se han identificado varios filotipos que pertenecen a la misma familia. Estos forman parte del ranking de las bacterias más abundantes en distintos estudios realizados en suelos principalmente mediante la técnica de shotgun (Brewer, Handley, Carini, Gilbert, & Fierer, 2016). Según este estudio, DA101 se encuentra mayormente en suelos de pastizales que en suelos forestales y concluyeron que DA101 prefiere suelos que reciben cantidades elevadas de insumos de carbono lábiles. En suelos de cultivo se observo una abundancia entre 1.1% y 7.5%.

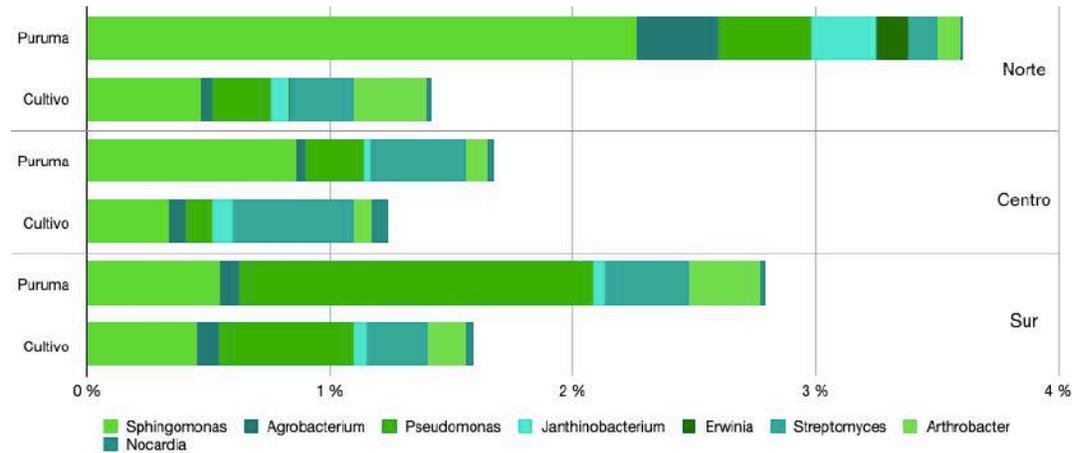
*Sphingomonas* fue abundante en el norte (2%), este genero se distribuye extensamente en la naturaleza, y tiene un rango amplio de aplicaciones dentro la biotecnología como la biorremediación (Bourne, Riddles, Jones, Smith, & Blakeley, 2001). En las muestras de suelos de cultivo se observó una equitativa abundancia en todas las muestras (en promedio 0.5%).

*Pedobacter* fue abundante en Sullcataca (altiplano norte) en suelo puruma (2%), un genero Gram negativo que habita diferentes tipos de ecosistemas (Zhou et al., 2012), este genero ha sido identificado como un contaminante en sistemas de agua ultra puras, (Gallego, García, & Ventosa, 2006).

*Flavisolibacter* (1.5-1.8%) y ligeramente *Hymenobacter* (0.8-1.2%) fueron abundantes en el norte, este ultimo llego a 3.42% en el centro. Es un genero que habita una amplia gama de nichos ecológicos y varios tipos de ambientes extremos (minas de uranio) (Ohn, Ten, Kim, Cho, & Jung, 2018), son bacterias cultivables y muy sensibles a numerosos antibióticos (Krieg et al., 2010),

*Segetibacter* (3.2%), *Candidatus solibacter* (2.7%) y *Devosia* (3.5%) fueron abundantes en Cañaviri (altiplano central), estas bacterias, propias de los suelos, son inmóviles con excepción de *Devosia* que posee flagelo. *Candidatus solibacter* es un genero muy estudiado, ejerce diferentes funciones dependiendo de las condiciones del ambiente, empleadose también como biorremediador de minas de sulfuro (Johnson, Yajie, & Okibe, 2008).

Dos especies muy estudiadas y ubicuas en diferentes tipos de hábitat fueron abundantes en el sur, Bacillus en Saitoco (2.9%) y Pseudomonas en Irpani (2.4%). Estos géneros son ampliamente estudiados, varios géneros han llegado a tener gran importancia en diferentes ámbitos como la biorremediación, en la industria, en la agricultura y otros. En cuanto al género Pseudomonas algunas especies han sido reportadas como fitopatógenos, particularmente en plantas agrícolas. Además de estos géneros fitopatógenos ya mencionados, se identificó otros géneros en menor abundancia presentes en las muestras (Figura 16).



**Figura 16.** Géneros bacterianos potencialmente fitopatógenos presentes en los suelos de cada zona agroproductiva.

Se observó mayor abundancia de fitopatógenos en suelos Puruma en comparación con los suelos en cultivo, esta abundancia reducida de fitopatógenos tuvo variación entre las diferentes muestras, lo que indica la diferencia en el manejo del suelo en cada sitio de muestreo, sumado con la utilización de agroquímicos o pesticidas que modificarían la abundancia y desarrollo de estos microorganismos.

En cuanto a las etapas de crecimiento de la quinua, las proporciones y abundancias de estos fitopatógenos no tuvieron un comportamiento general (Anexo 4), sino más bien tuvieron comportamientos diversos; mientras en algunos sitios se observa aumento en las abundancias en otros existe disminución, otros se reducen en fase lechosa y aumenta de nuevo en fase de pre-cosecha. En forma general se observa que en la zona sur del altiplano existe una mayor abundancia de estos fitopatógenos, mientras en norte y centro las abundancias son menores pero muy similares entre sí.

#### **D. Grupos taxonómicos según la fase de crecimiento de la Quinua.**

En cuanto a las diferencias taxonómicas durante las fases de crecimiento de la quinua, se observaron varias formas de comportamientos entre los diferentes taxones. Para los reinos, en el norte del altiplano se observó una leve

disminución de bacterias, por el contrario las arqueas y los grupos no identificados tuvieron un ligero incremento.

**Tabla 12.** Abundancia a nivel de reino en las muestras del altiplano norte.

Sitio	Etapa	Arquea	Bacteria	No asignado
<i>Sullcataca baja</i>	Pre-siembra	0,007 %	99,973 %	0,020 %
	Fase lechosa	0,028 %	99,945 %	0,026 %
	Pre-cosecha	0,030 %	99,932 %	0,038 %
<i>Choquenaira</i>	Pre-siembra	0,025 %	99,930 %	0,045 %
	Fase lechosa	0,029 %	99,938 %	0,033 %
	Pre-cosecha	0,024 %	99,925 %	0,051 %

En la zona central del altiplano las muestras de cada sitio mostraron comportamientos contrarios, mientras que en la planicie de Cañaviri las bacterias incrementaban a través de las fases fonológicas, en la serranía disminuían. Esto significaría que las comunidades bacterianas tienen un desarrollo diferente en las pendientes. Sin embargo, en el altiplano Sur se observa comportamientos diferentes, en la planicie de Saitoco las bacterias disminuyeron y de igual modo en laderas, en la planicie y laderas de Irpani las bacterias incrementaron. Por lo tanto la variación no está relacionada con las pendientes o con el contorno del suelo, en termino no son factores que sean directamente relevantes a la diversidad de microorganismos.

Una aspecto particular se observó en los grupos no asignados, de manera general este grupo tiende a aumentar en abundancia durante las fases de crecimiento. Esto indica que la abundancia de los OTUs no identificados ha ido creciendo durante el desarrollo de la planta, esto significaría que en este grupo se encuentran microorganismos promotores o solo simbióticos con la quinua, ya que se reduce la abundancia de los microorganismos identificados. En este caso para conocer a estos grupos e identificarlos se requieren de otros estudios.

**Tabla 13.** Abundancia a nivel de reino en las muestras del altiplano central.

Sitio	Etapa	Arquea	Bacteria	No asignado
<i>Cañaviri (planicie)</i>	Pre-siembra	0,439 %	99,535 %	0,025 %
	Fase lechosa	0,028 %	99,957 %	0,014 %
	Pre-cosecha	0,013 %	99,966 %	0,021 %
<i>Cañaviri (serranía)</i>	Pre-siembra	0,045 %	99,936 %	0,019 %
	Fase lechosa	0,029 %	99,934 %	0,037 %
	Pre-cosecha	0,036 %	99,928 %	0,036 %

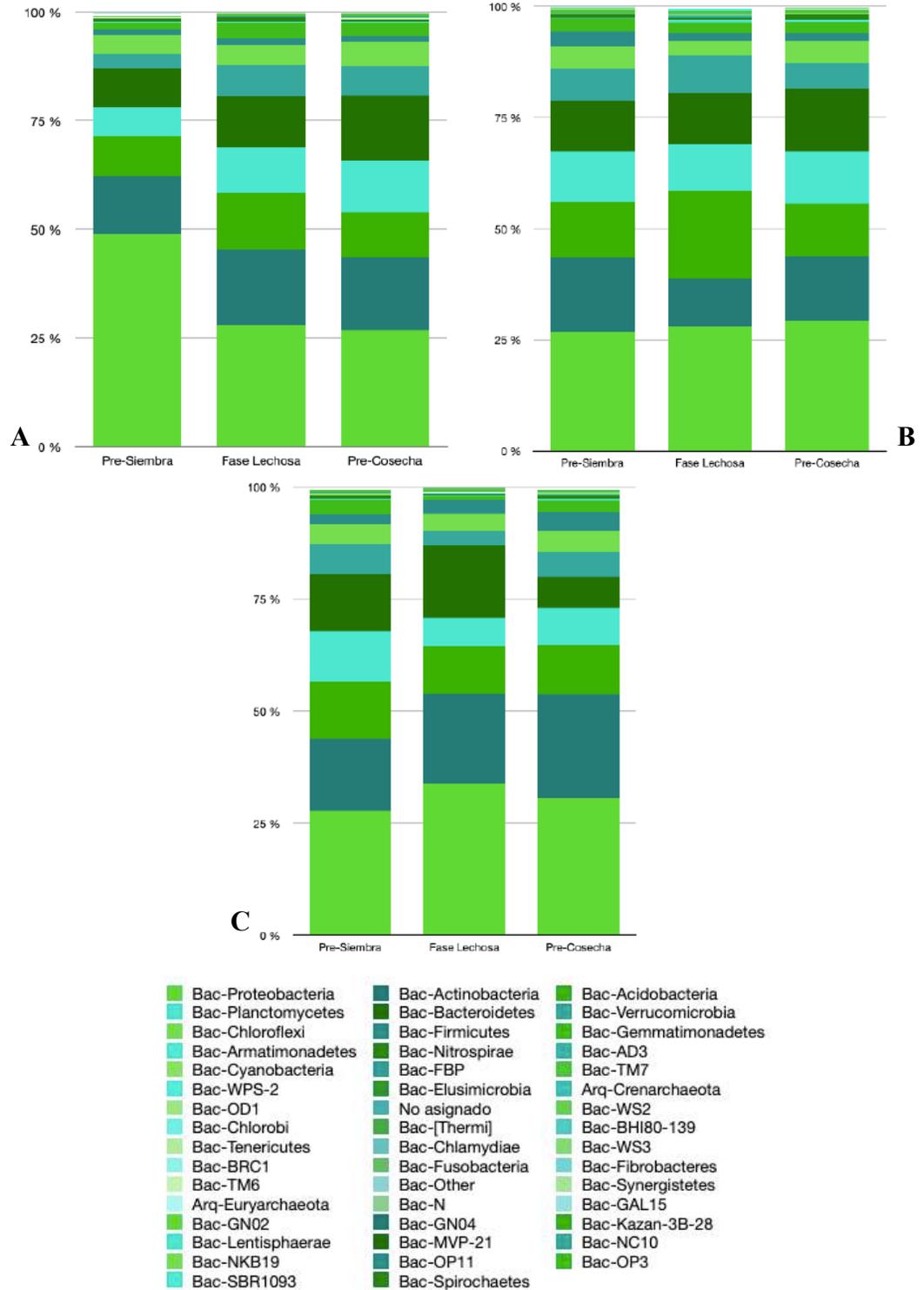
Sitio	Etapa	Arquea	Bacteria	No asignado
<i>Patacamaya</i>	Pre-siembra	0,028 %	99,936 %	0,037 %
	Pre-cosecha	0,099 %	99,839 %	0,062 %

**Tabla 14.** Abundancia a nivel de reino en las muestras del altiplano sur.

Sitio	Etapa	Arquea	Bacteria	No asignado
<i>Irpani (planicie)</i>	Pre-siembra	0,043 %	99,928 %	0,029 %
	Pre-cosecha	0,013 %	99,946 %	0,041 %
<i>Irpani (serrania)</i>	Pre-siembra	0,034 %	99,833 %	0,133 %
	Fase lechosa	0,030 %	99,944 %	0,026 %
	Pre-cosecha	0,010 %	99,963 %	0,026 %
<i>Saitoco (planicie)</i>	Pre-siembra	0,023 %	99,917 %	0,060 %
	Fase lechosa	0,000 %	100,000 %	0,000 %
	Pre-cosecha	0,089 %	99,872 %	0,038 %
<i>Saitoco (serrania)</i>	Pre-siembra	0,016 %	99,971 %	0,013 %
	Pre-cosecha	0,045 %	99,906 %	0,050 %

A nivel de filo algunos grupos tuvieron variación de las abundancias durante las etapas de desarrollo (Figura 17). Por ejemplo, Proteobacterias disminuyeron en el norte, mostraron un leve crecimiento en el centro y en el sur aumento durante la fase lechosa pero disminuyó antes de la cosecha. Actinobacterias tuvo un mayor crecimiento en el sur, seguido del norte, pero en el centro disminuyó en la Fase lechosa. Acidobacterias se mantuvo aproximadamente a los mismos niveles, solo en el centro aumentó en la fase lechosa pero disminuyo antes de la cosecha. Bacteriodetes mantuvo un crecimiento constante durante las fases en el norte y centro, pero en el sur aumento en la fase lechosa y disminuyo hasta la cosecha.

En los filios con menor abundancia la variación de estas durante el desarrollo de la quinua no fue muy pronunciada en comparación con los ya mencionados, solo en la muestra de Sullcataca Baja tuvieron una menor abundancia debido a la mayor presencia de proteobacterias.

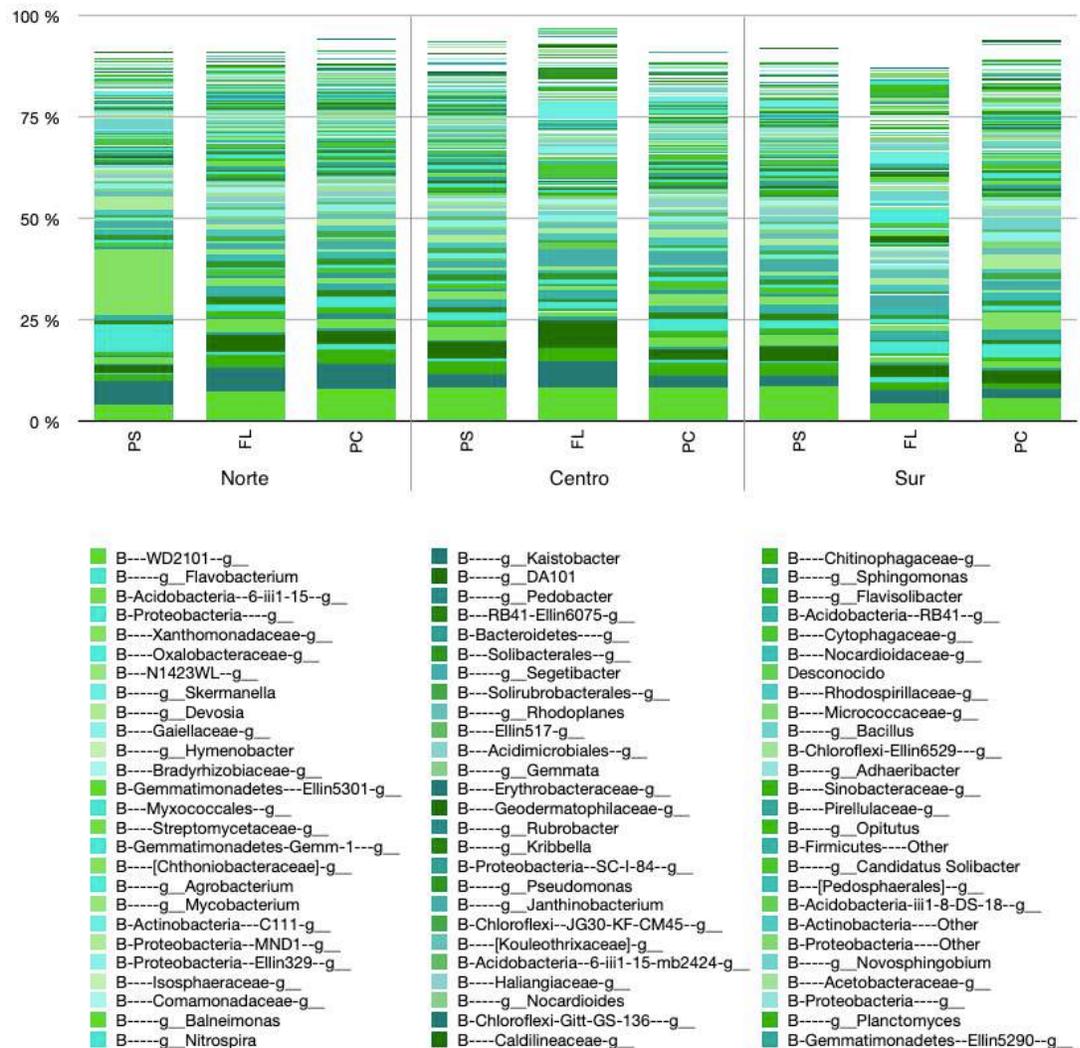


**Figura 17.** Abundancias a nivel de filo de las tres zonas agroproductivos del altiplano boliviano por fase de desarrollo de la quinua. **(A)** Zona norte, **(B)** Zona central, **(C)** Zona sur.

En cuanto a los géneros identificados durante el desarrollo de la quinua, se observó de manera general que las abundancias llegan a obtener proporciones casi equitativas durante las fases, es decir, las variaciones no fueron muy pronunciadas en el transcurso de las fases de desarrollo (Figura 18). Sin embargo, existen algunas observaciones que realizar; por ejemplo, en la zona norte, en la fase de Pre-Siembra se observó dos géneros muy abundantes, aunque no fueron identificados hasta este nivel taxonómico, uno pertenece a la familia de las Xanthomonadaceae (28,5%) y el otro pertenece a la familia de las Sphingomonadaceae (8,7%), ambos del filo Proteobacteria. En las fases posteriores y en todos los sitios de muestreo mantuvieron un nivel muy semejante que promedia  $2,1\% \pm 0,9\%$ .

El genero WD2101 (Phycisphaerae) fue uno de los más abundantes en cuanto nos referimos a todas las muestras, este tuvo un ligero aumento en las etapas de desarrollo de 4,2% a 8,2% (zona norte), aunque en el Sur disminuyo bastante en la fase lechosa de 8,6% a 4,6% y un incremento antes de la cosecha a 5,8%. Otro género abundante fue Kaistobacter, principalmente en el Norte y DA101 tuvo incremento en Fase lechosa en el Norte y Centro.

WD2101 es un ejemplo de bacteria no cultivable que es muy abundante en todo tipo de ecosistemas, se ha identificado que su principal función es degradar polisacaridos (Dedysh et al., 2021). Esta involucrado en la degradación de varios heteropolisacaridos, ademas se cree que están involucrados en la degradación de desechos de musgos (esfagno) (Ivanova, Kulichevskaya, Merkel, Toshchakov, & Dedysh, 2016).



**Figura 18.** Abundancias a nivel de género de las tres zonas agroproductivos del altiplano boliviano por fase de desarrollo de la quinua.

Adicionalmente se observó la presencia de ciertas bacterias conocidas por sus características en ayudar directa o indirectamente en el crecimiento de las plantas, denominadas Rizobacterias Promotoras de Crecimiento Vegetal (PGPR). En el cual existen varios grupos que no necesariamente promueven el crecimiento en todas las clases de plantas sino en ciertas especies (Moreno Reséndez, Carda Mendoza, Reyes Carrillo, Vásquez Arroyo, & Cano Ríos, 2018). Estos grupos están formados según el tipo de mecanismo o acción que efectúa a favor de las plantas.

Pseudomonas: algunas especies dentro este genero están identificados como patógenos en ciertos cultivos como tomate, endivia y brócoli, pero otras especies como *P. fluorescens* son PGPR. La actividad enzimática las convierte en un grupo de microorganismos importante, debido a que son responsables de la degradación aeróbica de muchos compuestos en diversos ecosistemas y son

empleadas en distintas maneras dentro la agricultura (Pérez Álvarez, Coto Arbelo, Echemendía Pérez, & Ávila Quezada, 2015).

Mesorhizobium: es un genero simbiótico de plantas tuvo mayor abundancia en el norte seguido del Centro y Sur; Bradyrhizobium tuvo la mayor abundancia en la Fase lechosa en la Zona Sur (1%), en los demás casos se mantenía entre el 0,02% y el 0,1%; Bacillus y Enterobacter han sido evaluadas y reportadas como agentes de biocontrol contra patógenos de plantas (Moreno Reséndez et al., 2018), pocas especies dentro estos grupos han sido reportados como fitopatogenos. Bacillus estuvo entre 0,3% y el 0,8% en el Norte, incrementando en cada fase fonológica, en el Sur llego casi al 3% en la ultima fase.

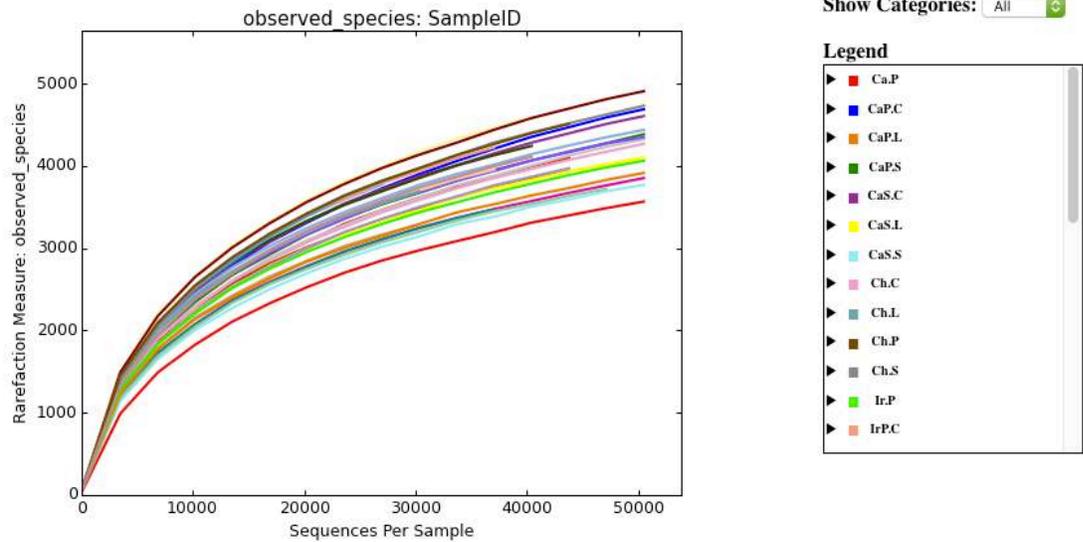
Paenibacillus: tuvo mayor abundancia en el Sur (0,49% en fase lechosa); Acinetobacter: una bacteria muy importante en el suelo que contribuye a su mineralización estuvo presente en las tres zonas mostrando un mismo comportamiento, de mayor abundancia en Pre-Siembra, reducción en fase lechosa y aumento en la Pre-Cosecha.

Sphingomonas aumentó en Fase lechosa, mientras que Cellulomonas disminuye en la misma fase. Otros géneros PGPR identificados en poca abundancia fueron: Burkholderia, Flavobacterium, Agrobacterium, Arthrobacter, Erwinia y Streptomyces.

#### **D. Análisis de diversidad alfa ( $\alpha$ ) y beta ( $\beta$ ).**

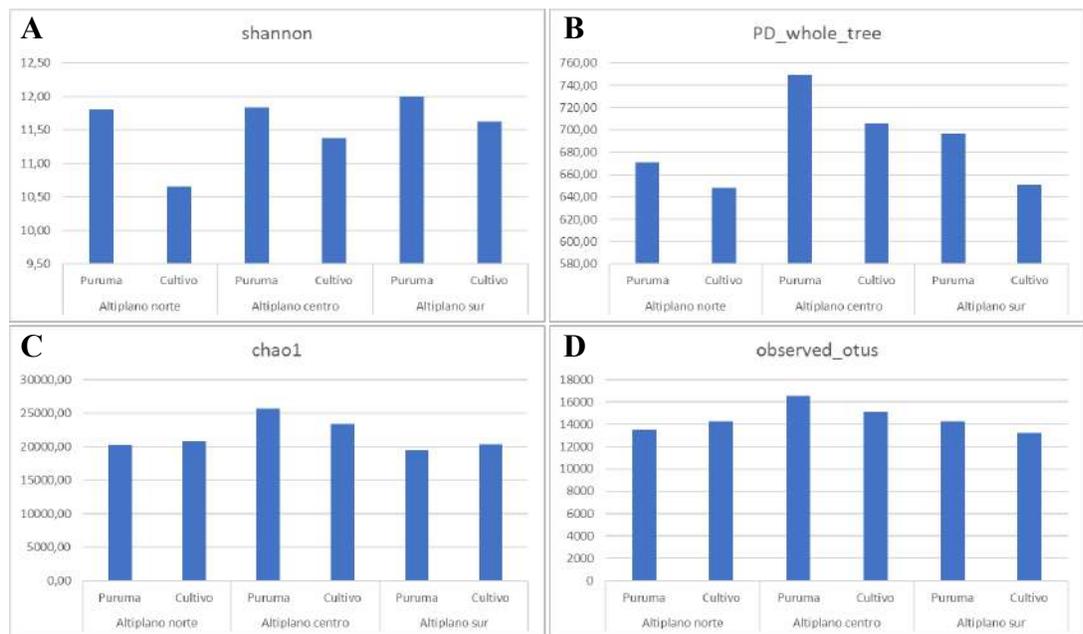
Para evaluar la riqueza y diversidad de las comunidades bacterianas en los suelos de cultivo se obtuvieron las curvas de rarefacción, la riqueza de especies (observed\_otus), el estimador de riqueza Chao1, el indicador de diversidad filogenética (PD\_whole\_tree) y la diversidad de Shannon.

Las curvas de rarefacción permiten observar el grado de cobertura alcanzado en la secuenciación y comparar visualmente la diversidad entre las distintas muestras. En la Figura 19 se muestran las curvas de rarefacción con el número de OTUs generados para cada muestra de suelo en función del número de secuencias. Con respecto a la forma de las curvas, varias de las muestras presentaron curvas que se aproximan a una asíntota horizontal indicando que la secuenciación cubrió una parte importante de la riqueza de especies en estas muestras. No obstante algunas de las muestras no alcanzaron el plateau, lo cual indica que se debe aumentar el esfuerzo de secuenciación para lograr una óptima cobertura de la diversidad.



**Figura 19.** Curvas de rarefacción por muestras para el número de especies observadas (Observed\_species).

Los índices o valores de diversidad fueron obtenidos siguiendo el protocolo de Qiime, con los comandos ya mencionados, estos valores finales son sobreestimados de acuerdo al manual del software, esta sobreestimación de los índices serán corregidas en versiones futuras. Es por eso que los valores sobrepasan las referencias teóricas según cada indicador.



**Figura 20.** Índices de diversidad alfa por región altiplánica. (A) Índice de Shannon, (B) Diversidad filogenética, (C) Estimador de riqueza chao1, (D) Numero de OTUs observados.

De manera general, en la Figura 20 se observa que el altiplano central posee una mayor diversidad microbiana en comparación a las zonas norte y sur, además la zona con la menor diversidad y riqueza microbiana es la zona norte del altiplano (Datos por muestra en Anexo 5). Los índices de diversidad de Shannon y diversidad filogenética PD\_whole\_tree muestran que la diversidad es mayor en suelos Puruma que en los suelos de cultivo en las 3 regiones del altiplano (Figura 20A y 20B).

PD\_whole\_tree determina la biodiversidad a travez del tiempo, en este caso observando los resultados del altiplano central indica que existe mayor diversidad filogenética, es decir mayor ramificación en el árbol filogenético de las especies en esa comunidad, caso contrario a lo que se observa en la zona norte que presenta índices menores a las otras zonas; esto se debe a varios factores pero principalmente a la mayor actividad antropogénica en la zona; adicionalmente se debe tener en cuenta que en una de los puntos de muestreo (Patacamaya) viene de un terreno en el que se realizan manejos particulares en el suelo ya que es un sitio de estudios agronómicos; por otro lado se observo que el efecto del pastoreo de ganados puede agregar sentido a estos resultados, ya que en la zona norte el pastoreo es muy reducido seguido de la zona sur y terminando en la zona central donde el pastoreo es mayor (Genin, 2006).

El índice de Shannon contempla la diversidad de especies (riqueza de especies) y la cantidad de individuos por especie. En este caso Shannon indica que la diversidad es mayor en la Zona Sur, es decir existe mayor riqueza de especies, mientras que en el Norte la riqueza es mucho menor.

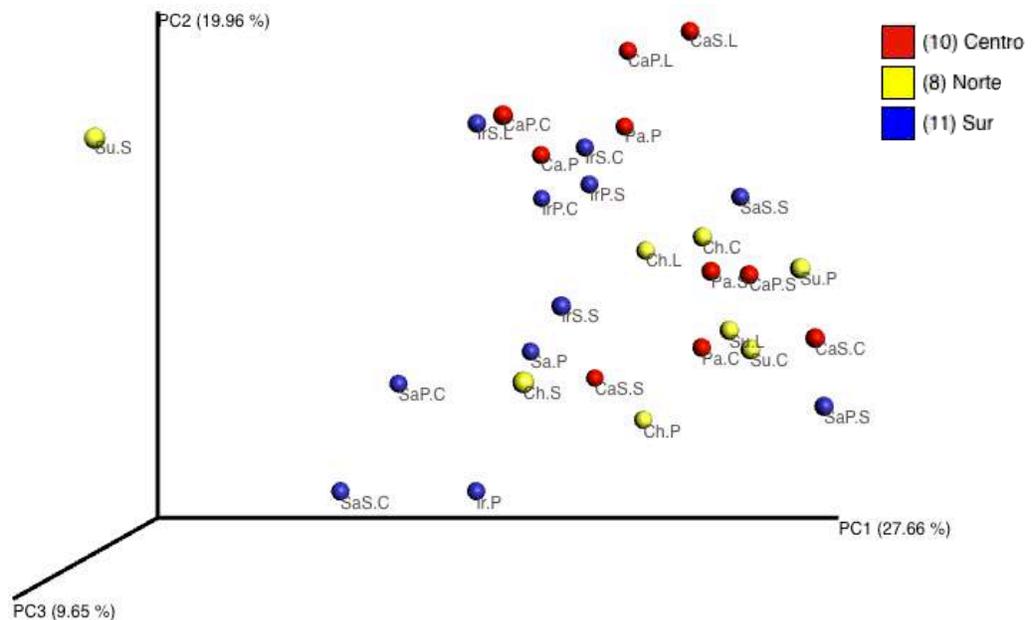
El índice Chao1 indica que la riqueza es mayor en los suelos de cultivo que en los suelos en descanso para las zonas norte y sur (Figura 20C), contrariamente a lo que se observa en la región central del altiplano. Chao1 toma en cuenta los singletons (un individuo) y doubletons (dos individuos) presentes en cada especie. En los suelos de cultivo la presencia de especies con pocos individuos es mayor en comparación con los suelos puruma, esto quiere decir que las comunidades en estas zonas poseen más especies con menor cantidad de individuos, durante el descanso las comunidades se van generando sin alteración por el manejo del suelo. Cuando el efecto de la preparación del suelo o la re-siembra tiene lugar estas comunidades se ven afectadas y por lo tanto su número se reduciría, mientras el suelo esta en descanso las bacterias poco frecuentes están en desarrollo (Gotelli & Colwell, 2011).

La cantidad de observaciones o individuos identificados en los suelos Puruma es mayor a los identificados en los suelos de cultivo en las regiones centro y sur del altiplano, por el contrario, en el altiplano norte la cantidad de individuos identificados en suelo Puruma es menor a la de suelo de cultivo (Ecología de Comunidades y Sistemas 2015).

Para el análisis de diversidad Beta, Qiime calcula matrices de distancia, que son luego representadas en gráficas PCoA y así visualizar las disimilitudes entre las muestras comunitarias. La métrica empleada fue tanto Weighted como Unweighted para encontrar agrupaciones que puedan responder con la hipótesis planteada.

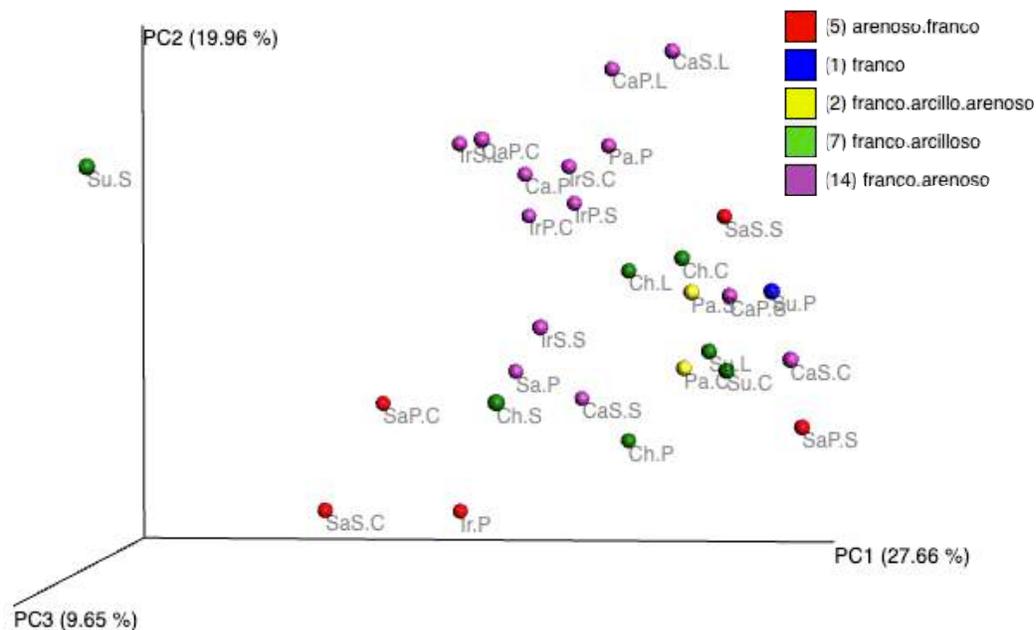
Unifrac, semejante a PD\_Whole\_tree en la diversidad alfa, no solo indicara que tan diferentes son las muestras entre sí, sino también mostrara si existe una relación evolutiva entre estas. Weighted Unifrac es útil para examinar las diferencias en la estructura comunitaria; mientras Unweighted Unifrac es más sensible a las diferencias en las características de baja abundancia. Ambos son útiles para interpretar los grupos aunque hay casos en los que ambos, o solo uno, o ninguno pueden arrojar diferencias significativas entre los grupos experimentales (C. A. Lozupone, Hamady, Kelley, & Knight, 2007). Y esto es lo que se observó en los resultados obtenidos con dos principales variables tomadas en cuenta para este análisis.

En general, en cuanto a la región altiplánica, no se observó una conformación de grupos solida, respecto a las variables espaciales y temporales como: Zona y Fase fonológica. En la gráfica 21 se observa una mezcla general de todas las muestras. Algo que sí es posible distinguir es la ligera diferenciación de las muestras de la zona Sur, esto indica que existe una disimilitud entre comunidades del Sur y las comunidades del centro y norte conjuntamente.



**Figura 21.** Análisis PCoA por zona agroproductiva (Weighted Unifrac). Se indican los porcentajes de variación en las 3 coordenadas principales para todas las muestras coloreadas por zona.





**Figura 23.** Análisis PCoA por el tipo de suelo (Weighted Unifrac). Se indican los porcentajes de variación de las 3 coordenadas principales. Agrupamiento claro para el tipo franco-arenoso.

Las gráficas para Unweighted Unifrac se muestran en Anexo 6, en el que también se puede observar claramente agrupamiento por el tipo de suelo.

En cuanto al análisis de diversidad beta, que requieren análisis bioestadísticos más detallados para poder comprender la estructura microbiana y tener conclusiones más precisas, quedan pendientes para posteriores estudios, en los que se deben tener mayor información sobre variables pertinentes.

#### **E. Relaciones entre las comunidades microbianas y las características de los suelos.**

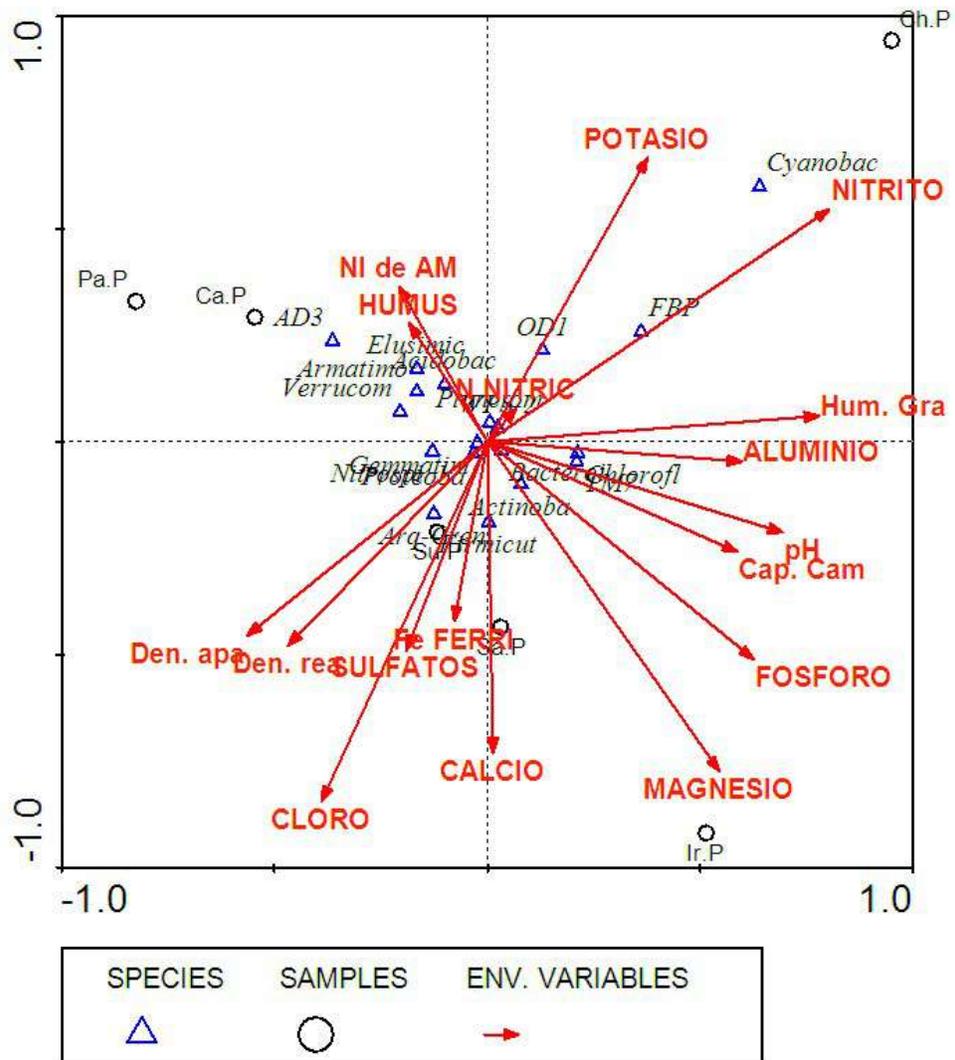
El análisis CCA para las muestras Puruma muestra la relación potencial de las variables ambientales y las comunidades bacterianas en cada punto de estudio. En este análisis se tomaron en cuenta a las bacterias más abundantes, representadas con un triángulo como símbolo, los analitos en flechas rojas que indican su tendencia, y los sitios en círculos.

La mayoría de las bacterias se concentran en la parte central del gráfico, donde los valores o tendencias de los analitos representan un promedio o equilibrio entre los distintos sitios. Esto significa que los microorganismos están correlacionados con la mayoría de los analitos.

En la Figura 24 se observa la correlación con las muestras Puruma, donde Choquenaira (Ch.P) presenta una mayor correlación con las concentraciones de Potasio y Nitrito en comparación a los otros sitios, además tiene una correlación fuerte con la presencia de Cyanobacterias. En el Sur, Irpani presenta mayor

correlación con magnesio y fósforo en comparación con los otros sitios, en especial con Patacamaya (Pa.P) y Cañaviri (Ca.P) en el que se observa una correlación muy negativa con estos analitos. Esto concuerda con un estudio en el desierto de Atacama donde Cyanobacteria, aunque en muy poca abundancia, tiene correlación muy positiva con el potasio, el pH y otros (Rasuk et al., 2015).

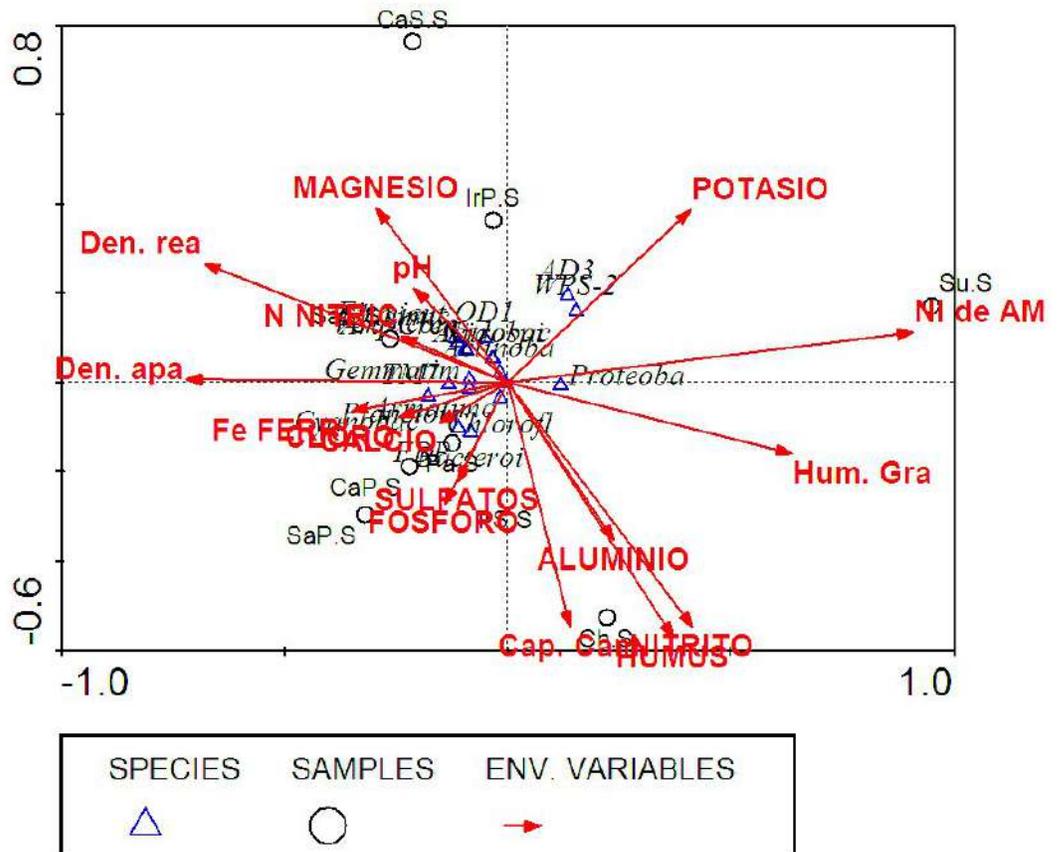
También se evidencia que los sitios de la zona central (Pa.P y Ca.P) tienen disimilitud con los sitios de la zona sur (Ir.P y Sa.P). Es decir las tendencias de los analitos son prácticamente contrarias. Algo similar ocurre entre los sitios de la zona norte, donde Choquenaira difiere de Sullcataca baja, se observa que los dos sitios están en cuadrantes diferentes. En este sentido los sitios de las zonas central y sur se muestran respectivamente en los mismos cuadrantes.



**Figura 24.** Gráfica de Correlación Canónica de las muestras puruma de las tres zonas agroproductivas. Las flechas rojas indican la tendencia de las variables ambientales.

En cuanto al análisis CCA con las muestras de suelos de cultivo (Figura 25), se ha mostrado una correlación muy diferente. En este caso no se observa correlaciones semejantes entre las zonas del altiplano, es decir que las correlaciones entre sitios de la misma zona son contrarias, así se observan muestras de diferentes zonas en un mismo cuadrante.

Esta diferencia aporta a los resultados de diversidad alfa, en que el manejo de la tierra cambia la estructura de las comunidades microbianas, las propiedades fisicoquímicas del suelo. Durante el descanso de los suelos, las estructuras de las comunidades bacterianas retomarían su estructura 'normal', teniendo en cuenta tanto las características fisicoquímicas como del tipo de suelo.



**Figura 25.** Gráfica de Correlación Canónica de las muestras de suelos de cultivo de las tres zonas agroproductivas. Las flechas rojas indican la tendencia de las variables ambientales.

Este manejo, que es particular de cada sitio, revela una estructura bacteriana particular y diferente, sobre todo se observa que tiene un efecto mucho más fuerte en las propiedades fisicoquímicas de los suelos, donde los cambios en las tendencias es evidente. Aunque la mayoría de las bacterias han permanecido en la parte central del gráfico indicando un equilibrio general, ciertamente se

observó cómo sus abundancias fueron diferentes entre los diferentes sitios y etapas de desarrollo de la quinua.

Esta investigación sugiere que las comunidades microbianas del suelo pueden desempeñar un papel importante en la agricultura. Sin embargo, este resultado no es universal; Por el contrario, reflexiona sobre la importancia relativa de las comunidades microbianas en comparación con los factores abióticos para aumentar la productividad de los cultivos, también resalta la simbiosis entre microorganismos y plantas.

El análisis metagenómico es una disciplina que ha empezado a tomar mucha relevancia para los estudios de microorganismos en diferentes tipos de ecosistemas, y a pesar de ser una ciencia reciente, ha tenido grandes avances tecnológicos particularmente en la bioinformática, en el cual se están empleando algoritmos que emplean inteligencia artificial y redes neuronales.

## **VII. Conclusiones.**

Se ha descrito el protocolo estándar del análisis y procesamiento metagenómico dirigido al gen 16S rRNA en comunidades microbianas en muestras de suelos de cultivo de quinua. Todas las muestras obtenidas fueron procesadas, a pesar de no contar con la totalidad de los puntos estudiados debido a dificultades técnicas del proyecto principal COSUDE, pero que representaron las 3 zonas agroproductivas; se tuvo que descartar una muestra de las 30 muestras procesadas, debido a los sesgos ya mencionados. La etapa de procesamiento dentro de Qiime y empleando como referencia Green Genes, una base de datos muy aplicada y muy extendida, se ha desarrollado según las factibilidades técnicas adquiridas para este estudio.

Se determinó la diversidad y abundancia de las comunidades presentes en suelos ambientales, empleando la secuenciación de nueva generación a diferentes niveles taxonómicos. Además, se determinó que la diversidad en la zona norte del altiplano es la más baja en comparación a las otras zonas, la mayor diversidad y riqueza de especies se hallan en centro y sur respectivamente. Esto concuerda con la hipótesis inicial de que el rendimiento en las zonas central y sur son favorecidas por la diversidad microbiana, aunque no se determinó el efecto del cambio climático de manera certera.

Adicionalmente se ha observado aumento en las abundancias de las secuencias que no fueron identificadas con Green Gens, lo que sugiere que existen nuevos representantes potenciales de filos bacterianos aun no descritos que pueden tener alguna relación tanto simbiótica como promotora en el desarrollo de la quinua.

Por otro lado, se observó de forma general que los suelos en descanso recuperan una mayor diversidad filogenética en comparación a los suelos de cultivo. Sin embargo, dentro de estos se ha observado el aumento de fitopatógenos, aunque se ha visto que en suelos de cultivo su menor abundancia puede ser resultado del manejo de los suelos incluyendo el uso de agroquímicos y similares. También se observó mayor

abundancia de fitopatogenos en la zona sur del altiplano, pero no de manera significativa en comparación las otras zonas teniendo en cuenta que esta zona posee mayores abundancias.

Se completo el análisis multivariante con las muestras de suelos puruma y con muestras de suelos de cultivo donde se observó que algunos parámetros fisicoquímicos como la concentración de nitrito, potasio y humus contribuyen a explicar cómo la comunidad bacteriana en Choquenaira seguida de Sullcataca baja y Cañaviri difieren más que entre los otros sitios. Este estudio no pretende dar los valores absolutos de la presencia de bacterias y de las correlaciones observadas, ya que este tipo de estudio es el primero en su tipo y se deben tomar una atención más detallada en los sesgos encontrados.

## VIII. Bibliografía

- ANAPQUI. (2017). Asociación Nacional de Productores de Quinoa. Retrieved from ANAPQUI - DO QUINUA REAL website:
- Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., & Wingett, S. (2012). FastQC. In.
- Angulo, M., Mollinedo, O., Claros, M., & Ortuño, N. (2013). Efecto de diferentes cepas provisionarias en plantas de quinoa (*Chenopodium quinoa* Willd.) en condiciones de invernadero. 185-255.
- Aragón Oraquine, O., Arias Irusta, B. O., Chacolla Arias, E., & Tito Villca, L. R. (2014). *Atlas de riesgo agropecuario y cambio climático para la soberanía alimentaria* (V. H. V. Mamani Ed. 1 ed.): Ministerio de Desarrollo Rural y Tierras.
- Aronesty, E. (2013). Comparison of Sequencing Utility Programs. *The Open Bioinformatics Journal*, 7(1), 1–8.
- BERNARDET, J.-F., SEGERS, P., VANCANNEYT, M., BERTHE, F., KERSTERS, K., & VANDAMME, P. (1996). Cutting a Gordian Knot: Emended Classification and Description of the Genus *Flavobacterium*, Emended Description of the Family Flavobacteriaceae, and Proposal of *Flavobacterium hydatis* nom. nov. (Basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *International Journal of Systematic and Evolutionary Microbiology*, 46(1), 128-148. doi:<https://doi.org/10.1099/00207713-46-1-128>
- Billings, S., Hirmas, D., Sullivan, P., Lehmeier, C., Bagchi, S., Min, K., . . . Richter, D. (2018). Loss of deep roots limits biogenic agents of soil development that are only partially restored by decades of forest regeneration. *Elem Sci Anth*, 6, 34. doi: 10.1525/elementa.287
- Bokulich, N. A., Subramanian, S., Faith, J. J., Gevers, D., Gordon, J. I., Knight, R., . . . Caporaso, J. G. (2012). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nature Methods*, 10(1), 57–59.
- Bourne, D. G., Riddles, P., Jones, G. J., Smith, W., & Blakeley, R. L. (2001). Characterisation of a gene cluster involved in bacterial degradation of the cyanobacterial toxin microcystin LR. *Environmental Toxicology*, 16(6), 523-534. doi:<https://doi.org/10.1002/tox.10013>
- Brewer, T. E., Handley, K. M., Carini, P., Gilbert, J. A., & Fierer, N. (2016). Genome reduction in an abundant and ubiquitous soil bacterium ‘*Candidatus Udaeobacter copiosus*’. *Nature Microbiology*, 2, 16198. doi:10.1038/nmicrobiol.2016.198
- Caporaso, J. G., Bittinger, K., Bushman, F. D., DeSantis, T. Z., Andersen, G. L., & Knight, R. (2009). PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics*, 26(2), 266–267.

- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., . . . et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7(5), 335–336.
- Cárdenas, J. E., Urquizo, O. V., Cárdenas, M. I., Fernández, R., & Vladimir, O. (2015). Degradación de suelos y tipología de productores para la sostenibilidad del cultivo de quinua (*Chenopodium quinoa* Wild) en el intersalar del departamento de Oruro, Bolivia. *APTHAPI*, 1, 58-68.
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., . . . Tiedje, J. M. (2014). Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Research*, 42, D633-D642. doi:10.1093/nar/gkt1244
- DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., . . . Andersen, G. L. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology*, 72(7), 5069–5072.
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19), 2460-2461.
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16), 2194–2200.
- Escalante, A. E., Barbolla, L. J., Ramírez-Barahona, S., & Eguiarte, L. E. (2014). The study of biodiversity in the era of massive sequencing. *Revista Mexicana de Biodiversidad*, 85(4), 1249-1264.
- Faith, D. P., & Baker, A. M. (2007). Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. *Evolutionary bioinformatics online*, 2, 121--128.
- Farag, I. (2015). In Silico Analysis of the Metabolic Potential and Niche Specialization of Candidate Phylum "Latescibacteria" (WS3). *PLoS ONE*, 10. doi:10.1371/journal.pone.0127499
- Gabriel, J., Pereira, R., Angulo, A., Magne, J., & Gandarillas, A. (2012). *LA MEJORA GENETICA DE PAPA EN BOLIVIA: LOGROS, LECCIONES*.
- Gallego, V., García, M. T., & Ventosa, A. (2006). *Pedobacter aquatilis* sp. nov., isolated from drinking water, and emended description of the genus *Pedobacter*. *International Journal of Systematic and Evolutionary Microbiology*, 56(8), 1853-1858. doi:https://doi.org/10.1099/ijs.0.64176-0
- Genin, D. (2006). Campos nativos de pastoreo y producción animal en la puna semiárida y árida andina. *Science et changements planétaires / Sécheresse*, 17, 265-274.
- Gilbert, J. A., Jansson, J. K., & Knight, R. (2014). The Earth Microbiome project: successes and aspirations. *BMC biology*, 12, 69. doi:10.1186/s12915-014-0069-1

- Glasel, J. A. (1995). Validity of nucleic acid purities monitored by 260nm/280nm absorbance ratios. *BioTechniques*, 18(1), 62-63.
- Gomez-Montano, L., Jumpponen, A., Gonzales, M. A., Cusicanqui, J., Valdivia, C., Motavalli, P. P., . . . Garrett, K. A. (2013). Do bacterial and fungal communities in soils of the Bolivian Altiplano change under shorter fallow periods? *Soil Biology and Biochemistry*, 65, 50-59.
- Gotelli, N. J., & Colwell, R. K. (2011). Estimating species richness. *Biological diversity: frontiers in measurement and assessment*, 12, 39--54.
- Hernández-León, R., Velázquez-Sepúlveda, I., Orozco-Mosqueda, M. C., & Santoyo, G. (2010). *Metagenómica de suelos: grandes desafíos y nuevas oportunidades biotecnológicas* (Vol. 79).
- Hiruma, K., Gerlach, N., Sacristan, S., Nakano, R. T., Hacquard, S., Kracher, B., . . . Schulze-Lefert, P. (2016). Root Endophyte *Colletotrichum tofieldiae* Confers Plant Fitness Benefits that Are Phosphate Status Dependent. *Cell*, 165(2), 464-474. doi: 10.1016/j.cell.2016.02.028
- Huang, W., Umbach, D. M., & Li, L. (2005). Accurate anchoring alignment of divergent sequences. *Bioinformatics*, 22(1), 29–34.
- Huberman, J. A. (1995). Importance of measuring nucleic acid absorbance at 240 nm as well as at 260 and 280 nm. *BioTechniques*, 18(4), 636.
- Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J., Chinwalla, A., . . . White, O. (2012). The Human Microbiome Project (HMP) Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* 486: 207-214. *Nature*, 486, 207-214. doi:10.1038/nature11234
- Illumina. 16S Metagenomic Sequencing Library Preparation. In Illumina (Ed.), (pp. 28).
- Illumina. Advantages of paired-end and single-read sequencing. Retrieved from <https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/paired-end-vs-single-read.html>
- Inc., I. Advantages of paired-end and single-read sequencing. Retrieved from <https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/paired-end-vs-single-read.html>
- J. F, S., & Russell, D. (2001). *Molecular Cloning: A Laboratory Manual (3-Volume Set)* (Vol. 1).
- Jenway. Troubleshooting Guide for Low and Micro Volume Measurements of Nucleic Acids with the Genova Plus/Nano. In (pp. 4): [bibby-scientific.com](http://bibby-scientific.com).
- Johnson, D., Yajie, L., & Okibe, N. (2008). "Bioshrouding" - A novel approach for securing reactive mineral tailings. *Biotechnology letters*, 30, 445-449. doi:10.1007/s10529-007-9574-4

- Juan Risi, W. R., Mauricio Pacheco. (2015). *Producción y mercado de la quinua en Bolivia*.
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2012). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, *41*(1), e1–e1.
- Koppen, W. (1936). Das geographische System der Klimat. *Handbuch der Klimatologie*, 46.
- Krieg, N., Staley, J., Brown, D., Hedlund, B., Paster, B., Ward, N., . . . Whitman, W. (2010). *Bergey's Manual® of Systematic Bacteriology: Volume Four The Bacteroidetes, Spirochaetes, Tenericutes (Mollicutes), Acidobacteria, Fibrobacteres, Fusobacteria, Dictyoglomi, Gemmatimonadetes, Lentisphaerae, Verrucomicrobia, Chlamydiae, and Planctomycetes*. doi:10.1007/978-0-387-68572-4
- Laja, G. (2005). *PLAN DE DESARROLLO MUNICIPAL LAJA*. SERVICIO DEPARTAMENTAL DE AUTONOMIAS DE LA PAZ Retrieved from <http://autonomias.gobernacionlapaz.com/sim/municipio/pdm/laja2001-2005.pdf>.
- Lozupone, C., & Knight, R. (2005). UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Applied and Environmental Microbiology*, *71*(12), 8228–8235.
- Lozupone, C. A., Hamady, M., Kelley, S. T., & Knight, R. (2007). Quantitative and Qualitative Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. *Applied and Environmental Microbiology*, *73*(5), 1576-1585. doi:doi:10.1128/AEM.01996-06
- Magurran, A. E. (1988). *Ecological Diversity and Its Measurement*.
- McDonald, D., Price, M. N., Goodrich, J., Nawrocki, E. P., DeSantis, T. Z., Probst, A., . . . Hugenholtz, P. (2011). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *The ISME Journal*, *6*(3), 610–618.
- Miranda, R. R. (2015). Precios de la Quinua Real en Challapata. Retrieved from <http://laquinua.blogspot.com/>
- Mocali, S., & Benedetti, A. (2010). Exploring research frontiers in microbiology: The challenge of metagenomics in soil microbiology. *Research in microbiology*, *161*, 497-505. doi:10.1016/j.resmic.2010.04.010
- Monasterio, M., Andressen, R., & Terceros, L. (2007). Regímenes climáticos del altiplano sur de Bolivia: una región afectada por la desertificación. *Revista geográfica venezolana, ISSN 1012-1617, Vol. 48, N° 1, 2007, pags. 11-32, 48*.
- Montesinos, E. A. E. (2016). Adaptación del cultivo de quinua (*Chenopodium quinoa willd*) al cambio climático en los Andes del Perú. *19*(37), 15-23.

- Montesinos, E. A. E., & Montesinos, E. A. E. (2017). Adaptación del cultivo de quinua (*Chenopodium quinoa willd*) al cambio climático en los Andes del Perú. *Revista del Instituto de Investigación de la Facultad de Ingeniería Geológica, Minera, Metalúrgica y Geográfica*, 19(37).
- Moreno Reséndez, A., Carda Mendoza, V., Reyes Carrillo, J. L., Vásquez Arroyo, J., & Cano Ríos, P. (2018). Rizobacterias promotoras del crecimiento vegetal: una alternativa de biofertilización para la agricultura sustentable. *Revista Colombiana de Biotecnología*, 20, 68-83.
- Nair, H. P., Vincent, H., & Bhat, S. G. (2014). Evaluation of five in situ lysis protocols for PCR amenable metagenomic DNA from mangrove soils. *Biotechnol Rep (Amst)*, 4, 134-138. doi:10.1016/j.btre.2014.09.008
- Noss, R. F. (1990). Indicators for Monitoring Biodiversity: A Hierarchical Approach. *Conservation Biology*, 4(4), 355–364.
- Ohn, J.-E., Ten, L. N., Kim, B.-O., Cho, Y.-J., & Jung, H.-Y. (2018). *Hymenobacter rufus* sp. nov., a bacterium isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, 68(9), 2983-2989. doi:https://doi.org/10.1099/ijsem.0.002934
- Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., & Hugenholtz, P. (2018). A proposal for a standardized bacterial taxonomy based on genome phylogeny. *bioRxiv*.
- Patacamaya, G. (2010). *PLAN DE DESARROLLO MUNICIPAL DE PATACAMAYA*. Servicio Departamental de Autonomías de La Paz Retrieved from <http://autonomias.gobernacionlapaz.com/sim/municipio/pdm/patacamaya2006-2010.pdf>.
- Peet, R. K. (1974). The Measurement of Species Diversity. *Annual Review of Ecology and Systematics*, 5(1), 285–307.
- Pérez Álvarez, S., Coto Arbelo, O., Echemendía Pérez, M., & Ávila Quezada, G. (2015). *Pseudomonas fluorescens* Migula, ¿control biológico o patógeno?: *Pseudomonas fluorescens*, biological control or pathogen? *Revista de Protección Vegetal*, 30, 225-234.
- Polz, M. F., & Cavanaugh, C. M. (1998). Bias in template-to-product ratios in multitemplate PCR. *Appl Environ Microbiol*, 64(10), 3724-3730.
- Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, 5(3), e9490.
- PROGRAMA CONTROL DE CALIDAD DE MUESTRAS*. (2010). Retrieved from <http://www.bancoadn.org/docs/programa-control-calidad-muestras.pdf>
- Pylro, V. S., Roesch, L. F. W., Ortega, J. M., do Amaral, A. M., Tótoia, M. R., Hirsch, P. R., . . . Committee, T. B. M. P. O. (2014). Brazilian Microbiome Project: Revealing

- the Unexplored Microbial Diversity—Challenges and Prospects. *Microbial Ecology*, 67, 237-241. doi:10.1007/s00248-013-0302-4
- Ramírez Muñoz, M., & González Pazmiño, J. R. (2015). Caracterización del microbioma Bacteriano de la concha negra (*Anadara tuberculosa*, Sowerly, 1833) en el santuario de los manglares de Tumbes. (Bachelor Thesis), Universidad Estatal Península de Santa Elena, Retrieved from <http://repositorio.upse.edu.ec/handle/46000/2118>
- Rasuk, M., Fernández, A., Kurth, D., Contreras, M., Novoa, F. F., Poiré, D., & Farías, M. (2015). Bacterial Diversity in Microbial Mats and Sediments from the Atacama Desert. *Microbial ecology*, 71. doi:10.1007/s00248-015-0649-9
- Rodríguez-Herrera, R., Aguilar-González, C., Ayala-Labarríos, L., Rocha-Revilla, J., Padilla-García, V., & Espinosa-Hernández, T. (2009). Detección de microorganismos mediante métodos moleculares. *AQM Revista de Divulgación Científica*, 1(1).
- Rodríguez-Santiago, B., & Armengol, L. (2012). Tecnologías de secuenciación de nueva generación en diagnóstico genético pre- y postnatal. *Diagnóstico Prenatal*, 23(2), 56-66. doi:10.1016/j.diapre.2012.02.001
- Ruiz, K. B., Biondi, S., Oses, R., Acuña-Rodríguez, I. S., Antognoni, F., Martínez-Mosqueira, E. A., . . . et al. (2013). Quinoa biodiversity and sustainability for food security under climate change. A review. *Agronomy for Sustainable Development*, 34(2), 349–359.
- Salinas\_Garci\_Mendoza, G. (2002). *Ajuste Plan de Desarrollo Municipal Salinas Garcí Mendoza*. Retrieved from [http://vpc.planificacion.gob.bo/uploads/PDM\\_S/04\\_ORURO/040801%20Salinas%20de%20Garc%C3%AC%20Mendoza.pdf](http://vpc.planificacion.gob.bo/uploads/PDM_S/04_ORURO/040801%20Salinas%20de%20Garc%C3%AC%20Mendoza.pdf).
- Sarco S., T. C., Aliaga S., Condori J., Garcia M., Geets M., Mathjis E., RaesD., Cusicanqui J., Mamani F., Miranda R., Saavedra C., Yucra E. (2011). Dinámica de comercialización de quinoa (*Chenopodium quinoa* Willd) en tres ferias del altiplano Boliviano. *Compendio de Trabajos de Investigación Proyecto Quinagua - UMSA*, 138-148.
- Schirmer, M., D'Amore, R., Ijaz, U. Z., Hall, N., & Quince, C. (2016). Illumina error profiles: resolving fine-scale variation in metagenomic sequencing data. *BMC Bioinformatics*, 17(1), 125. doi:10.1186/s12859-016-0976-y
- Schwartz, E., Van Horn, D. J., Buelow, H. N., Okie, J. G., Gooseff, M. N., Barrett, J. E., & Takacs-Vesbach, C. D. (2014). Characterization of growing bacterial populations in McMurdo Dry Valley soils through stable isotope probing with 18O-water. *FEMS Microbiology Ecology*, 89(2), 415-425. doi:10.1111/1574-6941.12349
- SempÈrÈ, T., HÈrail, G., Baby, P., Marocco, R., Oller, J., & Barrios, L. (1991). Elaltiplano boliviano - una provincia de cuencas intramontanas de antepals relacionadas con el acortamiento cortical enla regiÛn del oroclino boliviano.

- Sica\_Sica, G. (2010). *Ajuste Plan de Desarrollo Municipal Sica Sica*. SERVICIO DEPARTAMENTAL DE AUTONOMIAS DE LA PAZ Retrieved from <http://autonomias.gobernacionlapaz.com/sim/municipio/pdm/sica-sica2006-2010.pdf>.
- Simpson, E. H. (1949). Measurement of Diversity. *Nature*, *163*(4148), 688–688.
- Tahon, G., Tytgat, B., Lebbe, L., Carlier, A., & Willems, A. (2018). *Abditobacterium utsteinense* sp. nov., the first cultivated member of candidate phylum FBP, isolated from ice-free Antarctic soil samples. *Systematic and Applied Microbiology*, *41*. doi: 10.1016/j.syapm.2018.01.009
- Techtmann, S. M., & Hazen, T. C. (2016). Metagenomic applications in environmental monitoring and bioremediation. *J Ind Microbiol Biotechnol*, *43*(10), 1345-1354. doi: 10.1007/s10295-016-1809-8
- Vázquez-Baeza, Y., Pirrung, M., Gonzalez, A., & Knight, R. (2013). EMPeror: a tool for visualizing high-throughput microbial community data. *GigaScience*, *2*(1).
- Wang, Y., Gao, Z.-M., Li, J.-T., Bougouffa, S., Tian, R. M., Bajic, V. B., & Qian, P.-Y. (2016). Draft genome of an Aerophobetes bacterium reveals a facultative lifestyle in deep-sea anaerobic sediments. *Science Bulletin*, *61*, 1176-1186. doi:10.1007/s11434-016-1135-6
- Wikipedia. (2019). Metagenómica --- Wikipedia, La enciclopedia libre. Retrieved from <https://es.wikipedia.org/w/index.php?title=Metagen%C3%B3mica&oldid=114524562>
- Yoon, J., Jang, J.-H., & Kasai, H. (2013). *Algisphaera agarilytica* gen. nov., sp nov., a novel representative of the class Phycisphaerae within the phylum Planctomycetes isolated from a marine alga. *Antonie van Leeuwenhoek*, *105*. doi:10.1007/s10482-013-0076-1
- Zamora, A., Malaver, N., & Ramos, J. (2012). ANÁLISIS FUNCIONAL DE MICROORGANISMOS: UN ESTIMADOR DE DIVERSIDAD Y ESTRUCTURA COMUNITARIA (Vol. 32).
- Zhou, Z., Jiang, F., Wang, S., Peng, F., Dai, J., Li, W., & Fang, C. (2012). *Pedobacter arcticus* sp. nov., a facultative psychrophile isolated from Arctic soil, and emended descriptions of the genus *Pedobacter*, *Pedobacter heparinus*, *Pedobacter daechungensis*, *Pedobacter terricola*, *Pedobacter glucosidilyticus* and *Pedobacter lentus*. *International Journal of Systematic and Evolutionary Microbiology*, *62*(Pt\_8), 1963-1969. doi:<https://doi.org/10.1099/ijs.0.031104-0>

## **Anexos.**

### **Anexo 1.** Protocolo de extracción de ADN mediante kit comercial.

La muestra de 0.25g fue transferida a un tubo de lisis que contenía una solución acuosa de tiocianato de guanidina y partículas de granate el cual debe ser agitado en vortex para dispersar la muestra en la solución. Se agregó 60 µl de solución de lisis que contiene Dodecilsulfato sódico (SDS) y otros agentes de disrupción necesarios para completar la lisis celular, posteriormente se invirtió repetidamente por 10-12 minutos y luego los tubos se centrifugaron a 10000 g por 30 segundos.

Se transfirió el sobrenadante a un tubo nuevo, se agregó 200 µl de la solución precipitadora IRT (Inhibitor Removal Technology®) que contiene un reactivo que precipita material orgánico e inorgánico sin ADN que incluye sustancias húmicas, desechos celulares y proteínas, se incubó a 4° C durante 5 minutos y se centrifugó a 10000 g por un minuto.

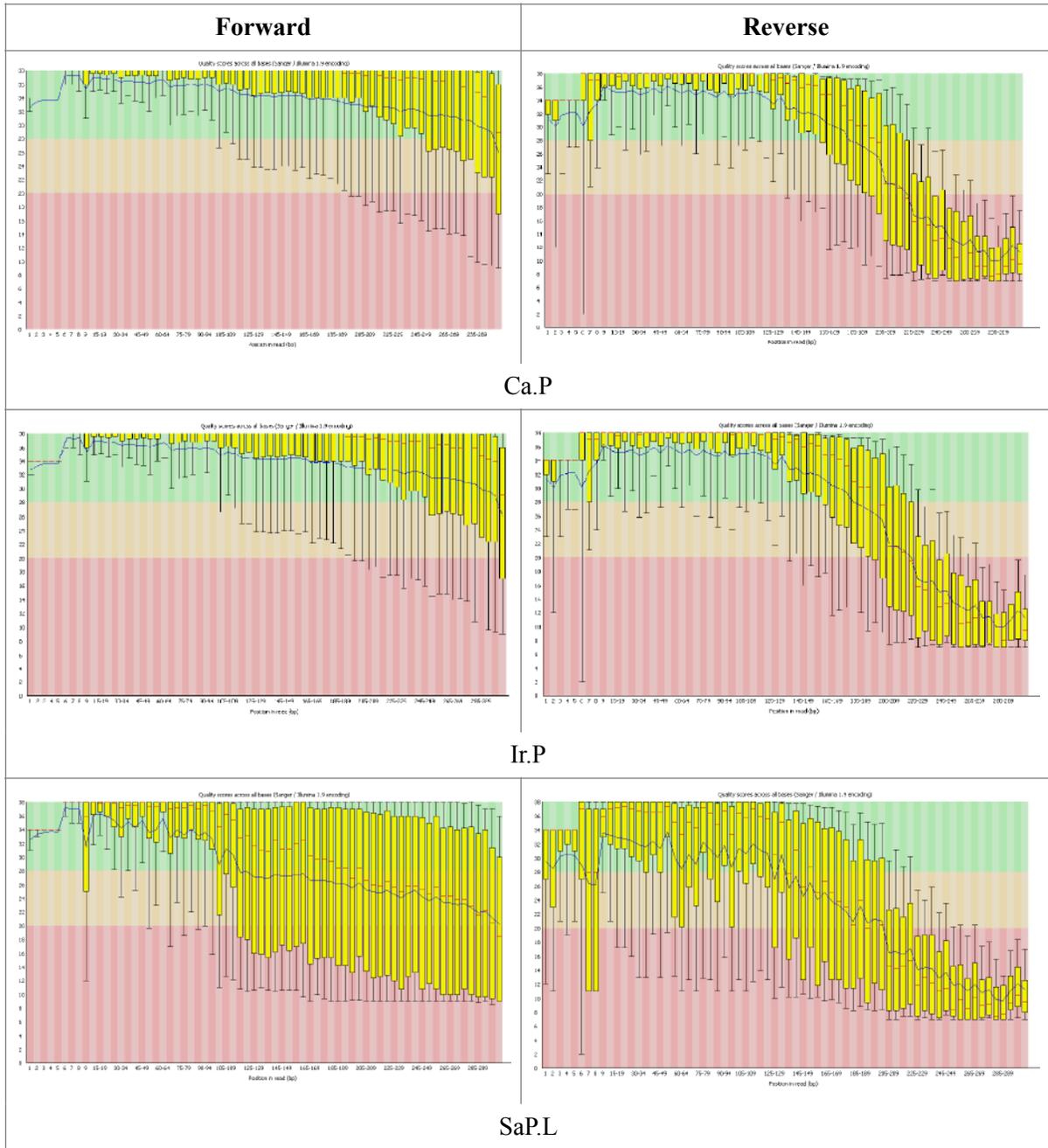
Nuevamente se transfirió hasta 600 µl del sobrenadante a un tubo nuevo y se agregó 200 µl de la solución IRT y se agitó en vortex brevemente, luego se incubó a 4° C durante 5 minutos y se centrifugo a 10000 g por un minuto.

Se transfirió hasta 750 µl del sobrenadante a un tubo nuevo, se agregó hasta 1200 µl de solución salina de alta concentración hasta llenar el tubo, esto permitirá que el ADN se una a la membrana de sílice de la columna-filtro dejando pasar cualquier material orgánico o inorgánico que aun pueda estar presente. Así, se cargó 675 µl de esta en el tubo de centrifugado y se procedió a centrifugar a 10000 g por un minuto. Se desechó el filtrado y se adicionó 675 µl del sobrenadante al tubo con la columna-filtro y se repitió el centrifugado desechando el filtrado y cargando el resto del sobrenadante para centrifugar por ultima vez.

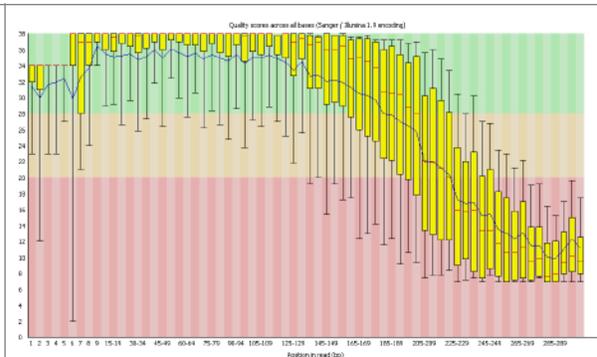
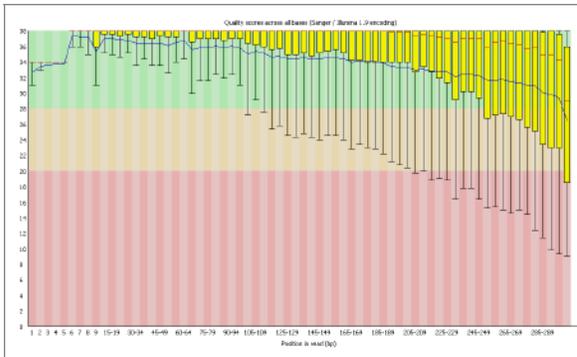
El ADN unido a la membrana de sílice fue lavado agregando 500 µl de solución de lavado a base de etanol, luego se centrifugó a 10000 g por 30 segundos permitiendo eliminar la sal residual, el ácido húmico y otros contaminantes, al tiempo que permite que el ADN permanezca unido a la membrana de sílice. Posteriormente se descartó el filtrado y se centrifugó una vez más a 10000 g por un minuto, para remover la solución salina residual.

Por ultimo se transfirió la columna-filtro a un tubo nuevo y se agregó justo al centro de la membrana de sílice 100 µL de tampón de elución (10mM Tris) para liberar el ADN de la membrana centrifugando a 10000 g durante 30 segundos, se descartó la columna-filtro dejando el tubo con el ADN precipitado para finalmente ser almacenado a 4° C.

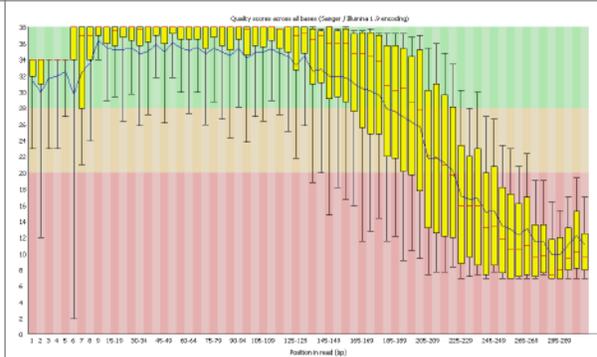
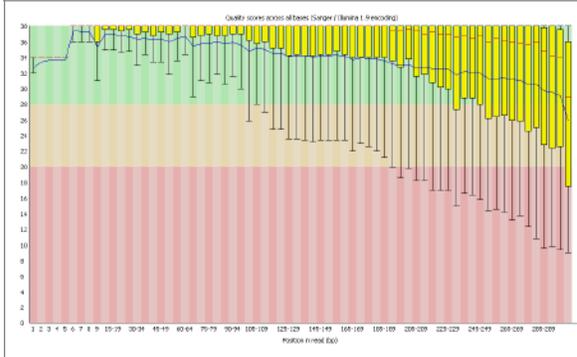
**Anexo 2.** Distribución de la calidad de las bases en las lecturas PairedEnd (forward y reverse) por cada muestra procesada.



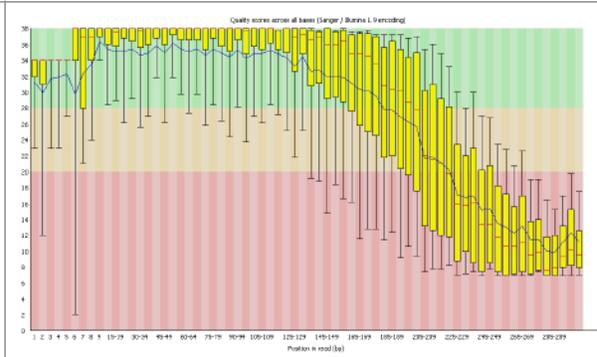
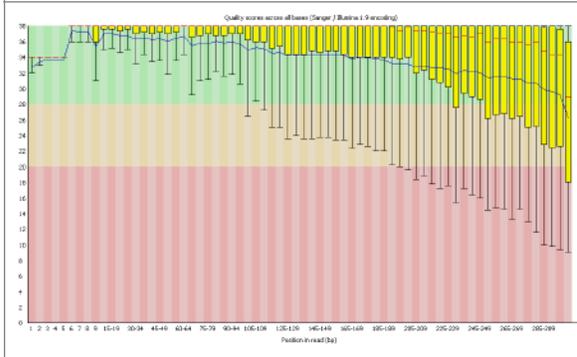




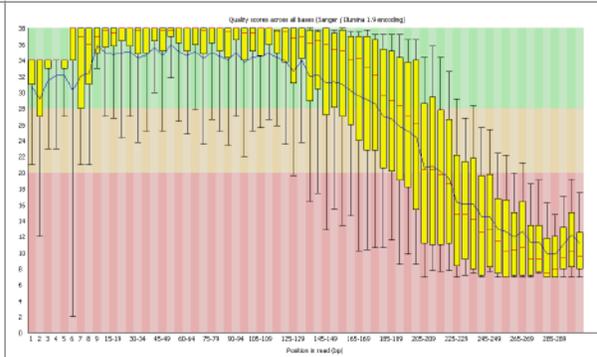
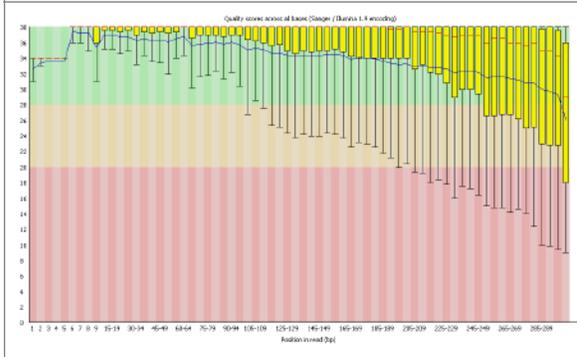
Su.L



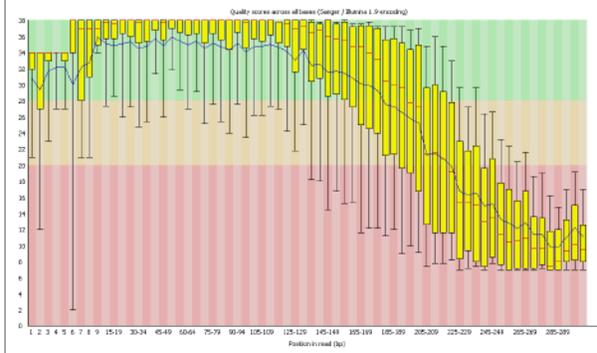
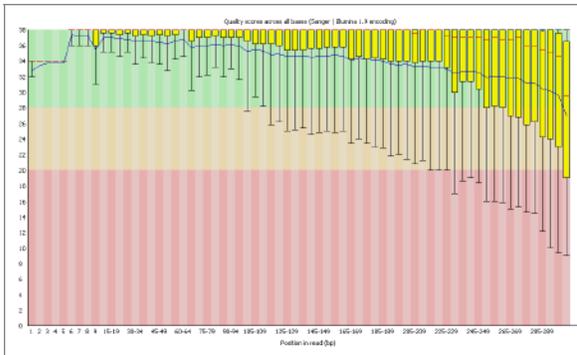
Ch.L



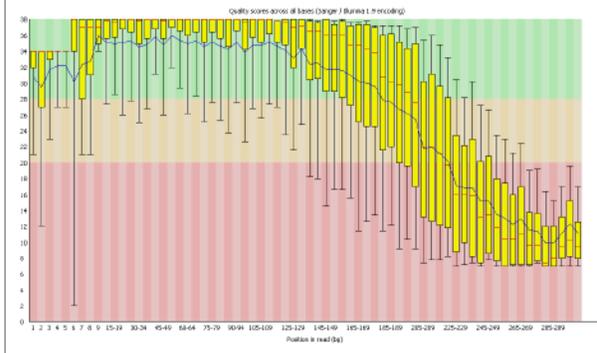
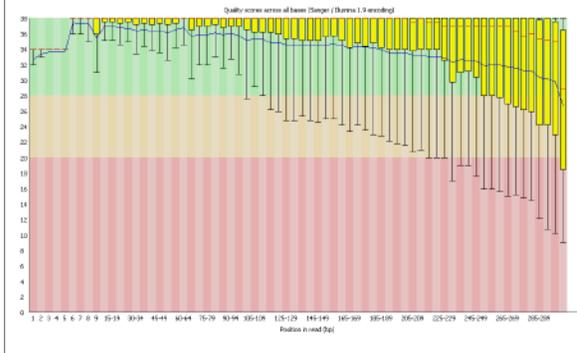
Pa.C



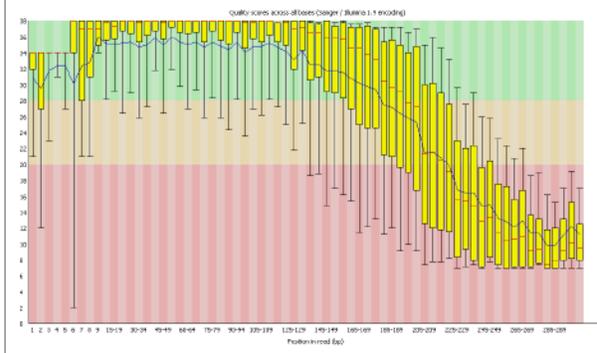
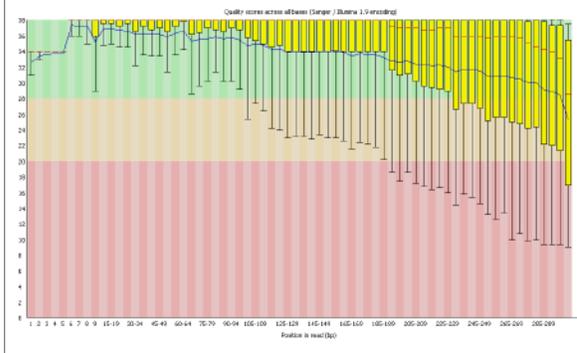
Ca.P



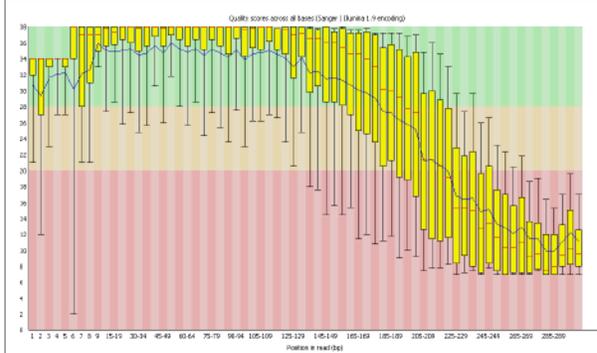
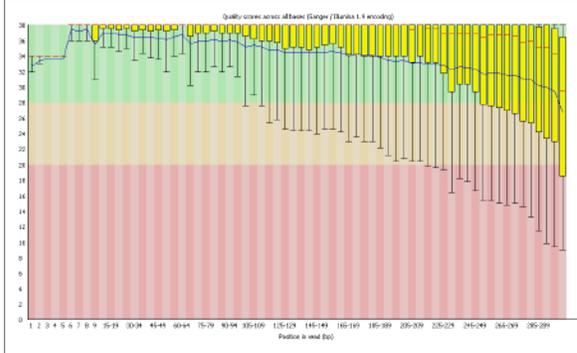
CaS.C



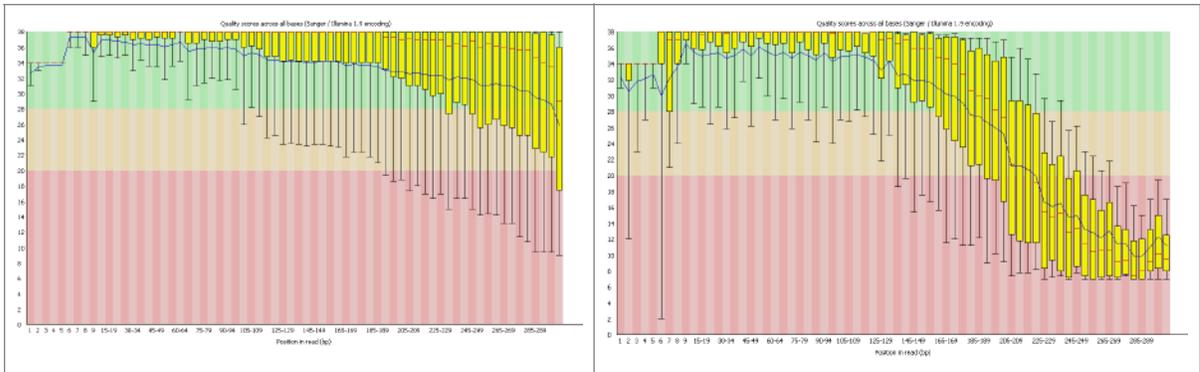
Su.C



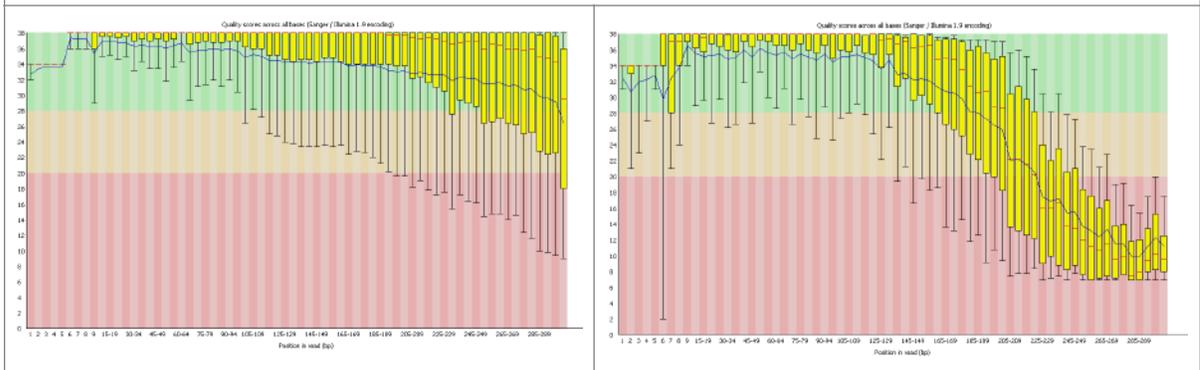
Ch.C



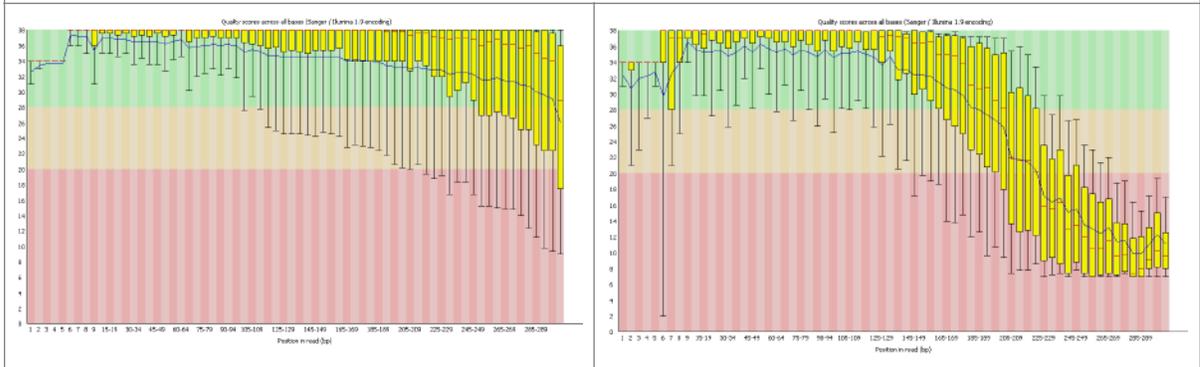
Sa.P.C



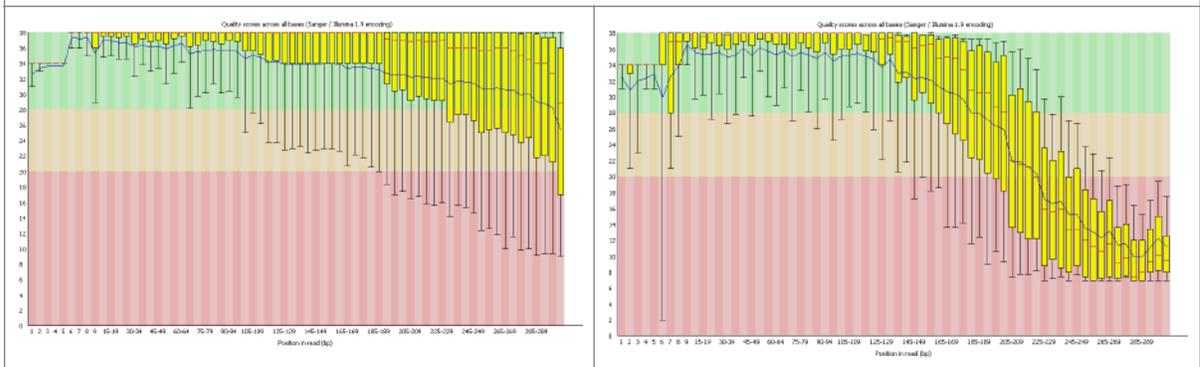
Ir.S.C



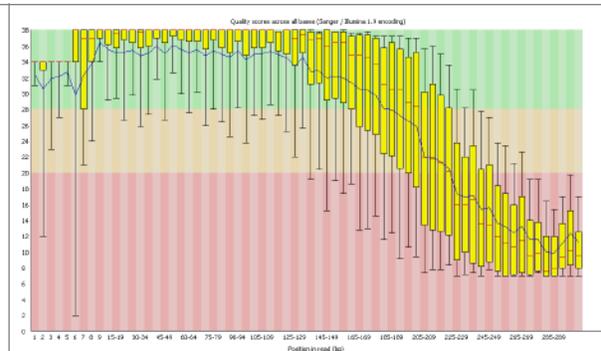
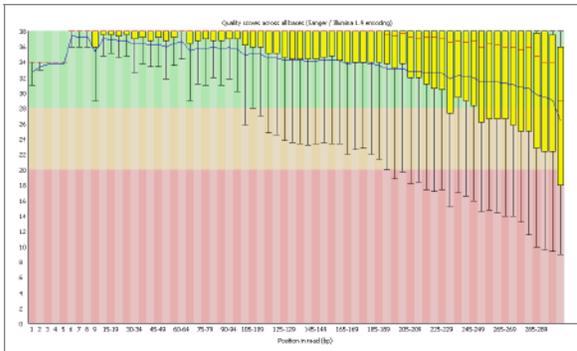
Sa.S.C



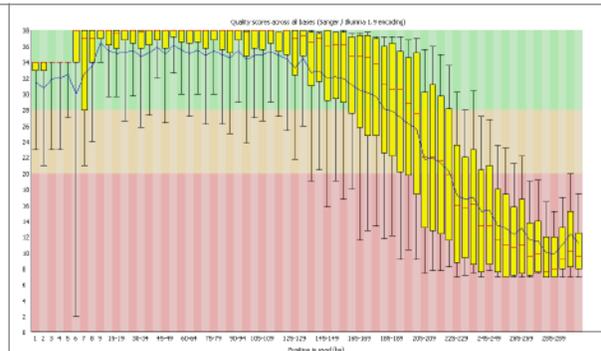
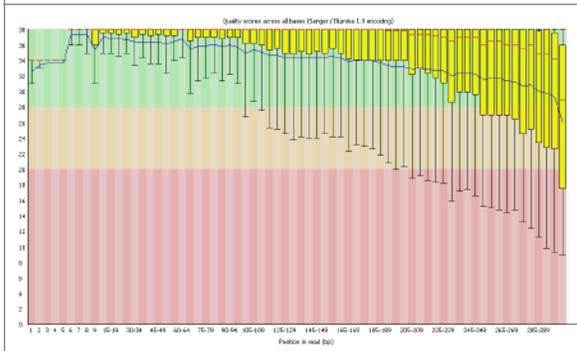
Ir.P.C



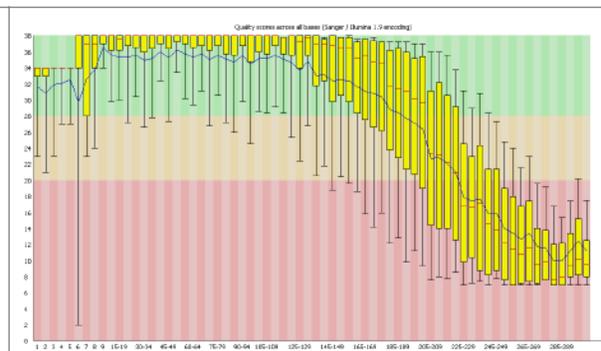
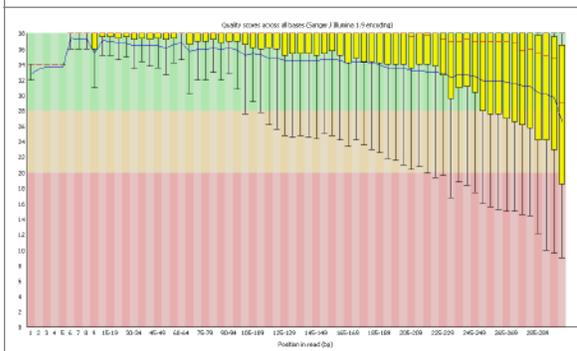
Sa.S.S



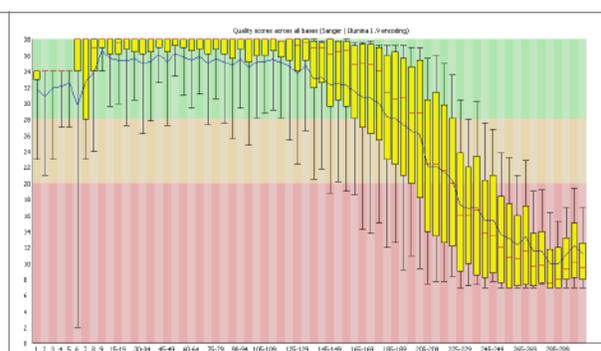
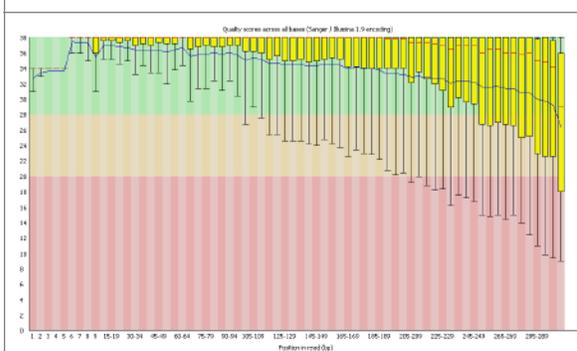
IrS.S



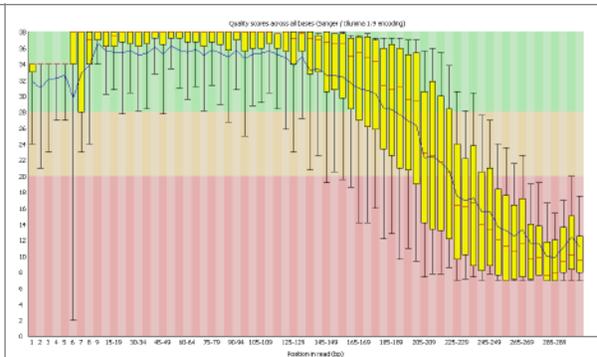
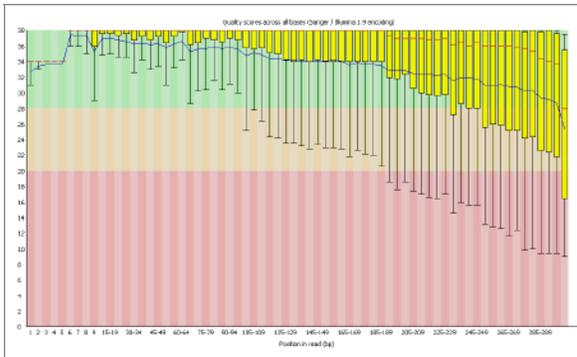
SaP.S



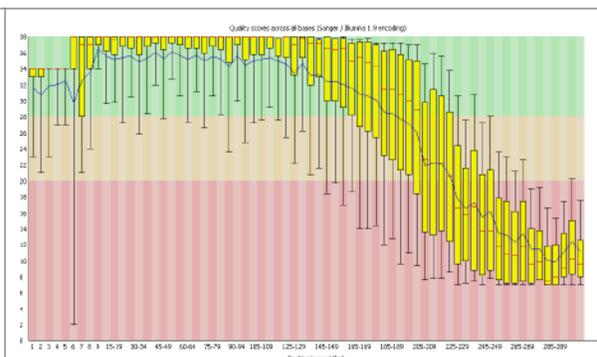
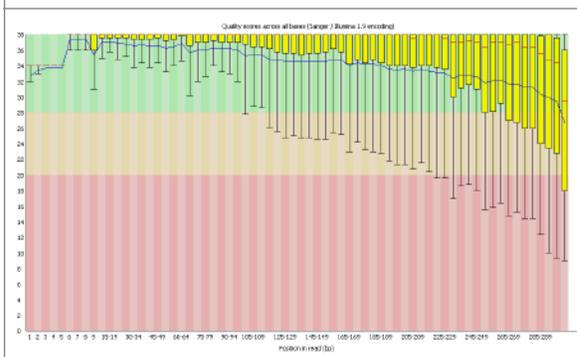
IrP.S



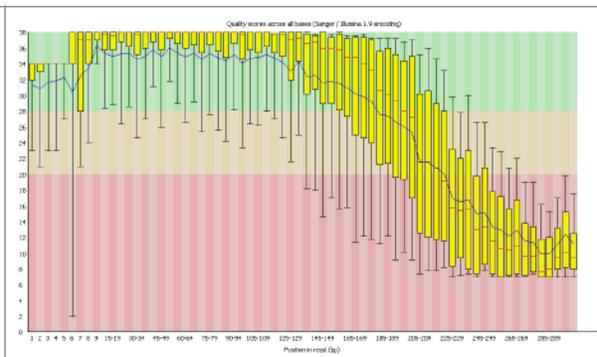
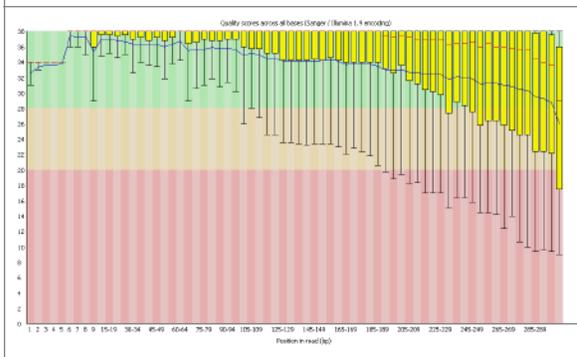
CaS.S



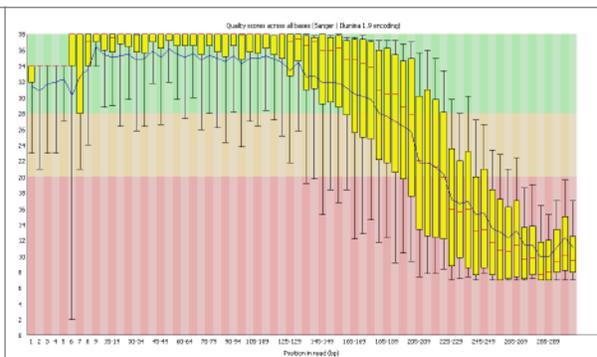
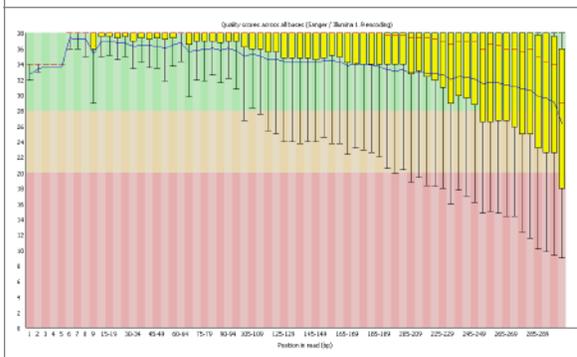
Su.S



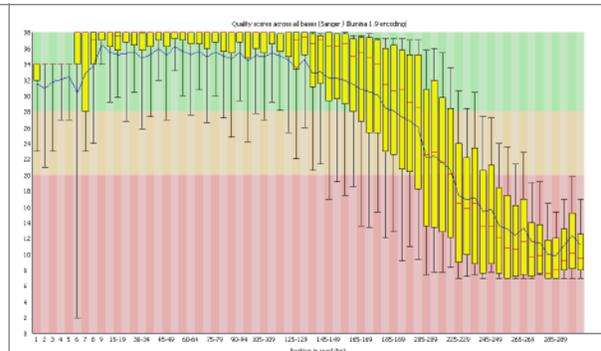
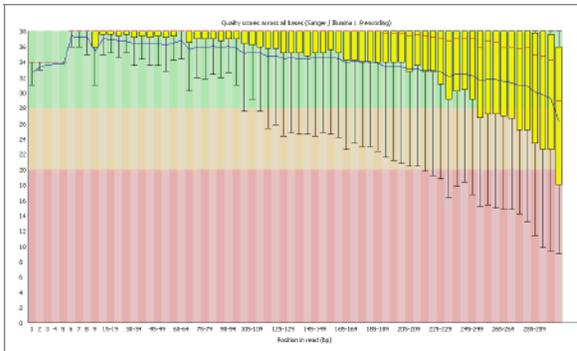
Ca.P.S



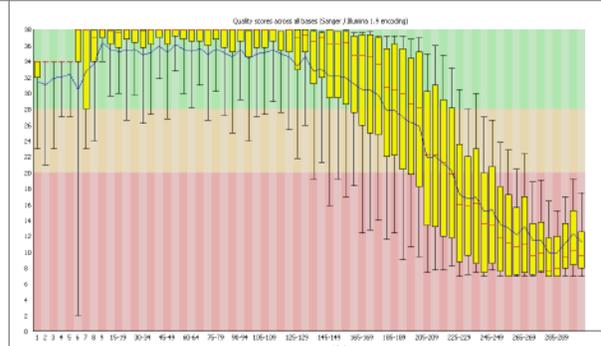
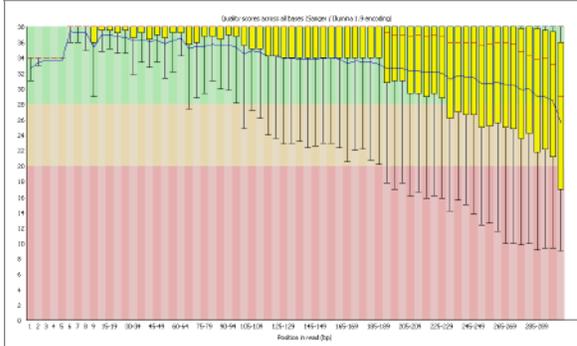
Pa.S



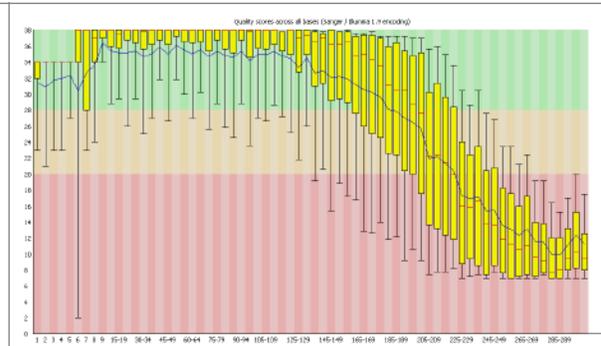
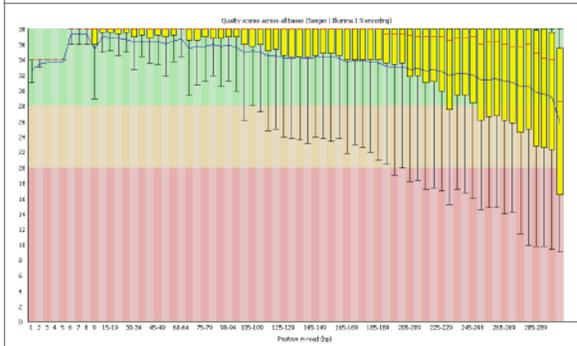
Ch.S



Pa.P

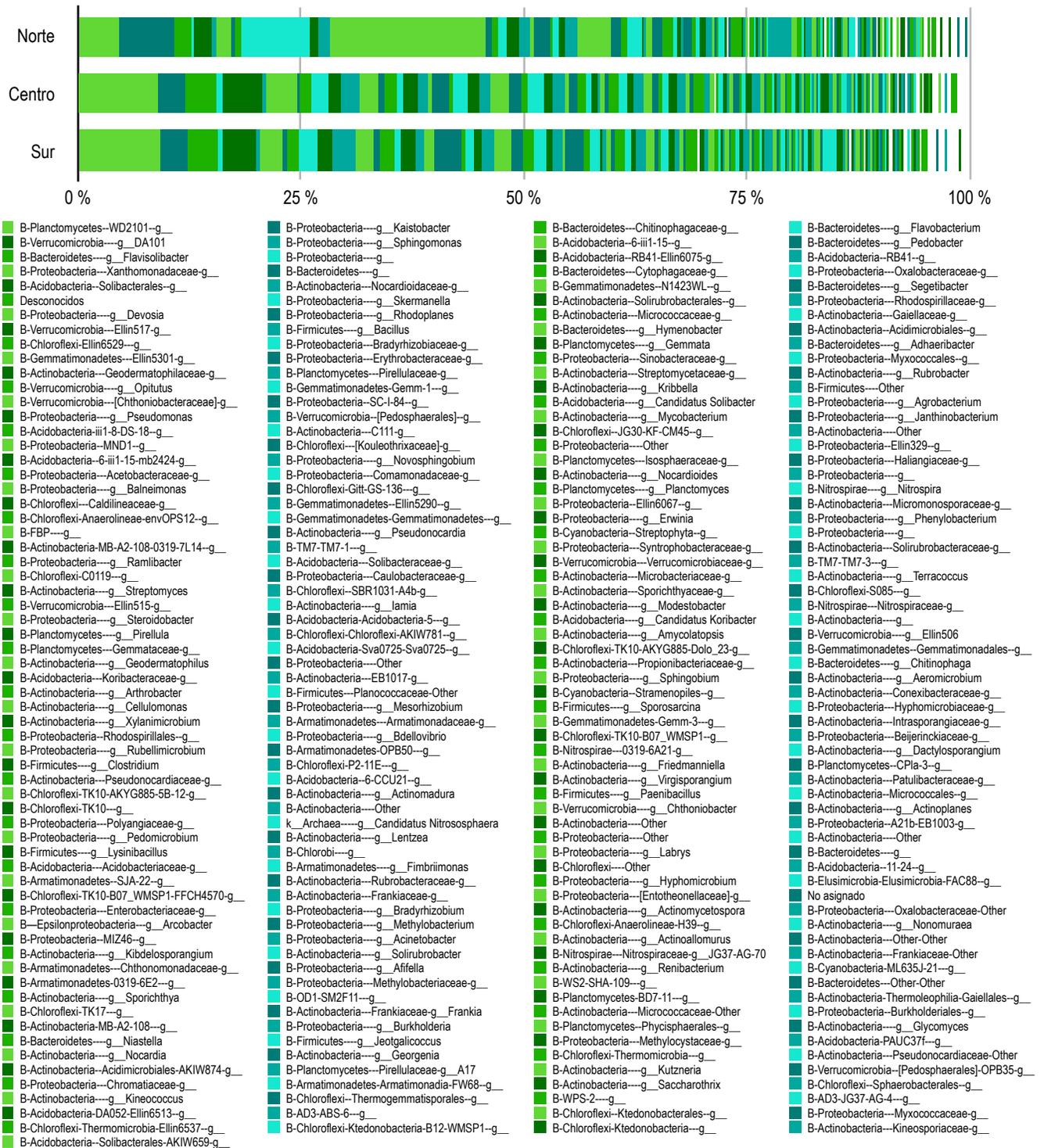


Ch.P

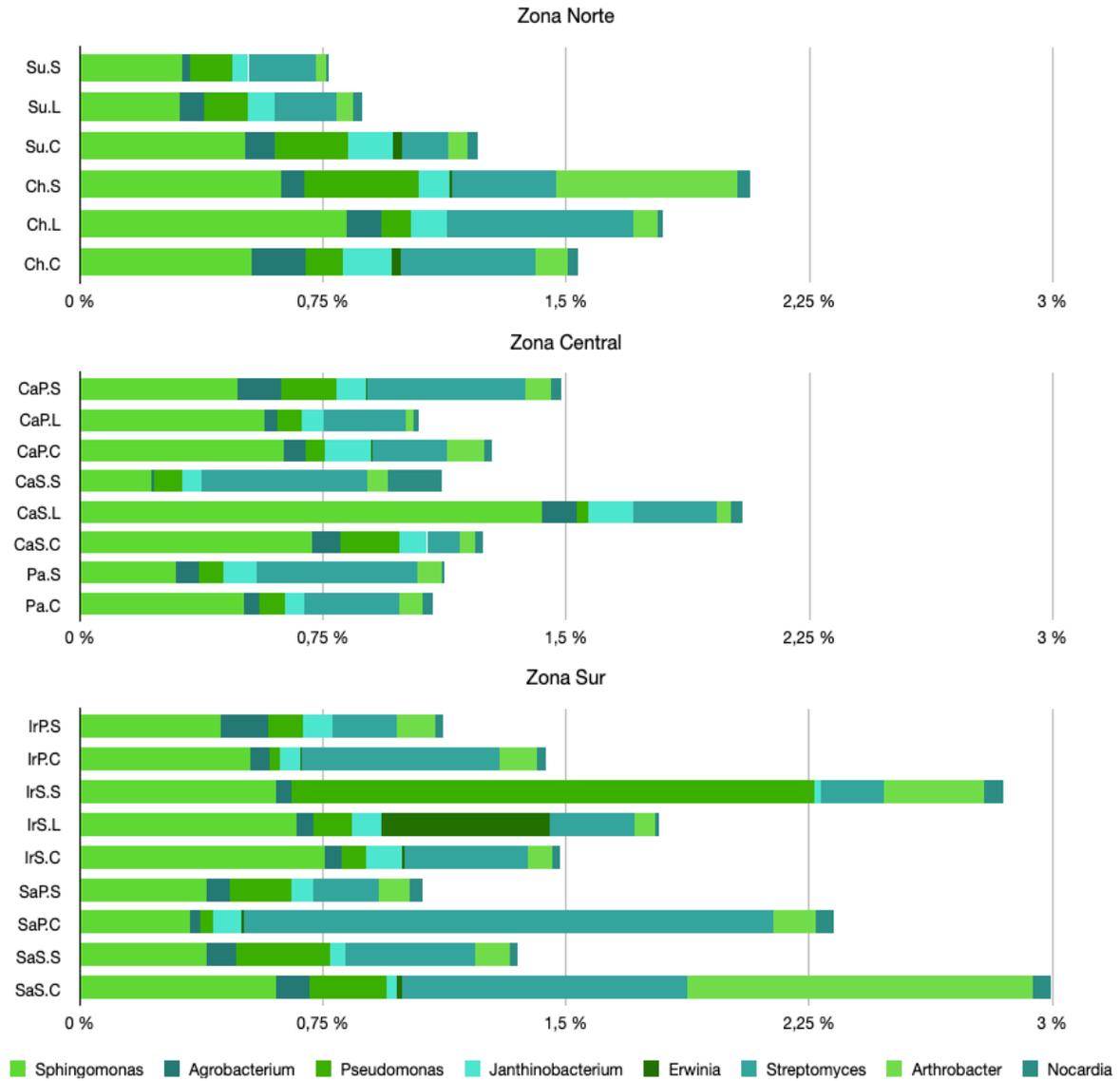


CaS.L

### Anexo 3. Abundancias de los géneros más abundantes en suelos de cultivo por zona agroproductiva.



**Anexo 4.** Géneros bacterianos potencialmente fitopatógenos presentes en los suelos de cada zona agroproductiva por fase de crecimiento de la quinua (.S: pre-siembra, .L: fase lechosa, .C: pre-cosecha).



**Anexo 5. Índices de diversidad alfa de las 30 muestras procesadas.**

<b>Localidad</b>	<b>shannon</b>	<b>PD_whole_tree</b>	<b>chaol</b>	<b>observed_otus</b>
Cañaviri_P	12,07	806,22	26278,57	17609
Irpani_P	12,07	781,38	21579,11	16851
Saitoco_bajo_FL	6,85	26,26	956,11	136
Irpani_alto_FL	11,71	691,09	26864,95	16319
Cañaviri_bajo_FL	11,78	764,06	24691,93	17284
Saitoco_P	11,93	611,83	17504,87	11664
Laja_P	11,59	609,15	19966,67	12542
Laja_FL	11,87	618,92	21598,43	12916
Choquenaira_FL	11,67	538,94	19241,26	11077
Patacamaya_PC	12,37	842,62	25534,56	17278
Cañaviri_bajo_PC	11,70	732,60	27246,95	16661
Cañaviri_alto_PC	11,91	725,11	25060,60	15532
Laja_PC	12,16	755,97	28857,61	16890
Choquenaira_PC	11,69	701,47	26193,47	15676
Saitoco_bajo_PC	11,36	520,58	16768,43	10755
Irpani_alto_PC	12,01	641,16	23524,68	14394
Saitoco_alto_PC	11,67	624,09	18756,89	13252
Irpani_bajo_PC	12,05	650,18	23857,46	14374
Saitoco_alto_PS	11,34	543,07	19861,79	11248
Irpani_alto_PS	11,91	759,24	20785,74	15234
Saitoco_bajo_PS	11,96	797,45	25017,06	16999
Irpani_bajo_PS	12,22	800,80	27446,81	18428
Cañaviri_alto_PS	11,29	706,65	21632,66	14734
Laja_PS	9,51	545,95	17974,25	12331
Cañaviri_bajo_PS	11,94	778,55	28959,44	17081
Patacamaya_PS	11,46	705,36	25182,29	15496
Choquenaira_PS	11,81	750,47	23550,50	16157
Patacamaya_P	11,59	692,76	25008,42	15566
Choquenaira_P	12,03	732,84	20492,60	14526
Cañaviri_alto_FL	11,94	750,99	25861,55	17140



